Ulrich Schmidt

Professionelle Videotechnik

Grundlagen, Filmtechnik, Fernsehtechnik, Geräte- und Studiotechnik in SD, HD, DI, 3D

6. Auflage

Springer Vieweg

Professionelle Videotechnik

Ulrich Schmidt

Professionelle Videotechnik

Grundlagen, Filmtechnik, Fernsehtechnik, Geräte- und Studiotechnik in SD, HD, DI, 3D

6. Auflage

Springer Vieweg

Prof. Dr. Ulrich Schmidt
Hochschule für Angewandte Wissenschaften Hamburg
Department Medientechnik
Hamburg, Deutschland

ISBN 978-3-642-38991-7 ISBN 978-3-642-38992-4 (eBook)
DOI 10.1007/978-3-642-38992-4

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Springer Vieweg
© Springer-Verlag Berlin Heidelberg 2000, 2003, 2004, 2005, 2009, 2013
Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Springer Vieweg ist eine Marke von Springer DE. Springer DE ist Teil der Fachverlagsgruppe Springer Science+Business Media.
www.springer-vieweg.de

Vorwort

Die Videotechnik ist gegenwärtig von einem tiefgreifenden Übergang von analoger zu digitaler Technik betroffen. Der Wandel ist ähnlich revolutionär wie die Veränderung, die vor einigen Jahren bei den Printmedien stattfand. Heute werden fast alle Bearbeitungsschritte der Druckvorstufe mit Standardcomputern durchgeführt, in ähnlicher Weise halten auch im Videobereich Digitalsignale und Computer Einzug. Die Signale verlieren als Digitaldaten ihre videotechnischen Spezifika, sie werden auf Standardfestplatten gespeichert und können auch in digitaler Form bis zum Konsumenten gebracht werden. Dabei ist es möglich, die Distributionswege der analogen Fernsehausstrahlung, Kabel, Satellit und terrestrische Sendewege für die Digitalübertragung ebenso zu nutzen, wie schnelle Computer-Datennetze. Die moderne Videotechnik ist durch eine enge Verknüpfung mit den Bereichen Computertechnik und Telekommunikation gekennzeichnet.

Das vorliegende Buch dient der Darstellung der Videotechnik in dieser Zeit des Umbruchs. Die Aspekte der aktuellen Videotechnik werden hier umfassend behandelt: auf analoger ebenso wie auf digitaler Ebene.

Das Buch ist sowohl für Studierende der Video- und Medientechnik als auch als Nachschlagewerk gedacht. Es basiert auf dem Buch, das im Jahre 1996 unter dem Titel »Digitale Videotechnik« im Franzis-Verlag erschienen ist. Es wurde völlig neu konzipiert, erheblich erweitert und trägt heute den passenderen Titel »Professionelle Videotechnik«.

Nach wie vor ist das Ziel, den gesamten Bereich der professionellen Videotechnik zu erfassen. Dazu werden zunächst die analogen und digitalen Videosignalformen, bis hin zu einer erweiterten Darstellung der MPEG-Codierung, ausführlich erörtert. Daran schließt sich nun ein eigenes Kapitel über Signalübertragungsformen an. Hier ist ein größerer Abschnitt über das digitale Fernsehen (DVB) hinzugekommen sowie eine Betrachtung zu neuen Distributionswegen, deren Entwicklung stark vom Internet geprägt ist.

Die beiden folgenden Kapitel behandeln die Bildaufnahme- und Wiedergabesysteme. Bei ersterem ist eine breitere Darstellung der Filmabtaster und der Bildwandler aufgenommen worden. In letzterem nimmt das Thema Plasmadisplays nun breiteren Raum ein. Im darauf folgenden Kapitel über die Aufzeichnungsverfahren sind die neuen MAZ-Formate, wie Betacam SX, D8, D9 etc. hinzugekommen sowie ein Abschnitt über die DVD. Im Kapitel Videosignalbearbeitung ist der Themenbereich der nichtlinearen Editingsysteme ausgeweitet worden, der Bereich Studiosysteme ist dagegen nicht mehr enthalten. Dieses Thema nimmt nun erheblich mehr Umfang ein und befindet

sich in einem eigenen Kapitel. Hier werden zunächst Betrachtungen zum Einsatz datenreduzierter Signale im Produktionsbereich vorgestellt. Daran schließt sich das Thema Netzwerke für Videostudios an. Es folgt die Darstellung der Postproduktions- und Produktionsbereiche, von Studios über Newsroom-Systeme bis zu SNG- und Ü-Wagen und zum Schluss eine ausführliche Darstellung des virtuellen Studios.

Aufgrund der thematischen Breite können nicht alle Bereiche erschöpfend behandelt werden, es sollte aber möglich sein, alle Aspekte im Gesamtkontext der Videotechnik einordnen zu können.

Mein Dank für die Unterstützung bei der Erstellung der Buches gilt Jens Bohlmann und Roland Greule für die Klärung nachrichten- und lichttechnischer Fragen, Torsten Höner und Anette Naumann für ihre Hilfe bei der Bildbearbeitung und Korrektur sowie Karin Holzapfel, Walter Folle und Tilmann Kloos für die Erstellung von Fotografien. Für die gute Zusammenarbeit mit dem Springer-Verlag bedanke ich mich bei Herrn Lehnert und Frau Cuneus.

Bremen, im März 2000 Prof. Dr. U. Schmidt

Vorwort zur 6. Auflage

Der stetige Wandel der Medien- und Videotechnik ist im Jahre 2013 davon geprägt, dass SDTV-Systeme nun auf breiter Front von HD-Systemen abgelöst sind und diese wiederum eine enge Verbindung zur digitalen Filmtechnik eingehen. Im Zuge dessen steigt das Interesse an noch höher aufgelösten Bildern, die unter dem Stichwort 4k propagiert werden. Hinzu kommt die Bestrebung, auch mit immer höherer zeitlicher Auflösung zu arbeiten. Im Laufe der Zeit wurden diese Sektoren vor allem bei den Kameras erschlossen. Eine weiterhin dominante Entwicklung ist zudem der immer stärkere Einsatz von IT-Technologie, vor allem der filebasierte Austausch von Videodaten in Netzwerken.

Die sechste Auflage des vorliegenden Buches wurde entsprechend in allen Kapiteln aktualisiert. Gegenüber der fünften Auflage gab es insbesondere starke Entwicklungen bei den Kameras und den Displays. Bei Letzteren kamen OLED-Displays als neue Flachbildschirme hinzu. Die Stereoskopie, d. h. die Nutzung der Videosysteme zur Aufnahme und Abbildung von 3D-Szenen, nimmt nun noch breiteren Raum ein. Neu aufgenommen wurde das Thema HD-Produktionsrichtlinien für den Fernsehbereich.

Bremen, im April 2013 Prof. Dr. U. Schmidt

Inhaltsverzeichnis

1	**Entwicklungsgeschichte**	**1**
2	**Das analoge Videosignal**	**16**
2.1	Licht und Wahrnehmung.	16
2.1.1	Fotometrische Größen.	19
2.1.2	Das menschliche Sehvermögen	21
2.2	Der Bildaufbau	30
2.2.1	Das diskretisierte Bild	31
2.2.2	Zeilenzahl und Betrachtungsabstand	33
2.2.3	Das Zeilensprungverfahren	37
2.3	Das monochrome Videosignal	41
2.3.1	Das BAS-Signal	41
2.3.2	Signalanalyse	53
2.3.3	Videosignalübertragung im Basisband.	59
2.4	Das Farbsignal	66
2.4.1	Farbe und Wahrnehmung	67
2.4.2	Farbmischung	70
2.4.3	Farbmetrische Größen	72
2.4.4	Das RGB-Signal	81
2.4.5	Das Komponentensignal	82
2.4.6	Das Farbsignal mit Farbhilfsträger	87
2.4.7	Farbfernsehnormen	93
2.4.8	Farbsignalhierarchie	104
2.5	Signalkontrolle	107
2.5.1	Betriebsmesstechnik	108
3	**Das digitale Videosignal**	**118**
3.1	Digitale Signalformen	118
3.2	Grundlagen der Digitaltechnik	119
3.2.1	Digitalisierung	120
3.2.2	D/A-Umsetzung	126
3.2.3	Digitale Signalübertragung im Basisband	127
3.3	Digitale SD-Videosignale	133
3.3.1	Digital Composite Signal	133
3.3.2	Digital Component Signal	135
3.3.3	Schnittstellen	142
3.4	Digitale HD-Signale	148
3.4.1	Bildaufbau digitaler HD-Signale	149
3.4.2	Digitale HD-Schnittstellen	154

3.5 Videodatenreduktion ... 158
3.5.1 Grundlagen ... 159
3.5.2 Redundanzreduktion ... 161
3.5.3 DCT ... 163
3.5.4 Wavelet-Transformation ... 170
3.5.5 DPCM ... 172
3.6 JPEG ... 176
3.7 Der DV-Algorithmus ... 178
3.7.1 DV-DIF ... 180
3.8 MPEG ... 182
3.8.1 MPEG-Videocodierung ... 183
3.8.2 MPEG-Audio ... 200
3.8.3 MPEG-Systems ... 210
3.9 ITU-H.26x und SMPTE-VC-Standards ... 216
3.9.1 H.264/AVC ... 216
3.9.2 H.265/HEVC ... 221
3.9.3 SMPTE VC-1 ... 222
3.9.4 SMPTE VC-2 ... 223
3.9.5 SMPTE VC-3, Avid DNxHD und Apple ProRes ... 224

4 **Fernsehsignalübertragung** ... 226

4.1 Analoge Modulationsverfahren ... 226
4.1.1 Multiplexverfahren ... 226
4.1.2 Amplitudenmodulation (AM) ... 228
4.1.3 Frequenzmodulation (FM) ... 231
4.2 Fernsehbegleitsignale ... 233
4.2.1 Fernsehbegleitton ... 233
4.2.2 Signale in der Vertikalaustastlücke ... 235
4.3 Analoge Übertragungsverfahren ... 239
4.3.1 Terrestrische Ausstrahlung ... 240
4.3.2 Satellitenübertragung ... 243
4.3.3 Kabelfernsehen ... 249
4.3.4 Verbesserte analoge Fernsehsysteme ... 252
4.4 Digitale Modulationsverfahren ... 259
4.4.1 Phase Shift Keying PSK ... 260
4.4.2 Quadraturamplitudenmodulation QAM ... 261
4.4.3 Amplitude Phase Shift Keying APSK ... 262
4.4.4 Orthogonal Frequency Division Multiplex OFDM ... 263
4.5 DVB ... 267
4.5.1 Fehlerschutz ... 270
4.5.2 Digitales Satellitenfernsehen DVB-S ... 274
4.5.3 Digitales Kabelfernsehen DVB-C ... 277
4.5.4 Digitales terrestrisches Fernsehen DVB-T ... 279
4.5.5 Data Broadcasting ... 285
4.5.6 DVB-Endgeräte ... 286
4.5.7 Neue digitale Videodienste ... 291
4.6 Internet und IPTV ... 292

4.6.1 Internet-Grundlagen ... 293
4.6.2 Netzwerke ... 293
4.6.3 Internet-Protokolle ... 294

5 Filmtechnik ... 296

5.1 Film als Speichermedium ... 296
5.1.1 Die Filmschwärzung ... 297
5.1.2 Farbfilm ... 299
5.2 Filmformate ... 300
5.2.1 Bildfeldgrößen ... 302
5.2.2 Filmkennzeichnung und Konfektionierung ... 305
5.3 Filmeigenschaften ... 307
5.3.1 Belichtung und Schwärzung ... 307
5.3.2 Kennlinie und Kontrastumfang ... 308
5.3.3 Farbfilmeigenschaften ... 312
5.3.4 Die Lichtempfindlichkeit ... 313
5.3.5 Das Filmkorn ... 314
5.3.6 Das Auflösungsvermögen ... 315
5.4 Filmentwicklung und -kopie ... 317
5.4.1 Filmentwicklung ... 317
5.4.2 Der Kopierprozess ... 318
5.4.3 Die Lichtbestimmung ... 319
5.4.4 Blenden, Titel- und Trickarbeiten ... 321
5.5 Filmton ... 322
5.5.1 Das Lichttonverfahren ... 323
5.5.2 Mehrkanaltonverfahren ... 325
5.6 Der digitale Film ... 330
5.6.1 Die Bildauflösung ... 331
5.6.2 Die Grauwertauflösung ... 333
5.6.3 Die Farbqualität ... 336
5.7 Filmabtaster ... 337
5.7.1 Filmabtastung bildpunktweise ... 339
5.7.2 Filmabtastung zeilenweise ... 342
5.7.3 Filmabtastung bildweise ... 344
5.7.4 Gradations- und Farbkorrektur ... 345
5.8 Filmbelichtung ... 348
5.8.1 CRT-Belichter ... 348
5.8.2 Laserbelichter ... 349
5.9 Filmdatenspeicherung ... 351

6 Bildaufnahmesysteme ... 353

6.1 Bildwandler ... 354
6.1.1 Röhrenbildwandler ... 356
6.1.2 CCD-Bildwandler ... 359
6.1.3 CMOS-Bildwandler ... 368
6.1.4 Plenoptischer Bildwandler ... 371
6.1.5 Sensorempfindlichkeit ... 372

6.1.6 Rauschen und Dynamik ... 373
6.1.7 Global Shutter ... 376
6.1.8 Defokussierung ... 377
6.1.9 Farbbildwandler ... 378
6.1.10 Bildformatwechsel ... 381
6.2 Kameraoptik ... 383
6.2.1 Grundlagen der Optik ... 383
6.2.2 Das optische System der Kamera ... 396
6.3 Die Videokamera ... 404
6.3.1 Das elektronische System der Kamera ... 405
6.3.2 Die Digitalkamera ... 413
6.3.3 Der Weißabgleich ... 414
6.3.4 Automatikfunktionen ... 415
6.3.5 Bildstabilisierungssysteme ... 416
6.3.6 Der Kamerasucher ... 418
6.3.7 Die Studiokamera ... 419
6.3.8 Die EB/ EFP-Kamera ... 425
6.3.9 Semiprofessionelle Kameras ... 428
6.3.10 Amateurkameras ... 429
6.4 Digitale Cinematographie ... 430
6.4.1 Die Filmkamera ... 430
6.4.2 Anforderungen an digitale Filmkameras ... 435
6.4.3 Digitale Filmkameras ... 437
6.5 Kamerasupport ... 453
6.5.1 Stereoskopische Aufnahmen ... 457

7 Bildwiedergabesysteme ... 460

7.1 Bildwiedergabe mit Kathodenstrahlröhren ... 461
7.1.1 Das Funktionsprinzip ... 461
7.1.2 Farbbildröhren ... 464
7.1.3 Signalverarbeitung ... 467
7.2 Flüssigkristall-Displays ... 470
7.2.1 LCD-Funktionsprinzip ... 470
7.2.2 Pixelansteuerung und -beleuchtung ... 474
7.3 Selbstleuchtende Flachbildschirme ... 477
7.3.1 Plasmadisplays ... 477
7.3.2 OLED-Displays ... 479
7.3.3 FED und SED ... 482
7.4 Videomonitore ... 483
7.4.1 Produktionsmonitore ... 485
7.4.2 Heimmonitore und Fernsehempfänger ... 489
7.4.3 3D-Displays ... 492
7.5 Großbildprojektion ... 494
7.5.1 Aktive Großbilderzeugung ... 495
7.5.2 Passive Großbilderzeugung ... 497
7.6 Kinoprojektion ... 506
7.6.1 Filmprojektion ... 506

7.6.2 Digitalprojektion im Kino ... 509
7.6.3 Digital Cinema ... 512
7.6.4 3D-Kino ... 516

8 Bildaufzeichnungsgeräte ... 520

8.1 Entwicklungsgeschichte ... 520
8.2 Grundlagen der Magnetaufzeichnung ... 521
8.2.1 Das magnetische Feld ... 521
8.2.2 MAZ-Grundprinzip ... 524
8.2.3 Magnetband- und Kopfführung ... 530
8.3 Analoge Magnetbandaufzeichnung ... 538
8.3.1 Signalverarbeitung bei der FM-Aufzeichnung ... 541
8.3.2 Zeitbasiskorrektur (TBC) ... 543
8.4 Audio-Aufzeichnung in Analog-Recordern ... 545
8.5 FBAS-Direktaufzeichnung ... 550
8.5.1 Das 2"-Quadruplex-System ... 550
8.5.2 Das 1"-A- und C-Format ... 552
8.5.3 Das 1"-B-Format ... 554
8.6 Colour-Under-Aufzeichnung ... 556
8.6.1 Signalverarbeitung ... 557
8.6.2 Colour-Under-Formate ... 559
8.7 Komponentenaufzeichnung ... 568
8.7.1 Signalverarbeitung ... 569
8.7.2 Komponentenaufzeichnungsformate ... 571
8.8 Digitale Magnetbandaufzeichnung ... 576
8.8.1 Signalverarbeitung ... 578
8.8.2 Digitale Komponentenformate ohne Datenreduktion ... 583
8.8.3 Digitale Compositesignal-Aufzeichnung ... 585
8.8.4 Digitale Komponentenformate mit geringer Datenreduktion ... 587
8.8.5 Digitale Komponentenaufzeichnung mit DV-Datenreduktion ... 591
8.8.6 Digitale Komponentenaufzeichnung mit MPEG-Datenreduktion ... 596
8.8.7 Digitale HD-MAZ ... 599
8.9 Bandlose digitale Signalaufzeichnung ... 604
8.9.1 Magnetplattensysteme ... 605
8.9.2 Optische Speichermedien ... 612
8.9.3 Festwertspeicher ... 623

9 Videosignalbearbeitung ... 624

9.1 Bildmischer ... 624
9.1.1 Mischerkonzepte ... 625
9.1.2 Bildübergänge ... 628
9.1.3 Stanzverfahren ... 630
9.1.4 Funktionen in Digitalmischern ... 637
9.1.5 Mischeraufbau ... 638
9.1.6 Bildmischer-Typen ... 643
9.2 Videografik- und Effektgeräte ... 646
9.2.1 Schriftgeneratoren ... 646

9.2.2 Digitale Effektgeräte ... 648
9.2.3 Grafik und Animation ... 650
9.3 Elektronische Schnittsteuersysteme ... 652
9.3.1 Grundfunktionen ... 652
9.3.2 Timecode ... 655
9.3.3 Linear Editing ... 662
9.3.4 Fernsteuersignale ... 665
9.3.5 Schnittsteuersysteme ... 667
9.4 Nonlineare Editingsysteme ... 670
9.4.1 Grundlagen ... 670
9.4.2 NLE-Betrieb ... 673
9.4.3 Filmschnitt ... 681
9.5 Compositingsysteme ... 687
9.5.1 Grundfunktionen ... 687
9.5.2 Ausführung von Compositingsystemen ... 693
9.6 Computeranimation ... 695

10 Videostudiosysteme ... 698

10.1 Signale im Produktionsbereich ... 699
10.1.1 Analoge Signalformen ... 699
10.1.2 Digitale Signalformen ... 700
10.1.3 Fileformate ... 713
10.2 Signalverteilung ... 724
10.2.1 Verteilung von Analogsignalen ... 724
10.2.2 Digitale Signalverteilung ... 728
10.2.3 Videosignalschnittstellen ... 730
10.3 Videodaten in Netzwerken ... 736
10.3.1 Systembetrachtung Signale und Netze ... 737
10.3.2 Netzwerke aus dem Telekom-Bereich ... 742
10.3.3 Netzwerke aus dem Computerbereich ... 744
10.4 Produktionsanforderungen ... 753
10.4.1 Produktionsempfehlungen der EBU ... 753
10.4.2 Technische Richtlinien-HDTV für ARD, ZDF und ORF ... 755
10.5 Produktionseinheiten ... 762
10.5.1 Produktionsbereiche ... 762
10.5.2 Postproduktion ... 773
10.5.3 Sendekomplexe und -abwicklung ... 776
10.5.4 Vernetze Produktion ... 778
10.5.5 SNG-Fahrzeuge ... 782
10.5.6 Ü-Wagen ... 787
10.6 Das virtuelle Studio ... 794
10.6.1 Technische Ausstattung des virtuellen Studios ... 796
10.6.2 Kameratracking ... 800
10.6.3 Produktion im virtuellen Studio ... 806

11 Literaturverzeichnis ... 811

12 Sachverzeichnis ... 819

1 Entwicklungsgeschichte

Die Begriffe „Video“ (lat.: ich sehe) und Film bezeichnen Illusionsmedien, mit denen bei den Menschen das Gefühl erzeugt werden kann, dass ihnen eine visuelle Realität dargeboten wird. Zur Erzeugung einer besonderen Faszination ist dabei die Zeitgebundenheit der Präsentation entscheidend. In Verbindung mit dem Ton spricht man von zeitbasierten Medien, die sich hier von Gemälden, Fotografien, Texten etc. abgrenzen [54].

Die Illusion wird heute fast ausschließlich mit Sequenzen von Einzelbildern erzeugt, die bestimmte Vorgänge der Realität für kurze Zeit festhalten, während die fixierten Zeitbereiche durch Dunkelphasen oder kurze Pausen unterbrochen sind, sodass man mit dem Experimentalfilmer Werner Nekes die Frage stellen kann „Was geschah zwischen den Bildern?“ [161].

Illusion und Traum faszinierten die Menschen schon vor langer Zeit, entsprechend finden sich schon früh Vorläufer moderner Medien die von Nekes gesammelt wurden. Viele Verfahren entstanden zunächst direkt bei der Rezeption, wo sich bereits eine starke Wirkung allein durch die Betrachtung in dunklen Räumen ergibt. So wurden bewegte Schattenbilder oder Scherenschnitte dargestellt (Abb. 1.1) und schon die Bewegung der zugehörigen Lichtquelle, und sei es nur eine Kerze, erzeugte das Gefühl animierter Bilder.

Ein wichtiger Schritt war in diesem Zusammenhang die Entwicklung der Laterna Magica, der Wunderlampe, die im 17. Jahrhundert einen größeren Bekanntheitsgrad erlangte als der Physiker Huygens sie wissenschaftlich beschrieb. Sie ist der Vorläufer der heutigen Projektoren (Abb. 1.2) und etablierte sich im Laufe der Zeit als wissenschaftliches Instrument, aber auch als professionelles Unterhaltungsmedium. Dabei wurde die Illusion dadurch unterstützt, dass die Laterna nicht sichtbar aufgestellt wurde und auf ungewöhnliche Projektionsflächen, wie z.B. Rauch, projiziert wurde. Ein durch sie dargestellter,

Abb. 1.1. Bewegtes Schattenspiel [161]

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4_1, © Springer-Verlag Berlin Heidelberg 2013

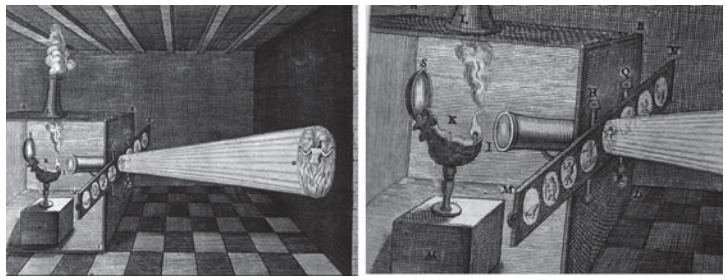


Abb. 1.2. Laterna Magica [161]

ständig bewegter Uhrzeiger kann als erster Langfilm betrachtet werden. Im 17. Jahrhundert wurde auch die Nachbildwirkung entdeckt und die Verschmelzung von Farben an einem schnell rotierenden Farbrad beschrieben.

Der Nachbildwirkung, d. h. der Effekt, dass man z. B. auf einer weißen Fläche die Wirkung eines schwarzen Musters auch nach Wegnahme desselben noch recht lange empfinden kann, wird meist als Grundlage der Bewegtbildwahrnehmung im Kino und Fernsehen betrachtet. Es ist aber nur ein Teil davon, denn es kommt der stroboskopische Effekt hinzu, der beschreibt, dass die zusammengehörigen Teilbilder durch Dunkelphasen unterbrochen werden. Ohne diese kann z. B. eine rotierende Lichtquelle nicht in ihrer Bewegung wahrgenommen werden sondern erscheint als heller Kreis. Die Wirkung zeigt sich auch auf der linken Seite von Abbildung 1.4, in der oben eine bewegte und unten eine stillstehende Bildfolge dargestellt ist. Rechts sieht man die Vergrößerung der bewegten Folge. Der stroboskopische Effekt wird durch die Schlitze erzeugt. Es wird deutlich, dass das Bild im Bereich der Schlitze viel klarer zu sehen ist.

Im 18. Jahrhundert entwickelten sich Konus- und Zylinderanamorphosen, die als Vorläufer von Kinotechnologien wie Cinemascope betrachtet werden können [25]. Die Bilder werden zunächst verzerrt hergestellt und werden erst bei der Darstellung über Spiegelung etc. in der gewünschten Form sichtbar. Im 19. Jahrhundert wurde dann der Stroboskopeffekt gezielt genutzt, um die Illusion längerer bewegter Bildsequenzen zu erzeugen. Im Jahre 1832 entstand das Phenakistiskop, das Lebensrad, welches es erlaubt, mehrere Bilder, die

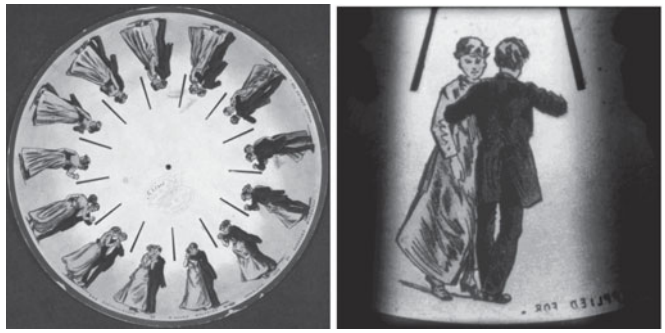


Abb. 1.3. Phenakistiskop [161]

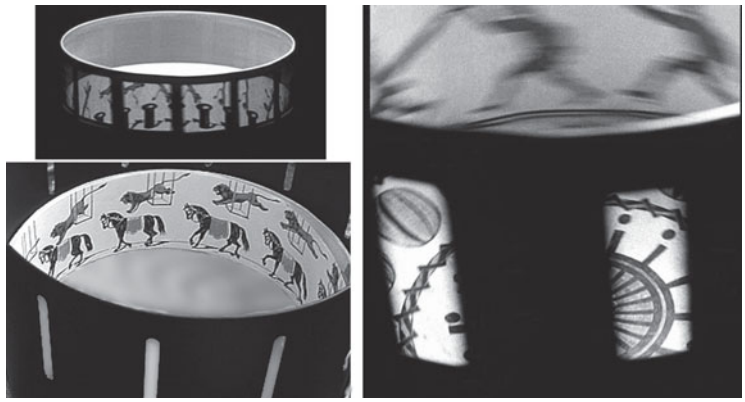


Abb. 1.4. Zoetrop [161]

eine geringe Bewegungsdifferenz aufweisen, über Schlitze zu betrachten, die den Lichtstrom unterbrechen (Abb. 1.3). Ebenso arbeitete das ein Jahr später vorgestellte Zoetrop, die Wundertrommel, bei der sich der Vorteil bot, dass in seinem Inneren verschiedene Papierstreifen eingesetzt werden konnten [161]. Somit war ein Programmwechsel möglich (Abb. 1.4). Mit gezeichneten Bildern ähnlicher Art wurde vorher bereits das Daumenkino entwickelt. Hier zeigte sich, dass es die Darbietung von mindestens etwa 16 Bildwechseln pro Sekunde erfordert, damit die Bildfolge von Menschen als eine zusammenhängende Bewegung wahrgenommen wird.

Zur gleichen Zeit gab es dann auch auf der Aufnahmeseite wichtige Entwicklungen. Hier ist insbesondere das Jahre 1826 zu nennen, aus dem die erste erhaltene Photographie, also eine nichtflüchtige, realitätsgetreue Abbildung stammt. Das Datum bezieh sich auf die Fixierung bzw. Speicherung des Bildes, die Abbildung selbst konnte schon um 1000, nämlich mit der Camera obscura (dunkle Kammer), erzeugt werden, die den heutigen Kameras ihren Namen gab (Abb. 1.5). In der Camera Obscura wird ein kleines Loch in einem stark abgedunkelten Raum benutzt, um aus der Vielzahl von Lichtstrahlen verschiedener Richtung nur wenige auszufiltern, sodass das zugehörige Objekt, von dem sie ausgehen, sichtbar wird. Das Bild ist umso schärfer, je mehr ausge-

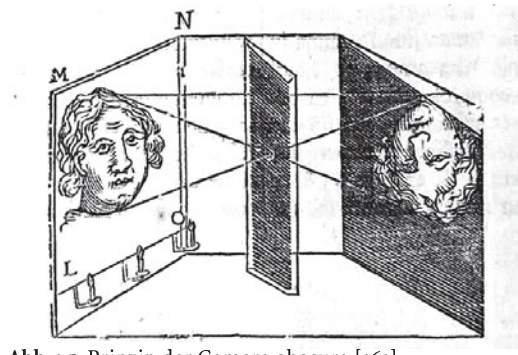


Abb. 1.5. Prinzip der Camera obscura [161]

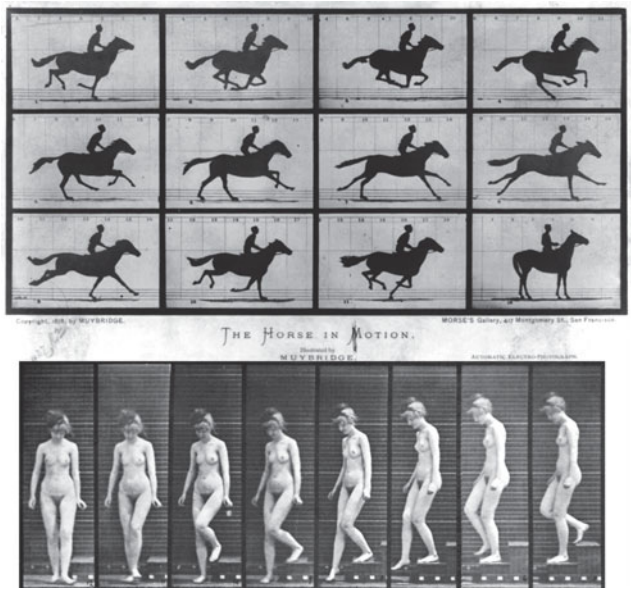


Abb. 1.6. Bewegungsstudien von Muybridge

blendet wird, d. h. je kleiner das Loch ist, jedoch sinkt damit natürlich die Lichtstärke der Abbildung, was später durch Einbringung einer Linse in das Loch verbessert wurde.

Zunächst wurde versucht, realitätsnahe Abbildungen durch Abzeichnen der Camera-Obscura-Bilder zu gewinnen. Im Jahre 1826 gelang es dann J. N. Niépce, das Bild auf einer Zinnplatte festzuhalten, was allerdings noch mit einer Belichtungszeit von acht Stunden verbunden war. Nach dem Tode von Niépce wurde die Entwicklung von Daguerre weitergeführt und es entstanden die sogenannten Daguerreotypen, ebenfalls Unikate, die noch nicht vervielfältigt werden konnten.

Um 1838 experimentierte der Engländer Talbot mit Papier als Trägermaterial, das er durch Chlorsilber lichtempfindlich machte. Das nasse Papier musste mehr als zwei Stunden belichtet werden, bevor die Umrisse der Abbildung als Negativ erschienen, d. h. dass helle Teile des abzubildenden Gegenstands dunkel wiedergegeben wurden und umgekehrt. Das Negativpapier konnte anschließend durch Wachs transparent gemacht werden, sodass nach Durchleuchtung und Schwärzung eines zweiten Chlorsilberpapiers das Positiv erschien. Die Erfindung wurde Photo Drawing oder auch Photo Graphics genannt, woraus der Name Fotografie entstand, der das wesentliche Merkmal, nämlich die nicht flüchtige Speicherung des Bildes, bezeichnet. Zu diesem Zeitpunkt waren bereits zwei wesentliche Bestandteile der modernen Fotografie entwickelt, nämlich das Negativ-Positiv-Verfahren und die Verwendung lichtempfindlicher chemischer Schichten auf Silberbasis. Durch die Entwicklung lichtstarker Objektive und die Verbesserung der lichtempfindlichen Schichten konnte im Laufe der Zeit die Belichtungsdauer weiter gesenkt werden. Ein wesentlicher Schritt wurde

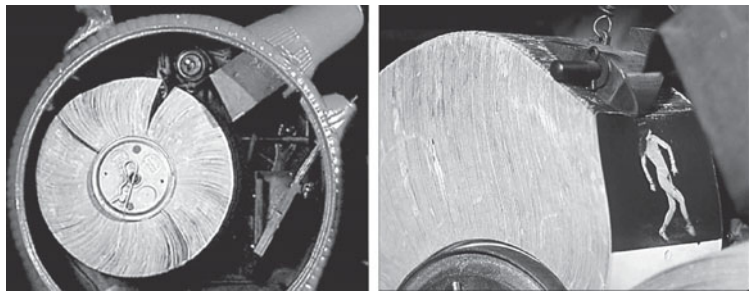


Abb. 1.7. Kinora mit gedruckten Bewegtbildsequenzen [161]

dabei 1860 getan, als man begann, ein feuchtes Bindemittel auf eine Glasplatte aufzutragen, die mit lichtempfindlichen Silbersalzen überzogen war. Die Platte wurde nur sehr kurz belichtet und dann sofort einer Entwicklung unterzogen, d. h. die belichteten Stellen wurden chemisch gewandelt, wodurch das noch unsichtbare, latente Bild erheblich verstärkt wurde. Anschließend wurden die nicht gewandelten Substanzen in einem Fixierprozess entfernt. Mit diesen zwei weiteren Bestandteilen des fotografischen Prozesses war es nun möglich, Belichtungszeiten im Sekundenbereich zu erreichen. Durch die Verwendung von Gelatine als Bindemittel wurde das Verfahren schließlich vereinfacht, sodass es möglich war, trocken zu arbeiten.

Damit konnte E. Muybridge im Jahre 1872 seine spektakulären Bewegungsstudien durchführen. Sie entstanden aber noch nicht durch eine schnelle zeitliche Bildaufnahme, sondern durch eine räumliche Anordnung von anfänglich 12 und später mehreren Kameras. Er bekam einen Preis dafür, dass er den Bewegungsverlauf eines galoppierenden Pferdes analysieren konnte (Abb. 1.6). Mit dem von ihm entwickelten Zoopraxiskop konnten die Bilder zeitlich aufeinander folgend dargestellt und im Bewegungszusammenhang, also z. B. als Darstellung eines an einem Ort galoppierenden Pferdes gesehen werden. Damit war bereits eine mediale Verbindung von Raum und Zeit genutzt, ähnlich den spektakulären modernen Effekten die z. B. beim Film „Matrix“ verwendet wurden. Auch das Daumenkino konnte nun mit Fotografien realisiert werden, sodass es mit Geräten wie der Kinora möglich war, in recht langen Bewegtbildsequenzen zu blättern (Abb. 1.7).

Die letzte entscheidende Vereinfachung ergab sich schließlich um 1888 durch die Verfügbarkeit von Nitrozellulose als flexiblem Schichtträger. Damit wurde die Fotografie massentauglich und es war die Basis der Filmtechnik geschaffen. Die Popularisierung der Fotografie begann mit der Kodak-Box von G. Eastman, die mit Rollfilm geladen wurde, sodass die Handhabung sehr vereinfacht war. Der ab 1889 verfügbare Rollfilm und die nun auf unter 1/16 Sekunde verkürzten Belichtungszeiten ermöglichten es dann schließlich, die für einen fließenden Bewegungseindruck erforderlichen mehr als 15 Bilder pro Sekunde in Echtzeit aufzuzeichnen. Mit dieser Bewegungsaufzeichnung, der Kinematographie, war ein neues Medium geboren.

Gemessen an der menschlichen Wahrnehmung war die Fotografie quasi echtzeitfähig geworden. Neben dem Rollfilm war dafür ein Apparat erforder-

Abb. 1.8. Filmsequenz der Gebrüder Lumière

lich, der den Film schnell genug transportierte und in den Transportpausen automatisch belichtete. Die Entwicklung eines solchen Apparates geschah in den Laboratorien von Thomas Alva Edison, der im Jahre 1891 den Kinematographen und das Kinematoskop als Geräte für die Aufnahme und Wiedergabe von Bewegtbildsequenzen zum Patent anmeldete. Der Filmtransport wurde dabei mit Hilfe einer Perforation im Film ermöglicht, die mit vier Löchern pro Bild definiert war. Das Kinematoskop war kein Projektionsgerät und damit nur für die Einzelbetrachtung geeignet. Eine Vorrichtung zur Projektion wurde erst in der folgenden Zeit entwickelt.

Im Jahre 1895 war dann der Cinematograph der Gebrüder Lumière einsatzbereit, bei dem die Funktionen von Kamera und Projektor in einem Apparat vereinigt waren. Mittels eines Greifers wurden die Filmbilder vor das Bildfenster gezogen und nach kurzem Stillstand automatisch weitertransportiert. Obwohl die Brüder Skladanowski in Berlin bereits am 1. November eine solche Filmvorführung vor Publikum gaben, gilt die erste öffentliche Filmvorführung mit dem Gerät der Brüder Lumière am 28.12.1895 heute als Geburtsstunde des Mediums Film. Zum ersten Mal war die Massentauglichkeit des Bewegtbildverfahrens als wesentliches Bestimmungsmerkmal erreicht, sodass sich die Gruppenrezeption als besonderes Spezifikum dieses Mediums etablieren konnte. Abgesehen von der Trennung von Kamera und Projektionsgerät hat sich das Grundprinzip der Kinematographie seither nicht verändert: Der perforierte Filmstreifen wird bei der Aufnahme und Wiedergabe schrittweise transportiert und steht bei Belichtung bzw. Projektion still. Während des Transports wird der Lichtweg abgedunkelt. Die technische Entwicklung wurde dabei auch von Oskar Meßter vorangetrieben, der mit dem Malteserkreuz ein hochwertiges Schaltwerk für den intermittierenden Filmtransport einsetzte. Meßter gilt als Begründer der deutschen Filmindustrie und arbeitete als Techniker, Regisseur und Produzent. Bereits ab 1895 begann durch die Brüder Pathé die Filmproduktion in großem Stil, durch die Brüder Lumière wurden die ersten Wochenschauen produziert sowie der bekannte Film „L'Arrivée d'un train à La Ciotat", die scheinbar dokumentarische Aufnahme eines einfahrenden Zuges, der in eindrucksvoller Weise auf die Kamera zufährt (Abb. 1.8). Ein Jahr später war mit der Doppelbelichtung bereits der erste Filmtrick entdeckt und 1902 wurde

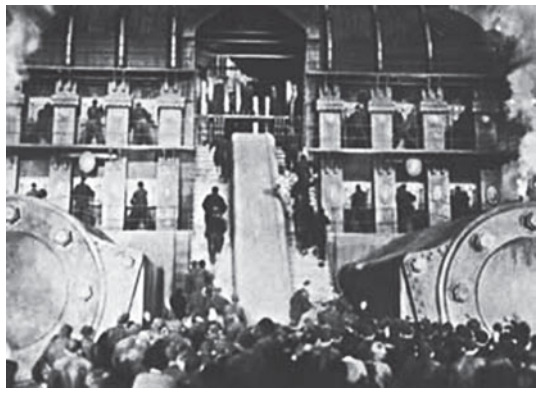


Abb. 1.9. Filmsequenz aus „Metropolis“

von Georges Méliès ein 16-Minuten-Film voller Spezialeffekte produziert. Der erste Animationsfilm, bei dem einzelbildweise belichtet und dabei Objekte in kleinen Schritten vor der Kamera bewegt werden, entstand 1907.

Im Jahre 1909 wurde nach einer internationalen Vereinbarung der 35 mm-Film als Standardformat festgelegt. Ab 1910 begann die Konzentration im Filmgeschäft und 1911 wurde in Hollywood, einem Vorort von Los Angeles in den USA, ein Filmstudio eröffnet, dem innerhalb eines Jahres viele weitere Studios folgten, sodass sich dieser Ort innerhalb kürzester Zeit zum Zentrum der US-Filmindustrie entwickelte. Die Studios erreichten eine monopolartige Stellung und bestimmten die Rechte über Kameras und Vorführsysteme ebenso wie das Verleihgeschäft. In großen, technisch gut ausgestatteten Anlagen wurde in sehr arbeitsteiliger Form produziert, um möglichst hohe Gewinne zu erzielen. Zusammen mit dem Starkult entstand hier die sogenannte Traumfabrik, die bis heute ihre Funktion beibehalten hat und den Weltfilmmarkt dominiert.

Auch in Deutschland entwickelte sich in den 20er Jahren mit der UFA in Babelsberg ein Filmkonzern, der ähnliche Produktionsweisen verwendete. Hier entstanden die großen deutschen Filme, wie Fritz Langs Metropolis, der sehr viele tricktechnische Aufnahmen enthält (Abb. 1.9). Hier konnte auch der Übergang zum Tonfilm mit vollzogen werden, der in Deutschland initiiert und zum Ende der 20er Jahre schließlich auf Druck aus den USA durchgesetzt wurde. Zu Beginn der 20er Jahre war es auch gelungen, die Schallsignaländerungen in Lichtintensitätsänderungen umzusetzen und auf Film aufzuzeichnen. Die Patente an dem Verfahren wurden in die USA verkauft, und von dort aus wurde die Durchsetzung des Tonfilms derart forciert, dass bereits zu Beginn der 30er Jahre die Ära der sogenannten Stummfilme beendet war. Die Bezeichnung Stummfilm bezieht sich auf das Fehlen der direkt aufgenommenen Dialoge. Doch waren auch vor der Einführung des Tonfilms die Filmvorstellungen oft von Erzählern und Musikern begleitet, die direkt auf die dargestellten Bildsequenzen reagierten und eine besondere Form eines Live-Erlebnisses erzeugten, die auch heute noch ihre besonderen Reize hat.

Erheblich länger als beim Tonfilm dauerte die Entwicklung der Farbfilmtechnik. Nachdem ab 1870 zunächst mit einer nachträglichen Kolorierung der Schwarzweißbilder von Hand begonnen worden war, wurde es später mög-

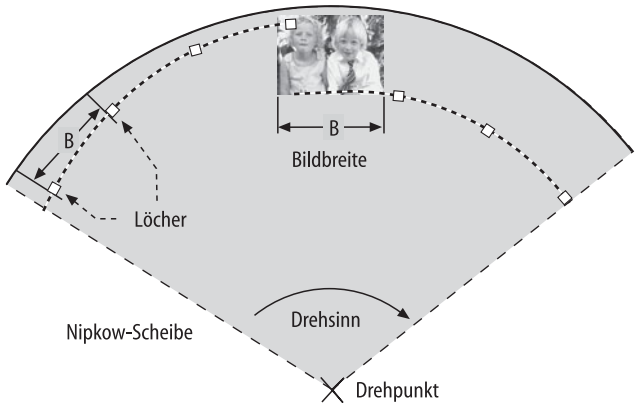


Abb. 1.10. Nipkow-Scheibe

lich, die lichtempfindlichen Emulsionen durch den Zusatz bestimmter Farbstoffe farbsensitiv zu machen. Größere Bedeutung erhielt der Farbfilm durch das Technicolor-Verfahren. Dabei wurden drei unterschiedlich farbsensitive Streifen belichtet und die Auszüge wurden übereinander gedruckt. Der erste abendfüllende Farbfilm nach dem Technicolor-Verfahren entstand 1935. Dieses Verfahren erforderte Spezialkameras und war kostspielig. Preisgünstigere Farbfilme, bei denen auch aufnahmeseitig alle farbsensitiven Anteile auf nur einem Filmstreifen untergebracht werden konnten, standen erst ab 1948 zur Verfügung, nachdem die chromagene Entwicklung nutzbar war, die auf Erkenntnissen über die Bildung von Farbstoffen beim Entwicklungsprozess beruht.

Die Farbfilmtechnik ist bis heute weitgehend gleich geblieben. Die Filmtontechnik wurde hingegen fortlaufend verbessert. Der Ton ist für die emotionale Wirkung des Films von sehr großer Bedeutung, entsprechend wurde bereits in den 40er Jahren mit Mehrkanalsystemen experimentiert, die das Räumlichkeitsgefühl der Audiowiedergabe steigern. Diese Entwicklung ist heute eng mit dem Namen Dolby verknüpft. Neben der Entwicklung von Rauschunterdrückungssystemen gelang es den Dolby Laboratories, beim Dolby-Stereo-System vier Tonkanäle in zwei Lichttonspuren unterzubringen. 1992 wurde von Dolby schließlich das sechskanalige Dolby-Digital-Verfahren eingeführt, das wiederum abwärtskompatibel zu Dolby Stereo ist und heute die größte Bedeutung unter den digitalen Kinotonformaten hat.

Neben der Einführung der Digitaltechnik im Tonbereich gewann in den 90er Jahren auch die digitale Bildbearbeitung immer größeren Einfluss. Zum Ende des Jahrtausends waren dann die Computersysteme so leistungsfähig, dass längere Spielfilmsequenzen in hoher Auflösung digital gespeichert und aufwendig bearbeitet werden konnten. Die Digitaltechnik spielt nicht nur im Produktionsbereich eine immer größere Rolle, sondern auch bei der Filmdistribution und der Wiedergabe. Für die Verteilung stehen hochwertige Datenreduktionsverfahren und hoch auflösende Projektoren zur Verfügung. Durch den Einsatz von digitalen High-Definition-Kameras ist schließlich eine vollständige digitale Infrastruktur im Kinobereich möglich geworden (Digital Cinema), die immer mehr dem Produktions- und Distributionssystem im Videobereich ähnelt.

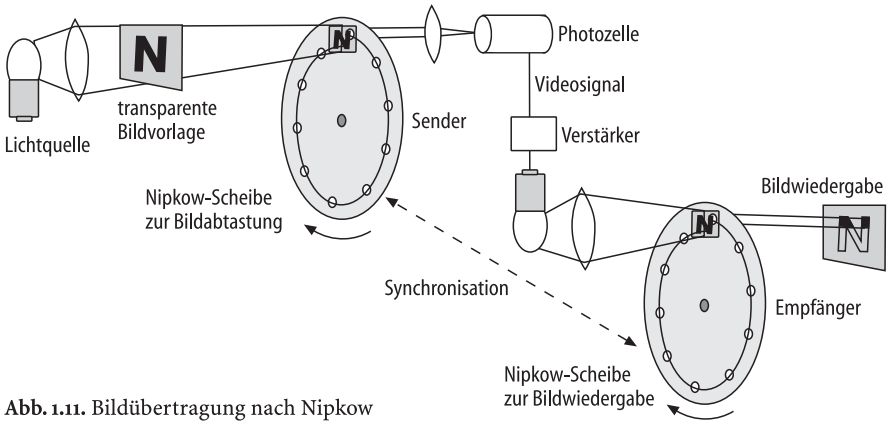


Abb. 1.11. Bildübertragung nach Nipkow

Die Grundlagen von Fernsehen und Video, d. h. der elektronischen Form der Bewegtbildübertragung, wurden zu einer Zeit geschaffen als der Zellulosefilm gerade begann, eine größere Verbreitung zu finden. Zu Beginn des 20. Jahrhunderts entstand dann die Fernsehtechnik als ein flüchtiges Medium, das als wesentlichen Unterschied zum Film die Eigenschaft hat, dass zu jedem Zeitpunkt nur die Information über einen einzelnen Bildpunkt und nicht über das ganze Bild vorliegt. Aus diesem Grunde spielte die Filmtechnik als Speichermedium bis zur Einführung von Videorecordern auch im Bereich der Fernsehproduktion eine entscheidende Rolle.

Einige für die Fernsehentwicklung wesentliche Erfindungen wurden bereits im 19. Jahrhundert gemacht. Im Jahre 1873 entdeckte C. May die Lichtempfindlichkeit des Selens, welche es ermöglicht, elektrische Ströme in Abhängigkeit von der Lichtintensität zu steuern. Um 1875 erforschte Kerr die Beeinflussbarkeit von polarisiertem Licht durch elektrische Felder (Kerrzelle) und Carey entwickelte die Idee zur Zerlegung des Bildes in einzelne Bildpunkte. Die parallele Übertragung der zugehörigen elektrischen Signale und die Ansteuerung eines entsprechenden Lampenrasters ist jedoch mit zu großem Aufwand verbunden. Sehr wichtig für die Entwicklung der Fernsehtechnik ist daher die Nutzung des Nachbildeffekts, d. h. die Erkenntnis, dass bei genügend schneller Abtastung die Bildpunktinformationen nicht gleichzeitig vorliegen müssen, sondern auch nacheinander übertragen werden können.

Für ein entsprechendes Realisierungskonzept erhielt im Jahre 1884 Paul Nipkow ein Patent. Nach Nipkow wird das Bild zeilenweise abgetastet. Dazu dient eine runde, drehbare Lochscheibe mit einer Anzahl von Löchern, die der Anzahl des in Zeilen zerlegten Bildes entspricht. Die Löcher sind spiralförmig angeordnet, sodass die Abtastung der zweiten Zeile genau dann beginnt, wenn das erste Loch das Bildfeld verlassen hat (Abb. 1.10). Die Helligkeitsinformationen der abgetasteten Zeile werden fortwährend von einer hinter dem Loch angebrachten Fotozelle in elektrische Signale umgesetzt, wobei sich bei großer Helligkeit entsprechend hohe Ströme ergeben.

Im Empfänger steuert das übertragene elektrische Signal die Helligkeit einer schnell reagierenden Lampe (Abb. 1.11). Hinter der Lampe befindet sich die

Abb. 1.12. Bildübertragung mit 30 und 96 Zeilen [8]

gleiche Nipkow-Scheibe wie im Sender. Falls beide Scheiben mit gleicher Umdrehungszahl laufen und zwar so, dass der Beginn der ersten Zeile im Sender und im Empfänger übereinstimmt, entspricht das durch die Scheibe transmittierte Licht dem abgetasteten Bild. Wichtige Aspekte der heutigen Videosysteme sind hier bereits anzutreffen: Die Bilderfassung geschieht zeilenweise, die parallel vorliegenden Informationen werden seriell übertragen, und es besteht die Notwendigkeit der Synchronisation von Sender und Empfänger.

Im Jahre 1897 wurde dann ein Bauteil entwickelt, das später einen wesentlichen Technologiesprung für die Fernsehentwicklung darstellen sollte, nämlich die braunsche Röhre von K. F. Braun, mit der Elektronenstrahlen sehr schnell gesteuert werden können. Das gilt für Strahlintensität und auch für die räumliche Ablenkung der Strahlen und so schlugen M. Dieckmann und G. Glage bereits 1906 vor, die Röhre, die im gleichen Jahr von R. v. Lieben als elektronische Verstärkerröhre entwickelt worden war, zur Darstellung von Schriftzeichen zu verwenden.

Aus der Basis dieser Vorarbeiten gelangte das Fernsehen in den 20er-Jahren in sein eigentliches Entwicklungsstadium. Ein bedeutender Zeitpunkt ist hier das Jahr 1923, in dem V. K. Zworykin (Abb. 1.13) ein Patent für das Ikonoskop, einen elektronischen Aufnahmebildwandler, bekam, der sich dadurch auszeichnete, dass er als erster das bis heute verwendete Speicherprinzip verwendete, um die erforderlichen Lichtempfindlichkeit zu erzielen. Kameras auf Basis dieses Bildwandlers wurden aber erst Mitte der 30er Jahre praxisreif. Die ersten Fernsehübertragungen mussten daher noch auf mechanischer Basis mit der Nipkowscheibe realisiert werden. Live-Übertragungen unter dem Einsatz der Nipkowscheibe waren sehr aufwendig und erforderten wegen der kleinen Löcher in den Scheiben sehr viel Licht, daher diente meist konventioneller Film als Zwischenstufe vor der Bildwandlung. Fast alle Beiträge wurden zunächst auf Film aufgezeichnet und über Filmabtaster mit der Nipkow-Scheibe umgesetzt. K. A. Karolus machte 1924 Fernsehversuche mit einer Nipkow-Bildzerlegung in 45 Zeilen. Auf der fünften großen Funkausstellung im Jahre 1928 führte D. v. Mihaly einen Fernsehempfänger mit 30 Zeilen vor, während Karolus schon eine Bildauflösung von 96 Zeilen zeigen konnte (Abb. 1.12). Dabei war er auch in der Lage, das Licht mit Hilfe einer Kerrzelle zu modulieren und über ein Spiegelrad auf eine Leinwand zu projizieren.

Abb. 1.13. Zworykin mit Ikonoskop-Röhren und Ikonoskop-Kamera [8]

In England experimentierte J. L. Baird mit einem 30-Zeilensystem und nachdem im Juli 1929 die erste vorläufige deutsche Fernsehnorm mit 30 Zeilen bei 12,5 Bildwechseln in der Sekunde festgelegt war, konnte er mit Förderung des Reichspostministers H. Bredow in Deutschland das erste tägliche Programm ausstrahlen. Im selben Jahr begann auch die BBC in England mit ersten Versuchssendungen. Wie in Deutschland geschah dies durch Nutzung von Hörfunksendern, die im Mittelwellenbereich mit 9 kHz Bandbreite arbeiteten.

Nachdem P. T. Farnsworth bereits 1929 in den USA ein voll elektronisches Bildübertragungsverfahren zeigen konnte, wurde in den 30er-Jahren die Bildübertragung mittels Nipkow-Scheibe endgültig von der Röhrentechnologie abgelöst. A. Wehnelt und von M. v. Ardenne entwickelten die Intensitätsmodulation des Elektronenstrahls für Wiedergabegeräte. Im Jahre 1931 stellte von Ardenne sein erstes elektronisches Bildübertragungssystem vor. Auf der Sendeseite verwendete er den Lichtpunktabtaster, der ein Filmbild mit Hilfe eines gleichmäßig leuchtenden Bildschirms in ein elektrisches Signal umsetzt, auf der Wiedergabeseite die Braunsche Röhre. Die elektronische Kamera, die von Ardenne noch fehlte, wurde dann 1933 von der amerikanischen Firma RCA auf Basis des Ikonoskops von Zworykin gebaut [43].

Die weitere Entwicklung bezog sich auf die Steigerung der Bildauflösung. Im Jahre 1934 wurde in Deutschland eine Norm mit 180 Zeilen bei 25 Bildern festgelegt, die aber schon 1937 von einer 441-Zeilnennorm mit 50 Halbbildern pro Sekunde abgelöst wurde. In England arbeitete man zu dieser Zeit mit 405 Zeilen. Ab dem Jahr 1935 fand in Deutschland ein regelmäßiger Sendebetrieb statt und die Firma Telefunken begann mit der Herstellung eines eigenen Ikonoskops, sodass die olympischen Spiele 1936 mit Hilfe von zwei elektronischen Kameras übertragen werden konnte (Abb. 1.13). Aus Mangel an Empfangsgeräten wurde das Programm meist in sogenannten Fernsehstuben betrachtet, in denen ein größeres Publikum an den Fernsehübertragungen teilnehmen konnte. Mit dem Beginn des zweiten Weltkriegs wurde 1939 der Sendebetrieb in Deutschland eingestellt. Die Weiterentwicklung des Fernsehens fand nun vor allem in den USA statt, wo bereits im Jahre 1941 die bis heute gültige US-Fernsehnorm mit 525 Zeilen festgelegt wurde.

Im Jahre 1935 wurde das erste Magnettonbandgerät, das AEG K1, vorgestellt. Die zugehörige Technologie gelangte nach dem Krieg in die USA, wo 1948 die

Abb. 1.14. Fernsehsendungen in den 50er-Jahren [8]

Firma Ampex das erste US-Magnettonbandgerät vorstellte und dann bald auch zum Videorecorder weiterentwickeln konnte.

Nach dem Krieg begann 1946 die BBC mit der Ausstrahlung eines Programms unter Nutzung der 405-Zeilennorm, Frankreich verwendete 819 Zeilen. In Deutschland wurde der Fernseh-Versuchsbetrieb ab 1948 aufgenommen. Dabei entschied man sich für eine Zeilenzahl von 625, die durch Koordination von W. Gerber 1950 von den meisten europäischen Ländern allgemein akzeptiert wurde (Gerber-Norm). Nur England und Frankreich schlossen sich hier erst später, bei Einführung des Farbfernsehens, an. Der Grund für die Wahl von 625 Zeilen lag darin, dass bei der Adaption von US-Geräten an das europäische 50-Hz-Netz die Zeilenfrequenz nahezu beibehalten werden konnte, denn aus $525 \cdot 60$ Hz ergibt sich nahezu der gleiche Wert wie aus $625 \cdot 50$ Hz.

Unter der Medienhoheit der alliierten Besatzungsmächte wurden 1949 die ersten Hörfunksender im Rahmen des öffentlich-rechtlichen Rundfunks zugelassen. Im Jahre 1950 gründete sich dann die Arbeitsgemeinschaft der Rundfunkanstalten Deutschlands, kurz ARD, die von den Rundfunkanstalten aus neun Bundesländern getragen wurde. Federführend bei der Wiederaufnahme des Fernsehbetriebs war der Nordwestdeutsche Rundfunk (NWDR) der ab 1950 mit Testsendungen begann. Am 25.12.1952 konnte schließlich der regelmäßige Fernsehrundfunk wieder aufgenommen werden und bereits am nächsten Tag war auch die Geburtsstunde der bis heute wichtigsten deutschen Nachrichtensendung, der Tagesschau [8]. Die landesweite Verbreitung des Programms wurde bis 1957 durch den Ausbau eines Richtfunknetztes erreicht, das die Rundfunkanstalten über den Sternpunkt Frankfurt verbindet [150]. Über Richtfunkstrecken konnte 1953 auch die erste spektakuläre Live-Übertragung, nämlich die der Krönungsfeierlichkeiten der Queen Elisabeth von England realisiert werden (Abb. 1.14).

Während also Ende 1952 in Deutschland das Fernsehen seine Geburtsstunde nach dem Kriege erlebte, wurde ein Jahr später in den USA bereits mit Farbfernsehsendungen nach der noch heute gültigen NTSC-Norm begonnen (National TV System Committee). Japan und viele Staaten Südamerikas übernahmen NTSC, aber in Europa wurde das Verfahren wegen der schlechten Farbstabilität (Never The Same Color) nicht akzeptiert. In Frankreich wurde als Alternative im Jahre 1957 SECAM (séquentiel couleur à mémoire) und in Deutschland das PAL-Verfahren (Phase Alternation Line) eingeführt. Dieses 1963 bei Telefunken entwickelte System ist farbstabil und mit weniger Problemen behaftet als SE-

Abb. 1.15. Kamera, Richtfunk- und Satellitenübertragung in den 60er-Jahren [8]

CAM, sodass viele Staaten das bis heute gültige Verfahren übernahmen. Die regelmäßige Ausstrahlung von PAL-Sendungen begann in Deutschland 1967.

Nachdem ab 1952 in größerem Stil mit dem Bau von Fernsehgeräten begonnen wurde, konnte um 1957 in Deutschland eine Zuschauerzahl von einer Million erreicht werden. Diese stieg in den folgenden Jahren stark an und man begann, über die Einrichtung eines zweiten Fernsehkanals nachzudenken. Im Jahre 1961 war vorgesehen, ein sehr regierungsnahes Programm auszustrahlen, das sogenannte Adenauer-Fernsehen. Dieses wurde jedoch gerichtlich untersagt und so entstand das unabhängige Zweite Deutsch Fernsehen (ZDF), das 1963 seinen Sendebetrieb aufnahm. In den 60er Jahren begann auch die Entwicklung der Satellitenübertragungstechnik. 1962 gelang die erste Fernsehübertragung von den USA nach Europa und 1965 war der erste Fernsehkanal via Satellit betriebsbereit (Abb. 1.15).

Das auf der deutschen Funkausstellung von 1967 in Betrieb genommene Farbfernsehen bekam einen Nachfrageschub durch die olympischen Spiele 1972 und die Fußball-Weltmeisterschaft 1974. Diese Ereignisse waren Anlass für viele Zuschauer, sich ein farbtaugliches Fernsehempfangsgerät zu kaufen. Zu dieser Zeit konnten bereits drei Programme empfangen werden, da ab 1967 die dritten Programme ausgestrahlt wurden, die von den verschiedenen Landesrundfunkanstalten betrieben werden und inhaltlich einen regionalen Charakter haben.

Das Programmangebot erhöhte sich in Deutschland schließlich im Jahre 1984 erheblich durch die Zulassung privater Sender. Es entstand das sogenannte Duale System mit gebührenfinanzierten öffentlich-rechtlichen und werbefinanzierten privaten Programmen. Die technische Verbreitung wurde durch die gleichzeitige Einführung des Kabelfernsehens ermöglicht. Ein Breitbandkabelnetz verbindet heute fast alle Haushalte Deutschlands und wird von mehr als der Hälfte der Zuschauer genutzt, um mehr als 50 Programme leitungsgebunden anstelle der bis dahin dominanten Funkübertragung zu empfangen. Seit den 90er-Jahren verwenden auch immer mehr Haushalte den Satellitenempfang, der heute bei mehr als 30% Nutzungsanteil liegt.

Bezogen auf die Bildqualität wurde in den 80er Jahren an einem Ersatz für das PAL-Signal gearbeitet. Dies geschah bereits mit Blick auf eine höhere Bildauflösung (PALplus und HD-MAC). Die Entwicklung dieser noch analogen Systeme wurde jedoch durch die Digitaltechnik überholt, insbesondere nachdem die Möglichkeit sehr effizienter Datenreduktion (MPEG) deutlich wurde.

Seit etwa 2000 ist das digitale Fernsehen (Digital Video Broadcasting, DVB) in Europa auf allen Verbreitungswegen, auch auf dem terrestrischen Weg als DVB-T, verfügbar, das das analoge Fernsehen gegenwärtig vollständig ablöst.

Während das PAL-Signal bis zur Einführung von DVB der Sendestandard war, wurde es im Produktionsbereich schon früher durch eines mit höherer Qualität ersetzt. Zunächst trat auf der Produktionsseite ein erheblicher Wandel ein, als ab 1960 erstmals elektronische Magnetbandaufzeichnungsverfahren (MAZ) zur Verfügung standen. Nach den ersten Geräten, die mit Magnetbändern von ca. 5 cm Breite arbeiteten, wurden bis zum Ende der 80er-Jahre die kostspieligen MAZ-Formate B und C verwendet, noch mit offenen Spulen und Bändern von 2,5 cm Breite und auf Basis eines PAL-codierten Signals. Als sich mit Betacam SP die Verfügbarkeit eines preiswerteren MAZ-Systems auf Kassettenbasis abzeichnete, wurde in der 90er Jahren der Wechsel zu dem bei diesem Format verwendeten Komponentensignal vollzogen und als Standard im Produktionsbereich eingeführt.

Mitte der 80er-Jahre waren auch bereits erste digitale MAZ-Systeme verfügbar und im Laufe der Zeit waren alle Studiogeräte auf digitaler Basis erhältlich. Gegenwärtig arbeiten die weitaus meisten Geräte des Produktionsbereichs digital. Der Datenaustausch erfolgt meist in Echtzeit über das seriell digitale Interface. Die Arbeitsabläufe ähneln dabei zwar manchmal noch denen der überkommenen Struktur, doch man spricht davon, dass diese erste Phase der Digitalisierung abgeschlossen ist, während nun die zweite Phase aktuell ist, die von der weitgehenden Vernetzung der Produktionskomplexe geprägt ist. Dem Einsatz dieser Systeme kommt die schon in den 90er Jahren begonnene Verwendung von datenreduzierten Signalen entgegen, die die Belastung der Netzwerke gering hält und gleichzeitig eine ökonomische Speicherung auf Festplattensystemen (Servern) mit zunehmender Unabhängigkeit von spezifischen Magnetbandformaten erlaubt.

Bis in die 80er-Jahre des 20. Jahrhunderts hinein gab es neben der Fernsehtechnik nur die Studio- bzw. Broadcasttechnik mit sehr speziellen, aufwendigen Geräten, denn die Videotechnik war zu Beginn ihrer Entwicklung immer ein Teilgebiet der Fernsehtechnik. Ab 1975 begann sich daneben ein Marktsegment zu entwickeln, das durch die Verfügbarkeit von preiswerten Videorecordern (insbesondere das Format VHS) für die Heimanwendung geprägt ist. Damit entstand auch ein Markt für die Programmverteilung mittels Kassetten, auf denen v. a. Spielfilme angeboten wurden und ein sogenannter semiprofessioneller Produktionssektor, in dem Schulungs- und Präsentationsvideos für einen kleineren Zuschauerkreis realisiert werden. Die Entwicklung für die Heimanwenderseite (Consumer-Sektor) ist von preiswerten Geräten, z. B. auch Kameras, geprägt, die einfach zu handhaben sind. Diese Aspekte wirken auf die professionelle Seite zurück, ein Einfluss, der heute immer stärker wird. Zum Beispiel wurde die Verwendung von Magnetbandkassetten anstelle offener Spulen zunächst für die semiprofessionellen Systeme (U-Matic) entwickelt und dann für den professionellen Bereich übernommen. Das gleiche gilt für Bildstabilisierungsverfahren, die zunächst nur das Verwackeln bei Aufnahmen mit kleinen Camcordern verhindern sollten. Eine besonders starke Rückwirkung auf die professionelle Seite gab es durch die Digitalisierung der Heimanwen-

dergeräte ab 1995. Hier tritt besonders das DV-Format hervor, das eine so gute Qualität aufweist, dass es in den Varianten DVCPro und DVCAM schnell für den professionellen Sektor adaptiert wurde, mit dem Ergebnis, dass DVCPro ein akzeptiertes Broadcastformat und DVCAM ein sehr verbreitetes Digitalformat im (semi-)professionellen Bereich wurde.

Die Digitalisierung beseitigte für die Heimanwender auch die letzte Hürde für eigene Videoproduktionen, nämlich die Nachbearbeitung. Aufgrund der inzwischen stark angestiegenen Leistungsfähigkeit von Heim-PCs konnten die Videosignale nun im Rechner geschnitten und vertont werden, wobei die Verbindung zwischen PC- und Videogeräten mit Firewire- oder USB-Schnittstellen erheblich vereinfacht wurde. Hier waren auch bereits früh Standards für den Datenaustausch gefragt, die als proprietäre Formate der großen PC-Systementwickler, namentlich Apple (QuickTime) und Microsoft (AVI), entstanden und bis heute eine große Rolle spielen. Zwischen Heim- und Studiosystemen entwickelte sich der sogenannte Prosumermarkt mit Camcordern in der Preisklasse um 5000 Euro, die bei guter Leistungsfähigkeit ein Zehntel der Geräte im Broadcastbereich kosten.

Die Verfügbarkeit immer effizienterer Datenreduktionsverfahren förderte schließlich einen sehr wichtigen Trend: Die Entwicklung hochauflösender Videosysteme (High Definition, HD). Dabei wird die Bildauflösung horizontal und vertikal etwa verdoppelt und statt des bei Standardauflösung (SD) verwendeten Bildseitenverhältnisses 4:3 das Breitbildformat 16:9 benutzt. Damit kann sich einerseits für die Betrachter im Heim ein kinoähnliches Sehgefühl einstellen, während andererseits der Kinobereich immer stärker von digitalen HD-Videosystemen beeinflusst wird (Digital Cinema). Diese Entwicklung wurde erheblich dadurch forciert, dass nun preisgünstige Displays (LCD und Plasma) und Projektoren zur Verfügung stehen, die diese Bilder adäquat wiedergeben können. Nach dem Jahre 2000 wurde fast nur noch in HD-taugliche Studiogeräte investiert. Kurze Zeit später wurde HD auch für die Heimanwender verfügbar. Mit dem DV-kompatiblen HDV-System gab es hier ein Format, das eine ähnliche Rolle spielt wie das DV-Format im Bereich der Standard-Definition-Signale.

Ein wichtiger Aspekt für die Vereinfachung der Produktionsabläufe (Medienhandhabung) im professionellen und auch im Heimanwenderbereich ist die zunehmende Verfügbarkeit bandloser Speichersysteme auf Disc- und Flashkartenbasis. Sie erlauben einen Szenenzugriff ohne Spulzeiten. Die DVD (Digital Versatile Disc) ersetzte z. B. die VHS-Kassette als Abspielmedium. In engem Zusammenhang mit dem bandunabhängigen, schnellen Zugriff auf die Bildsequenzen steht auch die Vereinfachung der Postproduktion. Heute können mit Standard-PC-Systemen Bildsequenzen in Kinoqualität verarbeitet werden. Der Schnitt und die Effektbearbeitung sind ebenso möglich wie die Farbkorrektur und die Nachvertonung. Insgesamt ist sowohl im professionellen als auch im Heimanwenderbereich eine zunehmende Konvergenz unterschiedlicher Medientypen von Film über Fernsehen bis hin zum gering aufgelösten Handy-TV zu beobachten. Diese Medien- und auch die Technologiekonvergenz mit IT-Systemen und Netzwerken sind kennzeichnend für die zukünftige Entwicklung.

2 Das analoge Videosignal

Zur Umsetzung von veränderlichen optischen Bildinformationen in ein elektrisches Signal wird zunächst eine zweidimensionale Abbildung erzeugt. Das Bild entspricht einer flächigen Anordnung sehr vieler Leuchtdichtewerte, die sich zeitlich dauernd ändern. Die hohe Informationsdichte muss reduziert werden. Dies soll möglichst so geschehen, dass der dabei entstehende Fehler dem menschlichen Auge als irrelevant erscheint. Eine erste Reduktion ist bereits die Abbildung auf zwei Dimensionen, weiterhin wird die Bildinformation durch zeitliche und räumliche Diskretisierung reduziert. Zur Bestimmung weiterer für den Bildeindruck relevanter Parameter, müssen einige Grundlagen über das Licht und die menschliche Wahrnehmung bekannt sein.

2.1 Licht und Wahrnehmung

Licht kann als elektromagnetische Welle beschrieben werden, in der sich die elektrische und magnetische Feldstärke periodisch ändert. Die Ausbreitungsgeschwindigkeit der Welle beträgt im Vakuum $c = 2{,}998 \cdot 10^8$ m/s (Lichtgeschwindigkeit). In dichteren Medien ist die Geschwindigkeit um den Faktor 1/n kleiner, n ist dabei der Brechungsindex (Abb. 2.1). Die Wellenform wiederholt sich periodisch in Raum und Zeit, zeitlich mit der Periodendauer T. Die Anzahl der Perioden pro Sekunde entspricht der Frequenz $f = 1/T$, bzw. der Kreisfrequenz $\omega = 2\pi f$. Die räumliche Periode ist durch die Wellenlänge λ gekennzeichnet. Zwischen der Wellenlänge λ und der Frequenz f gilt die Beziehung $c = \lambda \cdot f$.

Es gibt elektromagnetische Wellen mit einer Vielzahl von Frequenzen, das Spektrum reicht dabei von technischem Wechselstrom über Radiowellen bis zur Gammastrahlung (Abb. 2.2). Auch Licht lässt sich als elektromagnetische Welle beschreiben. Sichtbares Licht umfasst nur einen kleinen Teil des elektro-

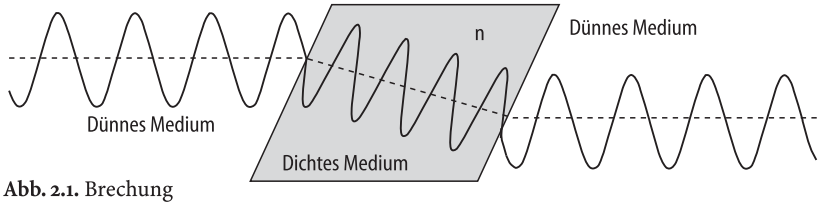


Abb. 2.1. Brechung

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4_2, © Springer-Verlag Berlin Heidelberg 2013

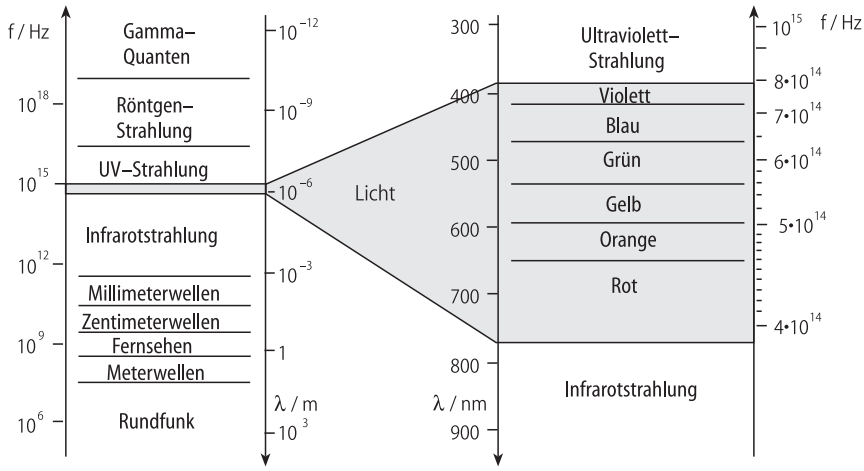


Abb. 2.2. Spektrum elektromagnetischer Wellen

magnetischen Spektrums, zugehörige Wellenlängen erstrecken sich über einen Bereich von 380 bis 780 nm. Die Farbempfindung des Auges hängt von der Wellenlänge ab. Licht mit einer Wellenlänge von 400 nm erscheint blau, 700 nm rot, dazwischen liegen die Farben des Regenbogens (Spektralfarben).

Die Lichtenergie resultiert aus atomaren Prozessen. Ein Elektron kann im Atom nur bestimmte ausgewählte Energiezustände einnehmen. Mit genügend äußerer Energie, zum Beispiel aus Erwärmung, kann das Elektron in einen höheren Energiezustand gelangen, es kehrt aber nach kurzer Zeit in den Ausgangszustand zurück und gibt die vorher aufgenommene Energie als elektromagnetische Strahlung wieder ab (Abb. 2.3). Diese Strahlungsenergie E ist sehr genau definiert und der Frequenz f proportional. Es gilt: $E = h \cdot f$ mit der Planckschen Konstanten $h = 6{,}6 \cdot 10^{-34}$ Ws². Durchlaufen alle an der Lichtemission beteiligten Elektronen dieselbe Energiedifferenz, dann entsteht monochromatisches Licht mit nur einer Wellenlänge (Laser). In den meisten Fällen sind jedoch viele verschiedene Energieniveaus beteiligt und es entsteht ein Gemisch aus vielen Frequenzen und Polarisationsrichtungen.

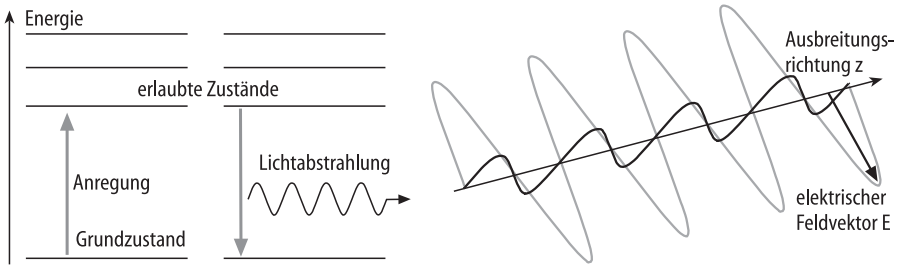


Abb. 2.3. Energieniveaus und Polarisation der elektromagnetischen Welle

Tabelle 2.1. Remissionsgrade verschiedener Materialien

schwarzer Samt	R = 0 %	grüne Blätter	15 ... 30 %
matte schwarze Farbe	1 .... 5 %	helle Haut	25 ... 35 %
schwarzes Papier	5 ... 10 %	weißes Papier	60 ... 80 %
Mauerwerk	10 ... 15 %	weißes Hemd	80 ... 90 %
Normalgrau (Fotokarte)	18 %	frischer Schnee	93 ... 97 %

Der Begriff Polarisation bezieht sich auf die Tatsache, dass die Lichtwelle eine Transversalwelle ist. Der Schwingungsvektor des elektrischen Feldes E steht senkrecht auf der Ausbreitungsrichtung z (Abb. 2.3). Behält der Vektor für das gesamte betrachtete Licht die gleiche Schwingungsebene bei, so spricht man von linear polarisiertem, bei gleichmäßiger Drehung um die Ausbreitungsrichtung von zirkular polarisiertem Licht [97]. Polarisiertes Licht entsteht, wenn aus einem Gemisch verschiedener Polarisationsrichtungen nur eine herausgefiltert wird. Das kann bei der Reflexion an metallischen Oberflächen, durch doppelbrechende Kristalle oder Polarisationsfolien geschehen. Polarisiertes Licht wird bei der Strahlteilung in Projektoren oder im LC-Display genutzt, denn es kann in seiner Intensität beeinflusst werden, indem es durch einen Polarisator geschickt wird, der eine andere Richtung hat als die Lichtpolarisation. Bei gekreuzten Polarisationsrichtungen kommt kein Licht durch.

Das Licht wird von den Gegenständen, auf die es trifft, zurückgeworfen. Die Reflexionseigenschaften hängen mit der Oberflächenbeschaffenheit der Gegenstände zusammen. Reflexion im engeren Sinne liegt vor, wenn der Strahlrückwurf ideal gerichtet ist, dagegen spricht man von Remission, wenn der Rückwurf ideal diffus ist und das Licht nach allen Seiten gestreut wird (Abb. 2.4 links). Da die viele Oberflächen eher rau sind, wird meist mit dem Remissionsgrad R gerechnet. Tabelle 2.1 zeigt Remissionsgrade verschiedener Materialien.

Bei dem gerichteten Rückwurf, der eigentlichen Reflexion, wird das Licht als Strahl betrachtet. Trifft dieser auf einen transparenten Gegenstand, so wird ein Teil von der Oberfläche reflektiert ein anderer Teil wird das Material durchdringen. Bei schrägem Einfall auf die Oberfläche ändert dieser Anteil seine Richtung, der Strahl wird gebrochen. Falls der Lichtstrahl dabei von einem optisch dichteren in ein dünneres Medium gelangt, wird er vom Einfallslot weg gebrochen und umgekehrt. Es gilt das Brechungsgesetz $\sin \varepsilon / \sin \varepsilon' = n' / n$ (Abb. 2.4 Mitte). Für den ersten Fall, bei $n > n'$, kommt es ab einem Grenzwinkel $\varepsilon_{gr} = \arcsin (n'/ n)$ zu dem Umstand, dass kein Licht mehr transmittiert werden kann. Dieser Umstand wird Totalreflexion genannt. Nahezu die gesamte Lichtenergie wird zurückgeworfen, was man sich bei der optischen Signalübertragung mittels Lichtwellenleitern zu Nutze macht (Abb. 2.4 rechts).

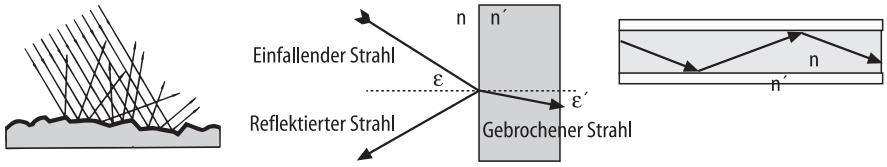


Abb. 2.4. Remission, Reflexion und Totalreflexion

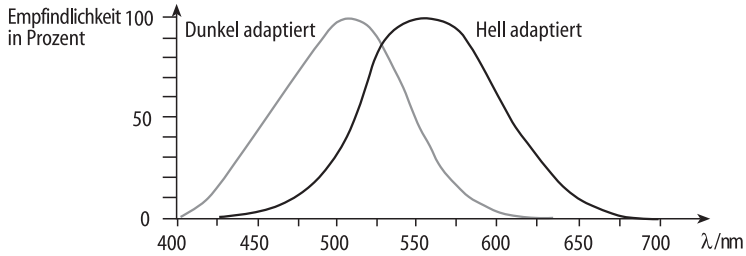


Abb. 2.5. Spektrale Hellempfindlichkeit des menschlichen Auges

2.1.1 Fotometrische Größen

Die Fotometrie dient der Messung von Lichtintensitäten und Helligkeit. Bei der Definition fotometrischer Größen wird zwischen energetischen Größen und solchen unterschieden, die unter Einbeziehung des spektralen Hellempfindlichkeitsgrades $V(\lambda)$ als visuelle Größen (mit dem Index v) definiert sind. Der relative spektrale Hellempfindlichkeitsgrad $V(\lambda)$ kennzeichnet die Frequenzabhängigkeit der Augenempfindlichkeit. Die Empfindlichkeit ist in der Mitte des sichtbaren Spektralbereichs (Grün: ca. 550 nm) maximal, zu den Rändern (Rot, Blau) hin fällt sie ab. Abbildung 2.5 zeigt die spektrale Hellempfindungskurven für Nacht- und Tagsehen (skotopisches und photopisches Sehen).

Die Lichtenergie pro Zeit wird mit dieser Bewertung als Lichtstrom Φ_v bezeichnet [Einheit: Lumen (lm)]. Ein Watt Strahlungsleistung entspricht im Maximum des Hellempfindlichkeitsgrades, bei $\lambda = 555$ nm für das photopische Sehen, einem Lichtstrom von 683 lm. Für die Fotometrie geht man vereinfachend von dem idealen, sogenannten Lambertstrahler aus, der durch eine gleichförmige Abstrahlung in den Halbraum oberhalb der leuchtenden Fläche gekennzeichnet ist. Wird der Lichtstrom hier auf einen Raumwinkel Ω mit der Einheit Steradiant (sr) bezogen, so ist die Lichtstärke I_v definiert, als

$$I_v = \Phi_v/\Omega$$

mit der Einheit Candela (cd), wobei gilt: 1 cd = 1 lm/sr. Die Lichtstärke relativ zu einer leuchtenden Fläche A_1, ergibt die Leuchtdichte L_v. Unter Einbeziehung eines Betrachtungswinkels ε_1 relativ zur Flächennormalen gilt (Abb. 2.6):

$$L_v = I_v/(A_1 \cos \varepsilon_1).$$

Die Einheit der Leuchtdichte ist cd/m², die alte Einheit ist ein Apostilb (asb), wobei gilt: 1 cd/m² = 3,14 asb. Farbbildröhren für Fernsehanwendungen erreichen bis 100 cd/m². Tabelle 2.2 zeigt die Leuchtdichten weiterer Lichtquellen.

Tabelle 2.2 Leuchtdichten im Vergleich

TV-Umfeldleuchtdichte	ca. 10 cd/m²	bedeckter Himmel	500 cd/m²
Bildschirmweiß (Röhre)	ca. 80 cd/m²	klarer Himmel	4000 cd/m²
helles Material im Raum	100 cd/m²	elektrische Lampen	ca. 10^4 cd/m²
Bildschirmweiß (LCD)	200 cd/m²	Lampenfaden	ca. 10^7 cd/m²

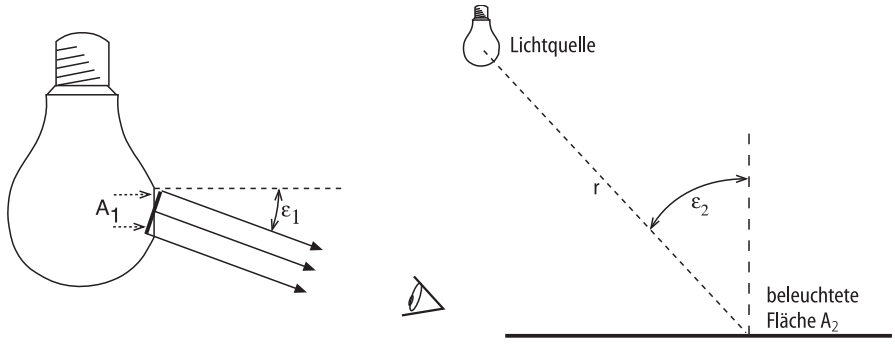


Abb. 2.6. Zur Definition der Leuchtdichte und der Beleuchtungsstärke

Auf der Empfängerseite interessiert als fotometrische Größe vor allem der Lichtstrom pro beleuchteter Fläche A_2, der als Beleuchtungsstärke E_v bezeichnet wird. Unter Einbeziehung des Winkels ε_2 zwischen Strahlrichtung und Flächennormale gilt:

$$E_v = \Phi_v / A_2 = (I_v \cos \varepsilon_2 \Omega_0)/r^2.$$

Dabei beschreibt r den Abstand zwischen der Lichtquelle und A_2 (Abb. 2.6). Die Einheit der Beleuchtungsstärke ist $lm/m^2 = lx$ (Lux). Die Beleuchtungsstärke definiert also den Lichtstrom, der unter einem bestimmten Winkel auf eine Fläche auftrifft.

Wenn ein Objekt nicht selbst leuchtet, sondern angestrahlt wird, folgt die Leuchtdichte einerseits aus der Beleuchtungsstärke, mit der es bestrahlt wird, und andererseits aus den Reflexions- bzw. Remissionseigenschaften des Objektes. Bei idealer Remission gilt folgende Beziehung zwischen Leuchtdichte L und Beleuchtungsstärke E:

$$L = R \cdot E/\pi.$$

Das Verhältnis von geringster zu größter Leuchtdichte in einer Szene bestimmt den Kontrastumfang L_{min} / L_{max}. Er kann mit dem sogenannten Spot-Fotometer bestimmt werden, einem Leuchtdichtemessgerät, das einen sehr kleinen Öffnungswinkel aufweist. Bei Fernsehsystemen kann nur ein relativ geringer Kontrastumfang von ca. 1:50 verarbeitet werden. Beleuchtungsstärke und Remissionsgrad sind zwei Parameter, die zur Bestimmung der Empfindlichkeit von Videokameras benutzt werden (s. Kap. 6.2).

Die angegebenen Formeln der Fotometrie gelten für den vereinfachten Fall einer gleichmäßigen Lichtverteilung. In den meisten Fällen reichen sie aus, um einfache Abschätzungen durchzuführen. So kann z. B. ermittelt werden, welchen Lichtstrom ein Projektor in einem Kino erzeugen muss, wenn eine Leinwand (bei R = 0,9) der Größe 6 m · 10 m mit einer typischen Leuchtdichte von L = 48 cd/m^2 ausgeleuchtet werden soll: Mit $E = L \cdot \pi/R$ ist dafür eine Beleuchtungsstärke von E = 167 lx erforderlich. Um diese auf 60 m^2 Leinwand hervorzurufen, braucht man eine Lichtquelle (den Projektor), die einen Lichtstrom $\Phi_v = E_v \cdot A_2 = 10\,000$ lm hat.

2.1.2 Das menschliche Sehvermögen

Das menschliche Sehvermögen beruht auf der Umsetzung von Lichtenergie in Nervenreize, die vom Gehirn verarbeitet werden. Abbildung 2.7 zeigt, von oben gesehen, einen Querschnitt durch das menschliche Auge. Es hat eine Kugelform mit etwa 24 mm Durchmesser und ist mit einer transparenten Substanz gefüllt, die einen Brechungsindex von 1,33 aufweist. Das Licht durchdringt die Horn- und Bindehaut und fällt auf die Pupille, die als Blende wirkt und abhängig vom Lichteinfall zwischen 1,5 und 8 mm Durchmesser veränderlich ist. Damit ist die Anpassung (Adaption) an verschiedene Helligkeiten möglich. Das Licht geht dann durch die Linse, die mit Hilfe von Muskeln in ihrer Form veränderbar ist, sodass unterschiedlich entfernte Gegenstände scharf auf die Netzhaut (Retina) abgebildet werden können (Akkommodation). Die Abbildungseigenschaft des Auges entsteht durch die Linse und aufgrund der Krümmung der Hornhaut.

Die lichtempfindliche Schicht in der Netzhaut besteht aus ca. $120 \cdot 10^6$ Zellen, unterscheidbar nach Stäbchen und ca. $6 \cdot 10^6$ Zapfen. Die Zapfen dienen bei hohen Leuchtdichten ($L > 10$ cd/m^2) als Rezeptoren der Helligkeits- und Farbwahrnehmung und für das scharfe Sehen. Die Stäbchen sind sehr empfindlich und registrieren die schwachen Helligkeitswerte. Ab $L < 0{,}08$ cd/m^2 spielen die Zapfen keine Rolle mehr, allerdings ist dann auch keine Farbunterscheidung mehr möglich (Nachtsehen). Abbildung 2.7 zeigt die horizontale Verteilung von Stäbchen und Zapfen in der Netzhaut. Die Zapfen sind am stärksten in der Fovea centralis, dem kleinen Bereich des deutlichen Sehens konzentriert.

Der Mensch erfasst seine Umgebung durch ständiges Abtasten der Objekte mit der fovea centralis und obwohl der Bereich des schärfsten Sehens sehr klein ist, kann doch ein größerer Bereich scharf gesehen werden. Bei der Abtastung akkommodiert und adaptiert sich das Auge ständig für jede neue Blickrichtung. Der optische Gesamteindruck ergibt sich aus dieser Abtastung und der Gedächtnisleistung des Gehirns.

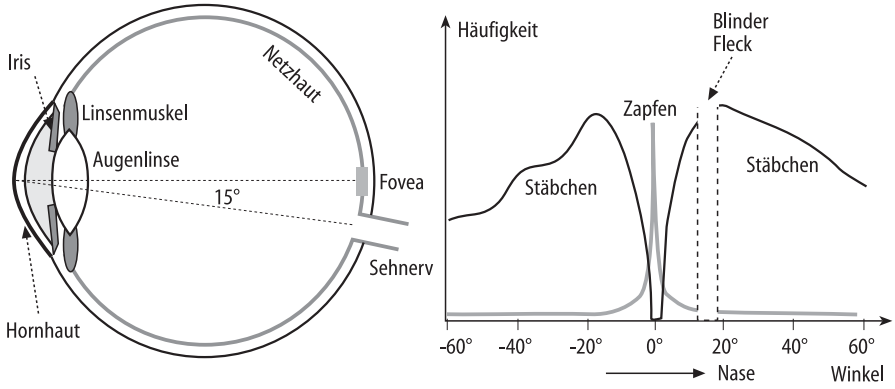


Abb. 2.7. Aufbau des Auges und Verteilung von Stäbchen und Zapfen auf der Netzhaut

Abb. 2.8. Hochpasswirkung an Graukanten und Simultankontrast

Am Ort des sogenanten blinden Flecks gibt es keine Rezeptoren, dort leitet der Sehnerv die Reize zum Zentralnervensystem und zum Gehirn. Der Lichtreiz führt zu einer neuronalen Erregung. Für die ca. 10^8 Rezeptoren stehen aber nur etwa 10^6 Neuronen zur Verfügung, sodass die Erregung einer Nervenfaser von mehreren Rezeptoren herrührt, die auf der Netzhaut zusammen das rezeptive Feld bilden. Im Bereich des schärfsten Sehens gibt es allerdings eine 1:1-Zuordnung zwischen Rezeptoren und Neuronen. Während der kleine Bereich der fovea centralis somit das scharfe Sehen bestimmt ist der große Bereich außerhalb davon von den Reizen der Stäbchen bestimmt. Sie sind lichtempfindlicher als die Zapfen. Die Empfindlichkeit wird aufgrund der Zusammenfassung zu den rezeptiven Feldern noch weiter erhöht, was dazu führt, dass schnell bewegte oder schwach beleuchtete Objekte aus den Augenwinkeln eher und mit einer hohen Bewegungsauflösung wahrgenommen werden. Damit lässt sich z. B. erklären, warum auf einem Röhrenmonitor, bei Betrachtung eines Fernsehbildes aus den Augenwinkeln, eine stärkere Flimmererscheinung auftritt als bei zentraler Sicht auf das Gerät.

Das rezeptive Feld ist veränderlich, so kann z. B. die Lichtempfindlichkeit auf Kosten der Sehschärfe erhöht werden. Durch den Lichtreiz entsteht eine Erregung in Form von elektrischen Impulsen, wobei sich der Grad der Erregung aus der Impulsfrequenz ergibt. Nun sind die rezeptiven Felder so gestaltet, dass sie konzentrische Bereiche aufweisen, wobei in einem Bereich eine Erregung und im anderen eine sogenannte laterale Hemmung auftritt. Die laterale Hemmung wirkt wie ein Hochpass, der die hohen Ortsfrequenzen verstärkt, die zu den feinen Bilddetails und scharfen Helligkeitsunterschieden gehören. Damit lassen sich die Erscheinung der Machschen Streifen erklären, d. h. die subjektiv empfundene Hervorhebung der Übergänge an einer Grautreppe sowie die Tatsache, dass die Schärfewahrnehmung bei großen Kontrasten hoch ist. Die Hochpasswirkung erklärt auch den sogenannten Simultankontrast, d. h., dass dem Auge ein mittleres Grau in dunkler Umgebung heller erscheint als in heller Umgebung (Abb. 2.8).

2.1.2.1 Die Ortsauflösung

Der Gesichtssinn umfasst horizontal einen Winkel von ca. 180°. Aufgrund der beschriebenen Zuordnung von Rezeptoren und Neuronen können die Menschen allerdings nicht gleichzeitig alle Punkte in diesem Umkreis scharf sehen, sondern müssen ihre Konzentration auf kleinere Bereiche lenken. Ein mittlerer Sehwinkel, das deutliche Sehfeld oder der Bereich des »Normalsehens« beträgt

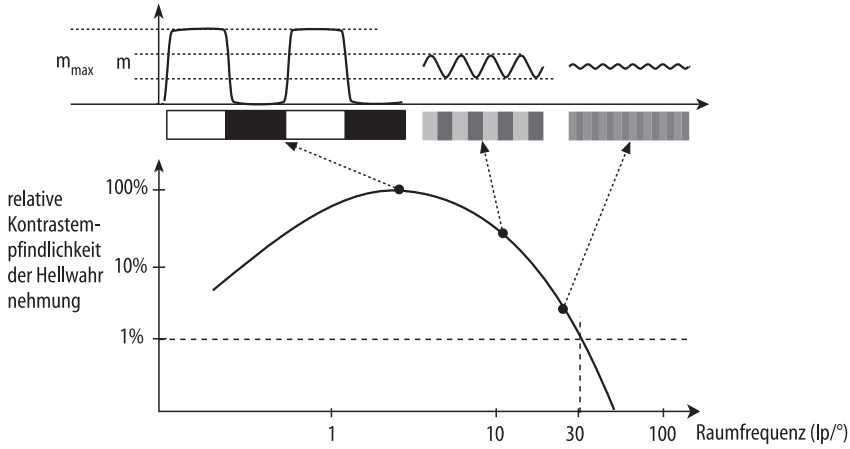


Abb. 2.9. Örtliche Auflösung der S/W-Wahrnehmung im Zusammenhang mit der MTF

vertikal etwa 30°. Wenn ein Objekt bequem »mit einem Blick« erfassbar sein soll, sollte sich der vertikale Betrachtungswinkel auf etwa 10° beschränken. Der Bereich des deutlichen Sehens mit der fovea centralis umfasst nur einige Grad, wobei die maximale Auflösung etwa eine Winkelminute beträgt.

Das Auflösungsvermögen beschreibt die Wahrnehmbarkeit feiner Details. Als Maß dafür kann untersucht werden, wie viele Schwarz/Weiß-Linienpaare pro Grad im Blickfeld als getrennt wahrgenommen werden können. Abbildung 2.9 zeigt, wie die menschliche Wahrnehmung der Helligkeitsunterschiede bzw. des Kontrasts von der Raumfrequenz abhängt. Es wird deutlich, dass Kontraste am deutlichsten zwischen 2 und 10 Linienpaaren gesehen werden. Bei höheren Liniendichten, d. h. bei höherer Ortsfrequenz, nimmt die Unterscheidbarkeit ab und findet eine Grenze bei ca. 30 lp/°, d. h. 60 Linien/° oder einer Winkelminute.

Der Verlauf der Kurve in Abbildung 2.9 kann als Modulationstransferfunktion (MTF) beschrieben werden. Eine MTF gibt an, wie die Modulation

$$m = (s_{max} - s_{min}) / (s_{max} + s_{min})$$

bei verschiedenen Liniendichten im Verhältnis zur maximalen Modulation m_{max} steht (Abb. 2.9). Es gilt:

$$MTF = m/m_{max}.$$

Diese Größe wird allgemein für technische Systeme verwendet, die das Auflösungsvermögen begrenzen.

Die Ortsauflösung ist ein wichtiges Merkmal, das die Schärfewahrnehmung von Bildern bestimmt. Es ist aber nur ein Teilaspekt. Aufgrund der oben beschriebenen Hochpasswirkung, die durch die rezeptiven Felder bewirkt wird, ist es für die Schärfewahrnehmung ebenfalls wichtig, wie kontrastreich das Bild wiedergegeben wird. Dieser Umstand lässt sich beim Vergleich zweier Bilder mit verschiedener Auflösung und unterschiedlichem Kontrast zeigen.

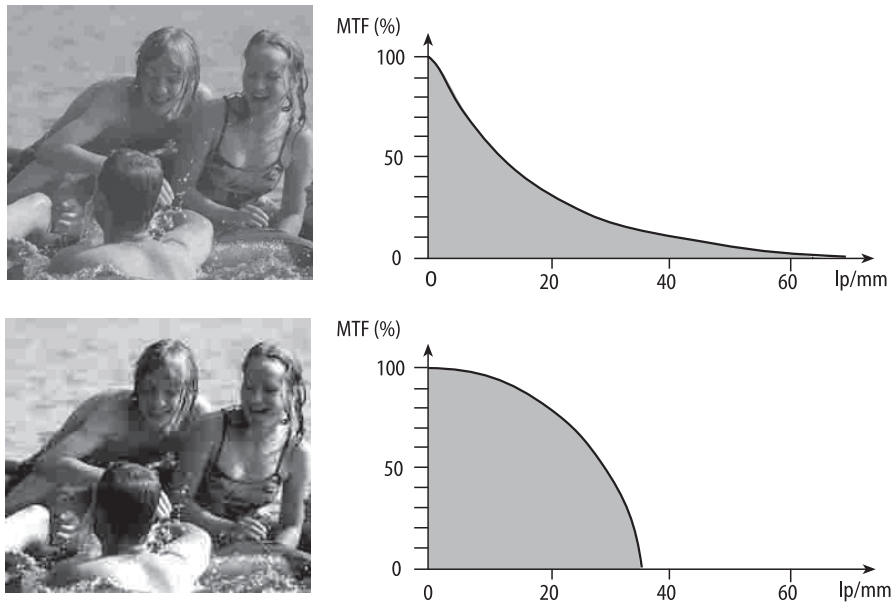


Abb. 2.10. Bilder mit unterschiedlichem Kontrast und die zugehörigen MTF

Abbildung 2.10 zeigt den Vergleich anhand zweier Bilder, von denen das obere horizontal doppelt so viele Bildpunkte, bzw. Auflösung hat wie das untere. Beim unteren Bild wurden die Kontraste an den groben Umrissen durch Filter verstärkt. Diese Verstärkung führt zu erhöhter Modulation mit entsprechender Auswirkung auf die Modulationstransferfunktion bei tiefen Frequenzen.

Die MTF für das obere Bild weist bei hohen Ortsfrequenzen noch Werte auf, bei denen die MTF für das untere Bild bereits null ist. Im Bild wird dies an den feineren Strukturen in den Haaren deutlich. Die MTF des oberen Bildes zeigt aber bei tiefen Frequenzen relativ kleine Werte, während die MTF-Werte des unteren Bildes durch die Filterung bei den tiefen Ortsfrequenzen verstärkt sind. Obwohl es eine geringere Auflösung hat, wird daher das untere Bild auf den ersten Blick als schärfer empfunden. Der Schärfeeindruck korrespondiert mit der Fläche unter der Kurve. Vereinfacht gesagt: je größer die Fläche, desto schärfer das Bild.

Die Trennbarkeit von Linien bzw. die Kontrastempfindlichkeit des Auges hängt weiterhin auch vom Modulationsgrad der Bildvorlage und von der Beleuchtungsstärke ab. Die Auflösungsleistung des Gesichtssinns ist darüber hinaus von der zeitlichen Auflösung bestimmt. Sie ist bei ruhenden Bildvorlagen am größten. Werden bewegte Objekte betrachtet, so können wir sie weniger scharf erkennen, allerdings ist der Mensch in der Lage, den Bewegungen bis zu einem gewissen Grad zu folgen (Tracking). Je länger wir uns auf ein Objekt konzentrieren können, desto mehr Einzelheiten sind unterscheidbar, aber auch beim Fixieren werden kleinere Augenbewegungen ausgeführt, die verhindern, dass das Bild aufgrund einer Gewöhnung verblasst.

2.1.2.2 Die Zeitauflösung

Eine für die Bewegtbildübertragung überaus wichtige Frage ist natürlich, wie viele Bilder pro Sekunde übertragen werden müssen, damit der Gesichtssinn einen in den Bildern dargestellten Bewegungsablauf als zusammenhängend empfindet. Hierzu lässt sich schon mit dem Daumenkino feststellen, dass ca. 20 Bilder pro Sekunde ausreichen, um einen gleichmäßig erscheinenden Bewegungsablauf zu erzielen. Die Bildfrequenz beim Film beträgt entsprechend 24 Hz.

Der mit Dunkelpausen verbundene Bildwechsel führt zu einer ständig wiederholten Anpassung der rezeptiven Felder mit wechselnder Erregung und Hemmung der Neuronen, was für die Wahrnehmung sehr anstrengend ist und als sehr unangenehmes Flackern empfunden wird. Das Flackern wird als Großflächenflimmern bezeichnet, da es alle Bildpunkte zugleich betrifft und das Auge mit sehr vielen Zellen gleichzeitig darauf reagiert. Die sogenannte Flimmer-Verschmelzungsfrequenz, bei der diese störende Empfindung verschwindet, liegt oberhalb 50 Hz. Die Flimmerempfindlichkeit steigt mit der Bildhelligkeit weiterhin ist sie vergleichsweise hoch wenn der Bildeindruck von den Blickseitenbereichen herrührt.

Aus diesem Grund werden Filmbilder im Kino bei der Wiedergabe zwei- oder dreimal projiziert, sodass sich eine Verdopplung der Dunkelpausen bei einer Flimmerfrequenz von 48 Hz bzw. 72 Hz ergibt. Der kleinere Wert ist für das Kino bereits akzeptabel, da hier die Rezeption in abgedunkelten Räumen erfolgt und auf der Leinwand nur geringe Leuchtdichten von unter 50 cd/m^2 verwendet werden. Bei Wiedergabesystemen, die so gestaltet sind, dass der Beobachter bei hoher Helligkeit große Aufmerksamkeit aufbringt und bei denen das Bild einen großen horizontalen Sehwinkel einnimmt (z. B. Computerdisplays), sollte die Großflächenflimmerfrequenz oberhalb 70 Hz liegen.

Angesichts der gestiegenen technischen Möglichkeiten wird vorgeschlagen, auch in künftigen Kino- und ggf. auch Fernsehstandards mit 80, 100 oder 120 Bildern/s zu arbeiten (HFR, High Frame Rate). Dies gilt insbesondere für stereoskopischen Anwendungen. Die stereoskopische Wahrnehmung, die im übernächsten Abschnitt dargestellt ist, orientiert sich an klaren Konturen im Bild. Diese erfordern kurze Belichtungszeiten, was aber bei geringer Bildwiederholrate mit ruckenden Bewegungen der Objekte im Bild verknüpft sein kann.

Fernseh- und Videosysteme sind unkritischer gegenüber ruckender Wahrnehmung bewegter Objekte im Bild als das Kino, denn sie arbeiten mit Halb- oder Vollbildern, die sowohl bei der Aufnahme als auch bei der Wiedergabe mit 50 oder 60 Bildern/s verarbeitet werden. Die Kinobildwiedergabe ist bezüglich der Bewegungsauflösung dagegen kritischer einzustufen, denn auch wenn die Bilder bei der Projektion mit 48 Hz gezeigt werden, stehen bei der Aufnahme doch nur 24 Bilder/s zur Verfügung. Untersuchungen zum ruckenden Bildeindruck, durchgeführt mit einem weißen Balken, der horizontal vor einem weißen Gitter auf dunklem Grund bewegt wird, zeigen, dass bei 24 fps Objektgeschwindigkeiten von höchstens 0,2 Bildhöhen/s tolerierbar sind, wenn die Artefakte als nicht störend wahrgenommen werden sollen [153]. Daher sollten Panoramaschwenks, die das ganze Bild betreffen, für das Kinos nicht schneller als mit 0,2 Bildhöhen/s ausgeführt werden. Videosysteme, die aufnahmeseitig mit 50 oder 60 Bildphasen arbeiten, erlauben danach ca. 0,3 BH/s.

2.1.2.3 Die Helligkeitsempfindung

Die Hellempfindung eines Objekts ist vom Adaptionszustand des Auges und der Umfeldbeleuchtung bestimmt. Das Auge ist in der Lage, sich an Leuchtdichten anzupassen, die einen Bereich von elf Zehnerpotenzen umfassen, allerdings ist die Adaption mit verschiedenen Zeitkonstanten bis zu 30 Minuten verbunden. Bei adaptiertem Auge können auf einem geeigneten Videodisplay ca. 200, im Allgemeinen ca. 500 Helligkeitswerte unterschieden werden.

Der Zusammenhang zwischen Hellempfindung L_e und der Leuchtdichte L als Reizgröße ist im mittleren Leuchtdichtebereich logarithmisch. Das Weber-Fechnersche oder auch psychophysische Grundgesetz lautet:

$$L_e = \log (c \cdot L),$$

wobei c eine Konstante ist.

Bei der Betrachtung von Fernsehbildschirmen spielt die Vermeidung von Streulichtern eine große Rolle, da sie die Kontrastempfindung einschränken. Trotzdem sollte dafür gesorgt werden, dass der Raum, in dem das Monitorbild betrachtet wird, nicht völlig abgedunkelt ist. Die Umfeldleuchtdichte sollte ca. 10% der Spitzenleuchtdichte auf dem Schirm betragen, damit dem Auge die dunkelsten Bildpartien tatsächlich als schwarz erscheinen.

2.1.2.4 Die dreidimensionale Wahrnehmung

Der Mensch ist in der Lage, seine Umwelt dreidimensional zu erfassen, also neben der Raumhöhe und der -breite auch die Raumtiefe. Film- und Videosysteme beschränken das Bild gewöhnlich auf zwei Dimensionen und legen darüber hinaus auch noch einen definierten Rahmen fest. Es können aber

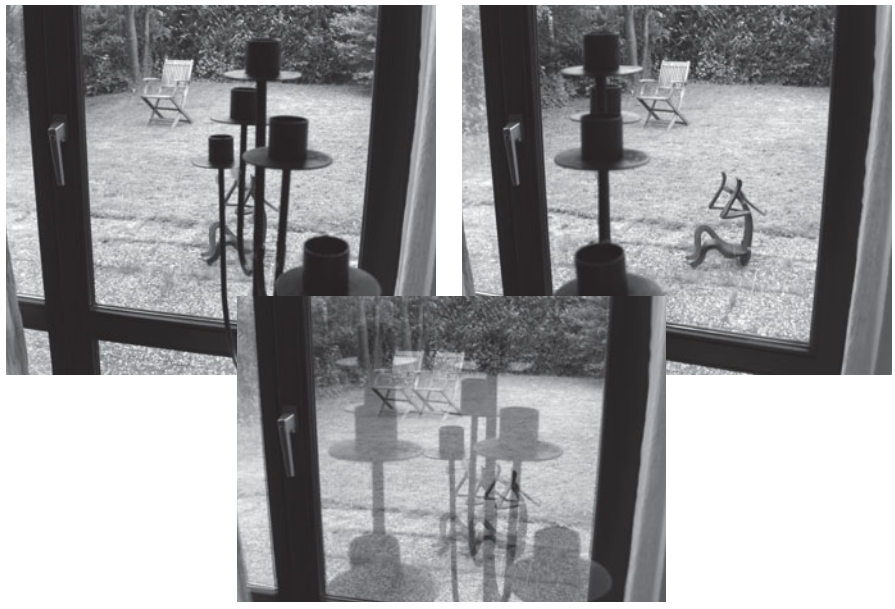


Abb. 2.11. Bilder aus zwei verschiedenen Perspektiven

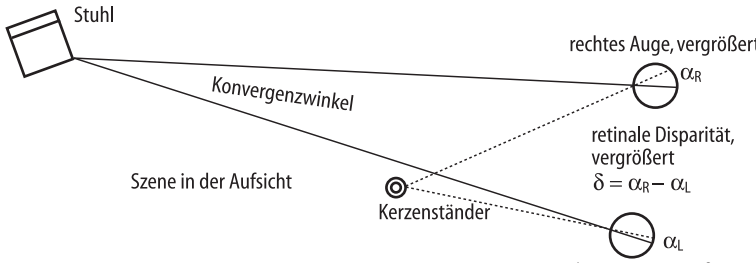


Abb. 2.12. Entstehung der Disparität im menschlichen Auge

auch technische Systeme entworfen werden, die den Menschen eine 3D-Wahrnehmung ermöglichen. Systeme, die Bilder wirklich dreidimensional abbilden, sind in der Entwicklung, aber extrem aufwendig. Daher wird in der Regel nur der Hauptmechanismus genutzt, der die Tiefenwahrnehmung des Menschen bestimmt, nämlich das binokulare Sehen, d. h, das Sehen mit zwei Augen, die aufgrund ihres Abstandes zu zwei verschiedenen Ansichten eines vom Gesichtssinn fixierten Objekts führen. Diese Ansichten entsprechen zwei leicht unterschiedlichen Perspektiven, die vom Gehirn zu einem Bild verschmolzen werden, das dann auch die Raumtiefe repräsentiert. In Abbildung 2.11 werden die Perspektivunterschiede z. B. an der Relation zwischen dem Kerzenständer im Vordergrund und dem Stuhl im Hintergrund deutlich.

Technische Systeme, die diesen Mechanismus nachbilden, nennt man stereoskopische Systeme (Stereoskopie = Raum-Sicht). Sie werden auch als s3D-Systeme bezeichnet. Wie in der Natur nutzen sie zwei zusammengehörige Bilder oder Bildfolgen, die als Halbbilder bezeichnet werden. Aufnahmeseitig braucht man für die stereoskopischen Systeme zwei synchrone Kameras, deren optische Achsen in erster Näherung einen Abstand haben sollten, der dem der menschlichen Augen, also ca. 65 mm, entspricht. Wiedergabeseitig müssen die beiden Bilder den Augen getrennt zugeführt werden.

Zur Erklärung der binokularen Raumtiefenwahrnehmung kann man zunächst annehmen, dass der Mensch mit beiden Augen einen Objektpunkt, z. B. an einem Stuhl, fixiert (Abb. 2.12). Der Winkel zwischen den beiden Blickachsen wird als Konvergenzwinkel bezeichnet und schon der Grad der dafür erforderlichen Spannung der Augenmuskulatur wird vom Gehirn als Hinweis bei der Raumtiefenwahrnehmung benutzt, die Stereopsis genannt wird. Der fixierte Punkt wird nicht zweimal gesehen, sondern es findet im Gehirn eine Fusion statt. Weitere Punkte, die auch fusioniert werden, heißen korrespondierende Punkte. Alle korrespondierenden Punkte liegen auf einer Fläche im Raum, die als Horopter bezeichnet wird. Andere Objektpunkte, die in der Raumtiefe getrennt sind, können nicht fusionieren, da sie in beiden Augen relativ zum fixierten Punkt verschiedene Informationen erzeugen, man spricht von retinaler Disparität (Abb. 2.12). Aus dieser, also der Winkeldifferenz $\alpha_R - \alpha_L$, erhält das Gehirn schließlich die Information über die Tiefenordnung der Punkte.

Während die Disparität im Auge auftritt, wird die Abweichung der beiden Ansichten in Abbildungen vor den menschlichen Augen als Parallaxe bezeichnet. Bei der Betrachtung der Leinwand entspricht sie dem Grad der Verschie-

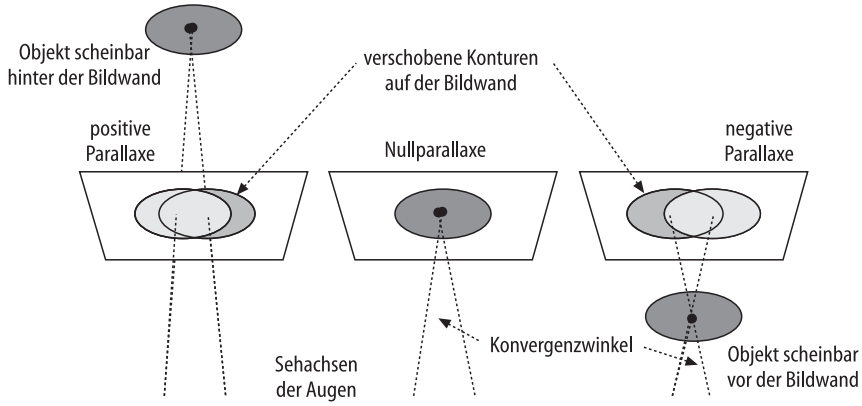


Abb. 2.13. Darstellung der positiven und negativen parallaktischen Verschiebung

bung, mit dem die Konturen der übereinander gezeigten Bilder dargestellt sind (Abb. 2.13). Sie wird in Abbildung 2.11 im mittleren Bild deutlich, das die Überlagerung zeigt, z. B. an der horizontalen Verschiebungen des Kerzenständers im Vordergrund und des Stuhls im Hintergrund. Je nach Tiefenwirkung differieren die Konturen oder Ansichten mehr oder weniger. Nicht verschobene Konturen liegen direkt übereinander und bilden die Nullparallaxe (Zero Parallax Setting, ZPS). In Abbildung 2.11 liegt sie auf dem Fensterrahmen im Mittelgrund.

Wenn die Nullparallaxe bei einer Projektion im Kino auf der Bildwand liegt, erscheinen die Objekte mit unverschobenen Konturen in einem Abstand, der dem Abstand zur Bildwand entspricht (Abb. 2.13). Die Blickachsen laufen dabei aufeinander zu und treffen sich auf der Bildwand, wobei der Konvergenzwinkel vom Betrachtungsabstand abhängt.

Die Abweichung von der Nullparallaxe verändert die Tiefenwahrnehmung. Die Abweichung kann positiv und negativ sein. Eine positive Parallaxe liegt vor, wenn die beiden Halbbilder so verschoben sind, dass sich die Sehachsen scheinbar hinter der Bildwand treffen (Abb. 2.13). Ein Objekt in dem dem rechten Auge zugeordneten Bild ist dann weiter nach rechts verschoben als im anderen. In Abbildung 2.11 wird das an dem Stuhl und der Figur deutlich, die im rechten Bild relativ zum Fensterrahmen weiter rechts liegt als im linken. Die meisten Objekte sollten in der Tiefe mehr oder weniger weit hinter der Bildwand positioniert sein, in der Regel sollte also die Parallaxe positiv sein.

Falls sich die Sehachsen vor der Bildwand kreuzen, liegt negative Parallaxe vor (Abb. 2.13), dann scheinen sich Objekte vor der Bildwand, nahe an den Zuschauern zu befinden. In Abbildung 2.11 wird die negative Parallaxe an dem im Vordergrund stehenden Kerzenständer sichtbar. Dieser befindet sich im rechten Bild relativ zum Fensterrahmen nämlich weiter links als im linken Bild. Die negative Parallaxe führt aufgrund der Objektnähe zum Zuschauer zu sehr spektakulären 3D-Erlebnissen. Beispiele für dramatische Darstellungen sind Waffen oder Tiere, die den Zuschauern scheinbar sehr nahe kommen. Sie sollten jedoch vorsichtig verwendet werden. Insbesondere ist zu vermeiden, dass diese Objekte vom Bildrand geschnitten werden, der das sogenannte

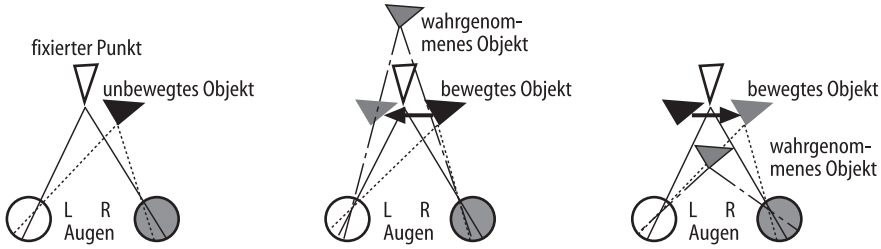


Abb. 2.14. Entstehung des Pulfrich-Effekts bei einseitig abgedunkeltem Auge

Scheinfenster bildet. Dies in der Natur nicht möglich und diese Scheinfensterverletzung zerstört die 3D-Illusion aufgrund eines kognitiven Konflikts, da die Menschen aus Erfahrung wissen, dass Vordergrundobjekte Rahmen im Hintergrund verdecken, aber nicht von ihnen beschnitten werden können. Objekte, die nur aus der Bildmitte scheinbar in den Zuschauerraum hinein ragen, sind unproblematisch. Das gilt auch für den Kerzenständer in Abbildung 2.11.

Die Tiefenhinweise aus dem beidäugigen Sehen werden bei der natürlichen Wahrnehmung durch Tiefenhinweise ergänzt, die schon beim einäugigen Sehen gewonnen werden können (monokular). Bei der Umsetzung in stereoskopischen Systemen kann es aber statt einer Ergänzung auch zu einem Widerspruch, einer sogenannten Rivalität der Eindrücke, kommen, die bei der Bildgestaltung möglichst vermieden werden sollten. Monokulare Tiefenhinweise sind die lineare Perspektive, wie z. B. sich in der Ferne verjüngende Linien, die in Abbildung 2.11 am Fensterrahmen und der veränderten Ansicht des Kerzenständers deutlich werden. Entscheidend ist auch die relative Größe, d. h. die Entfernungsabschätzung von Objekten aufgrund ihrer bekannten Größe, wie der Kerzenständer im Vordergrund in Relation zu einem Stuhl im Hintergrund. Hinzu kommen Aspekte wie die Verdeckung (Okklusion) eines entfernteren Objekts, wie bei der Figur in Abbildung 2.11, die im linken Bild vom Kerzenständer verdeckt ist. Weitere Gesichtspunkte sind u. a. der Objektschatten und die Akkommodation des Auges.

Zwei Tiefenhinweise ergeben sich nur bei Bewegung. Das sind die Bewegungsparallaxe, d. h. die Erscheinung, dass sich bei Vorbeifahrt die entfernteren Objekte scheinbar weniger schnell bewegen als die nahen und der Pulfrich -Effekt. Dieser beruht auf der Beobachtung Pulfrichs, dass eine Bewegung von einem abgedunkelten Auge etwas später wahrgenommen wird als vom anderen und bei Bewegung von links nach rechts näher erscheint (Abb. 2.14).

Die Betrachtung stereoskopischer Bilder kann mit einiger Anstrengung verbunden sein, denn die Verschmelzung der Halbbilder gelingt nicht allen Menschen gleich gut. Insbesondere fällt es einigen Menschen schwer, die Fokussierung der Augen (Akkommodation), die bei s3D-Wahrnehmung auf die Bildwand fixiert ist, von der Konvergenz zu entkoppeln, die sich mit der Raumtiefenwahrnehmung ändert, denn gewöhnlich verändern sich beide zusammen. Es entsteht der sogenannte Akkommodations-Konvergenz-Konflikt. Die Lösung des Konflikts gelingt manchen Menschen schlecht oder gar nicht und ist dann eine der Ursachen von Unwohlsein, der 3D-Sickness.

2.2 Der Bildaufbau

Um eine elektronische Übertragung von Bildern zu erreichen, werden diese auf eine Bildwandlerfläche abgebildet und in Signale umgesetzt. Die durch die optische Abbildung erzeugten Helligkeitswerte werden räumlich und zeitlich diskretisiert und stehen dann nach der Umsetzung als Informationen über die einzelnen Bildpunkte in reduzierter, jedoch immer noch sehr großer Zahl gleichzeitig bereit. Die parallele Übertragung aller Informationen wäre sehr unwirtschaftlich. Ein wesentlicher Gedanke ist daher, die Bildpunktinformationen seriell statt parallel zu übertragen (Abb. 2.15). Wenn die Abtastung des Bildes, die Umsetzung und der Bildaufbau bei der Wiedergabe schnell genug vor sich gehen, erscheint dem menschlichen Auge ein ganzes Bild, obwohl zu jedem Zeitpunkt nur ein Bildpunkt übertragen wird.

Um ein Fernsehprogramm an viele Zuschauer verteilen zu können, sind technische Normen für das Videosignal erforderlich. Die teilweise recht komplizierte Form des Signals lässt sich aus historischer Perspektive begründen. Die Videotechnik war zu Beginn reine Fernsehtechnik. Als Broadcastsystem ging es darum, die hohe Informationsdichte auf ökonomische Weise an möglichst viele Zuschauer zu übertragen. Ein wesentlicher Punkt ist dabei die Verfügbarkeit preiswerter Fernsehempfangsgeräte. Diese basierten zu Zeiten der Systemkonzeption auf der Bilddarstellung mittels der Kathodenstrahlröhre (Bildröhre, Cathode Ray Tube, CRT), die im nächsten Abschnitt dargestellt wird. Die Art des Bildaufbaus, die Form von Synchronsignalen und die Definition der Signaldarstellung ist direkt an dieses Display gekoppelt. Die Parameter sind aus Kostengründen eindeutig festgelegt. Das Display der Endgeräte hat somit bis heute einen großen Einfluss auf die Entwicklung von Videosystemen.

Die enge Kopplung des Videosignaldefinitionen an die Bildröhre stellte lange Zeit kein Problem dar, weil erst zu Beginn des 21. Jahrhunderts alternative Displaysysteme zu akzeptablen Preisen zur Verfügung standen. Die heute verfügbaren LCD- und Plasmabildschirme sind jedoch nicht mehr gut an die Standarddefinitionen des Videosignals angepasst.

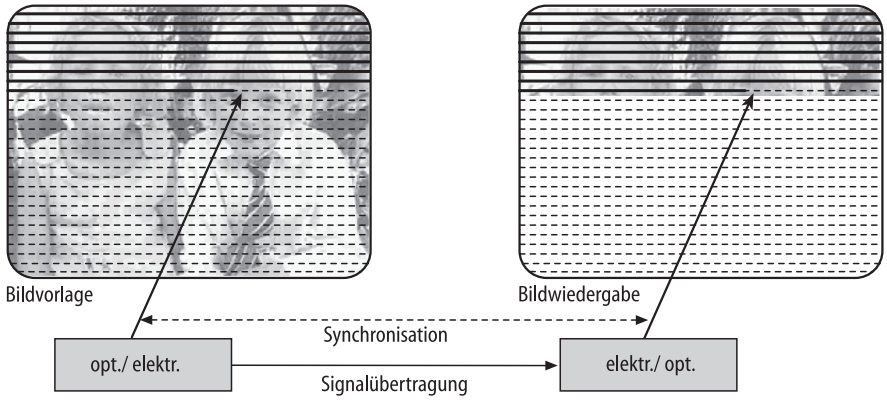


Abb. 2.15. Grundprinzip der Bildübertragung

2.2.1 Das diskretisierte Bild

Als sehr wichtiger Parameter des Videosignals wird zunächst die Bildwechselfrequenz und damit die zeitliche Diskretisierung festgelegt. Es gibt bis heute weltweit zwei dominierende Fernsehnormen, aus Europa und Amerika, die sich vor allem bezüglich der Bildwechselfrequenz unterscheiden. Man wählte für Europa mit 25 Hz und für Amerika mit 30 Hz jeweils die halben Netzfrequenzen. Damit sind einerseits die Bildwechsel schnell genug, um Bewegungen als fließend erscheinen zu lassen und andererseits können so bei der Wiedergabe störende Interferenzen mit den Stromversorgungen der TV-Geräte und bei der Aufnahme Interferenzen mit Beleuchtungseinrichtungen vermieden werden.

Analog zum Nipkow-Prinzip wurde festgelegt, dass jedes einzelne Bild zeilenweise abgetastet wird. Die Helligkeit eines jeden Bildpunktes ruft im Bildwandler ein elektrisches Signal hervor. Die Umsetzung der Information in serielle Form geschieht heute allerdings nicht mehr mit mechanischen, sondern mit elektronischen Mitteln. Aus dem räumlichen Nebeneinander der Bildpunktwerte wird ein zeitliches Nacheinander [194].

Eine einfache Möglichkeit zur Parallel-Seriell-Umsetzung bietet die Nutzung des Elektronenstrahls in einer Braunschen Röhre (Abb. 2.16). Beim Röhrenbildwandler (s. Kap. 6.2) werden Elektronen in einer Vakuumröhre mittels einer geheizten Kathode erzeugt und durch elektrische Felder zu einem Strahl gebündelt. Der Elektronenstrahl wird durch magnetische Felder abgelenkt und zeilenweise über eine lichtempfindliche Schicht geführt. Dabei ändert sich der Stromfluss in Abhängigkeit von der Bildpunkthelligkeit, da der Widerstand der Schicht von der Lichtintensität abhängt.

Auf die gleiche Weise wird auch auf der Wiedergabeseite gearbeitet. Die Intensität des Elektronenstrahls in der Wiedergaberöhre wird vom Videosignal gesteuert. Der Strahl wird in gleichem Rhythmus wie bei der Aufnahme über eine lichtemittierende Schicht geführt, die in Abhängigkeit von der Stromstärke der auftreffenden Elektronen mehr oder weniger stark zum Leuchten angeregt wird (Abb. 2.16).

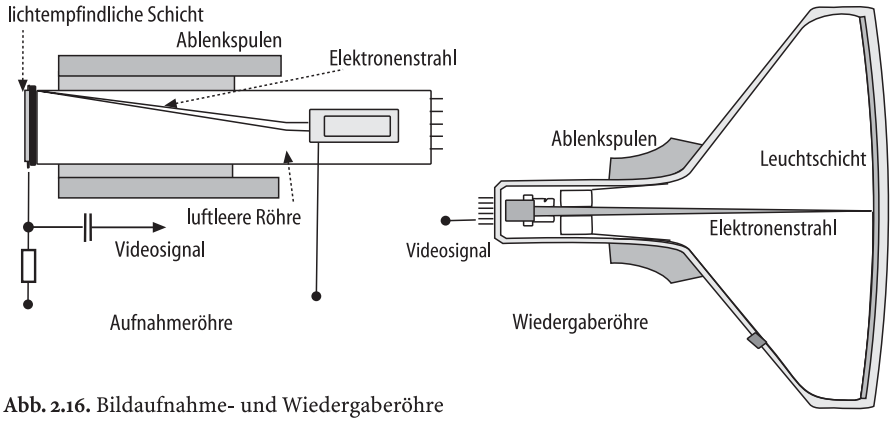


Abb. 2.16. Bildaufnahme- und Wiedergaberöhre

Abb. 2.17. Vergleich von Bildern mit 100 und 400 Zeilen

Aufgrund der Zeilenstruktur ist das Bild bei Röhrenbildwandlern nur in der Vertikalen diskretisiert oder gerastert, bei modernen Halbleiterbildwandlern (CCD, CMOS) ist auch jede Zeile in klar abgegrenzte Bildpunkte (Picture Elements, Pixel) zerlegt, sodass das Bild in beiden Dimensionen der Bildfläche (Zeilen und Spalten) und auch bezüglich der Zeit diskretisiert ist. Je mehr Bildpunkte vorhanden sind, desto schärfer erscheint das Bild (Abb. 2.17).

Die Zeilen- und Spaltenstruktur bildet eine Abtaststruktur mit einer räumlichen Periodizität und den Perioden x_T und y_T (Abb. 2.18), denen die Abtastfrequenzen $f_{Tx} = 1/x_T$ und $f_{Ty} = 1/y_T$ zugeordnet werden können. Die Strukuren der Bildvorlage selbst können mit Raum- oder Ortsfrequenzen f_x und f_y beschrieben werden. Große Bildstrukturen führen zu wenigen ortsabhängigen Wechseln, also kleinen Ortsfrequenzen und umgekehrt. Um eine störfreie Bildwandlung zu gewährleisten, muss für alle drei Dimensionen das Abtasttheorem erfüllt sein (s. Kap. 3.2.1), d. h. das abgetastete Signal muss derart auf Orts- und Zeitfrequenzen begrenzt sein, dass gilt:

$$f_x < f_{Tx}/2 = 1/(2\ x_T),\ f_y < f_{Ty}/2 = 1/(2\ y_T),\ f < f_T/2.$$

Die störungsfreie Verarbeitung der diskretisierten Daten erfordert bezüglich aller drei Dimensionen eine Tiefpass-Vorfilterung auf der optischen Seite, also vor der Wandlung, so wie eine Nachfilterung im Empfänger.

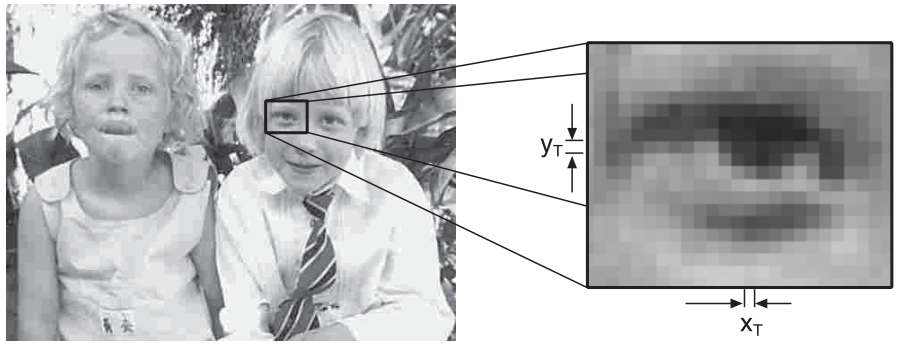


Abb. 2.18. Pixelstruktur des Bildes

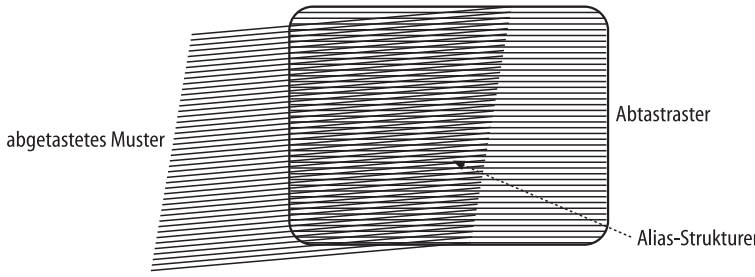


Abb. 2.19. Entstehung von Alias-Strukturen

Eine effektive Filterung ist aufwendig und wurde bei der TV-Konzeption zu Beginn der Fernsehentwicklung nicht vorgesehen. Als Vorfilterung dient nur die unvermeidbare Unschärfe, die aufgrund der endlichen Öffnung (Apertur) des abtastenden Elektronenstrahls entsteht, bzw. eine bewusst herbeigeführte Unschärfe bei Halbleiterbildwandlern. Eine elektronische Nachfilterung im Empfänger findet in Standard-TV-Systemen nicht statt.

Damit entstehen zwangsläufig Bildverfälschungen (Aliasanteile) zeitlicher Art wie Flimmern, ruckende Bewegungen oder sich scheinbar rückwärts drehende Speichenräder. Als räumliche Fehler treten Treppenstrukturen an Diagonalen auf und Alias-Strukturen, die im Original nicht enthalten sind. Letztere ergeben sich bereits, wenn man durch zwei Siebe oder regelmäßige Muster schaut. Eines entspricht dem Abtastraster, das zweite der abgetasteten Struktur (Abb. 2.19).

Die Filterung auf der Wiedergabeseite wird zum Teil von den Zuschauern vorgenommen, die einen Abstand zum Wiedergabegerät einnehmen, der größer ist als der eigentlich erforderliche Abstand, bei dem zwei Zeilen des Bildes gerade nicht mehr unterschieden werden können. Die Zuschauer führen also absichtlich eine Auflösungsverminderung herbei, um weniger von den Aliasanteilen gestört zu werden. Diese Betrachtung kann als Begründung für die Untersuchungsergebnisse von Kell et al. angeführt werden, die 1934 aufgrund von Versuchsreihen an Testpersonen den Faktor der Abstandsvergrößerung mit einem Wert von ca. 1,5 bestimmten. Der Kehrwert, der Kellfaktor $k \approx 0{,}67$ kann auch so interpretiert werden, dass mit einem Fernsehsystem, das mit 100 Zeilen arbeitet, nur etwa 67 Zeilen einer Bildvorlage aufgelöst werden.

2.2.2 Zeilenzahl und Betrachtungsabstand

Die Anzahl der Zeilen, in die das Bild zerlegt wird, wurde zu Beginn der Fernsehtechnik von kleinen Zahlen an immer weiter gesteigert und international verschieden festgelegt, z. B. auf Werte von 441, 405 oder 819 (s. Kap. 1). Schließlich etablierten sich dann die Zahlen 525 und 625, die bis heute für Standardfernsehsysteme gültig sind und zwar für den amerikanisch und den europäisch dominierten Teile der Erde bei Bildfrequenzen von 30 Hz und 25 Hz. Diese Zahlen lassen sich in Bezug auf die menschliche Bildwahrnehmung begründen.

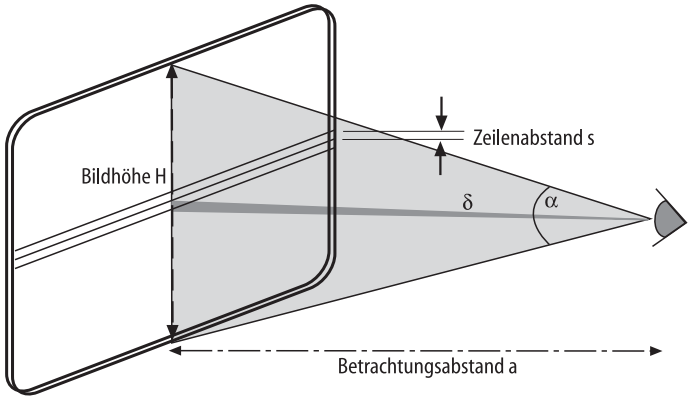


Abb. 2.20. Betrachtungsabstand und Zeilenauflösung

Zunächst wird der Betrachtungsabstands a bestimmt. Mit Abbildung 2.20 wird deutlich, dass der Betrachtungsabstand a aus dem Winkel $\alpha/2$ und der halben Bildhöhe H/2 folgt. Es gilt:

$$\tan \alpha/2 = H/2a, \text{ bzw. } a = H/(2 \tan \alpha/2).$$

Auf ähnliche Weise ergibt sich der Zeilenabstand s. Wenn der Betrachtungsabstand a so gewählt wird, dass zwei Zeilen gerade nicht mehr unterschieden werden können, so hängt s vom Auflösungsvermögen des Auges und damit vom zugehörigen Grenzwinkel δ ab, und es gilt:

$$\tan \delta/2 = s/2a.$$

Der Zeilenabstand s hängt schließlich wiederum mit der Zeilenzahl Z zusammen. Es gilt:

$$s = H/Z$$

und daraus folgt

$$Z = H/s = (\tan \alpha/2) / (\tan \delta/2).$$

Standardfernsehsysteme sind als sogenannte Televisionssysteme konzipiert. Die Zuschauer haben einen relativ großen Abstand zum Wiedergabegerät, das Bild ist mit einem Blick erfassbar, wobei für den vertikalen Sehwinkel α Werte zwischen 10° und 15° und für den horizontalen Sehwinkel β 13°–20° angegeben werden. Das Bildseitenverhältnis Breite/Höhe wird daher als B/H = 4/3 gewählt. Mit der Annahme $\alpha = 14°$ folgt für den Betrachtungsabstand:

$$a = H/(2 \tan \alpha/2) = 4 H.$$

Bei Berücksichtigung des Kellfaktors ergibt sich eine Abstandsvergrößerung um etwa den Faktor 1,5, also $a' = 6 H$.

Als Beispiel kann ein Fernsehgerät mit 50 cm sichtbarer Bilddiagonale herangezogen werden. Es hat eine Bildhöhe von 30 cm und mit $a = 4 H$ würde der Betrachtungsabstand 1,2 m betragen. Man kann leicht feststellen, dass die meisten Menschen einen Betrachtungsabstand $a' \approx 1{,}8$ m bevorzugen.

Das maximale Auflösungsvermögen des Auges liegt bei einer Winkelminute, häufig wird aber auch mit $\delta = 1{,}5'$ gerechnet. Mit den Annahmen $\alpha = 14°$ und $\delta = 1{,}4'$ folgt daraus für die Zeilenzahl Z:

$$Z = H/s = (\tan \alpha/2) / (\tan \delta/2) \approx 600.$$

Europäische Standardfernsehsysteme arbeiten mit 625 Zeilen, wovon maximal 575 sichtbar (aktiv) sind, in Amerika werden 525 mit 485 sichtbaren Zeilen verwendet. Daraus resultieren Zeilenfrequenzen, nämlich $625 \cdot 25$ Hz = 15625 Hz und $525 \cdot 30$ Hz = 15750 Hz, die nahezu gleich sind.

Abbildung 2.21 zeigt den Vergleich der Bildwirkung des gerade beschriebenen, sogenannten SDTV-Systems (Standard Definition TV) mit der von HDTV,

Abb. 2.21. Vergleich der Bildwirkung von Standard-TV und HDTV

Tabelle 2.3. SD- und HD-Bildaufbauparameter

	SD		HD	
Gesamtzeilenzahl	625	525	750	1125
Länderspezifika	EU	US/J	international	
Bildaufbau	i	i	p	i
Aktive Zeilen	576	486	720	1080
Bildseitenverhältnis	4:3		16:9	
Vollbildfrequenz	25 Hz	29,97 Hz	24 / 25 / 29,97 / 30 / 50 / 60 Hz	

dem High Definition Television. Nachdem die Definitionen für SDTV seit mehr als 50 Jahren gültig sind, wurde seit Beginn des 21. Jahrhunderts die Einführung von HD-Systemen betrieben. Dabei können die Zuschauer aufgrund der gesteigerten Bildauflösung einen kleineren Abstand zum Bildschirm einnehmen, sodass auch die Blickseitenbereiche das Bild erfassen. Das Blickfeld wird weitaus mehr gefangen genommen als bei SDTV-Systemen. Für diese Bildwirkung wird das Wort Telepräsenz verwendet.

Bei dem Übergang von SD- zu HD-Videosystemen wird die Zeilenzahl ungefähr verdoppelt. Außerdem wird das Bildseitenverhältnis von 4:3 auf 16:9 verändert, sodass sich bei einem Betrachtungsabstand von ca. 3 H ein kinoähnliches Sehgefühl einstellen kann. Der Informationsgehalt ist damit mehr als verfünffacht. Um das bei gegebenem Betrachtungsabstand nutzen zu können, sind große Displays erforderlich. Bei einem Betrachtungsabstand von a = 3 m ist mit a = 3 H eine Bilddiagonale von ca. 1,8 m erforderlich. Dafür können, wie im Kino, Projektionssysteme eingesetzt werden. Alternativ stehen mit dem Plasmadisplay und LC-Bildschirmen auch flache Direktsichtdisplays zur Verfügung, deren Bilddiagonalen immer weiter vergrößert werden (s. Kap. 7).

Der HD-Bereich ist heute etabliert. Da HD in analoger Form in der Praxis jedoch selten auftritt, werden die HD-Signale genauer im Kapitel über Digitalsignale dargestellt, in diesem Kapitel geht es um SD-Signale. Es sei nur kurz erwähnt, dass es im HD-Bereich zwei Normen gibt, die bezüglich der Bildauflösung beide international gültig sind. Tabelle 2.3 zeigt die Parameter zusammen mit denen der SD-Systeme. Während sich im HD-Bereich zunächst nur ein Format mit 1125 Zeilen brutto (Nettoauflösung 1920 x 1080) durchgesetzt hatte, kam durch die Vorgaben bei der Einführung des digitalen Fernsehens in den USA die zweite Variante mit 750 Zeilen brutto (Netto 1280 x 720) hinzu.

Neben den SD-Systemen arbeitet auch die HD-Variante mit 1080 Zeilen gewöhnlich im interlaced Mode (1080i) (s. u.), die Variante mit 720 Zeilen gibt es nur mit progressiver Abtastung (720p). Die Abtastformen interlaced und progressiv werden im nächsten Abschnitt erklärt.

Auch Computermonitore haben hohe Zeilenzahlen, denn auch hier beträgt der Betrachtungsabstand nur etwa a ≈ 3 H, z. B. a ≈ 70 cm bei einem 17"-Monitor (≈ 23 cm Bildhöhe). Damit ergibt sich ein Blickwinkel mit tan $\alpha/2$ = 0,15 und eine erforderliche Zeilenzahl Z ≈ 800. Der Kellfaktor wird hier nicht berücksichtigt, denn das statische Computerbild ist relativ frei von Aliasstrukturen. Computermonitore sollten mit Bildwechselfrequenzen > 70 Hz arbeiten, da bei großen Blickwinkeln das Großflächenflimmern stark in Erscheinung tritt.

2.2.3 Das Zeilensprungverfahren

Wenn alle Bildzeilen Zeile für Zeile nacheinander fortschreitend abgetastet werden, spricht man von progressivem Bildaufbau. Diese Form der Abtastung wird jedoch nicht bei Standard-TV-Systemen verwendet, denn wäre das der Fall, würde das Bild wegen der geringen Bildwechselfrequenz von 25 Hz stark flimmern. Zur Beseitigung des Großflächenflimmerns muss die Bildwechselfrequenz auf mindestens 50 Hz erhöht werden (s. Abschn. 2.1.2), was bei Beibehaltung der progressiven Abtastung zu doppelter Signalbandbreite führt. Eine andere Möglichkeit der Bildfrequenzverdopplung ist die zweimalige Wiedergabe eines jeden Bildes, bzw. die Unterbrechung des Lichtstroms, wie es in der Kinotechnik üblich ist. Dieses Verfahren setzt aber einen Bildspeicher, bzw. das gleichzeitige Vorliegen aller Bildpunkte voraus, was erst mit den sogenannten 100-Hz-Fernsehgeräten realisiert werden konnte.

Da es zu Beginn der Fernsehentwicklung unmöglich war, jeden Empfänger mit einem Speicher auszustatten, wurde das Zeilensprungverfahren eingeführt (2:1, engl.: Interlaced-Mode). Das Gesamtbild wird dabei bereits bei der Aufnahme so erzeugt, dass es in zwei Teilbilder aufgeteilt ist die ineinander verkämmt sind. Sie werden beim 50-Hz-System in der Phase der Gesamtbilddauer von 40 ms mit jeweils 20 ms Dauer abgetastet, wodurch sich die Flimmerfrequenz verdoppelt. Abbildung 2.22 zeigt das Grundprinzip des Zeilensprungverfahrens. Bei raumorientierter Zeilenzählung enthält das erste Teilbild die ungeraden, das zweite Teilbild die geraden Zeilen.

Das Gesamtbild (Frame) des Standardfernsehsystems mit 625 Zeilen wird in zwei Halbbilder (Fields) mit je 312,5 Zeilen zerlegt. Die örtlich zweite Zeile wird zeitlich als 313. übertragen. Aufgrund der ungeraden Zeilenzahl (625 bzw. 525) wird die letzte Zeile des 1. Halbbildes nur halb geschrieben, der Elektronenstrahl springt senkrecht nach oben und schreibt als zweite Hälfte die erste (Teil-) Zeile des 2. Halbbildes. Nach Beendigung des 2. Halbbildes wird der Elektronenstrahl schließlich von rechts unten nach links oben geführt, und das nächste Bild wird wiedergegeben. Während des Strahlrücksprungs wird eine Dunkelphase eingefügt, um den Strahl unsichtbar zu machen (s. Abschn. 2.2.5). Damit ist die Anzahl der Dunkelphasen und somit auch die Flimmerfrequenz verdoppelt.

Die ungerade Zeilenzahl wurde gewählt, damit der Strahlrücksprung für beide Halbbilder zwischen den gleichen Vertikalspannungsniveaus stattfinden kann. Bei gerader Zeilenzahl müssten diese Spannungen halbbildweise leicht

Abb. 2.22. Grundprinzip des Zeilensprungverfahrens

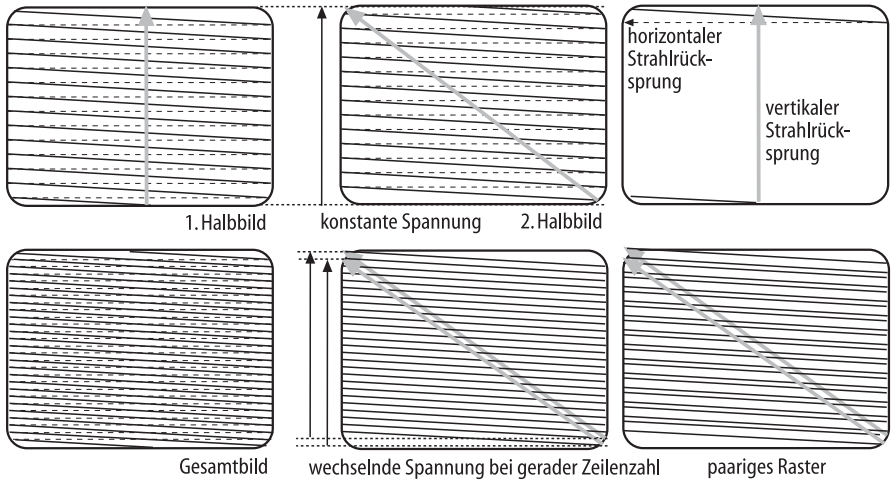


Abb. 2.23. Zeilensprungverfahren bei ungerader Zeilenzahl

verändert werden. Dabei besteht die Gefahr, dass sich eine paarige Rasterstruktur ergibt, die Zeilen des zweiten Halbbildes also nicht exakt zwischen denen des ersten liegen (Abb. 2.23).

In den Abbildungen 2.23 und 2.24 sind die Zeilen schräg dargestellt, um deutlich zu machen, dass aus Gründen einer möglichst einfachen technischen Realisierung die Ablenkspannungen für den Elektronenstrahl horizontal und vertikal als Sägezahnsignal ununterbrochen durchlaufen (Abb. 2.24). Die vertikalen Strahlrücksprünge sind dabei idealisiert (unendlich kurz) dargestellt, um das Bild erkennbar zu halten. Die Zeilen in den Dunkelphasen existieren nicht örtlich, aber zeitlich.

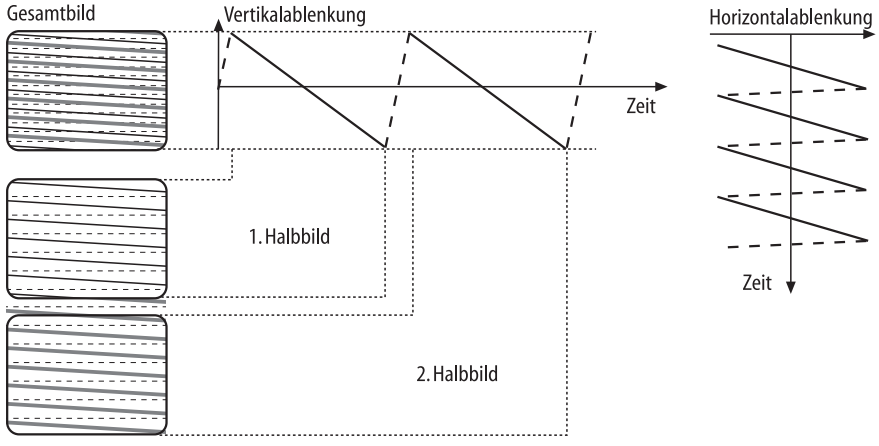


Abb. 2.24. Zeilensprungverfahren und Ablenkspannungen für Röhrenmonitore

2.2.3.1 Zeilensprungartefakte

Das Zeilensprungverfahren kann als effektives Beispiel für eine Informationsreduktion im Analogbereich betrachtet werden. Es ermöglicht bei Inkaufnahme geringer Bildverschlechterung die Halbierung der erforderlichen Übertragungsbandbreite bzw. Informationsmenge. Die Bildverschlechterung betrifft folgende Artefakte: Vertikal bewegte Bilder treten gegenüber progressiver Bildabtastung mit verringerter Vertikalauflösung auf. Dieser Verlust, der sogenannte Interlace-Faktor, ist bildinhaltsabhängig und wird mit durchschnittlich 30% angegeben. Zusammen mit dem Kellfaktor rechnet man mit einem Gesamtverlust der Auflösung von 50%. Bei unbewegten Bildvorlagen, aber bewegtem Auge, scheint das Bild vertikal zu wandern. Bei bewegten Bildvorlagen geben beide Halbbilder verschiedene Bildteile wieder, es entsteht ein Moiré an horizontalen Bildstrukturen. Dünne Linien, die nur eine Breite von einer Zeile haben, werden nur von einem Halbbild wiedergegeben und flackern mit 25 Hz. Dies wird besonders deutlich, wenn computergenerierte Bilder im Zeilensprungverfahren wiedergegeben werden. Daher werden die Grafiken für Videosysteme so gestaltet, dass horizontale Linien mit der Breite von nur einer Zeile nicht auftreten.

Neben den Artefakten bringt das Zeilensprungverfahren auch praktische Nachteile mit sich, die sich vor allem bei der digitalisierten Form des Videosignals zeigen. Hier werden die Halbbilder nicht durchgängig getrennt behandelt, sondern oft zu einem Vollbild zusammengesetzt. Wenn dann die Teilbilder aus einer Szene mit schnell bewegten Inhalten stammen, entstehen an den Bildkonturen regelmäßig ausgefranste Kanten. Abbildung 2.25 zeigt diese Fehler am Beispiel einer animiert bewegten Schrifteinblendung. Der Grund ist, dass die örtlich benachbarten Zeilen aus zwei zeitlich unterschiedlichen Bewegungsphasen stammen. Im europäischen Fernsehsystem mit 25 Bildern pro Sekunde dauert jedes Vollbild 40 ms. Die beiden Halbbilder haben also einen Zeitabstand von 20 ms. Wenn die Schrift z. B. in einer Sekunde, oder 50 Halbbilddauern, über den Bildschirm bewegt wird, unterscheiden sich die Strukturen der Halbbilder zeitlich um 1/50 und damit auch örtlich um zwei Prozent.

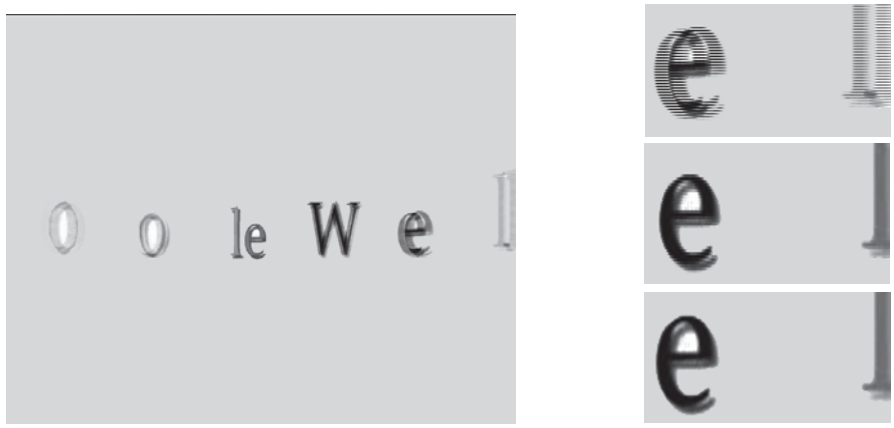


Abb. 2.25. Zeilensprungartefakte bei Zusammensetzung zweier Halbbilder

Das Problem der ausgefransten Kanten kann gelöst werden, indem auf der Aufnahmeseite statt der 50 Halbbilder 25 Vollbilder erzeugt werden, die erst nachträglich in zwei Teilbilder zerlegt werden. In diesem Fall spricht man von segmented frames. Das Verfahren kann erst in jüngerer Zeit mit Hilfe von Bildspeichern realisiert werden. Es wird im Fernsehbereich nicht verwendet, denn es passt nur zu Filmbildern, die mit 24 fps oder 25 fps erzeugt werden. Für die Fernsehtechnik, die z. B. auch für Sportübertragungen verwendet wird, ist in diesem Fall die Bewegungsauflösung zu gering, denn bei 25 Vollbildern wird eine Bewegungsphase nur alle 40 ms neu erfasst, während bei 50 Halbbildern die Bewegungsauflösung doppelt so hoch ist. Auch die Schrifteinblendung aus Abbildung 2.25 würde dann ruckender erscheinen.

Dass sich in modernen Videosystemen die Halbbildproblematik verschärft, zeigt sich auch an den Flachdisplays. Diese sind so entwickelt worden, dass sie u. a. als Computerdisplays nutzbar sind und arbeiten beim Bildaufbau mit progressiver Adressierung. Die Darstellung eines Zeilensprungsignals erfordert daher im Endgerät eine Beseitigung des Zeilensprungs (das sogenannte deinterlacing). Deinterlacing geschieht im einfachsten Fall durch Beschränkung auf ein Halbbild, dann ist allerdings die vertikale Ortsauflösung halbiert. Eine zweite Möglichkeit ist die Zeilenwiederholung (Abb. 2.25 rechts, Mitte). Die beste Qualität zeigt die Interpolation, also die Zusammenrechnung der Teilbilder (Abb. 2.25 rechts, unten). Um eine gute Qualität zu erzielen, ist dafür einiger Aufwand erforderlich [152]. In guten Systemen geschieht sie bewegungskompensiert, was für die Bewegungserkennung klare Strukturen, wie bei der Schrift, erfordert. Generell lässt sich Deinterlacing nicht verlustfrei realisieren – ein gewichtiger Grund bei der Entscheidung zwischen den beiden HD-Auflösungen 1080i und 720p [189].

2.2.3.2 Halbbilddominanz

Im Zuge der Digitalisierung der Videotechnik tritt auch immer häufiger das Problem der Festlegung der Halbbilddominanz auf, d. h. die Frage, welches Halbbild das erste ist. Abhängig vom Videosignal können die Zeilen des ersten Halbbildes oben (upper field) oder unten liegen (lower field). Digitalsysteme zum Schnitt, Compositing, DVD-Authoring etc. erlauben nun die Wahl der Halbbilddominanzen, um Material aus verschiedenen Quellen zusammen verwenden zu können. Das führt jedoch zu dem Problem, dass der Anwender sich damit auskennen muss. Hinzu kommt, dass die Begriffswahl sehr verwirrend ist. Die Einstellparameter werden für das erste Halbbild (first field) mit upper oder lower aber auch mit odd oder even, also ungerad- und geradzahlig bezeichnet. Diese Begriffe ergeben allerdings nur einen Sinn, wenn sie sich auf raumorientierte Zählung der aktiven Zeilen beziehen.

Neben der uneinheitlichen Bezeichnung gibt es auch eine sehr uneinheitliche Benutzung der Bezeichnungen in den Handbüchern zu der Software. Um die korrekte Einstellung der Halbbilddominanz gewährleisten zu können, ist es am sichersten, die bearbeiteten Sequenzen auf einem separaten Monitor darzustellen, der im interlaced mode arbeitet und anhand von Sequenzteilen mit hoher Bewegung festzustellen, ob sich ein ruckender Bewegungseindruck ergibt, der aus der Vertauschung der Halbbilder resultiert.

2.3 Das monochrome Videosignal

Die abgetasteten Bildpunkte haben verschiedene Helligkeiten. Helle Bildpunkte werden durch ein hohes, dunkle Punkte durch ein geringes Videosignal repräsentiert. Die zugehörigen Grauwerte und Parameter sind für SD-Signale in der Empfehlung ITU-R BT.470 standardisiert, in ITU-R BT.709 für HD-Signale. Die International Telecommunication Union (ITU) ist neben der Society of Motion Picture and Television Engineers (SMPTE) eines der wichtigsten Standardisierungsgremien im Video- und Fernsehbereich.

2.3.1 Das BAS-Signal

Die Abkürzung BAS steht für Bild-, Austast- und Synchron und beschreibt ein Signal, das aus diesen drei Anteilen zusammengesetzt ist. Zunächst sei hier der Bildsignalanteil beschrieben.

2.3.1.1 Das Bildsignal

Der Unterschied zwischen dem dunkelsten und dem hellsten Punkt entspricht in Europa und bei HD-Systemen einer Spannungsdifferenz zwischen 0 V und 0,7 V. Oft wird als Minimalwert nicht exakt 0 V, sondern ein Wert zwischen 0 und 0,01 V eingestellt, der als Schwarzabhebung bezeichnet wird. Im amerikanischen und japanischen 525-Zeilensystem werden hier andere Werte verwendet (s. Abschn. 2.3.1.2). Abbildung 2.26 zeigt den Signalverlauf für eine Bildzeile. Aufgrund der gleichmäßigen Abtastung der Zeile kann der Ortskoordinaten x auch eine Zeitkoordinate t zugeordnet werden. Der dargestellte Signalverlauf entspricht damit auch dem Verlauf über der Zeit.

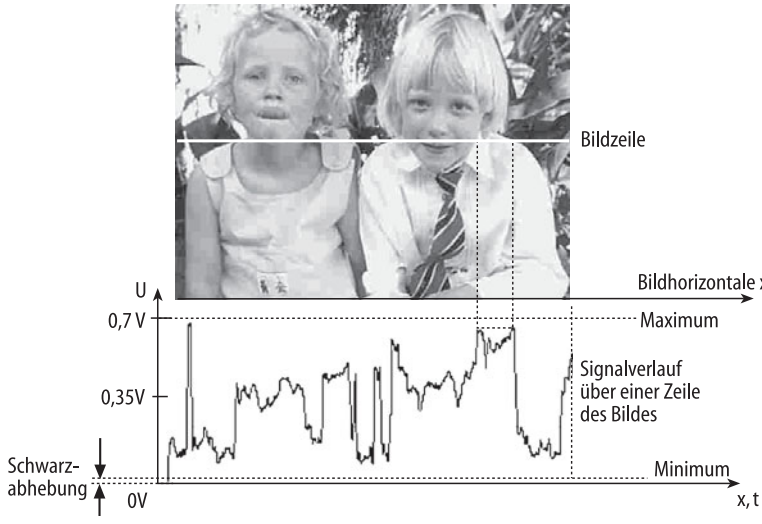


Abb. 2.26. Videosignalverlauf einer Bildzeile

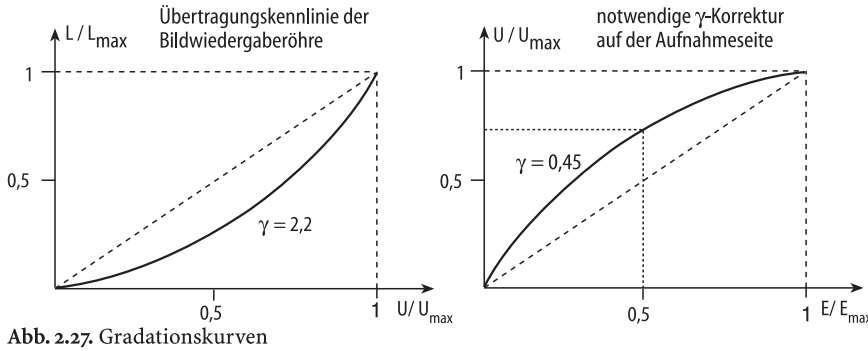


Abb. 2.27. Gradationskurven

Der Zusammenhang zwischen dem Signal (Spannung U) und der Helligkeit bei der Bildwiedergabe (Leuchtdichte L) ist nicht linear. Das Grauwertverhalten, die Gradation, kann durch eine Potenzfunktion beschrieben werden. Für SD-Systeme gilt:

$$L/L_{max} = (U/U_{max})^{\gamma_W}.$$

Dieses Gleichung resultiert aus der Steuerkennlinie der zur Bildwiedergabe eingesetzten Braunschen Röhren. Anstatt deren Nichtlinearität in jedem Empfänger auszugleichen, wurde beschlossen, die sogenannte γ-Vorentzerrung schon auf der Aufnahmeseite einzusetzen. Damit muss der γ-Wert für die Wiedergabe für alle verbindlich festgelegt werden. Er betrug zunächst $\gamma_W = 2{,}2$. Für die Aufnahmeseite folgte daraus, dass $\gamma_A = (2{,}2)^{-1} = 0{,}45$ eingestellt werden muss, um eine lineare »Über-Alles-Gradation« zu erzielen (Abb. 2.27). Auf der Aufnahmeseite gilt mit der Beleuchtungsstärke E und der Signalspannung U:

$$U/U_{max} = (E/E_{max})^{\gamma_A}.$$

Die Spezifikationen in der HD-Norm ITU-R BT.709 berücksichtigen einen Übergang zu einer linearen Grauwertzunahme bei Leuchtdichten, die kleiner als 1,8% des Maximalwertes sind. Für HD-Signale gilt:

$$U/U_{max} = 1.099\ (E/E_{max})^{0{,}45} - 0.099 \qquad \text{für } 1 \geq E/E_{max} \geq 0{,}018$$

$$U/U_{max} = 4{,}5\ (E/E_{max}) \qquad \text{für } 0{,}018 \geq E/E_{max} \geq 0$$

Die γ-Vorentzerrung wird bereits in der Kamera vorgenommen, da so auf den Einbau von γ-Ausgleichsschaltungen in Bildröhrenempfänger verzichtet werden kann. Allerdings muss sie dafür in die modernen Geräte (z. B. LCD) wieder implementiert werden. Im Vergleich zur linearen Übertragung ergibt sich mit γ-Vorentzerrung ein besserer Störabstand, da bei 50% Helligkeit bereits ein Signal mit $0{,}5^{0{,}45} = 73\%$ der Maximalspannung entsteht.

Eine Abweichung von den festgelegten γ-Werten ändert die Grauwertabstufungen. Während bei der Aufnahme die bewusste Beeinflussung der Grauwerte im Bild erlaubt ist (s. Kap. 6), müssten sich alle Hersteller von Wiedergabegeräten streng an die Vorgabe halten. Der Wert $\gamma_W = 2{,}2$ gilt heute nur noch für das 525-Zeilensystem. Für die anderen SD-Systeme ist in ITU-R BT.470 ein $\gamma_W = 2{,}8$ festgelegt. Damit ergibt sich eine »Über-Alles-Gradation« von ca. 1,2.

2.3.1.2 Austastlücken

In Europa wird mit 25 Bildern pro Sekunde und bei SD-Signalen im Zeilensprungverfahren gearbeitet. Damit beträgt die Bilddauer 40 ms (vertikale Periodendauer T_V) und die Halbbilddauer 20 ms. Während 40 ms werden 625 Zeilen geschrieben, woraus eine Zeilen- oder H-Periodendauer T_H = 64 µs bzw. die Zeilenfrequenz 15,625 kHz folgt. Die angegebene Zeit steht nicht vollständig zum Schreiben einer Zeile zur Verfügung, denn da die Signaldefinitionen eng an die Funktion der Bildröhre gekoppelt sind, muss berücksichtigt werden, dass, nachdem der Strahl am rechten Bildrand ankommt, er Zeit braucht, um wieder zum linken zurückspringen zu können. Gleichermaßen wird eine gewisse Zeit benötigt, damit er zum Bildanfang zurückgeführt werden kann nachdem er das Bildende erreicht hat.

Wegen des geringen Standes der Technik zu Zeiten der Systemkonzeption wurde für den horizontalen Strahlrücksprung ein recht großer Zeitraum von 12 µs reserviert. Damit kann die steile Flanke der Sägezahnablenkspannung relativ flach bleiben (Abb. 2.24) und es werden Störungen durch Einschwingvorgänge gering gehalten, was den Aufwand für die Konstruktion der Fernsehempfänger verringert. Der Zeitraum wird horizontale Austastlücke genannt, da der Elektronenstrahl beim Rücksprung nicht sichtbar sein darf und daher abgeschaltet (ausgetastet) wird (Abb. 2.28). Der Pegel des Austastwertes wird ist geringfügig kleiner als der Schwarzwert und ist mit 0 V festgelegt. Für das Bildsignal wird die Schwarzabhebung eingestellt, sodass der Schwarzpegel bei ca. 0–2% des Abstands zwischen Bildweiß- und Austastwert (BA), d. h. bei 0–2% des Betrages zwischen Austastwert und Maximalspannung (Weiß, entsprechend 0,7 V) liegt.

Für den Strahlrücksprung nach Beendigung eines Halbbildes wird jeweils eine Vertikalaustastlücke von 1,6 ms reserviert, d. h. für die Dauer von 25 Zei-

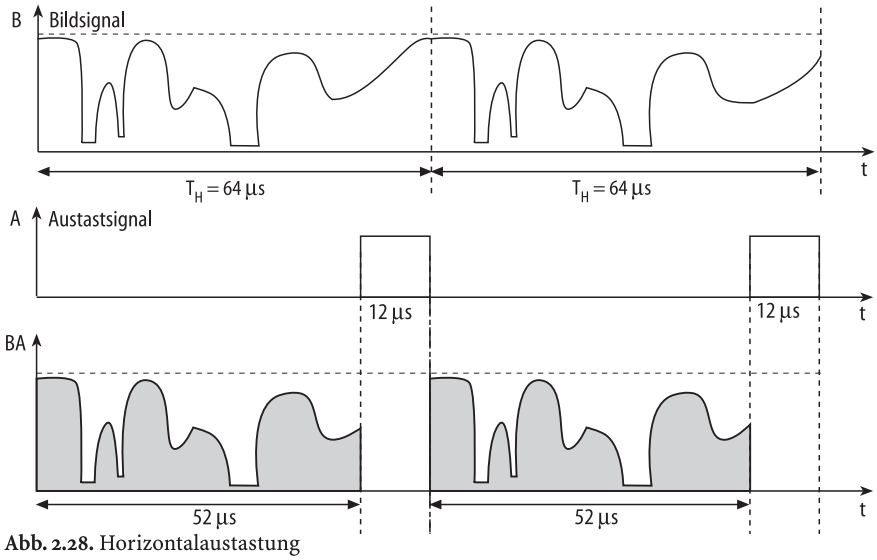


Abb. 2.28. Horizontalaustastung

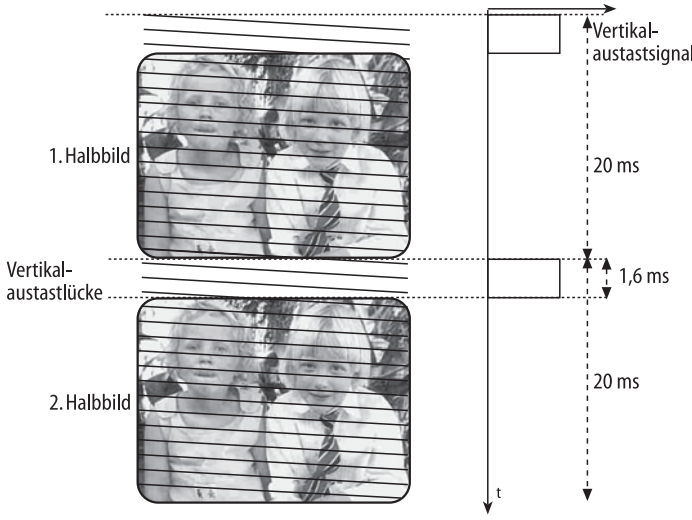


Abb. 2.29. Vertikalaustastung beim SD-Signal im 50 Hz-System

len pro Halbbild ist der Elektronenstrahl ausgeschaltet (Abb. 2.29). Im europäischen Fernsehsystem sind von den 625 Zeilen pro Bild also nur 575 Zeilen nutzbar. Die für das Bild nutzbare (aktive) Zeilendauer beträgt 52 µs, bei einer Gesamtzeilendauer von 64 µs.

Im 525-Zeilensystem wurde zunächst mit 30 Bildern pro Sekunde gearbeitet. Von den 525 Zeilen werden pro Halbbild 20 Zeilen ausgetastet, sodass 485 aktive Zeilen zur Verfügung stehen. Die Bildfrequenz wurde bei der Einführung der Farbfernsehtechnik um den Faktor 1000/1001 verringert, d. h. die Zeilendauer beträgt heute 63,55 µs. Die Austastlücke beträgt horizontal 10,9 µs, sodass als aktive Zeile 52,65 µs verfügbar sind.

Bei HD-Signalen existieren die beiden Standards mit 1080 (1080i) und 720 aktiven Zeilen (720p). Im 1080i-System beträgt nach neuerer Definition die aktive Zeilendauer international einheitlich 25,86 µs. Dabei wird brutto immer mit 1125 Zeilen gearbeitet, die vertikale Austastlücke umfasst also 45 Zeilen. Die Dauer der horizontalen Austastung ist, abhängig von der Bildfrequenz, international unterschiedlich. Sie ergibt sich aus der Differenz der Gesamtzeilendauer und der aktiven Zeilendauer (netto). Bei 50 Hz Halbbildwechselfrequenz beträgt die gesamte Zeilendauer $1/(1125 \cdot 25\ \text{Hz}) = 35{,}55\ \mu\text{s}$, die Austastlücke dauert somit 9,69 µs. Bei 60 Hz beträgt die Gesamtzeilendauer $1/(1125 \cdot 30\ \text{Hz}) = 29{,}63\ \mu\text{s}$ und die H-Austastung dauert 3,77 µs.

Im 720p-HD-System wird insgesamt mit 750 Zeilen gearbeitet, die vertikale Austastung umfasst also 30 Zeilen. Die aktive Zeilendauer beträgt hier 17,24 µs. Bei 50 Hz Vollbildwechseln dauert die gesamte Zeile $1/(750 \cdot 50\ \text{Hz}) = 26{,}67\ \mu\text{s}$ und die Horizontalaustastung damit 9,43 µs. Trotz der geringeren Zeilenzahl sind die Zeilendauern bei HD-Signalen mit 720p gegenüber 1080i in gleicher Größenordnung, da in der Regel mit höherer zeitlicher Auflösung durch 50 Voll- statt 50 Halbbilder gearbeitet wird.

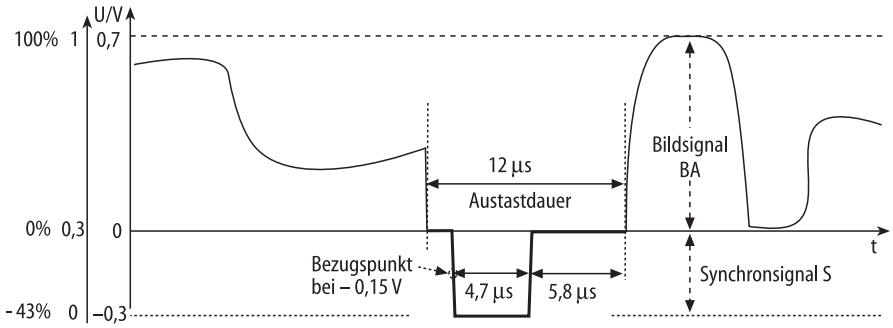


Abb. 2.30. Horizontal-Synchronimpuls im 625-Zeilensystem

2.3.1.3 Synchronsignale

Im Empfänger müssen die Zeilen in gleicher Weise geschrieben werden, wie sie in der Kamera erzeugt wurden, der Gleichlauf wird mit Synchronsignalen erreicht. Zur Horizontalsynchronisation wird nach jeder Zeile ein Rechtecksignal in der Horizontalaustastlücke positioniert. Dieser H-Synchronpuls liegt zwischen der sogenannten vorderen und hinteren Schwarzschulter (front/back porch), die im 625-Zeilensystem 1,5 µs bzw. 5,8 µs dauern, wahrend der Impuls selbst die Dauer von 4,7 µs hat (Abb. 2.30). Um eine sichere Synchronisation zu gewährleisten, hat der Synchronimpuls einen großen Wert, der bei SD-Systemen in Europa 3/7 und in Amerika 4/10 des Bildaussteuerbereichs BA beträgt. Er ist relativ zum Bildsignal negativ gerichtet, und damit leicht davon trennbar. Bezugsgröße für die Synchronisation ist die negativ gerichtete Impulsflanke bei halber Amplitude, also – 0,15 V (Abb. 2.30).

Bei HD-Signalen sind die Zeilenfrequenzen gegenüber SD etwa verdoppelt, das Timing für das Videosignal ist daher kritischer, sodass für analoge HD-Signale ein bipolares Synchronsignal verwendet wird (Tri Level Sync, TLS), das zweimal, mit je 300 mV, negativ und positiv auftritt. Die Bezugsgröße ist der Nulldurchgang (Abb. 2.31). Die beiden Teilimpulse dauern je 0,59 µs und der Nulldurchgang liegt um 2,59 µs vor dem Beginn der aktiven Zeile.

Neben der Information zum Zeilenwechsel braucht der Empfänger auch die Bildwechselinformation zur Vertikalsynchronisation. Um mit dem gleichen

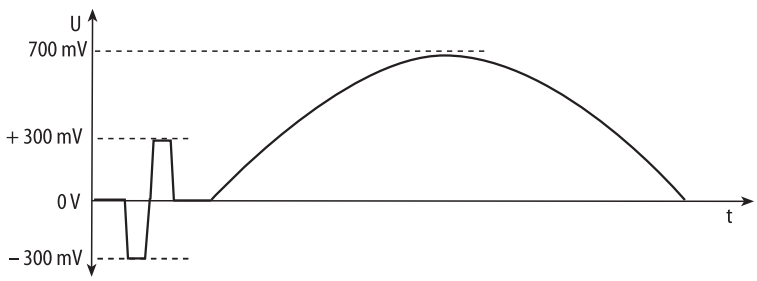


Abb. 2.31. Horizontal-Synchronimpuls bei HD-Systemen

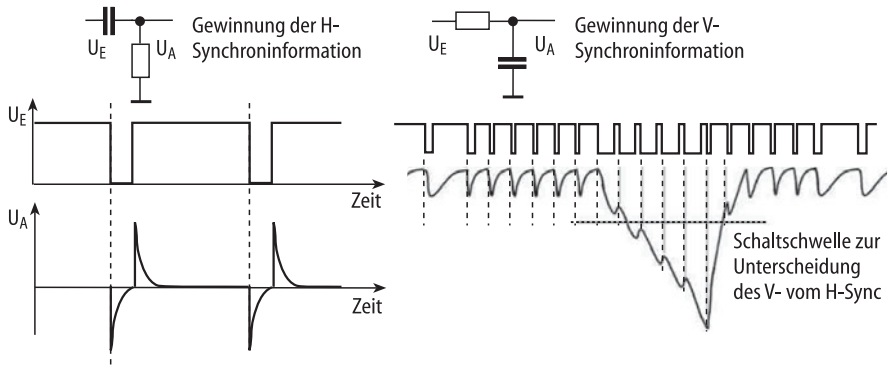


Abb. 2.32. Gewinnung der H- und V- Synchroninformationen mit Hilfe einfacher RC-Glieder

Spannungsniveau arbeiten zu können, wird das H-Synchronsignal zur Unterscheidung während der Vertikalaustastlücke deutlich verlängert. Beim 625-Zeilensystem beträgt die Impulsdauer 2,5 Zeilen (beim 525-Zeilensystem 3 Zeilen). Der Unterschied zwischen den Dauern der H- und V-Sync-Signale ist damit so groß, dass sie über einfache RC-Integrationsglieder getrennt werden können. Der Kondensator der RC-Kombination wird ständig durch die Synchronpulse geladen, aber nur bei Vorliegen des langen V-Pulses wird eine Kondensatorspannung erreicht, die ausreicht, um den Strahlrücksprung zum Bildanfang, also den Bildwechsel, auszulösen (Abb. 2.32 rechts). H- und V-Synchronimpulse werden zusammen als Synchronsignal S (auch Composite Sync) bezeichnet.

Damit die Zeilensynchronisation während der Bildsynchronisation nicht gestört ist, wird der V-Synchronpuls zweimal pro Zeile für je 4,7 µs unterbrochen, das Signal erscheint also als Folge von fünf Impulsen mit je 27,3 µs Dauer. Weiterhin werden dem V-Synchronsignal noch je fünf Vor- und Nachtrabanten (Pre- & Post-Equalizing Pulses) mit einer Dauer von jeweils 2,35 µs im Halbzeilenabstand hinzugefügt (Abb. 2.33). Die Vor- und Nachtrabanten dienen dazu, gleiche Anfangsbedingungen für die Integration zu schaffen, denn am Ende des ersten Halbbildes wird vor der Bildumschaltung nur eine halbe Zeile geschrieben, im Gegensatz zum Ende des zweiten Halbbildes, das mit einer ganzen Zeile endet. Ohne Trabanten wäre der Kondensator der RC-Kombination bei den verschiedenen Halbbildwechseln aufgrund der verschiedenen Entfernungen zum vorhergehenden H-Sync auf unterschiedliche Werte aufgeladen, und es ergäbe sich ein unsymmetrisches Halbbildschaltverhalten und damit eine paarige Rasterstruktur (Abb. 2.23).

Die Synchronsignale können ohne Oszilloskop auch mit Videomonitoren im sogenannten Puls-Kreuz-Modus überprüft werden. Hier wird die V-Ablenkung um die halbe Bildhöhe und die H-Ablenkung um ca. 1/3 Zeile verzögert, sodass die Synchronimpulse sichtbar werden (Abb. 7.25).

Die V-Austastlücke und die Vertikalsynchronimpulse von HD-Signalen sind in Abbildung 2.34 dargestellt. Hier werden im ersten Halbbild 22 und im zweiten 23 Zeilen ausgetastet, um die 1080 aktiven Zeilen zu erhalten. Der V-Sync-Hauptimpuls hat die Dauer von fünf Horizontalperioden, also bei 50 Hz Halb-

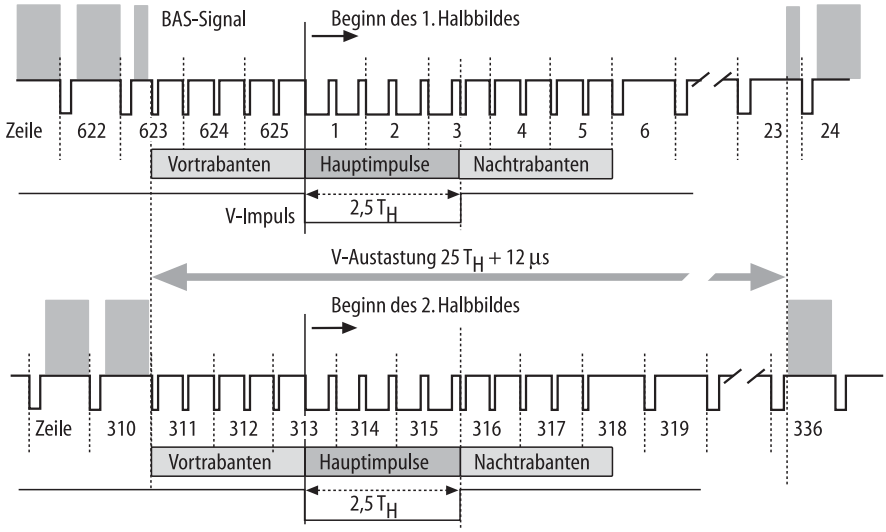


Abb. 2.33. Vertikal-Synchronimpulse im 625-Zeilensystem

bildwechselfrequenz 177,78 µs. Bei progressiv abgetasteten HD-Signalen wird die Impulsverdopplung in diesem Bereich weggelassen.

Die Abbildungen 2.33 und 2.34 zeigen auch die Festlegungen der Zeilenzählung. Zeile 1 des 625-Zeilen-Videobildes beginnt mit der Vorderflanke des Bildsynchronsignals für das erste Halbbild. Die zugehörigen Vortrabanten werden also am Ende des vorhergehenden Halbbildes übertragen. Die ersten 22,5 Zeilen eines Bildes fallen in die V-Austastlücke, das aktiv genutzte Bild beginnt mit der halben Zeile 23 und endet mit Zeile 310. Die nächsten 2,5 Zeilen enthalten wieder Vortrabanten und das zweite Halbbild beginnt mit Zeile 313, wo-

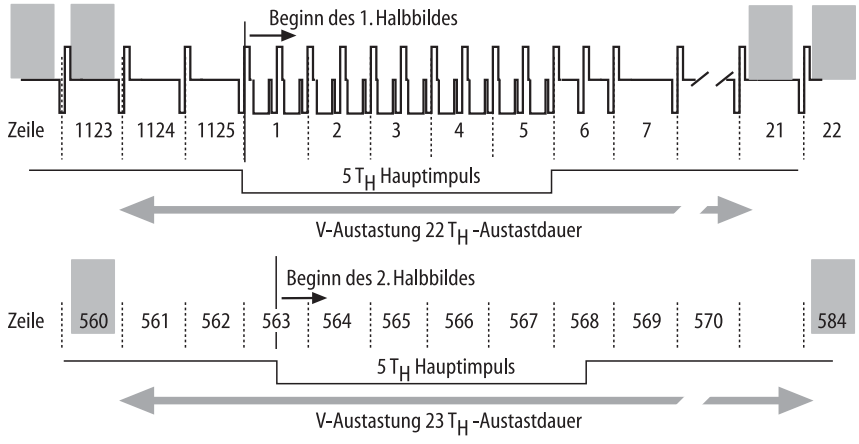


Abb. 2.34. Vertikal-Synchronimpulse bei HD-Systemen

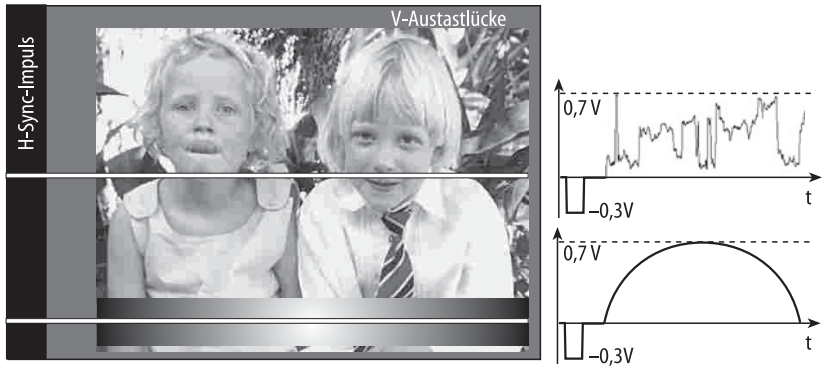


Abb. 2.35. Vergleich von Bild- und Austastbereich (625 Zeilen) und BAS-Signal zweier Zeilen

bei auch dieses mit der Austastlücke (Z. 313 ... 335) beginnt. Die Zählweise ist also zeit- und nicht raumorientiert, zwei aufeinander folgende Zeilen gehören zu verschiedenen Halbbildern. Das gilt auch für alle anderen Videosysteme, die mit Zeilensprung arbeiten.

2.3.1.4 Einzelheiten zum BAS-Signal

Bild-, Austast- und Synchronsignale werden zusammengefasst, um sie gemeinsam auf einer Leitung übertragen zu können. Alle Anteile des BAS-Signals sind für SD-Signale in den Abbildungen 2.30 und 2.35 dargestellt. In der beschriebenen Form stellt das BAS-Signal ein vollständiges monochromes Videosignal dar. Zusammengefasst: Die Bildhelligkeit wird durch die Signalspannung repräsentiert. Die Spannung des Gesamtsignals beträgt 1 V_{ss}, der Synchronboden liegt bei – 0,3 V, der Austastpegel bei 0 V und der Weißwert bei 0,7 V. Manchmal werden die Spannungen auch auf den Synchronboden bezogen, sodass der Austastwert relativ dazu bei 0,3 V und der Weißwert bei 1 V liegt. Die Angaben gelten für das 625-Zeilensystem und für HD-Systeme. Für ersteres veranschaulicht Abbildung 2.35 die zeitliche Relationen zwischen dem aktiven Bild und den Austastbereichen, wobei der schwarze Balken links das H-Synchronsignal darstellt.

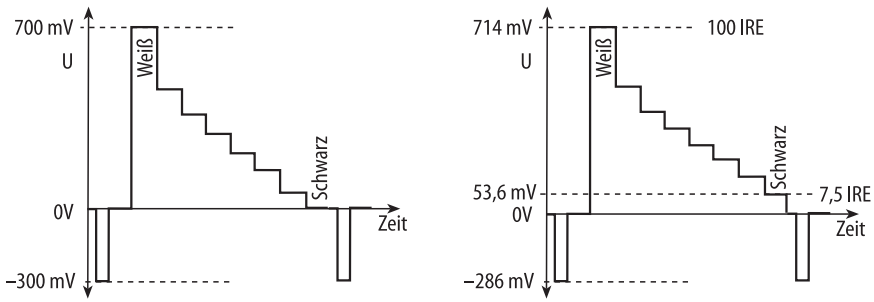


Abb. 2.36. Vergleich der Signalspannungsbereiche in Europa und Amerika

Im 525-Zeilensystem gibt es zwei Besonderheiten: Zum einen beginnt die Zählung der Zeilen bei den Vortrabanten und nicht beim Hauptimpuls. Zum anderen wird der Gesamtspannungsbereich von 1 V_{ss} nicht im Verhältnis 3/7 sondern mit 4/10 zwischen Synchronimpuls und Bildsignal aufgeteilt. Letzteres umfasst damit 714,3 mV, die 100 IRE-Einheiten mit je 7,14 mV zugeordnet werden. In den USA wurde dabei der Schwarzwert mit 7,5 IRE (Schwarzabhebung von 53,6 mV) festgelegt (Abb. 2.36), nicht aber in der japanischen Variante des 525-Zeilensystems, wo er bei 0 V bleibt.

2.3.1.5 Auflösung und Videogrenzfrequenz

Die Auflösung beschreibt den Grad der Darstellbarkeit von Bilddetails. Die Vertikalauflösung ist natürlich von der Zeilenanzahl abhängig. Aufgrund der Austastung von 25 Zeilen pro Halbbild können im europäischen SDTV-System von 625 Zeilen nur maximal 575 für das Bild genutzt werden. Alle aktiven Zeilen sind in voller Breite allerdings nur auf Studiomonitoren sichtbar, die im sogenannten Underscan-Modus betrieben werden können. Bei Röhrenmonitoren werden die Bildränder normalerweise horizontal und vertikal überschrieben damit keine schwarzen Bildränder erscheinen (Overscan, Abb. 2.37). Flachdisplays zeigen das volle Bild. Der Betrag des Überschreibens ist nicht genau festgelegt und variiert von Gerät zu Gerät. In Abbildung 2.37 beträgt er 7%. Bildwichtige Elemente sollten daher genügend Abstand zum Bildrand haben.

Die Bildauflösung für die Horizontale kann für 625 Zeilen folgendermaßen bestimmt werden: Bei Berücksichtigung des Kellfaktors k = 0,67 lassen sich vertikal maximal 0,67 · 575 = 385 Linien auflösen, dies entspricht 193 Paaren aus schwarzen und weißen Zeilen. Um die gleiche Auflösung auch Horizontal zu erreichen, sind für die Bildbreite unter Berücksichtigung des Bildseitenverhältnisses 4/3 · 385 = 514 auflösbare Spalten bzw. 257 Spaltenpaare erforderlich. Diese müssen in der aktiven Zeilendauer von 52 µs wenigstens als Sinus-Grundschwingung dargestellt werden können, woraus mit f_g = 257/(52 µs) eine Videogrenzfrequenz f_g = 4,94 MHz folgt. In der Praxis wird dieser Wert aufgerundet und mit f_g = 5 MHz gerechnet.

In einigen Fällen wird auch eine Maximalauflösung von 400 Linien angegeben. Dieser Wert gilt aufgerundet für die Vertikale, er wird aber auch für die volle Horizontalauflösung verwendet, was natürlich nur unter der Voraussetzung B:H = 1:1 gilt. Als Faustformel für den Zusammenhang von Grenzfre-

Abb. 2.37. Overscan- und Underscan-Darstellung

quenz (Bandbreite in MHz) und (Zeilen-)Auflösung Z_s gilt in diesem (etwas verwirrenden) Fall für beide Dimensionen: $Z_s = 80 \cdot f_g$.

Bei Reduktion der Grenzfrequenz auf Werte unter 5 MHz verringert sich dieser Wert. Dies geht aber nur zu Lasten der Horizontalauflösung, da vertikal natürlich weiter mit 625 Zeilen gearbeitet wird. Bei der Übertragung des Signals, z. B. durch einen VHS-Recorder mit einer Grenzfrequenz von 3 MHz, ergibt sich nach obiger Gleichung eine Horizontalauflösung von 240 Linien, was praktisch heißt, dass unter Berücksichtigung des Bildseitenverhältnisses tatsächlich etwa 320 Spalten oder 160 Spaltenpaare getrennt werden können.

Zur Auflösungsbestimmung kann das Testbild nach RMA (Radio Manufacturers Association) verwendet werden. Es enthält sogenannte Testbesen, d. h. schräg verlaufende Streifen, die immer enger zusammenrücken und ab einem bestimmten Punkt visuell nicht mehr getrennt werden können (Abb. 2.38). Die Besen sind mit Frequenzangaben und mit Zahlen zur Auflösungsangabe markiert. Die höchste TV-Auflösung von 5 MHz ist erreicht, wenn die schrägen Striche horizontal wie vertikal bis zum Wert 400 als getrennt erscheinen.

Als allgemeine Formel zur Berechnung der Videogrenzfrequenz gilt auch:

$$f_g = 1/2 \; k \; c \; Z^2 \; B/H \; f_v.$$

Der Faktor 1/2 ergibt sich aufgrund der Linienpaare die aus den Schwarz/Weiß-Wechseln entstehen (Abb. 2.39) und k ist der Kellfaktor. Z^2 B/H ist die Bildpunktzahl bei Berücksichtigung des Bildseitenverhältnisses B/H, und f_v ist die Vollbildfrequenz. Der Faktor c ergibt sich aus der aufgrund der Austastlücken verringerten Zeilenzahl und -dauer. Es gilt:

$$c = (1 - T_{Vaus} / T_V) \; / \; (1 - T_{Haus} / T_H).$$

Mit den Verhältnissen $T_{Haus} / T_H = 12/64 = 18{,}75\%$ und $T_{Vaus} / T_V = 1{,}6$ ms/20 ms = 8% gilt in Europa $c \approx 1{,}13$. Unter Verwendung von $k = 0{,}67$ und $f_v = 25$ Hz ergibt sich mit 625 Zeilen auch bei dieser Formel wieder die Grenzfrequenz $f_g = 4{,}94$ MHz. In Amerika gilt $T_{Haus} / T_H \approx 17\%$ und $T_{Vaus} / T_V = 7{,}6\%$ (Standard

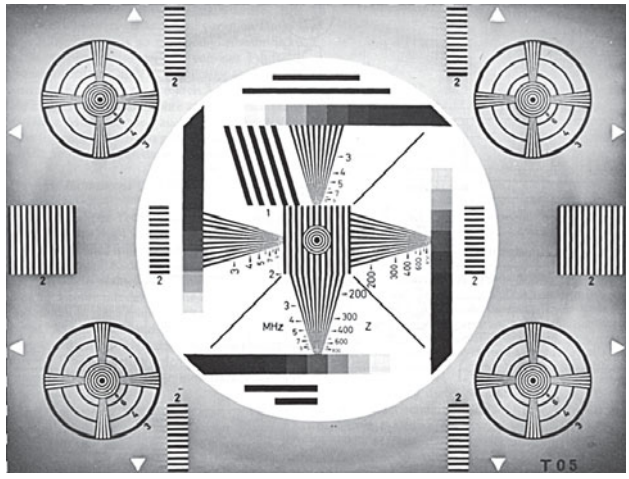


Abb. 2.38. RMA-Testbild

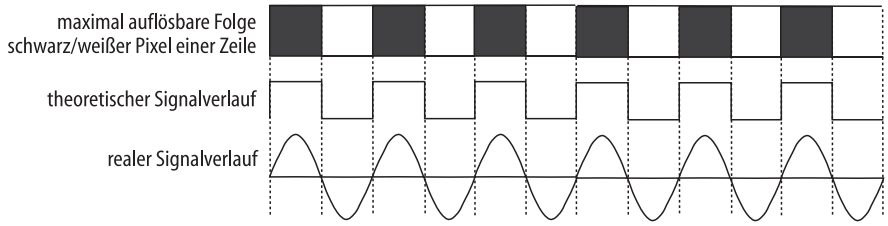


Abb. 2.39. Signalverlauf bei maximaler Horizontalauflösung

M, s. u), also c = 1,11. Mit k = 0,67 folgt hier f_g = 4,11 MHz. Die Bandbreite wird für 525-Zeilensysteme aufgerundet mit dem Wert f_g = 4,2 MHz angegeben. Wird das Bildseitenverhältnis ohne Veränderung der anderen Parameter auf 16:9 erhöht, wie es bei PAL-Plus (s. Kap. 4.4) geschieht, so verringert sich die Horizontalauflösung um den Faktor 4/3.

Beim Übergang zu HDTV wird außer dem Bildseitenverhältnis auch die Zeilenzahl verändert. Ausgehend von einer Verdopplung der H- und V-Auflösung und dem Faktor 1,33 für den Übergang zu 16/9 ergibt sich relativ zum SD-System mit $2 \cdot 2 \cdot 1{,}33 \cdot 5$ MHz eine erforderliche Grenzfrequenz von f_g = 26,7 MHz. Nominell ist die HD-Bandbreite für analoge Signale auf 30 MHz festgelegt.

Bei Computerbildübertragung wird der Kellfaktor meist außer acht gelassen, da sich der Betrachter nahe am Monitor befindet und die Artefakte wegen der geringen Bildbewegung nicht so stark ins Gewicht fallen. Bei einer Bilddarstellung mit 1024 H x 768 V aktiven Bildpunkten und f_V = 70 Hz ergibt sich hier eine Horizontalfrequenz von ca. 54 kHz und eine Grenzfrequenz von ca. 27,5 MHz für das S/W-Videosignal, wenn horizontal und vertikal relativ gleich große Austastlücken angenommen werden.

Die auflösungsbestimmende obere Grenzfrequenz ist ein wichtiger Parameter des Videosignals. Die untere Grenzfrequenz ist idealerweise 0 Hz und sollte in der Praxis höchstens bei 100 Hz liegen. Weitere wichtige Frequenzen sind die Zeilen- und die Bildfrequenzen. Erstere betragen bei den SD-Systemen mit 525 und 625 Zeilen 15,734 und 15,625 kHz. In einem HD-System mit 1125 Zeilen beträgt die Zeilenfrequenz bei 25 Hz Vollbildwechselfrequenz 28,125 kHz.

2.3.1.6 Internationale Standards

Das dargestellte BAS-Signal ist im Standard ITU-R BT.470 als CCIR B/G-Norm definiert worden. Der in Deutschland verwendetet B/G-Standard ist nur einer unter vielen internationalen Standards. Die Tabellen 2.4 und 2.5 zeigen eine Übersicht. Die Definitionen betreffen immer das Fernsehsignal und enthalten daher auch HF-Übertragungsparameter die länderspezifischer sind als die Bildaufbauparameter allein. Dieser Umstand verhindert, dass die beiden wesentlichen Videosysteme 525 und 625 Zeilen diesbezüglich nicht einfach mit den zugeordneten Buchstaben bezeichnet werden können.

In Tabelle 2.4 sind neben den europäischen Standards, die auf 625 Zeilen und 25 Vollbildern basieren, die Bildaufbauparameter des in Amerika und Japan gültige Standard M hervorgehoben, der auf 525 Zeilen und 30 Vollbildern/s basiert. Mit der Einführung der Farbfernsehtechnik wurde hier die Bildwechsel-

Tabelle 2.4. Verbreitete Fernsehstandards

Standard	B/G	D/K	I	L	M
Einsatzgebiet	Mitteleuropa	Osteuropa	Großbritan.	Frankreich	USA Japan
Zeilen/Vollbild	625	625	625	625	525
Halbbildfrequenz (Hz)	50	50	50	50	59,94
Zeilenaustastdauer T_{HA} (µs)	12	12	12	12	10,9
Halbbildaustastdauer (Zeilen)	$25 + T_{HA}$	$25 + T_{HA}$	$25 + T_{HA}$	$25 + T_{HA}$	$20 + T_{HA}$
Videobandbreite (MHz)	5	6	5,5	6	4,2
HF-Kanalbandbreite (MHz)	7/8	8	8	8	6
Bild/Tontr. Abstand (MHz)	5,5/5,74	6,5	6	6,5	4,5
Tonmodulation	FM	FM	FM	AM	FM

frequenz um den Faktor 1000/1001 von 30 Hz auf 29,97 Hz gesenkt. In Europa gültige Fernsehstandards weichen nur geringfügig voneinander ab. Differenzen gibt es in Europa vor allem wegen der Farbcodierung (PAL oder SECAM, s. Abschn. 2.4.7), aber auch aufgrund verschiedener Bild-Tonträgerabstände im Fernsehsignal, was z. B. zur Folge hat, dass mit einem gewöhnlichen deutschen PAL-Empfänger in Großbritannien zwar ein Bild, aber kein Ton empfangen werden kann. In Deutschland gilt der CCIR B/G-Standard.

Die Normung betrifft nur SD-Systeme, bei der HD-Übertragung wird in Deutschland mit 720p/50, d. h. mit 720 Zeilen netto bei progressiver Abtastung und 50 Hz Vollbildwechselfrequenz gearbeitet. SD-Systeme basieren dagegen immer auf dem Zeilensprungverfahren. Damit dieses einfach realisiert werden kann, wird immer eine ungerade Gesamtzeilenzahl verwendet (Abb. 2.23). Des Weiteren erfolgt die Normung so, dass es durch Reihenschaltung simpler Frequenzteiler ermöglich wird, die Bildfrequenz aus der Zeilenfrequenz abzuleiten, um die feste Verkopplung dazwischen sicherzustellen. Dazu wurde immer dafür gesorgt, dass sie Zeilenzahl als Produkt möglichst kleiner Primzahlen darstellbar ist. Z. B. beträgt die Zeilenzahl der alten englischen Norm A 405 = 3.3.3.3.5, Standard M hat 525 = 3.5.5.7 Zeilen, Standard B/G 625 = 5.5.5.5 Zeilen, der alte französische Standard E 819 = 3.3.7.13 und die erste HD-Norm hat 1125 Zeilen, die als 3.3.5.5.5 darstellbar sind.

Tabelle 2.5. Fernsehstandards verschiedener Länder

Europa				Außereuropäisch			
Land	Standard	Land	Standard	Land	Standard	Land	Standard
Belgien	B/H	Deutschland	B/G	Ägypten	B/G	Algerien	B/G
Dänemark	B/G	Finnland	B/G	Australien	B	Brasilien	M
Frankreich	E/L	Großbritannien	A/I	China	D	Indien	B
Italien	B/G	Luxemburg	B/G	Japan	M	Kanada	M
Niederlande	B/G	Norwegen	B/G	Phillipinen	M	Südafrika	I
Österreich	B/G	Polen	D	Sudan	B	Thailand	B
Portugal	B/G	Rußland	D/K	Türkei	B	USA	M
Spanien	B/G			Vietnam	D/K		

2.3.2 Signalanalyse

Die gebräuchlichste Art der Signaldarstellung ist die Wiedergabe des Signalverlaufs über der Zeit. Dieser Verlauf ist mit Oszilloskopen einfach darstellbar. Darüber hinaus ist es aber häufig erwünscht, die Signalamplituden in Abhängigkeit von der Frequenz zu beschreiben, d. h. ein Amplitudenfrequenzspektrum ermitteln zu können. Dieser Wechsel von der zeitabhängigen zur frequenzabhängigen Darstellung, und umgekehrt, ist mit der Fouriertransformation möglich [95].

Zunächst sei hier die Fouriertransformation als Bildung einer Fourierreihe für periodische Signale betrachtet. Jedes periodische Signal lässt sich als Summe (i. a. unendlich) vieler harmonischer Schwingungen (Sinus und Cosinus) so beschreiben, dass die Grundschwingung die gleiche Frequenz ω_o hat wie das periodische Signal und die weiteren Schwingungen (Oberschwingungen) ganzzahlige Vielfache davon sind. Es gilt:

$$s(t) = a_o + \sum_n [a_n \cdot \cos(n\,\omega_o t) + b_n \cdot \sin(n\,\omega_o t)].$$

Der Zählindex n läuft von 1 bis Unendlich.

Für die frequenzabhängige Darstellung ergibt sich daraus ein diskretes Spektrum, dessen Frequenzen bereits festliegen. Für das zu analysierende Signal muss lediglich bestimmt werden, mit welchen Amplituden die einzelnen Frequenzen vertreten sind, d. h. es müssen die Koeffizienten a_n und b_n für die sin- und cos-Anteile ermittelt werden. Falls das zu transformierende Signal in Form diskreter Messwerte vorliegt, die in regelmäßigen Abständen (Abtastintervall T_a) dem Signal entnommen wurden, lassen sich die Koeffizienten a_n und b_n (n = 0, 1, 2, ... N) folgendermaßen berechnen:

$$a_n = 2/K \cdot \sum_k y_k \cdot \cos(n\,\omega_o t_k)$$

$$b_n = 2/K \cdot \sum_k y_k \cdot \sin(n\,\omega_o t_k).$$

Dabei wurde angenommen, dass K diskrete Messwerte y_k mit einem Abtastintervall T_a vorliegen, wobei gilt: $K \cdot T_a = T_o$, $t_k = k \cdot T_a = k \cdot 2\pi/(K \cdot \omega_o)$. In diesem Fall sind höchstens $N \leq (K-1)/2$ Oberwellen berechenbar. Der Zählindex k läuft daher von 0 bis K–1. Dieser Zusammenhang drückt sich auch im Abtastheorem aus (s. Kap. 3).

Für einen kontinuierlichen Signalverlauf wird die Summe über die diskreten Punkte zu einem Integral, das sich über die Periode von t = 0 bis t = T erstreckt und es gilt:

$$a_o = 1/T \cdot \int_t s(t)\,dt,$$

$$a_n = 2/T \cdot \int_t s(t) \cdot \cos(n\,\omega_o t)\,dt,$$

$$b_n = 2/T \cdot \int_t s(t) \cdot \sin(n\,\omega_o t)\,dt.$$

Bei genauer Betrachtung enthält dieses Verfahren zur Analyse der Zeitfunktion auch seine Umkehrung, d. h. die Angabe der Vorgehensweise zur Zusammensetzung (Synthese) einer beliebigen periodischen Funktion aus harmonischen

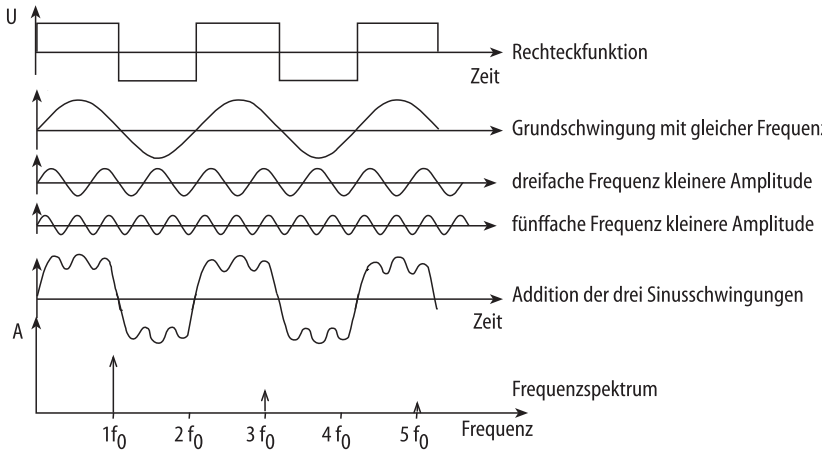


Abb. 2.40. Synthese und Amplitudenspektrum einer Rechteckfunktion

Schwingungen (Rücktransformation, s. u.). Wenn das ursprüngliche Signal durch Addition harmonischer Schwingungen exakt synthetisiert werden soll, ist im Allgemeinen eine unendlich große Anzahl von Schwingungen und Koeffizienten erforderlich. Näherungsweise gelingt die Synthese aber auch dann, wenn hochfrequente Anteile weggelassen werden. Abbildung 2.40 zeigt am Beispiel einer symmetrischen Funktion, wie sich bereits annähernd ein Rechteckverlauf ergibt, wenn nur Anteile bis zur 5. Harmonischen addiert werden.

Mit diesen Betrachtungen lässt sich bereits ein wesentlicher Teil des Videosignals analysieren, nämlich der rechteckförmige Verlauf des Synchronimpulses. Wenn als Bildsignal Schwarz vorliegt, so sind die Synchronimpulse allein vorhanden. Von der Periodendauer T = 64 µs belegt der Impuls die Dauer 4,7 µs, das Tastverhältnis T_i/T beträgt 4,7/64 (Abb. 2.41). Die Berechnung der Fourierkoeffizienten führt in diesem Fall zum Verschwinden der sin-Anteile (alle $b_n = 0$), es gilt:

$$a_o = T_i/T \qquad \text{für } n = 0$$

$$a_n = 2/(n\pi) \cdot \sin(n\pi \cdot T_i/T) \qquad \text{für } n > 0.$$

Die Funktion lässt sich also folgendermaßen darstellen:

$$s(t) = T_i/T + 2/\pi \cdot \sin(\pi \cdot T_i/T) \cdot \cos(\omega_o t) + 2/2\pi \cdot \sin(2\pi \cdot T_i/T) \cdot \cos(2\omega_o t) + \ldots$$

Die frequenzabhängige Darstellung der Fourierkoeffizienten in Abbildung 2.41 gibt das Amplitudenspektrum des BAS-Signals mit Bildinhalt Schwarz wieder, wenn das V-Sync-Signal unberücksichtigt bleibt. Das Spektrum zeigt einen Signalverlauf der von den Vielfachen der H-Sync-Frequenz f_o = 1/64 µs = 15,625 kHz geprägt ist, sowie von einer Einhüllenden, die periodische Nullstellen beim Kehrwert der Impulsdauer T_i aufweist. Zusätzlich zeigt sich, dass insgesamt die Amplituden der Oberschwingungen mit steigender Frequenz abnehmen. Der Verlauf der Einhüllenden entspricht der si-Funktion mit si x = sinx/x.

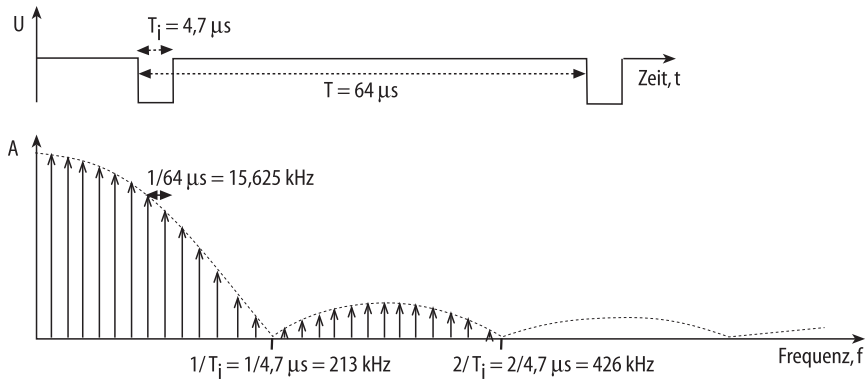


Abb. 2.41. Zeitverlauf und Amplitudenspektrum des H-Sync-Impulses im Videosignal

Für dieses Beispiel gilt aufgrund der oben angegebenen Fourierkoeffizienten:

$$a_n = 2T_i/T \cdot \sin(n\pi \cdot T_i/T) \,/\, (n\pi \cdot T_i/T).$$

Es wird deutlich, dass die Funktion immer dann verschwindet, wenn das Argument der sin-Funktion ein Vielfaches von π ist, bei $n = 1/T_i$, $2/T_i$, $3/T_i$ etc.

Mit komplexer Schreibweise lassen sich die sin- und cos-Anteile zur Vereinfachung folgendermaßen zusammenfassen:

$$s(t) = \sum_n c_n \cdot e^{jn\omega_0 t}.$$

Die Summe läuft von $-\infty$ bis $+\infty$. Der Zusammenhang mit den Koeffizienten a_n und b_n ist für $c_n > 0$ über folgende Beziehung gegeben:

$$c_n = (a_n - jb_n)/2.$$

Bei der Angabe der Koeffizienten werden die Integrationsgrenzen um T/2 verschoben, sodass sich das Integral über den Bereich $-T/2$ bis $+T/2$ erstreckt und für $n > 0$ gilt:

$$c_n = 1/T \cdot \int_t s(t) \cdot e^{-jn\omega_0 t}\, dt.$$

Für die Analyse nichtperiodischer Funktionen wird anstelle der Fourierreihe die Fouriertransformierte benutzt. Die Periodizität verschwindet, wenn die Periodendauer immer länger wird. In diesem Fall werden die Kehrwerte der Periodendauern immer kleiner, die Spektrallinien rücken immer enger zusammen und bilden im Grenzfall ein Kontinuum. Im Grenzfall geht $T \to \infty$, und es folgen als weitere Übergänge: $n \cdot f_0 \to f$, $1/T \to df$ und $c_n \cdot T \to S(f)$, wobei S(f) als Fouriertransformierte oder komplexe Spektraldichte bezeichnet wird.

Für die Fouriertransformierte S(f) und die Rücktransformation s(t) erstrecken sich die Integrale von $-\infty$ bis $+\infty$, und es gelten die Beziehungen:

$$S(f) = \int_t s(t) \cdot e^{-j\omega t}\, dt.$$

$$s(t) = \int_f S(f) \cdot e^{j2\pi ft}\, df.$$

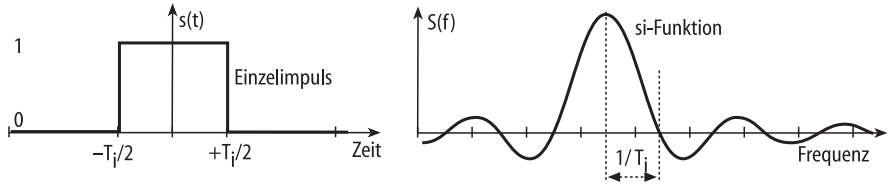


Abb. 2.42. Symmetrischer Rechteckimpuls und zugehörige Spektraldichte

Mit Hilfe der Fourieranalyse wird ein zeitabhängiges Signal frequenzabhängig dargestellt. Damit können Informationen gewonnen werden, die im Zeitbereich nicht erfassbar sind. Die Transformation ist symmetrisch und kann auch genau anders herum betrachtet werden. Die Zusammengehörigkeit von s(t) und S(f) sei hier durch die Schreibweise $s(t) \leftrightarrow S(f)$ gekennzeichnet.

Wird als Beispiel für eine nichtperiodische Funktion ein einzelner Rechteckimpuls nach Abbildung 2.42 herangezogen, so ergibt sich im Spektrum wieder die si-Funktion. Das zeigt die folgende Rechnung, die sich auf Abbildung 2.42 bezieht, in der aus formellen Gründen eine zweiseitige spektrale Darstellung gewählt ist, bei der negative Frequenzen auftreten. Mit dem Integral über den Bereich von $-T_i/2$ bis $+T_i/2$ gilt für dieses Beispiel:

$$S(f) = \int_t 1 \cdot e^{-j\omega t}\, dt = \int_t \cos(2\pi f t)\, dt - j \sin(2\pi f t)\, dt.$$

Aufgrund der Symmetrie verschwindet der Imaginärteil, und es gilt

$$S(f) = 1/\pi f \cdot \sin(\pi f T_i) = T_i \cdot \sin(a)/a \text{ mit } a = \pi f T_i.$$

Komplexere Signale können auch ohne Rechnung mit Sätzen über die Fouriertransformation erschlossen werden. Es gilt der Additionssatz:

$$s(t) = k_1 \cdot s_1(t) + k_2 \cdot s_2(t) \leftrightarrow S(f) = k_1 \cdot S_1(f) + k_2 \cdot S_2(f).$$

Es gilt der Faltungssatz:

$$s_1(t) * s_2(t) \leftrightarrow S_1(f) \cdot S_2(f),$$

wobei als Faltung zweier Zeitfunktionen das folgende Integral bezeichnet wird, das sich bezüglich τ von $-\infty$ bis $+\infty$ erstreckt:

$$s_1(t) * s_2(t) = \int_\tau s_1(t) \cdot s_2(t-\tau)\, d\tau.$$

Für die Faltung und auch allgemein gilt der Vertauschungssatz

$$s(t) \leftrightarrow S(f), \text{ und } \quad S(t) \leftrightarrow s(-f).$$

Damit können nun recht anschaulich einige konkrete Beispiele betrachtet werden, an denen die Symmetrie der Fouriertransformationen deutlich wird: Eine reine, zeitlich unendlich ausgedehnte Sinusschwingung erscheint im Frequenzspektrum nur als eine einzelne Linie. Mit dem Vertauschungssatz gilt umgekehrt, dass zu einem unendlich ausgedehnten Frequenzspektrum im Zeitbereich eine einzelne Linie (der sog. Dirac-Impuls) gehört (Abb. 2.43 a,b).

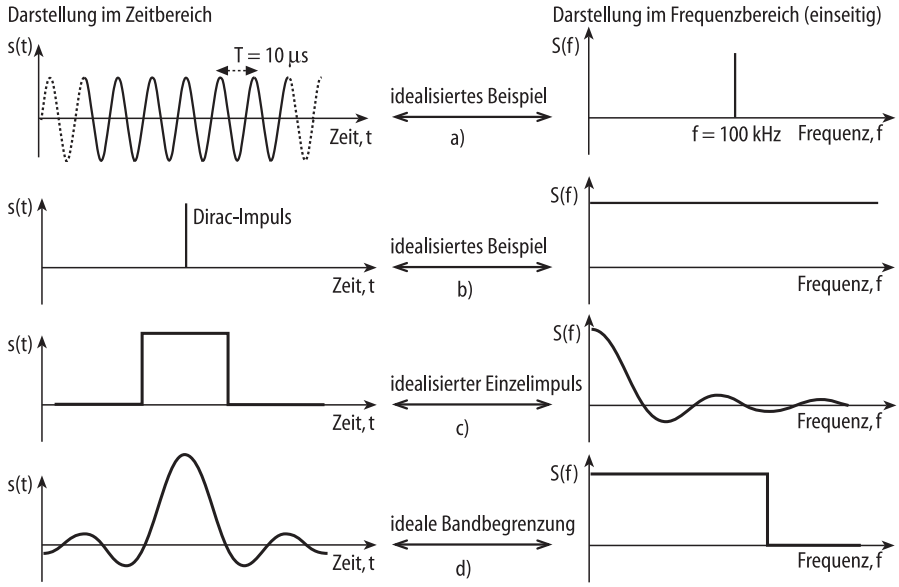


Abb. 2.43. Zeitfunktionen und zugehörige Spektraldichten

Ein einzelner Rechteckimpuls hat im Spektralbereich den Verlauf der si-Funktion (Abb. 2.43 c) und eine periodische Folge von Rechteckimpulsen erscheint gemäß dem Additionssatz als Überlagerung des idealen, unendlich ausgedehnten Linienspektrums mit dem Betrag der si-Funktion.

Die Anwendung des Vertauschungssatzes auf die Rechteckfunktion bedeutet, dass ihr Auftreten im Frequenzbereich einer idealen Frequenzbandbegrenzung entspricht. Damit geht im Zeitbereich wieder ein si-funktionsförmiger Verlauf einher (Abb. 2.43 d). Dieser Umstand ist z. B. für ein Digitalsignal von Bedeutung, das im Zeitbereich wegen der Umschaltung zwischen zwei Zuständen zunächst einen rechteckförmigen Verlauf hat. Wenn es aber über einen Kanal mit der oben angegebenen idealen Bandbegrenzung übertragen wird, so wird der Zeitverlauf in eine si-Funktion umgeformt (s. Kap. 3.2 und Abb. 2.43 d).

Bei der Betrachtung des Amplitudenspektrums des BAS-Signals zeigte sich bereits, dass dieses von den Linien des mit der 64 µs-Periode wiederholten Synchronimpulses dominiert wird. Dies gilt zunächst, wenn die eigentliche Bildsignalspannung dem Austastwert entspricht. Abbildung 2.44 zeigt das gemessene Spektrum eines solchen Signals. In der Übersicht ist der Verlauf der si-Funktion sichtbar. Die Nullstellen folgen aus dem Kehrwert der Impulsdauer von 4,7 µs mit 1/4,7 µs = 213 kHz. Im Gegensatz zur idealisierten Darstellung in Abbildung 2.41 enthält der Synchronanteil des BAS-Signals neben der Information über die H-Synchronisation hier auch den V-Sync-Anteil, der alle 20 ms wiederholt wird und einschließlich der Trabanten eine Dauer von 7,5 Horizontalperioden umfasst. Betrachtet man die V-Synchron-Impulsfolge vereinfacht als einen Rechteckimpuls, so beträgt die Dauer T_i = 480 µs, sodass die

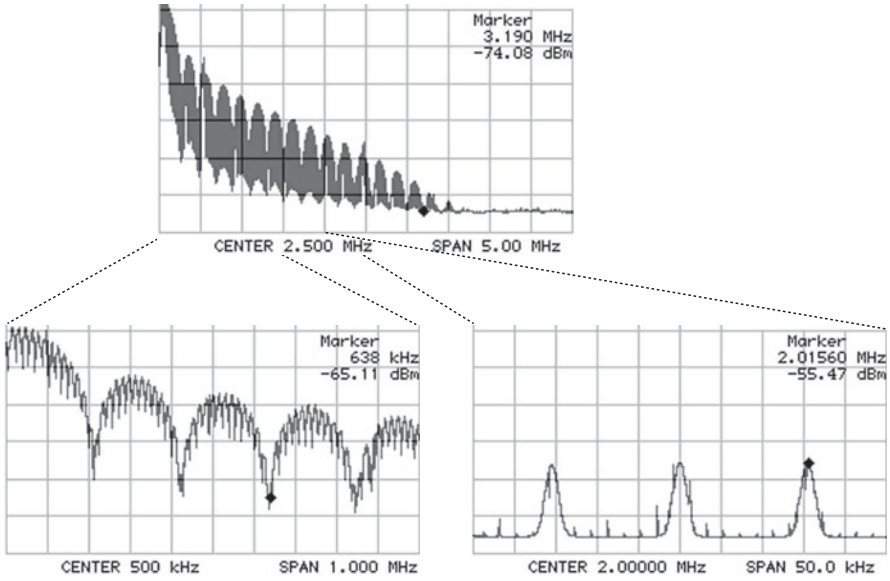


Abb. 2.44. Gemessenes Amplitudenspektrum des BAS-Signals mit Bildinhalt Schwarz

Nullstellen der daraus resultierenden si-Funktion bei $1/T_i$ = 2,08 kHz liegen. In Vergrößerung ist dieser Verlauf an den kleinen Spitzen in Abbildung 2.44 gerade erkennbar.

Abbildung 2.46 zeigt die Veränderung des Spektrums durch ein per Oszillator erzeugtes, sogenanntes Multiburst-Signal mit definierten Frequenzen von 0,5 MHz, 1 MHz, 2 MHz, 3 MHz, 4 MHz und 4,8 MHz sowie durch einen gewöhnlichen Bildinhalt. Dabei wird deutlich, dass auch mit Bildinhalt das BAS-Spektrum von einzelnen Linien im Abstand der Zeilenfrequenz dominiert wird. Wenn sich der Bildinhalt zeitlich ändert, verbreitern sich die Linien. Je mehr Bewegung im Bild auftritt, desto mehr füllt sich das Spektrum.

Mit Hilfe der genannten Sätze kann auch das Amplitudenspektrum des im PAL-Signal auftretenden Burst gedeutet werden (s. Abschn. 2.4.7). Der Burst

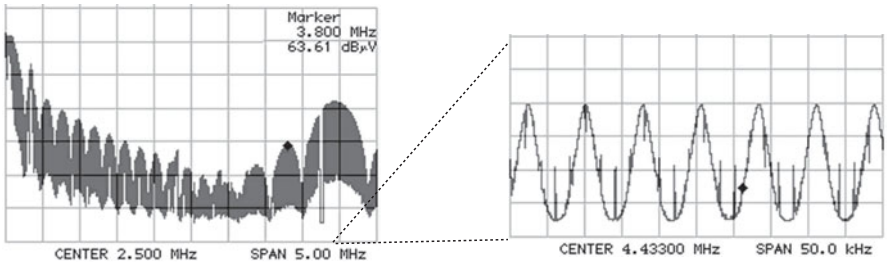


Abb. 2.45. Amplitudenspektren von Sync- und Burstsignal

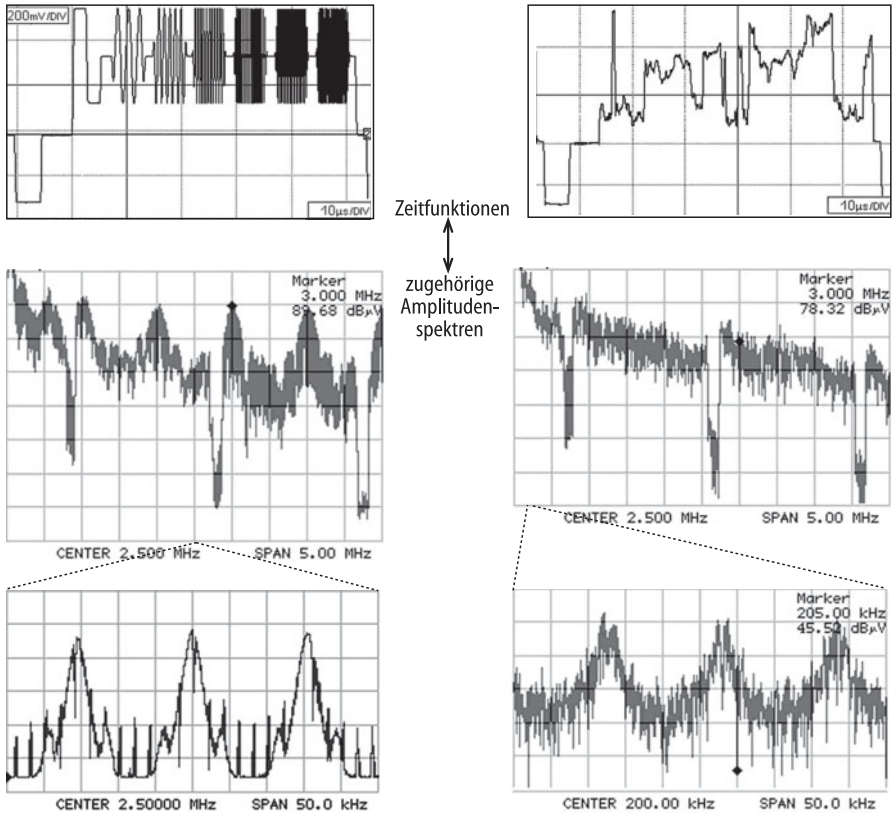


Abb. 2.46. Zeitverlauf und Amplitudenspektren von Multiburst- und Bildsignal

befindet sich in der H-Austastlücke und besteht aus einer cos-Schwingung mit f_{SC} = 4,43 MHz, die für eine Dauer von 10 Perioden, d. h. ca. 2,25 µs eingeschaltet wird. Abbildung 2.45 zeigt am Beispiel eines Black Burst Signals (Sync- und Burst), wie das Spektrum im Bereich um 4,43 MHz von f_{SC} und Nullstellen dominiert wird, die im Abstand des Kehrwerts der genannten Einschaltdauer, nämlich 443 kHz auftreten.

2.3.3 Videosignalübertragung im Basisband

Video- und Audiosignale können ohne weitere Manipulation übertragen werden, wenn das Übertragungsmedium eine dem Signal entsprechende Bandbreite aufweist. Für nicht frequenzversetzte Signale (Basisband) kommt dabei nur die Übertragung durch elektrische Leitungen in Frage. Eine Abstrahlung als elektromagnetische Welle ist im Basisband nicht möglich, da die untere Grenzfrequenz für Videosignale nahe bei o Hz liegt und tiefe Frequenzen zu ineffektiv abgestrahlt werden.

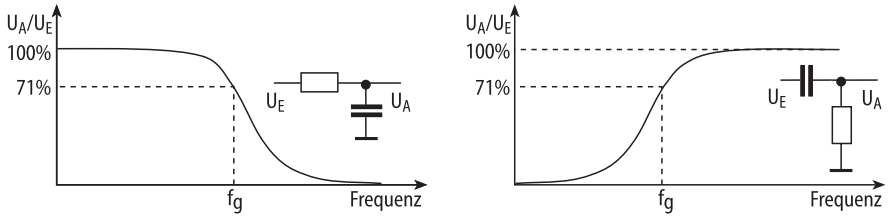


Abb. 2.47. Tief- und Hochpass

2.3.3.1 Signalverzerrungen

Bei der Signalübertragung werden die Signale verändert oder verzerrt. Verzerrungen sind unvermeidlich, sie werden (ggf. als Entzerrung) auch absichtlich herbeigeführt. Eine wichtige Art der Signalveränderung ist die Filterung. Einfache Filter lassen sich mit Kombinationen aus Widerständen und Induktivitäten (RL) oder Widerständen und Kapazitäten (RC) realisieren. Abbildung 2.47 zeigt Signalverläufe von RC-Kombinationen, die als Tief- und Hochpass wirken, also die hohen bzw. tiefen Frequenzen dämpfen. Diese einfachen Filter sind wenig steilflankig, die Grenzfrequenz für das Signal ist dort definiert, wo es auf $1/\sqrt{2}$ des Höchstwertes abgefallen ist.

RC-Kombinationen werden auch durch die Angabe einer Zeitkonstante τ charakterisiert, für die gilt: $\tau = RC$. τ entspricht dabei der Zeit, die vergeht, bis die Kapazität zu 63% aufgeladen oder um 63% entladen ist. In der Videotechnik werden RC-Filter, die durch feste Zeitkonstanten definiert sind, beispielsweise zur gezielten Hochfrequenzanhebung (Preemphasis) eingesetzt (s. Kap. 8.4). Abbildung 2.48 zeigt links eine starke Signalverformung durch einen RC-Hochpass und darunter eine schwächere, wie sie bei Videosignalen auftreten kann. Die zulässige Verformung durch die Hochpasswirkung wird als maximale Dachschräge des Weißimpulses im Prüfsignal angegeben.

Die Verbindungen aus Kapazität und Widerstand spielen elektrotechnisch eine große Rolle, denn durch die Kapazitäten lassen sich Signalverarbeitungs-

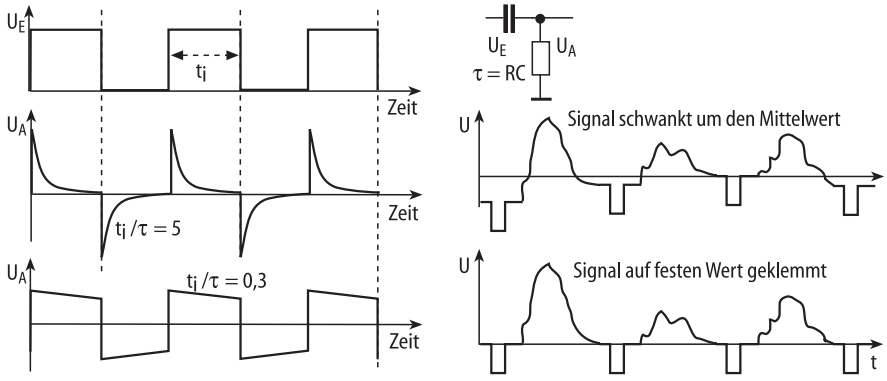


Abb. 2.48. Signalverformung durch Hochpass bzw. Kondensatorkopplung

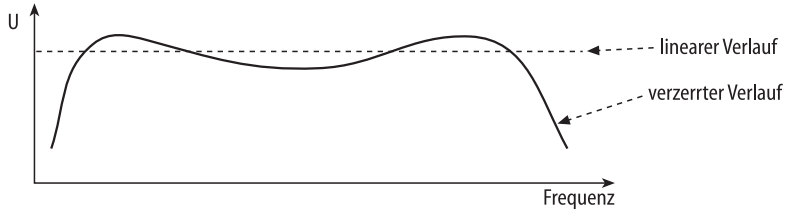


Abb. 2.49. Frequenzgang und lineare Verzerrungen

stufen gleichspannungsmäßig entkoppeln. Bei solchen Kopplungen ist aber keine Information über das Gleichspannungsniveau mehr übertragbar. Bei Videosignalen spielt diese Information aber eine große Rolle, denn der Austastwert soll immer auf festem Niveau liegen. In Abbildung 2.48 wird im rechten Teil deutlich, wie sich das Gleichspannungsniveau durch die Kondensatorkopplung in Abhängigkeit vom Bildsignal verschieben kann. Abhilfe lässt sich schaffen, indem hinter dem Kondensator eine elektronische Klemm-Schaltung dafür sorgt, dass das Gleichspannungsniveau auf den gewünschten Wert gehalten (geklemmt) wird.

Signalveränderungen, die entsprechend einer Filterung den Signalverlauf über der Frequenz betreffen, nennt man lineare Verzerrungen. Der Signalverlauf über der Frequenz wird als Amplitudenfrequenzgang bezeichnet (Abb. 2.49). Jede Abweichung vom linearen Verlauf stellt eine lineare Dämpfungs- oder Verstärkungsverzerrung dar. Verändert sich die Phasenlage des Signals, so spricht man von Phasenverzerrungen und Phasenfrequenzgang.

Neben den linearen treten auch nichtlineare Verzerrungen auf. Diese ergeben sich z. B. durch nichtlineare Verstärkerkennlinien und zeichnen sich dadurch aus, dass zusätzliche Frequenzen auftreten, die im ursprünglichen Frequenzband nicht enthalten waren (Abb. 2.50). Wird zum Beispiel ein unverzerrtes Sinussignal über einen Transistor mit nichtlinearer Kennlinie verstärkt, so erscheinen nach einer Fouriertransformation anstatt einer einzelnen Linie bei der Schwingungsfrequenz zusätzlich zur Grundfrequenz mit der Amplitude

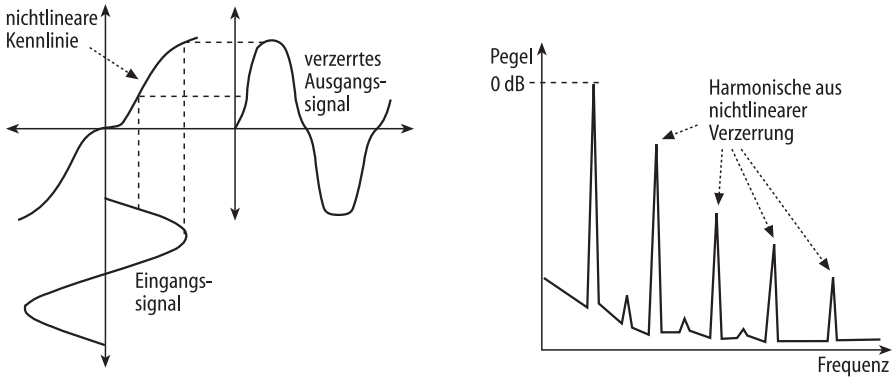


Abb. 2.50. Entstehung nichtlinearer Verzerrungen

A_1 Oberwellen mit nichtverschwindenden Amplituden A_2, A_3 etc. (Abb. 2.50). Zur quantitativen Bestimmung der nichtlinearen Verzerrungen, vor allem im Audiobereich, dient der Klirrfaktor k mit

$$k = \sqrt{(A_2^2 + A_3^2 + ...)} / A_1.$$

Neben der Verzerrung können Nutzsignale von verschiedenen Störsignalen, wie z. B. Netzeinstrahlungen (Brummen), beeinflusst werden. Nicht deterministische, zufällige Störsignale werden als Rauschen bezeichnet. Schon an einem einfachen elektrischen Widerstand R tritt das unvermeidliche thermische Rauschen auf, das durch die von der Wärme erzeugte, unregelmäßige Elektronenbewegung hervorgerufen wird. Die thermische Rauschleistung hängt neben der Temperatur T von der Übertragungsbandbreite B ab. Mit der Boltzmannkonstanten $k = 1{,}38 \cdot 10^{-23}$ Ws/K gilt die Beziehung:

$$P_r = k\ T\ B.$$

Für das Videosignal kann abgeschätzt werden, dass ein Störspannungsabstand von > 42 dB erforderlich ist, um eine gute Bildqualität zu erhalten.

2.3.3.2 Leitungen

Für die Übertragung von Videosignalen in ihrem ursprünglichen Frequenzbereich (Basisband) werden im professionellen Bereich Koaxialkabel verwendet. Als Steckverbindungen dienen verriegelbare BNC-Stecker (Abb. 2.51) (Bayonet Neill-Concelman). Im Heimbereich kommen aus Kostengründen auch einfache, abgeschirmte Leitungen mit Cinch-Steckern zum Einsatz oder die (mechanisch sehr instabile) 21-polige SCART-Steckverbindung, über die auch Audio- und weitere Signale geführt werden (Abb. 2.52).

Koaxialleitungen sind bei größeren Längen bis zu Grenzfrequenzen von etwa 500 MHz nutzbar, die frequenzabhängige Dämpfung bleibt bei hochwertigen Typen bis zu einer Frequenz von 100 MHz unter 7 dB pro 100 m Länge. Nur bei sehr großen Längen sind Leitungsentzerrer erforderlich, die den stärkeren Verlust von hochfrequenten gegenüber niederfrequenten Signalanteilen ausgleichen (Abb. 2.53). Bei der Basisbandübertragung wird die hohe Bandbreite des Übertragungsmediums schlecht genutzt. Sie ist relativ unwirtschaftlich und wird nur für kurze Übertragungsstrecken verwendet.

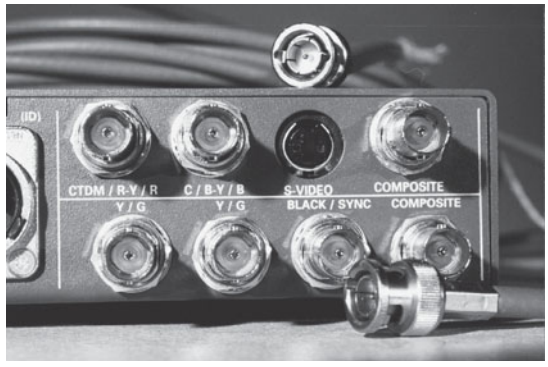


Abb. 2.51. BNC-Steckverbindungen

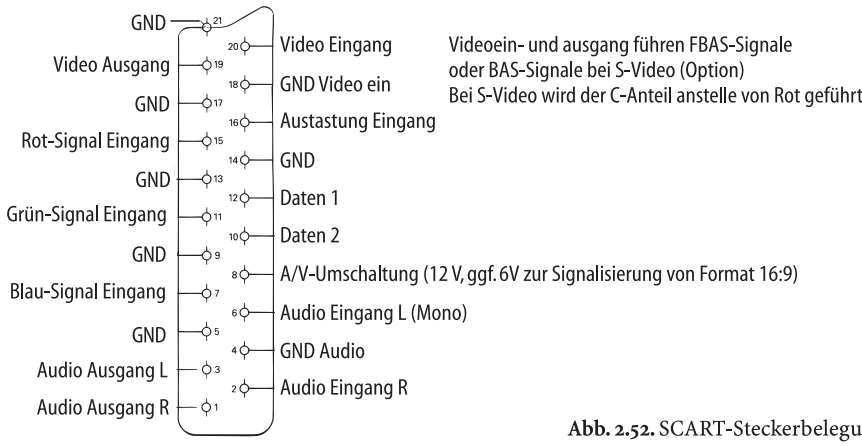


Abb. 2.52. SCART-Steckerbelegung

Elektrische Leitungen weisen einen ohmschen Widerstand auf und dämpfen das Signal. Aufgrund der bei hohen Frequenzen auftretenden Stromverdrängung aus dem Leiterinneren werden hohe Frequenzen stärker bedämpft als tiefe (Skineffekt). Die Leiter haben eine Induktivität und das Hin- und Rückleiterpaar bildet eine Kapazität. Bei gleichmäßiger Ausführung hat das Medium einen definierten Widerstand für die Wellen, die sich auf der Leitung ausbreiten. Unabhängig von der Leitungslänge gilt: $Z = \sqrt{L/C}$. In der Videotechnik werden Leitungen mit einem Wellenwiderstand von $Z = 75\ \Omega$ benutzt.

Wenn die Wellen zum Leitungsende gelangen und dort auf einen anderen Widerstand treffen, können sie, wie Licht beim Übergang zwischen verschiedenen Medien, reflektiert werden. Um dies zu vermeiden, müssen Videoleitungen immer mit dem Wellenwiderstand abgeschlossen werden. Aus dem gleichen Grund müssen auch Knickstellen in der Leitung vermieden werden, denn sie sind stets mit einer Änderung des Wellenwiderstands verbunden. Der 75-Ω-Leitungsabschluss wird durch Widerstände an den Geräten erreicht.

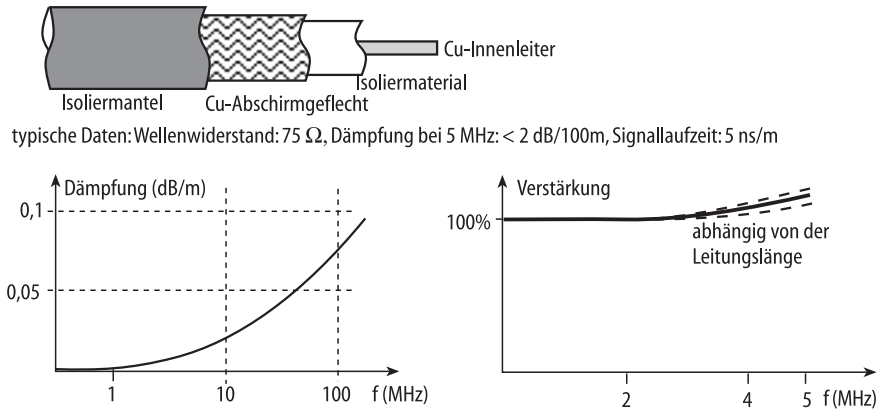


Abb. 2.53. Koaxialkabelaufbau, Leitungsdämpfung und Ausgleichsentzerrung

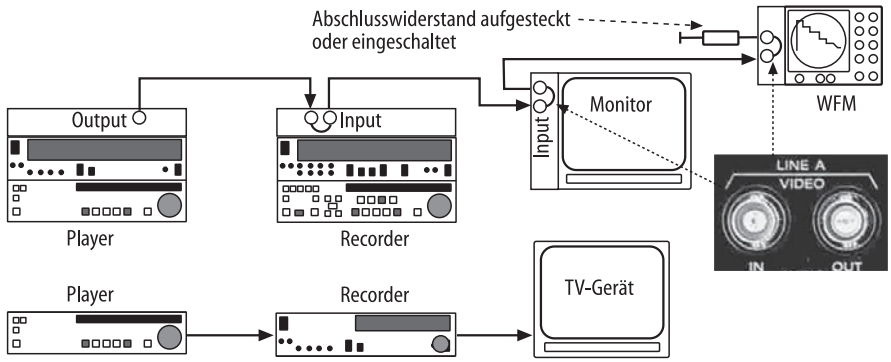


Abb. 2.54. Signalverteilung zu mehreren Geräten

Bei professionellen Geräten ist der Widerstand häufig abschaltbar und es sind zwei parallel geschaltete BNC-Buchsen vorhanden. Damit kann das Signal zu mehreren Geräten geführt werden (z. B. zum Recorder und zum Monitor). Der Abschlusswiderstand wird nur einmal, am Ende der Kette, eingeschaltet bzw. aufgesteckt (Abb. 2.54, oben). Bei fest eingebauten Widerständen würde sich der Widerstandswert durch die Parallelschaltung zweier angeschlossener Geräte halbieren. Bei Heimgeräten gibt es die Möglichkeit der Verteilung nicht, die Nutzer wären mit dem richtigen Abschluss überfordert. Hier wird eine Signalweitergabe, wie in Abbildung 2.54 unten gezeigt, nur mittels Durchlaufens der elektronischen Schaltungen in den Geräten ermöglicht, was bedeutet, dass ein fehlerhaftes Gerät den Fehler an die nachfolgenden Geräte weitergibt.

Die Ausbreitungsgeschwindigkeit des Signals beträgt ca. $2 \cdot 10^8$ m/s in der Leitung. Laufen zusammengehörige Signalanteile über verschiedene Wege, so können die Laufzeitdifferenzen unter Umständen zu störenden Phasenverschiebungen führen. Wird z. B. eine Kamera von einem zentralen Studiotaktgeber gesteuert, der 500 m entfernt ist, und wird das Videosignal über die gleiche Strecke zu einem Videomischer geführt, so bewirken die 1000 m Kabellängendifferenz eine Laufzeitdifferenz von 5 µs, also ca. 8% der Horizontalperiodendauer. Damit die Bilder bei der Mischung mit einem Signal aus einer Kamera, die über eine kurze Leitung angeschlossen ist, nicht um diesen Betrag

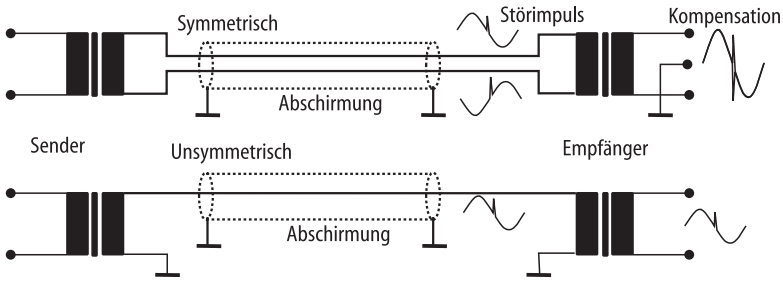


Abb. 2.55. Symmetrische und unsymmetrische Signalführung

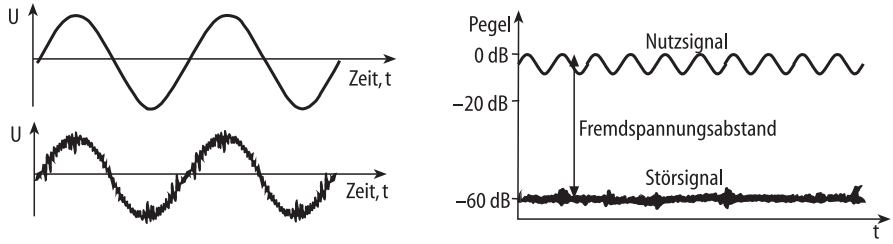


Abb. 2.56. Verrauschtes Signal

horizontal gegeneinander verschoben erscheinen, müssen die Laufzeiten durch einen H-Phasenabgleich elektronisch angeglichen werden.

Die Übertragung von Audiosignalen im Basisband mit 20 kHz Bandbreite ist bezüglich der Signalreflexion unkritisch. Audiosignale sind jedoch sehr anfällig gegen äußere Störsignale (Brummen), daher sollten sie symmetrisch geführt werden, sodass sich von außen induzierte Störungen aufheben können. Die symmetrische Signalführung wird mit einer Abschirmung und zwei Leitern realisiert, die gegenphasig beide das gleiche Signal tragen. Bei der Zusammenführung addieren sich, im Gegensatz zur unsymmetrischen Übertragung, die Nutzsignale, während sich Störsignale, die auf beiden Leitern gleichphasig auftreten weitgehend kompensieren (Abb. 2.55). Diese Art der Störreduktion wird nicht nur bei Audiosignalen verwendet, sondern z. B. auch bei der seriellen Datenübertragung mit der RS-422-Schnittstelle (s. Kap 9.3) oder bei der Parallelübertragung digitaler Videosignale. Zur symmetrischen Audiosignalführung werden 3-polige XLR-Steckverbindungen verwendet.

2.3.3.3 Verstärkung

Zum Ausgleich der auftretenden Dämpfungen sind Verstärker erforderlich. Ein einfach einzusetzender Verstärkertyp ist der Operationsverstärker. Dies ist ein Universalverstärker mit sehr hoher Leerlaufverstärkung (z. B. Faktor 10^4), die durch einfache Gegenkopplung angepasst werden kann.

Verstärkungsfaktoren, Spannungen, Leistungen etc. werden in der Nachrichtentechnik häufig als relative logarithmische Werte a mit der Einheit deziBel (dB) angegeben. Bei einer über Spannungsverhältnisse definierten Dämpfung oder Verstärkung $V_u = U_2/U_1$ wird das 20-fache des 10er-Logarithmus verwendet, bei Leistungsverhältnissen V_p das 10-fache, da die Leistung von der Amplitude zum Quadrat bestimmt ist. Für die Angabe absoluter Pegel ist der Bezug zu einer Referenzgröße erforderlich. Für Spannungen bzw. Leistungen gilt:

$$a = 20 \log U/U_{ref}, \text{ bzw. } a = 10 \log P/P_{ref}.$$

In der Audiotechnik ist der meistverwandte Bezugswert U_{ref} = 0,775 V (Bezugsindex u), der professionelle Audio-Studiopegel hat in Deutschland den Wert +6 dBu entsprechend 1,55 V. Semiprofessionelle Audiogeräte arbeiten oft mit – 10 dBV, die Einheit dBV ist auf den Wert U_{ref} = 1 V bezogen.

Auch beim Verhältnis von Signal- zu Rauschleistung P_S/P_N (Signal to Noise, S/N) werden Pegelangaben verwendet (Abb. 2.56). Es gilt:

$$S/N = 10 \log P_S/P_N.$$

2.4 Das Farbsignal

Der Begriff »Farbe« bezieht sich auf eine Sinneswahrnehmung, die durch einen Farbreiz aus der physikalisch beschriebenen Lichtstrahlung ausgelöst wird. Die Wahrnehmung des Farbtons hängt von der Frequenz der Strahlung ab. Farben können nicht absolut, sondern nur im Vergleich angegeben werden. Da verschiedene Farbreize die selbe Wahrnehmung hervorrufen können, wird unabhängig vom Farbreiz gleich aussehenden Farben die selbe Farbvalenz zugeordnet. Farbiges Licht kann aus weißem gewonnen werden, indem letzteres mit Hilfe eines Prismas in die Spektralfarben zerlegt wird, die sich auch am Regenbogen im Sonnenlicht zeigen (Abb. 2.57). Das Prisma hat die Eigenschaft, das Licht wellenlängenabhängig unterschiedlich stark zu brechen. Der Lichtweg kann hier auch umgekehrt betrachtet werden, d. h. dass aus der Vereinigung der Spektralfarben wieder Weiß entsteht. Zur Gewinnung von Weiß oder anderen Farben ist es dabei keine notwendige Bedingung, alle Spektralfarben einzusetzen. Um einen Großteil der Farben wiedergeben zu können, reicht es aus, eine additive Mischung aus den drei Grundfarben Rot, Grün und Blau vorzunehmen. Dieser Umstand wird beim Farbfernsehsystem ausgenutzt. Zur Übertragung des Farbsignals ist damit nicht für jede Spektralfarbe eine eigene Information erforderlich, sondern es reichen drei Kanäle aus, um einen großen Teil der Farben wiedergeben zu können, die bei Sonnenbeleuchtung von natürlichen Gegenständen reflektiert werden.

Die Farbigkeit eines Bildpunktes auf dem Röhrenbildschirm entsteht aus der Überlagerung des Lichts, das von drei rot-, grün- und blauleuchtenden Leuchtstoffen herrührt. Ein Bildpunkt entspricht einem Farbtripel. Wie im Falle der S/W-Wiedergabe werden die Leuchtstoffe von Elektronenstrahlen angeregt. In einer gewöhnlichen Farbbildröhre werden drei Strahlen gemeinsam zu den Farbtripeln geführt, wobei mit einer Schattenmaske dafür gesorgt wird, dass jeder Strahl immer nur eine Art von Leuchtstoff trifft. Damit kann die Intensität der jeweiligen Farbe über die Intensität des zugehörigen Elektronenstrahls gesteuert werden. Zur Ansteuerung der Farbbildröhren sind also drei elektrische Signale erforderlich, die als Farbwertsignale bezeichnet werden. Sie entstehen in drei Bildwandlern in der Kamera, wobei dem jeweiligen Bildwandler mit Hilfe optischer Filter nur einer der drei Farbauszüge zugeführt wird (Abb. 2.58). Zur Ermittlung genauerer Zusammenhänge müssen zunächst einige Fakten über Lichtquellen, Farbempfindung und Farbmischung betrachtet werden.

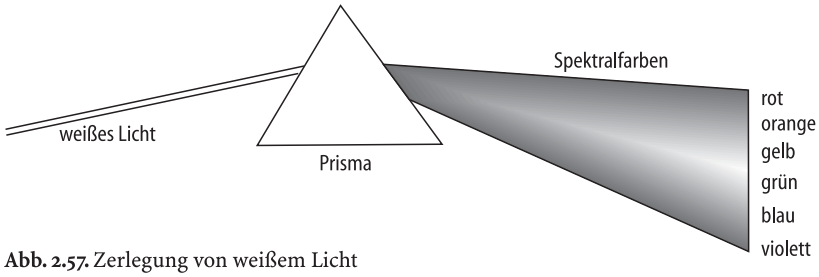


Abb. 2.57. Zerlegung von weißem Licht

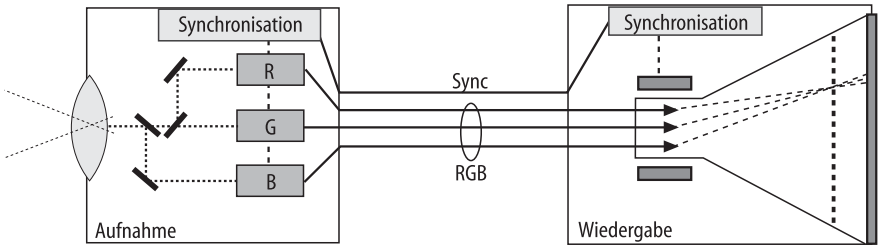


Abb. 2.58. Farbbild-Übertragungsprinzip

2.4.1 Farbe und Wahrnehmung

Das Licht als elektromagnetische Strahlung und einige Aspekte der Helligkeitswahrnehmung sind in Kap. 2.1 beschrieben. Hier werden nun die Gesichtspunkte dargestellt, die sich auf die Farbe beziehen.

2.4.1.1 Temperatur- und Linienstrahler

Erhitzte Körper, wie z. B. die Sonne, strahlen ein Wellenlängengemisch ab, das wir bei genügend hoher Temperatur als weiß empfinden. Bei der Untersuchung eines Gemischs zeigt sich, dass nicht alle Frequenzen gleichmäßig intensiv vertreten sind. Das Maximum des Spektrums verschiebt sich mit steigender Temperatur zu kürzeren Wellenlängen hin, die Strahlungsenergie steigt mit der vierten Potenz der absoluten Temperatur T.

Abbildung 2.59 zeigt Spektren von sogenannten schwarzen Strahlern, idealen Temperaturstrahlern, bei denen Absorption und Emission im Gleichgewicht stehen. Reale Strahler verhalten sich nicht ideal, qualitativ lassen sie sich aber durch die Spektralkurven der schwarzen Strahler hinreichend beschreiben. Das gesamte Spektrum eines Temperaturstrahlers lässt sich damit durch die Angabe der Farbtemperatur charakterisieren. Sonnenlicht hat etwa eine Farbtemperatur von 5800 Kelvin. Das Sonnenlicht wird in der Atmosphä-

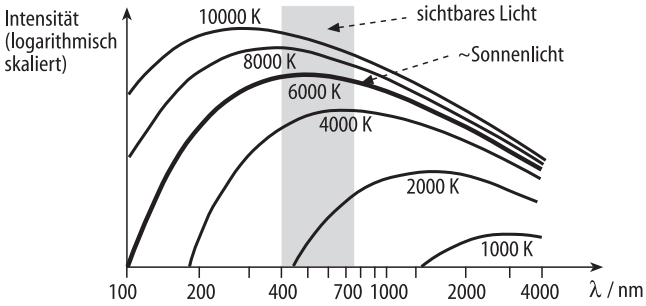


Abb. 2.59. Frequenzspektren schwarzer Strahler bei verschiedenen Temperaturen

Tabelle 2.6. Farbtemperaturen künstlicher und natürlichen Lichtquellen

Kerzenlicht	1850 K	Mondlicht	4100 K
25-W-Glühlicht	2500 K	Sonne früh/spät	4300 K
40-W-Glühlicht	2650 K	Sonne direkt	5800 K
Halogenlampe	3200 K	Normlicht D65	6504 K
HMI-Lampe	5600 K	Himmel, bedeckt	7000 K

re gestreut. Dabei tritt der Blauanteil verstärkt auf, sodass sich bei bedecktem Himmel die Farbtemperatur zu höheren Werten hin verschiebt. Entsprechend wurde für ein Referenzlicht ein Temperaturstrahler mit 6504 K zugrunde gelegt, es wird als Normlichtart D65 bezeichnet.

Eine Kunstlichtquelle (Glühlampe, Tungsten Lamp) erzielt nur kleinere Werte, bei professionellen Scheinwerfern werden ca. 3200 K erreicht (Tabelle 2.6). Wird eine Glühlampe nicht mit Nennleistung betrieben, so verändert sich mit der Leistungsaufnahme auch die Farbtemperatur.

Als Einheit für die Farbtemperatur wird neben Kelvin auch die besser an die Farbempfindung angepasste Einheit Micro Reciprocal Degrees M verwendet. Es gilt: $M = 10^6/T$. Die zugehörige Einheit mired (mrd) entspricht auch dem Kehrwert von Megakelvin $(MK)^{-1}$. Für das menschliche Auge sind Farbtemperaturdifferenzen ab 10 mrd sichtbar.

Kunstlicht lässt sich nicht nur mit Temperatur-, sondern auch mit Linienstrahlern erzeugen. Linienstrahler sind Gasentladungslampen, die ein Gemisch aus charakteristischen Frequenzen (Spektrallinien) aussenden (Abb. 2.60), das nicht kontinuierlich ist. Durch die Wahl des Gasgemischs kann das Intensitätsverhältnis zwischen den Frequenzen beeinflusst werden, z. B. derart, dass das Licht insgesamt eine Empfindung wie Tageslicht hervorruft.

2.4.1.2 Körperfarben

Körper, die nicht selbst Licht abstrahlen, reflektieren und absorbieren Licht aus externen Quellen in anderer Weise als ihre Umgebung und heben sich dadurch von ihr ab. Der Körper nimmt das Licht auf und strahlt es verändert als Körperfarbe wieder ab. Wenn nun der Reflexions- oder Absorptionsgrad von der Wellenlänge des Lichtes abhängt, erscheinen die Körper in weißem Licht farbig, da nicht mehr alle zur Erzielung des richtigen Weißeindrucks erforderlichen Frequenzen vertreten sind (subtraktive Mischung).

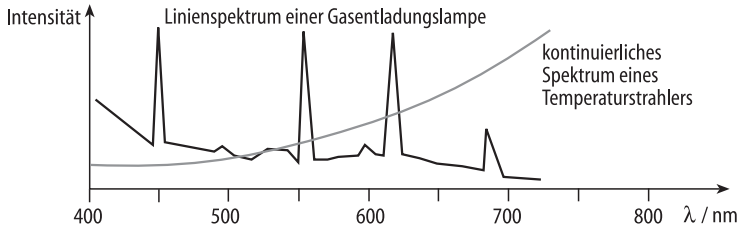


Abb. 2.60. Spektren von Gasentladungslampen und Lampen mit kontinuierlichem Spektrum

Ein gegenüber Bestrahlung mit weißem Licht veränderter Farbeindruck ergibt sich auch dann, wenn der Körper mit monochromatischem, bzw. nicht weißem Licht beleuchtet wird, denn er kann dann natürlich auch nur Licht der entsprechenden Farbe reflektieren; ein Körper, der kein Rotlicht reflektiert, erscheint in rotem Licht schwarz. Für die richtige Reproduktion des Farbeindrucks ist bei Videosystemen also die Beleuchtung und die Farbtemperatur von großer Bedeutung. Um mit Kunstlicht ähnliche Farben wie bei Tageslicht wiedergeben zu können, wurde die Normlichtart D65 definiert, die sich an der spektralen Verteilung des mittleren Tageslichts unter Einbeziehung von Streuung in der Atmosphäre mit der Farbtemperatur 6500 K orientiert.

2.4.1.3 Die Farbempfindung

Der Aufbau des menschlichen Auges ist in Kap. 2.1 beschrieben. Dort findet sich eine Einteilung der Photorezeptoren in Stäbchen und Zapfen. Die Zapfen des Auges lassen sich nach einem einfachen Modell in drei Gruppen einteilen, die jeweils besondere Empfindlichkeiten für die Farben Rot, Grün und Blau aufweisen. Alle Farbempfindungen kann man sich als unterschiedlich starke Reizung der einzelnen Gruppen vorstellen. Z. B. empfinden wir die Farbe Gelb, wenn gleichzeitig nur die rot- und grünempfindlichen Gruppen angeregt werden. Dieses Modell entspricht der additiven Mischung, bei der die Farben durch die Mischung der drei Farbwerte Rot, Grün, Blau erzeugt werden.

Die Farbwahrnehmung ist gegenüber der Helligkeitswahrnehmung mit geringerer räumlicher Auflösung verknüpft. Kleine farbige Flächen können schlechter unterschieden werden als feine Schwarz/Weiß-Muster. Abbildung 2.61 zeigt am Vergleich von abwechselnden schwarz/weißen Linien mit rot/grünen bzw. gelb/blauen Linien, dass die Kontrastempfindlichkeit für Farbmodulationen erheblich kleiner ist als für Helligkeitsmodulationen [136]. Das bedeutet für die Bildübertragung, dass es für feine Details ausreicht, nur die Helligkeitsinformation zu übertragen und auf die Farbe zu verzichten. Hier ergibt sich eine wichtige Möglichkeit zur Reduktion irrelevanter Information.

Wie an verschiedene Helligkeiten, so passt sich der menschliche Gesichtssinn auch an unterschiedliche Farbgemische an. Das von einem gewöhnlichen Blatt Papier reflektierte Licht wird als weiß empfunden, egal ob es mit Tages- oder Kunstlicht beleuchtet wird. Der Unterschied der Farbgemische wird uns nur deutlich, wenn wir beide Lichtarten gleichzeitig sehen, z. B. ein beleuchtetes

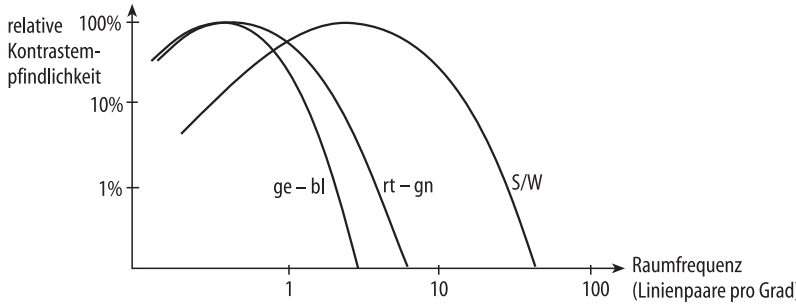


Abb. 2.61. Vergleich der Sehempfindlichkeit für farbart- und helligkeitsmodulierte Raumgitter

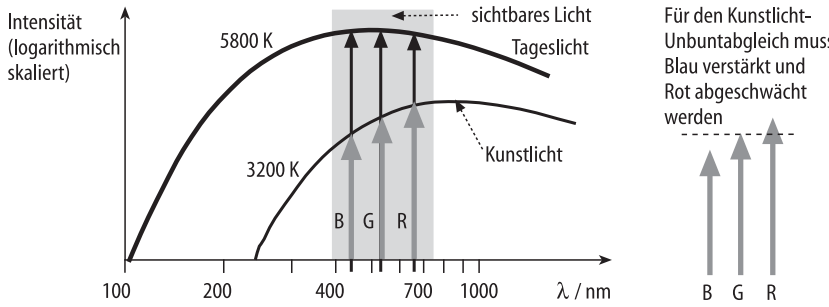


Abb. 2.62. Unbuntabgleich durch elektronische Verstärkung

Haus in der Dämmerung. Im Gegensatz zum menschlichen Auge passen sich technische Geräte zur Bildwandlung (außer bei Vorrichtungen zum automatischen Abgleich) nicht an verschiedene Farbgemische an. Hier muss für die jeweiligen Lichtverhältnisse ein Unbunt-Abgleich vorgenommen werden, d. h. die Geräte müssen so justiert werden, dass sich unabhängig von der Lichtart bei einer weißen Bildvorlage gleiche elektronische Signale ergeben (Abb. 2.62).

2.4.2 Farbmischung

Der Begriff »Farbe« lässt sich nach dem Helligkeitsanteil (engl.: luminance) und der Farbart differenzieren; der Begriff Farbart wiederum in den durch die Lichtwellenlänge bestimmten Farbton (engl.: hue) und die durch den zugemischten Weißanteil bestimmte Farbsättigung (engl.: saturation). Da der Begriff »Farbe« auch die Helligkeit umfasst, ist es korrekter von Buntheit zu sprechen, wenn umgangssprachlich »Farbe« gemeint ist. Bei Grauwerten spricht man demgemäß von »unbunten Farben«.

Untersuchungen der Abhängigkeit der Farbempfindung vom Farbreiz haben ergeben, dass es möglich ist, die meisten Körperfarben durch die Mischung von nur wenigen einfarbigen Lichtern nachzubilden. Bei der additiven Farbmischung werden dazu meist drei Farben verwendet. Diese Grundfarben heißen Primärvalenzen, wenn keine der Farben aus den anderen ermischbar ist. Allgemein sind dafür die drei spektral reinen Farben Rot (700 nm), Grün (546 nm) und Blau (436 nm) festgelegt.

Die Anteile der Farben an der Gesamtmischung werden als Farbwerte bezeichnet, sie ergeben zusammen die Farbvalenz. Bei Gleichheit der drei Farbwerte entsteht Weiß: R = B = G = 1. Bei gleichmäßiger Reduktion aller Farbwerte, z. B. auf R = B = G = 0,5, ergibt sich Grau. Die Mischung von jeweils nur zwei der drei Farben mit jeweils 100% Intensität ergibt die folgenden Kombinationen: Aus Rot und Grün bildet sich Gelb, aus Grün und Blau Cyan und aus Blau und Rot Magenta. Alle anderen Farben ergeben sich durch Veränderung der Intensitätsverhältnisse beim Mischen. Zwei Farben, die bei einer additiven Mischung Weiß ergeben, heißen Komplementärfarben. Zu Rot, Grün und Blau gehören also die Komplementärfarben Cyan, Magenta und Gelb (Abb.

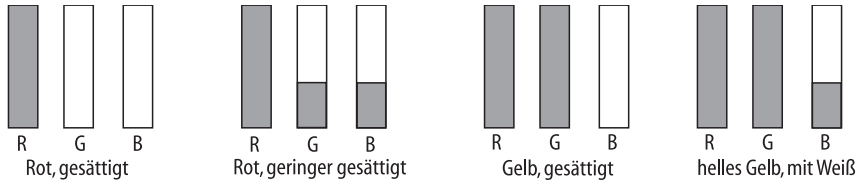


Abb. 2.63. Additive Mischung und Entsättigung

2.64). Wenn einer der Anteile im Farbgemisch nicht vertreten ist, entsteht eine gesättigte Farbe, die Sättigung sinkt durch Zumischung eines Grauwertes, d. h. durch Anhebung aller drei Komponenten um den gleichen Wert (Abb. 2.63).

Im Gegensatz zur additiven beruht die subtraktive Farbmischung auf dem Herausfiltern bestimmter Farben aus weißem Licht. So entsteht z. B. Grün, wenn ein Gelb- und ein Cyan-Filter eingesetzt werden (Abb. 2.64). Grundfarben für die subtraktive Mischung sind Cyan, Magenta und Gelb. Die subtraktive Mischung wird in der Drucktechnik angewandt, auch die Körperfarben entstehen subtraktiv.

Für die Farbfernsehtechnik interessiert vor allem die additive Mischung. Die Überlagerung von drei Lichtern entspricht der Addition von drei Farbreizfunktionen. Zur quantitativen Farbbestimmung werden zunächst die Farbreizfunktionen spezifiziert und Primärreizkurven festgelegt. Diese werden ermittelt, indem Beobachter ein Gemisch aus den Farbwerten Rot, Grün und Blau im Intensitätsverhältnis so einstellen, dass die entstehende Farbe gleich dem anzupassenden Farbreiz ist. Beim Vergleich von Farbwirkungen, die mit verschiedenen Primärreizkurvensystemen gewonnen werden, stellt man fest, dass sich trotz unterschiedlicher Ausgangsbasis die gleichen Farbempfindungen hervorrufen lassen. Dieser Umstand nennt sich Metamerie und ist für die Farbbildreproduktion der Fernsehtechnik von großer Bedeutung, denn auf der Wiedergabeseite soll der auf der Aufnahmeseite vorliegende Farbreiz aus drei Leuchtstoffen ermischt werden, deren Atome zur Lichtemission gebracht werden. Diese Leuchtstoffe (Phosphore) strahlen als Naturprodukte jedoch nicht die oben genannten spektral reinen Farben ab, sondern weisen eine eigene Frequenz und geringere Sättigung auf. Aufgrund der Metamerie ist es aber möglich, auch mit ihnen alle für die Reproduktion wesentlichen Farben darzustellen. Eine Klasse von metameren Farbreizen entspricht einer Farbvalenz.

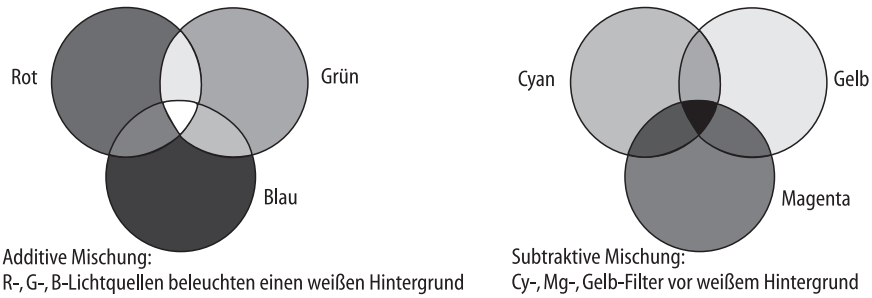


Abb. 2.64. Additive und subtraktive Mischung

2.4.3 Farbmetrische Größen

Die Farbmetrik dient der quantitativen Farbfestlegung. Eine Farbmetrik zu entwickeln, ist nicht einfach. Das zeigt sich schon an der Tatsache, dass es unmöglich ist, sich exakt über Farben sprachlich auszutauschen. Die Farbwahrnehmung hängt stark von äußeren Umständen, wie anderen Farben in der Nähe, aber auch kulturellen Einflüssen oder der menschlichen Gefühlswelt ab. Zwar ist der Mensch in der Lage, auf Farbtonänderungen sensibel zu reagieren, absolut kann er Farben jedoch schlecht angeben, am ehesten den Hautton als Erinnerungsfarbe. Die farbmetrischen Größen wurden dementsprechend auf der Basis von Farbvergleichen ermittelt.

Nachdem seit ca. 1860 das Farbwahrnehmungsmodell bekannt war, das auf der Vorstellung einer Reizung von rot-, grün- und blauempfindlichen Bereichen in der Netzhaut beruht, wurden um 1930 Experimente mit Menschengruppen durchgeführt, die diese Anregung quantitativ erfassten. Bezogen auf die Ergebnisse wurden die Gruppen als „Normalbeobachter“ bezeichnet. Dieser Bezeichnung wird noch der Wert 2° vorangestellt, weil die Farbvergleichsexperimente mit einem Sichtfeld vorgenommen wurden das auf einen Öffnungswinkel von 2° beschränkt war. Damit wurde dem Umstand Rechnung getragen, dass die farbrezeptiven Zapfen auf einen kleinen Raum beschränkt sind.

Die Experimente wurden so durchgeführt, dass auf einem zweigeteilten Schirm auf der einen Seite eine bestimmte Farbe dargestellt wurde und auf der anderen ein Farbgemisch aus rot-, grün- und blauleuchtenden monochromatischen Lichtquellen mit den Wellenlängen 700 nm, 546 nm und 436 nm. Die Probanden bekamen die Aufgabe, durch Variation der Intensität der drei Quellen das Lichtgemisch der Probefarbe möglichst weit anzunähern. Das Verfahren wurde dadurch optimiert, dass es möglich war, für Farben, bei denen der Angleich schlecht gelang, der Probefarbe noch Licht der anderen Grundfarben hinzuzumischen, was dann als negative Farbwerte gewertet wurde. Auf diese Weise entstanden die in Abbildung 2.65 dargestellten Spektralwertkurven der spektralen Primärvalenzen, die so normiert sind, dass sich für alle drei Anteile eine einheitliche Fläche unter der Kurve ergibt. Darin wird deutlich, dass sich die genannten negativen Anteile auf die Farbe Rot beziehen, dass sich also insbesondere bestimmte Blau-Grün-Töne ohne diese nicht ermischen ließen.

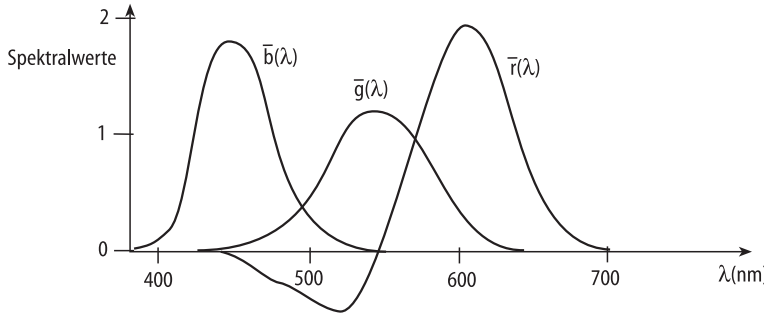


Abb. 2.65. Spektralwertkurven der spektralen Primärvalenzen

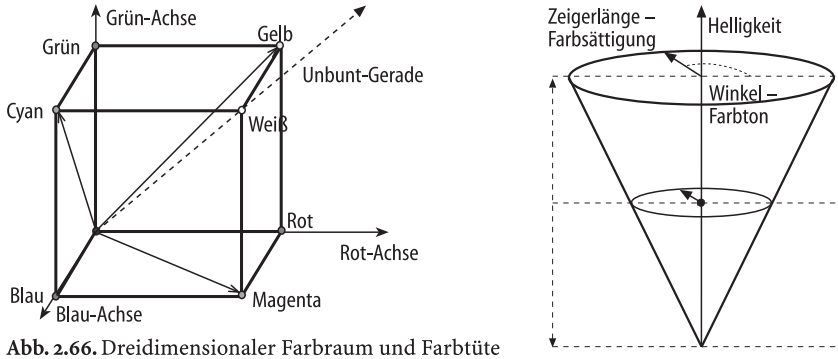


Abb. 2.66. Dreidimensionaler Farbraum und Farbtüte

Die Darstellung von Farben kann nun dreidimensional in einem Farbraum vorgenommen werden, dessen Achsen durch die Primärvalenzen R, G, B aufgespannt werden (Abb. 2.66). Vereinfacht lässt sich die Farbe auch in einer sogenannte Farbtüte darstellen, indem gemäß Abbildung 2.66 im dreidimensionalen Raum nach oben die Helligkeit und in der Ebene der Farbartkreis aufgetragen wird. Die Zeigerlage (Winkel) gibt hier den Farbton und die Zeigerlänge die Farbsättigung an.

Farbvalenzen mit gleicher Leuchtdichte unterscheiden sich nur in ihrer Farbart. Die Information über die Farbart allein lässt sich in einer Ebene darstellen, die sich aus einem diagonalen Schnitt durch den RGB-Würfel als sogenantes Primärvalenzfarbdreieck ergibt. Jeder Punkt in der Ebene gehört zu einer bestimmten Farbart und einem Farbort. Damit lässt sich die Farbmischung mit Hilfe der Schwerpunktregel angeben, d. h. der Farbort der Mischung ergibt sich als gemeinsamer Schwerpunkt, wenn den Farben an den Eckpunkten Gewichte zugeordnet werden. Farborte, die sich aus der Mischung positiver Farbwerte ergeben, liegen innerhalb des Farbdreiecks (innere Mischung) (Abb. 2.67). Um die Darstellung zu vereinfachen, werden die Farbwerte auf R + G + B normiert und es entstehen die Farbwertanteile r, g, b:

$$r = R / (R + G + B),\ g = G / (R + G + B),\ b = B / (R + G + B).$$

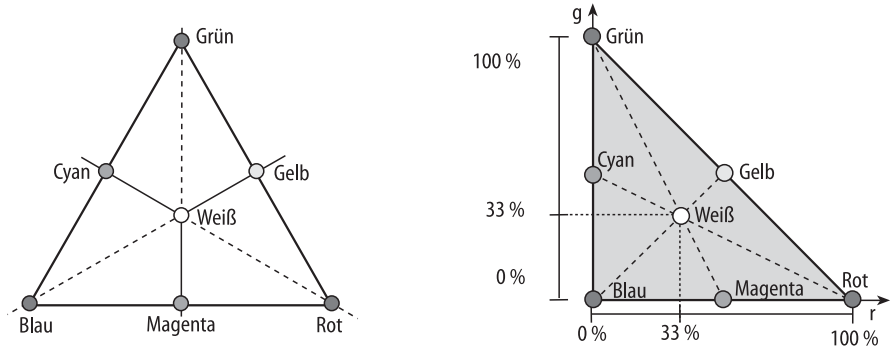


Abb. 2.67. Schnitt durch den dreidimensionalen Farbraum und r-g-Fläche

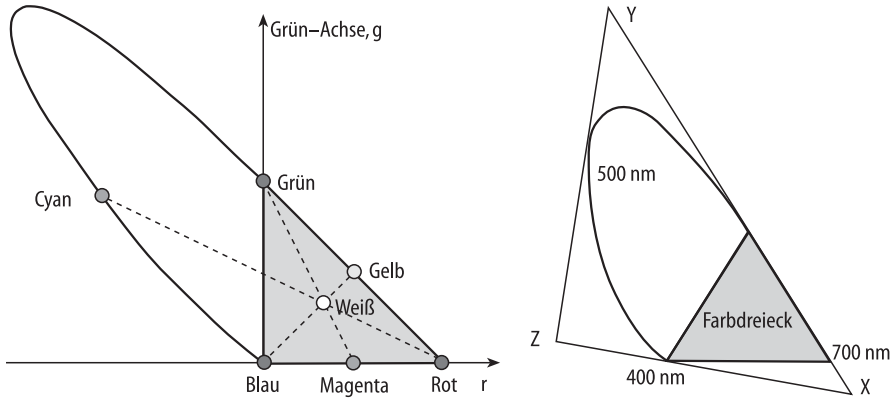


Abb. 2.68. Primärvalenzen im r-g- und XY-Koordinatensystem

Damit lässt sich der dreidimensionale Farbraum auf die zwei Dimensionen r und g reduzieren, denn die dritte ergibt sich aus r + g + b = 1 (Abb. 2.67).

Wenn in ein r/g-Koordinatensystem nun die Farbwerte der Normalbeobachter eingetragen werden, entsteht die in Abbildung 2.68 dargestellte hufeisenförmig begrenzte Fläche, deren Rand durch die maximal gesättigten Spektralfarben gebildet wird. Die zur Mischung benutzten Primärvalenzen bilden das graue Dreieck. Es wird auch hier deutlich, dass die außerhalb des Dreiecks liegenden Farben durch innere Mischung nicht nachgebildet werden können, bzw. die Bestimmung der Farbwerte einiger Spektralfarben negative Werte erfordert. Die Farbwerte der Spektralfarben werden als Spektralwerte bezeichnet.

Um die Darstellung weiter zu vereinfachen, wird ein Wechsel des Primärvalenzsystems vorgenommen, hin zu einem System, in dem keine negativen Werte verwendet werden müssen (Abb. 2.68). Das System wird dann als Normvalenzsystem bezeichnet und mit den Normfarbwerten X, Y, Z, gebildet. Der Wechsel des Primärvalenzsystems mit den Koordinaten R, G, B zum System X, Y, Z erfolgt als einfache lineare Koordinatentransformation.

Die Normfarbwerte X, Y, Z werden wiederum zu den Normfarbwertanteilen x, y, z normiert, sodass die Summe der drei Zahlen gleich eins ist:

$$x = X/(X + Y + Z),\ y = Y/(X + Y + Z),\ z = Z/(X + Y + Z).$$

Die Spektralwertkurven in diesem System werden als Normspektralwertkurven bezeichnet und sind in Abbildung 2.69 dargestellt. Die Normfarbwerte sind so gewählt, dass sich der gesamte Spektralfarbenzug in einem Quadranten des x/y-Koordinatensystems befindet (Abb. 2.70). Die Normfarbwerte liegen außerhalb des vom Spektralfarbenzug umschlossenen Gebietes und sind physikalisch nicht realisierbar. Sie werden als virtuelle Primärvalenzen bezeichnet. Die Lage des Koordinatensystems ist so bestimmt, dass die zu y gehörige Spektralwertkurve der spektralen Hellempfindlichkeitskurve $V(\lambda)$ des menschlichen Auges proportional ist. Damit ist der Normfarbwert Y proportional zur Leuchtdichte, was für die Farbfernsehtechnik von großer Bedeutung ist.

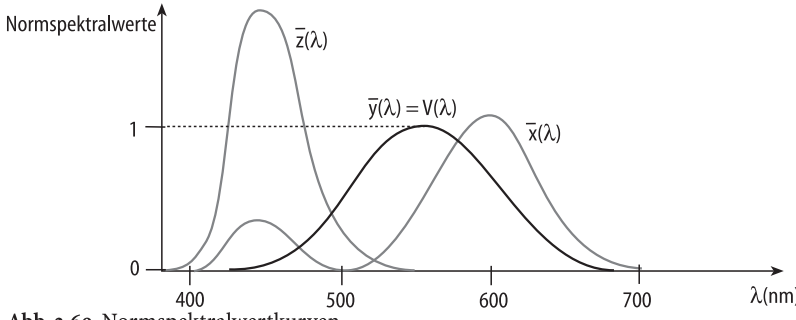


Abb. 2.69. Normspektralwertkurven

Die Normfarbtafel mit x/y-Koordinaten wurde von der internationalen Beleuchtungskommission (Commission Internationale d'Eclairage, CIE) als Bezug für die Farbmetrik genormt (Abb. 2.70). Farborte der monochromatischen Strahlung bilden hier den hufeisenförmigen Spektralfarbenzug am Rand. Das Hufeisen wird zwischen dem kurz- und langwelligen Ende durch die sogenannte Purpurgerade verbunden. Hier liegen die Purpurfarben, die nur aus dem Gemisch der Farben der Endpunkte entstehen (Blau und Rot) können.

Alle Farben liegen innerhalb des Kurvenzugs. Der Weißpunkt (Gleichenergieweiß E) hat die Koordinaten x = y = z = 0,33. Die Normlichtart D65 liegt in der Nähe dieses Wertes und hat die Koordinaten x = 0, 3127 und y= 0,3290. Jede Gerade, die vom Weißpunkt ausgeht, kennzeichnet eine Farbart, mit gleichem Farbton aber verschiedener Farbsättigung. Der Farbton wird durch den Ort auf dem Spektralfarbenzug angegeben, die Sättigung durch den Abstand

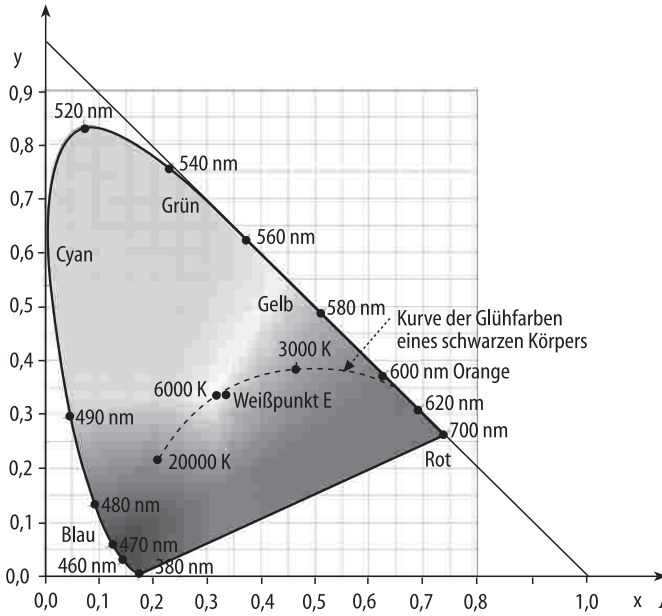


Abb. 2.70. CIE-Diagramm

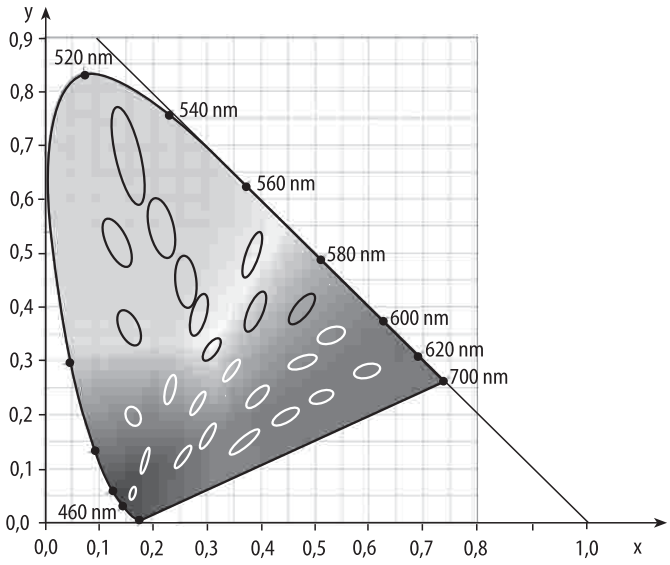


Abb. 2.71. McAdam-Ellipsen

vom Weißpunkt. Farborte von Mischfarben liegen auf einer geraden Verbindungslinie zwischen denen der Mischungskomponenten. Wenn die Mischung aus drei oder mehr Farben herrührt, können nur solche ermischt werden, die im Inneren des durch sie aufgespannten Dreiecks, bzw. Vielecks liegen. Eine große darstellbare Farbpalette ergibt sich also, wenn die Farborte möglichst in den „Ecken" des Hufeisens liegen, also bei Rot, Grün und Blau.

Es ist zu bedenken, dass das CIE-Diagramm nur die Farbart darstellt. Alle Farben lassen sich vollständig nur durch Hinzunahme der Helligkeitsinformation Y, also durch das Tripel xyY beschreiben, z. B. ist Braun ein dunkles Orange.

In Abbildung 2.70 wird deutlich, dass die Spektralanteile zwischen 490 nm und 580 nm (cyan/grün/gelb) über eine sehr weite Strecke des Kurvenzugs verteilt sind, während sich z. B. die Anteile für Blau (380 nm bis 480 nm) sehr zusammendrängen. Das deutet auf einen Nachteil des CIEXYZ-Farbmodells hin, nämlich dass die vom Menschen empfundenen Farbunterschiede nicht gleichabständig dargestellt werden.

Im Jahre 1942 hat McAdam diese Fähigkeiten zur Farbdifferenzierung untersucht. In Abbildung 2.71 sind die dort gefundenen Farbauflösungsfähigkeiten als sogenannte McAdam-Ellipsen eingetragen. Dabei entsprechen große Ellipsenflächen geringen Farbauflösungsfähigkeiten und umgekehrt. Die Lage der Ellipsen und die Ellipsenform lässt erkennen, dass der Bereich der Grüntöne sehr ausgedehnt erscheint und die Auflösungsfähigkeiten in Richtung Rot-Grün größer sind als in Richtung Gelb-Blau.

Um zu Farbräumen zu kommen die eine gleichabständige Darstellung der Farbwahrnehmung des Menschen erlauben, wurden 1976 zusätzlich zu CIEXYZ die Farbsysteme CIELAB und CIELUV definiert. Letzteres wird oft zur Darstellung der Farbräume von Displaysystemen verwendet. Der Übergang zwischen

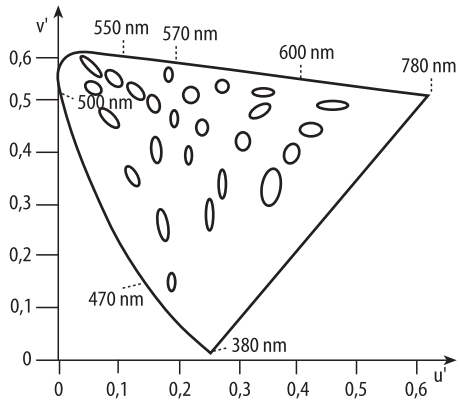


Abb. 2.72. Spektralfarbenzug und McAdam-Ellipsen im CIELUV-Koordinatensystem

den Farbräumen CIEXYZ und CIELUV ist mit den Größen L für den Helligkeitswert sowie u´ und v´ für die Buntheitsanteile durch die drei Beziehungen $L = 116 \cdot (y/Y_n)^{1/3} - 16$; $u' = 4X / (X + 15\,Y + 3\,Z)$ und $v' = 9Y / (x + 15Y + 3\,Z)$ gegeben.

Abbildung 2.72 zeigt den Verlauf der Spektralfarbenzugs im CIELUV-Farbraum. Im Vergleich zum CIEXYZ-System und an den dort wieder eingetragenen Ellipsen wird deutlich, dass sie ausgeglichener sind, da die Strecken im Gelbgrün-Bereich nun gestaucht sind während sie in den Bereichen um Rot bzw. Blau gedehnter dargestellt werden.

2.4.3.1 Farbmischung am Bildschirm

Für die Normung der Farbfernsehtechnik interessierte weniger das allgemeine Primärvalenzsystem sondern vor allem das Primärvalenzsystem des Displays, das sich aus den Valenzen der damals zur Wiedergabe eingesetzten Leuchtstoffe in Kathodenstrahlröhren ergab, denn damit können die Farben in der Realität ermischt werden. Um für jeden Empfänger den gleichen Farbeindruck zu erzielen, müssen die Leuchtstoffe genormt werden. In den 50er Jahren wurden zunächst für die USA Leuchtstoffe von der Federal Communications Commission (FCC) festgelegt. Die Normung erfordert auch die Festlegung eines Punktes der beschreibt, welche Art Weiß am Bildschirm entstehen soll. Tabelle 2.7 zeigt die Farborte gebräuchlicher Weißwerte. Je höher die zugeordnete Farbtemperatur ist, desto bläulicher erscheint der Weißton.

Tabelle 2.7. Farborte verschiedener Normbeleuchtungen

CIE-Normbeleuchtungen	x-Wert	y-Wert	Beschreibung
CIE-Normlicht C	0,3101	0,3162	Weißpunkt der NTSC-TV-Norm,
CIE-Normlicht E	1/3	1/3	Energiegleiches Weiß; X=Y=Z
D50 (5000K)	0,3457	0,3585	Weißpunkt für Wide-Gamut-RGB, Warmlicht
D65 (6500K)	0,312713	0,329016	Weißpunkt für sRGB und PAL/SECAM-TV
D9300 (9300K)	0,2848	0,2932	Weißpunkt für SD-Studiomonitore in Japan

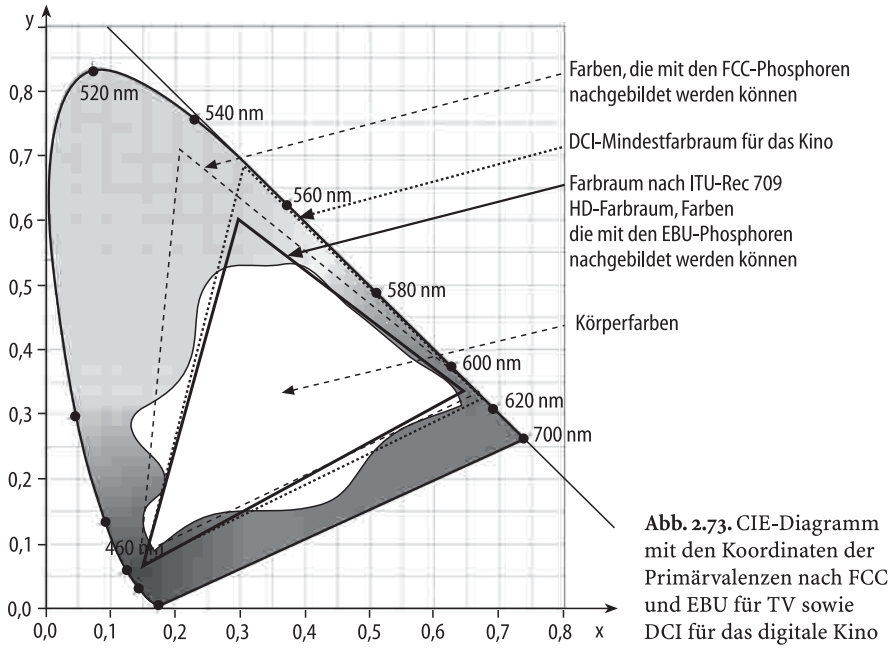


Abb. 2.73. CIE-Diagramm mit den Koordinaten der Primärvalenzen nach FCC und EBU für TV sowie DCI für das digitale Kino

Die FCC-Norm beinhaltete als Bezugsweiß die Normlichtart C. Moderne Videosysteme beziehen sich meist auf das Normlicht D65, das einen Weißeindruck hervorruft, der bei Beleuchtung mit Tageslicht aus der Atmosphäre mittags in Richtung Nord erscheint.

Von der European Broadcasting Union (EBU) wurden nicht nur ein anderer Weißwert sondern auch veränderte Leuchtstoffe genormt, die mit gleichem Strahlstrom bei der Wiedergabe eine höhere Leuchtdichte ermöglichen. Wenn diese nach FCC und EBU genormten Primärvalenzen in das CIE-Diagramm eingetragen werden, bilden sie die beiden in Abbildung 2.73 dargestellten Dreiecke. Die EBU-Primärvalenzen haben bezüglich CIExy folgende Koordinaten:

Rot: x = 0,64, y = 0,33; Grün: x = 0,29, y = 0,60; Blau: x = 0,15, y = 0,06.

Mit den drei Primärvalenzen lassen sich nur Farben darstellen, die innerhalb des jeweiligen Farbdreiecks liegen. Die EBU-Phosphore weisen insbesondere im Grünanteil eine geringere Sättigung auf. Die Qualität der Farbreproduktion wird dadurch aber kaum eingeschränkt, da die in der Natur am häufigsten vorkommenden Körperfarben (in Abb. 2.73 weiß gekennzeichnet) in beiden Fällen größtenteils innerhalb der durch die Farbwerte aufgespannten Dreiecke liegen. Beim Übergang von R, G, B auf X, Y, Z bezüglich der FCC-Primärvalenzen ergeben sich folgende Beziehungen:

$$X = 0{,}607\ R + 0{,}173\ G + 0{,}201\ B,$$

$$Y = 0{,}299\ R + 0{,}587\ G + 0{,}114\ B \text{ und}$$

$$Z = 0{,}000\ R + 0{,}066\ G + 1{,}118\ B.$$

Tabelle 2.8. xy-Koordinaten verschiedener Farb-TV- und Kino-Normen

TV-Norm		FCC 1953	EBU/PAL/SECAM	SMPTE C	Rec 709 / sRGB	DCI
Rot	x	0,670	0,640	0.630	0.640	0,680
	y	0,330	0,330	0.340	0.330	0,320
Grün	x	0,210	0,290	0.310	0.300	0,265
	y	0,710	0,600	0.595	0,600	0,690
Blau	x	0,140	0,150	0.155	0,150	0,150
	y	0,080	0,06	0.070	0,060	0,060
Weiß	x	0,310	0,3127	0,3127	0,3127	0,3140
	y	0,316	0,3290	0,3290	0,3290	0,3510

Allgemein gilt: XYZ = NPM · RGB, wobei XYZ und RGB Vektoren sind und NPM die sogenannte Normalized Primary Matrix, ist die sich aus den einem Videosystem zugeordneten Primärvalenzen und dem entsprechenden Weißpunkt ergibt. Der Luminanzanteil kann danach über folgende Gleichung gewonnen werden:

$$Y = 0{,}299\ R + 0{,}587\ G + 0{,}114\ B.$$

Diese Gleichung ist für die Farbfernsehtechnik sehr wichtig, denn sie dient zur Ermittlung eines SW-kompatiblen Signals aus den drei Farbwertsignalen.

Die genannte Umrechnung bezieht sich auf die FCC-Festlegungen von 1953. Obwohl in Europa andere Leuchtstoffe verwendet werden, wurde die Gleichung für den Y-Wert beibehalten und damit ein kleiner Farbfehler in Kauf genommen. Tatsächlich gilt dies heute auch für die USA, denn auch hier ging man, unter Beibehaltung der Beziehung, zu lichtstärkeren Leuchtstoffen über und benutzt heute Definitionen nach SMPTE C (Tabelle 2.8), die sich kaum von denen der EBU unterscheiden.

In beiden SD-Systemen werden also kleine Farbfehler akzeptiert. Erst bei der Definition der HD-Systeme wurden hier Korrekturen vorgenommen. Nach ITU-R BT. 709 gilt heute für HDTV bzw. HD-Videosignale die Beziehung [69]:

$$Y = 0{,}213\ R + 0{,}715\ G + 0{,}072\ B.$$

Für Computerdisplays gibt es keine allgemeine Festlegung, doch halten sich viele Monitorhersteller an den sRGB-Farbraum, der dem EBU-Farbraum sehr ähnlich bzw. der ITU Rec. 709 gleich ist (Tabelle 2.8).

Wenn bei Bildbearbeitungen das Ergebnis in fremde Farbräume übertragen werden soll, wie es z. B. bei der Bearbeitung von Filmdaten mit Hilfe von Videosystemen geschieht, so sind Farbraumtransformationen erforderlich. Diese sind in SMPTE RP 177 beschrieben [206]. Danach kann der Übergang von RGB zu XYZ und dann zum neuen RGB-Raum mit Hilfe der zugehörigen NPM bzw. der entsprechenden Inversen genutzt werden, also:

$$(RGB)_1 = NPM_1^{-1} \cdot NPM_2\ (RGB)_2.$$

Auch für das digitale Kino ist diese Transformation relevant, denn in den zugehörigen Spezifikationen der Digital Cinema Initiative (DCI) sind die Farben in XYZ-Koordinaten festgelegt. Das erlaubt einen problemlose Erweiterung

für zukünftige Entwicklungen. Die in Tabelle 2.8 genannten RGB-Farborte definieren den Minimalfarbraum, den die Projektoren bei der Umsetzung von XYZ in RGB mindesten erreichen sollen. Dieser Farbraum ist bereits größer als der HD-Farbraum nach Rec 709 der ITU. Damit, und ggf. mit künftig zu definierenden Farbräumen lässt sich darüber hinaus mit einem P7-Farbkorrekturschema arbeiten. Dieses hat sieben Parameter, da neben Weiß und RGB auch die Komplementärfarben Cyan, Magenta und Gelb genutzt werden. Damit kann einen sehr feine farbliche Anpassung erfolgen. Es ist aber zu beachten, dass ein P7-Farbraum zwangsläufig innerhalb des nativen Farbraums des Projektors liegt, der weiterhin durch drei Primärfarben definiert ist [14].

2.4.3.2 Farbmischkurven in der Kamera

In der Kamera oder im Filmabtaster wird das Bild mit Filtern in die RGB-Farbauszüge zerlegt. Die Filter sind meist als Interferenzfilter ausgeführt und haben Transmissionsverläufe, wie sie in Abb. 2.74 dargestellt sind. Zusammen mit der Strahlungsfunktion des Aufnahmelichts und den spektralen Empfindlichkeiten des Wandlers selbst ergeben sich daraus die in Abb. 2.75 dargestellten Farbmischkurven, wie sie für einen CCD-Wandler typisch sind. Diese müssen wiederum gleich den drei Spektralwertkurven des Bildschirmprimärvalenzsystems multipliziert mit der Strahlungsverteilung der Tageslichtart D65 (6500 K) sein, da der Unbuntabgleich so vorgenommen wird, dass bei Vorliegen gleich großer Farbwertsignale die Tageslichtart D65 wiedergegeben wird.

Abbildung 2.75 zeigt auch die Spektralwertkurven für das Primärvalenzsystem bezüglich der genormten EBU-Valenzen. Diese weisen negative Anteile auf, die sich nicht als Spektralempfindlichkeiten von Strahlteilerfiltern realisieren lassen. Eine Möglichkeit zur Vermeidung negativer Anteile ist der Übergang zu einem virtuellen Primärvalenzsystem, so wie es am Beispiel des Normvalenzsystems dargestellt wurde. Die Umrechnung zwischen einem virtuellen Primärvalenzsystem und dem des Bildschirms erfordert nur die o. g. lineare Transformation, die in der Kamera auf elektronische Weise mit den Farbwertsignalen durchgeführt werden kann.

Das virtuelle Primärvalenzsystem wird nun aus den Spektralwertkurven des Aufnahme-Bildwandlers gebildet und eine elektronische Schaltung, die sogenannte Matrix in der Kamera sorgt für die Transformation in das Bildschirmprimärvalenzsystem. Mit der Matrix ist es also möglich, quasi Farbmischkurven mit negativen Anteilen zu erzielen.

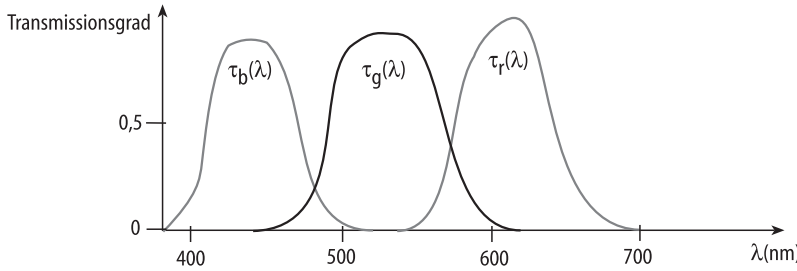


Abb. 2.74. Transmissionsverhalten von Farbfiltern in Kameras

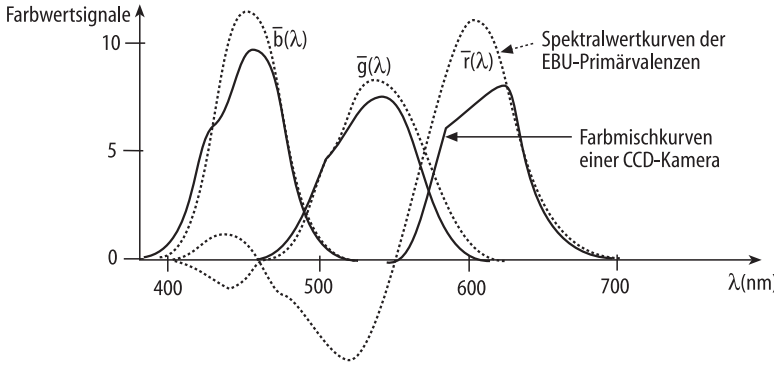


Abb. 2.75. Farbmischkurven der Kamera und Spektralwertkurven der EBU-Primärvalenzen

2.4.4 Das RGB-Signal

In sehr vielen professionellen Kameras werden mittels separater Wandler drei Farbwertsignale gebildet. Diese in elektrische Form gewandelten Signale müssen streng genommen als E'r, E'g und E'b bezeichnet werden. Das E kennzeichnet dabei das elektrische Signal und das Apostroph die Gamma-Vorentzerrung, die auch bei Farbsignalen direkt in der Kamera vorgenommen wird. Diese korrekte Bezeichnung wird jedoch selten verwandt, auch die elektrischen Signale werden meist abkürzend mit R, G, B, Y, etc. bezeichnet.

Im einfachsten Fall können zur Farbsignalübertragung drei Übertragungskanäle für jeweils eine der Grundfarben verwendet werden. Wie beim S/W-Signal wird in Europa ein Signalpegel von 100% wieder dem Spannungswert 0,7 V zugeordnet. Auch die Austast- und Synchronsignale des S/W-Verfahrens werden übernommen. Dieses RGB-Verfahren beansprucht mindestens drei Übertragungswege und damit die dreifache Bandbreite eines S/W-Systems (Abb. 2.76). Das Synchronsignal wird bei der RGB-Übertragung häufig separat geführt, wobei sich ein Bedarf von vier Signalleitungen ergibt. Alternativ wird das Synchronsignal allein auf der Leitung für Grün übertragen und dabei auch

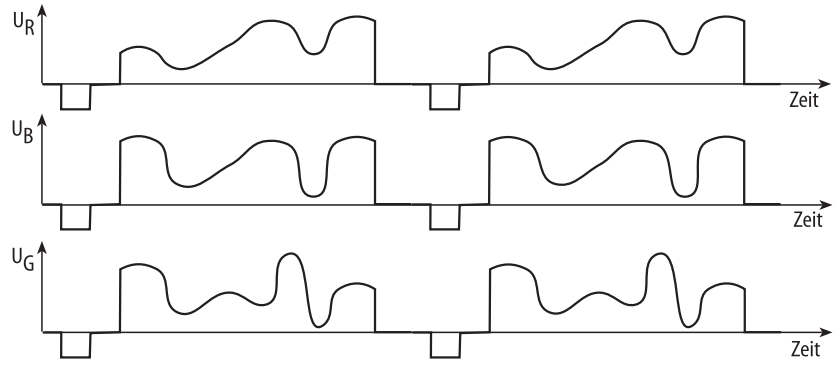


Abb. 2.76. RGB-Signal

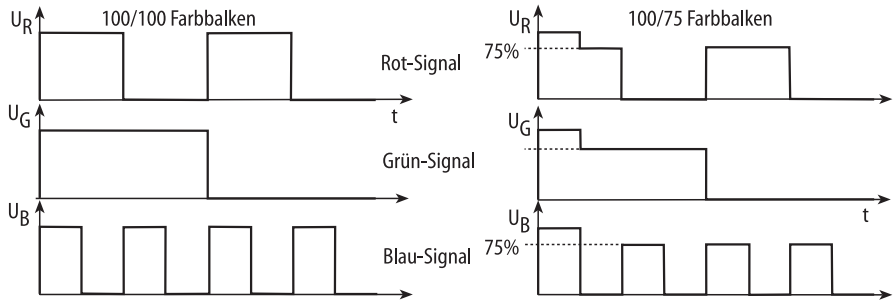


Abb. 2.77. 100/100- und 100/75-Farbbalken-Prüfsignal

als GBR oder RGsB bezeichnet oder auch auf allen drei Leitungen, wobei wie beim S/W-Signal ein negativ gerichteter Impuls von –0,3 V verwendet wird.

Die RGB-Übertragung bietet höchste Übertragungsqualität. Sie wird wegen des hohen Bandbreiten- und Leitungsbedarfs aber nur auf kurzen Strecken eingesetzt, z. B. zwischen Kameras und Einrichtungen der Bildtechnik oder zur Verbindung von Computern mit den zugehörigen Farbmonitoren. Bei der Übertragung über große Strecken soll das Farbvideosignal eine möglichst geringe Bandbreite beanspruchen, daher wird es auf Basis des RGB-Signals in vielfältiger Weise verändert.

Für alle elektrischen Farbsignalformen wird als Farbprüfsignal ein Farbbalken-Testbild verwendet, bei dem jede Zeile des Bildsignals die acht Farbkombinationen enthält, die sich ergeben, wenn die drei Signale für R, G, B jeweils zu 0% oder 100% eingeschaltet werden. Es entsteht so der 100/0/100/0-Farbbalken (Abb. 2.77), wobei die Balken nach sinkender Helligkeit angeordnet sind und damit bei S/W-Wiedergabe eine Grautreppe bilden. Neben Schwarz und Weiß und den Farben R, G, B, ergeben sich die Grundfarben der subtraktiven Farbmischung: Cyan, Magenta und Gelb. Oft wird auch das 100/0/75/0-Farbbalkensignal verwendet. Bei diesem Testsignal werden die Farben, außer bei der Mischung für Weiß, nur aus 75%-Pegeln ermischt, d. h. außer bei Weiß wird der Luminanz- und Chrominanzpegel auf 75% reduziert. Diese Maßnahme verhindert Übersteuerungen des Standard-Farbsignals und ermöglicht einen einfachen Vergleich von Chrominanz- und Weißwert bei PAL-codierten Signalen (s. Abschn. 2.4.6).

2.4.5 Das Komponentensignal

Bei der Einführung des Farbfernsehsystems wurden zwei wichtige Forderungen gestellt: Erstens sollte das Farbsignal schwarz/weiß-kompatibel sein, d. h. aus dem Farbsignal sollte das Leuchtdichtesignal für konventionelle S/W-Empfänger einfach ableitbar sein. Zweitens sollte es möglich sein, die Farbzusatzinformation im gleichen Kanal wie das monochrome Videosignal zu übertragen, ohne dass zusätzlich Bandbreite beansprucht wird. Damit die Forderungen erfüllt werden können, ist es erforderlich, das RGB-Signal zu verändern.

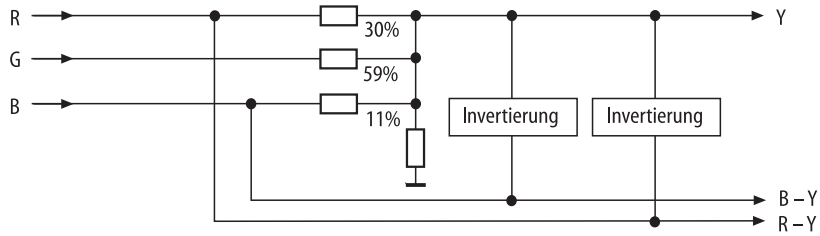


Abb. 2.78. Gewinnung des SD-Luminanzsignals und der Farbdifferenzsignale

Die erste Stufe ist die Gewinnung des S/W-kompatiblen Leuchtdichtesignals Y (Abb. 2.78). Die Beziehung für das Y-Signal (Luminanzsignal) folgt aus der Transformation zwischen den Koordinatensystemen RGB und XYZ bezüglich der Wiedergabeleuchtstoffe (s. Abschn. 2.4.3.1). Es gilt:

$Y = 0{,}299\ R + 0{,}587\ G + 0{,}114\ B$ für SD- und

$Y = 0{,}2126\ R + 0{,}7152\ G + 0{,}0722\ B$ für HD-Signale.

An dieser Stelle ist die verkürzte Bezeichnung RGB für die Farbwertsignale nicht ausreichend. Es muss erwähnt werden, dass hier, wie auch bei der unten angegebenen Differenzbildung, die γ-vorentzerrten Signale benutzt werden. Die korrekte Definition lautet für SD-Signale:

$$E_{Y}' = 0{,}299\ E_{R}' + 0{,}587\ E_{G}' + 0{,}114\ E_{B}'.$$

Die Faktoren für E_{R}', E_{G}', E_{B}' berücksichtigen die Helligkeitsbewertung für verschiedene Farben, sodass sich eine Umsetzung der Farb- in Grauwerte ergibt. Wenn nur das Luminanzsignal übertragen wird, erscheint ein Farbbalken als Grautreppe. Prinzipiell würde es ausreichen, neben dem Y-Signal zwei weitere Farbwertsignale, z. B. für Rot und Blau zu übertragen. Da der Y-Kanal die Information über die Helligkeit enthält, ist es aber günstig, dass die zwei weiteren Kanäle Informationen allein über die Farbart enthalten. Dann können Signalmanipulationen, wie z. B. eine Bandbreitenreduktion, vorgenommen werden, die allein die Farbart betreffen. Daher werden für Blau und Rot die Farbdifferenzkomponenten (R – Y) und (B – Y) gebildet. Die Farbdifferenzsignale enthalten keine Luminanzinformationen, sie verschwinden bei allen unbunten Farben da Farbkameras so eingestellt werden, dass die Signalanteile für RGB bei weißer Bildvorlage gleich sind. Auch bei einer grauen Fläche haben die RGB-Signalanteile jeweils gleiche Werte. Es gilt: R = G = B = Y, und (R – Y) = (B – Y) = 0 (Unbuntabgleich). Es werden die Differenzen von Y zu R und G gebildet, da der Grünauszug dem Y-Signal am ähnlichsten ist und sich bei G – Y nur geringe Pegel ergäben. Die Gewinnung des Leuchtdichtesignals Y und der Farbdifferenzsignale aus RGB und umgekehrt wird Matrizierung genannt, hier aber in einem anderen Sinne als bei der Kamera-Matrix. Im Prinzip ergibt sich Y einfach aus der Summation der RGB-Anteile unter Berücksichtigung der oben genannten Faktoren mit Hilfe entsprechender Spannungsteiler (Abb. 2.78). Aus der Abbildung wird auch deutlich, wie mit Hilfe des invertierten Y-Signals die Farbdifferenzsignale (R – Y) und (B – Y) gewonnen werden.

2.4.5.1 Pegelreduktion

Die Farbdifferenzsignale werden pegelreduziert damit sie keine höheren Spannungen aufweisen als das Luminanzsignal und um Übersteuerungen zu vermeiden. Dass bei den Differenzsignalen (R – Y) und (B – Y) höhere Spannungen als beim Y-Signal auftreten können, wird am Beispiel der Farben Gelb und Blau beim SD-Signal deutlich: Ein gesättigtes Gelb lässt sich aus je 100% Rot und Grün ermischen, der zugehörige Grauwert ergibt sich zu 89% und erscheint dem Auge fast so hell wie Weiß. Bei Gelb hat die SD-Farbdifferenzkomponente (R – Y) den Wert 0,11, und (B – Y) hat den Wert –0,89. Liegt im Bild dagegen allein ein gesättigtes Blau vor, so entstehen folgende Werte: Y = 0,11, weiterhin (R – Y) = – 0,11 und (B – Y) = 0,89.

Die Farbdifferenzwerte für die Komplementärfarben Gelb und Blau treten in gleicher Größe positiv und negativ auf. Die zugehörigen Spannungen ergeben sich aus der Multiplikation der Prozentwerte mit 0,7 V. Die Spannung für (B – Y) schwankt zwischen + 0,62 V (Gelb) und – 0,62 V (Blau), hat also mit 1,24 V_{ss} einen Wert, der 0,7 V_{ss} deutlich überschreitet. Eine Übersicht über die sich beim 100/100- und 100/75-Farbbalken ergebenden Spannungswerte für die Komponenten der SD-Signale zeigt Abbildung 2.79.

Ein einheitliches Komponentensignal soll bei allen drei Komponenten eine Maximalspannung von 0,7 V_{ss} aufweisen, daher werden die Farbdifferenzsignale zu C_R und C_B bzw. P_R und P_B reduziert. Die Bezeichnung ist hier uneinheitlich. Es wird empfohlen, C_R und C_B für digitale und P_R und P_B für analoge Komponenten zu verwenden. Die Reduktionsfaktoren sind in der EBU-Norm N 10 festgelegt, die sich auf einen 100/0/100/0-Farbbalken bezieht. Um z. B. die 1,24 V_{ss} der (B – Y)-Komponente im SD-Signal auf 0,7 V_{ss} zu reduzieren, ist der Faktor 0,564 erforderlich (Abb. 2.79). Für SD-Signale gilt:

$$P_R = 0{,}713\ (R - Y)$$

$$P_B = 0{,}564\ (B - Y)\ \text{(EBU-N 10, SD-Signal)},$$

die genauere Bezeichnung zeigt wieder, dass hier, wie auch in den folgenden Beziehungen, die γ-vorentzerrten Signale gemeint sind:

$$E_{PR}' = 0{,}713\ (E_R' - E_Y')$$

$$E_{PB}' = 0{,}564\ (E_B' - E_Y')\ \text{(SD-Signal)}$$

Bei einigen älteren Geräten (Betacam) werden Reduktionsfaktoren verwendet, die sich auf einen 75%-Farbbalken beziehen. Hier gilt: $P_R = 0{,}95\ (R - Y)$ und $P_B = 0{,}752\ (B - Y)$ (Betacam, veraltet).

HD-Signale sind bezüglich der Farbwerte und damit auch bezüglich Y anders skaliert als SD-Signale. Für HD-Signale gilt:

$$P_R = 0{,}6350\ (R - Y)$$

$$P_B = 0{,}5389\ (B - Y)\quad \text{(HD-Signal)}.$$

Für die Bildung des FBAS-Signals müssen die Pegel der Farbdifferenzsignale wiederum anders reduziert werden, da sie in modulierter Form dem Luminanzsignal überlagert werden (s. Kap. 2.4.6). Daher werden die Differenzsignale mit den Faktoren $c_u = 0{,}493$ und $c_v = 0{,}877$ reduziert. Die amplitudenredu-

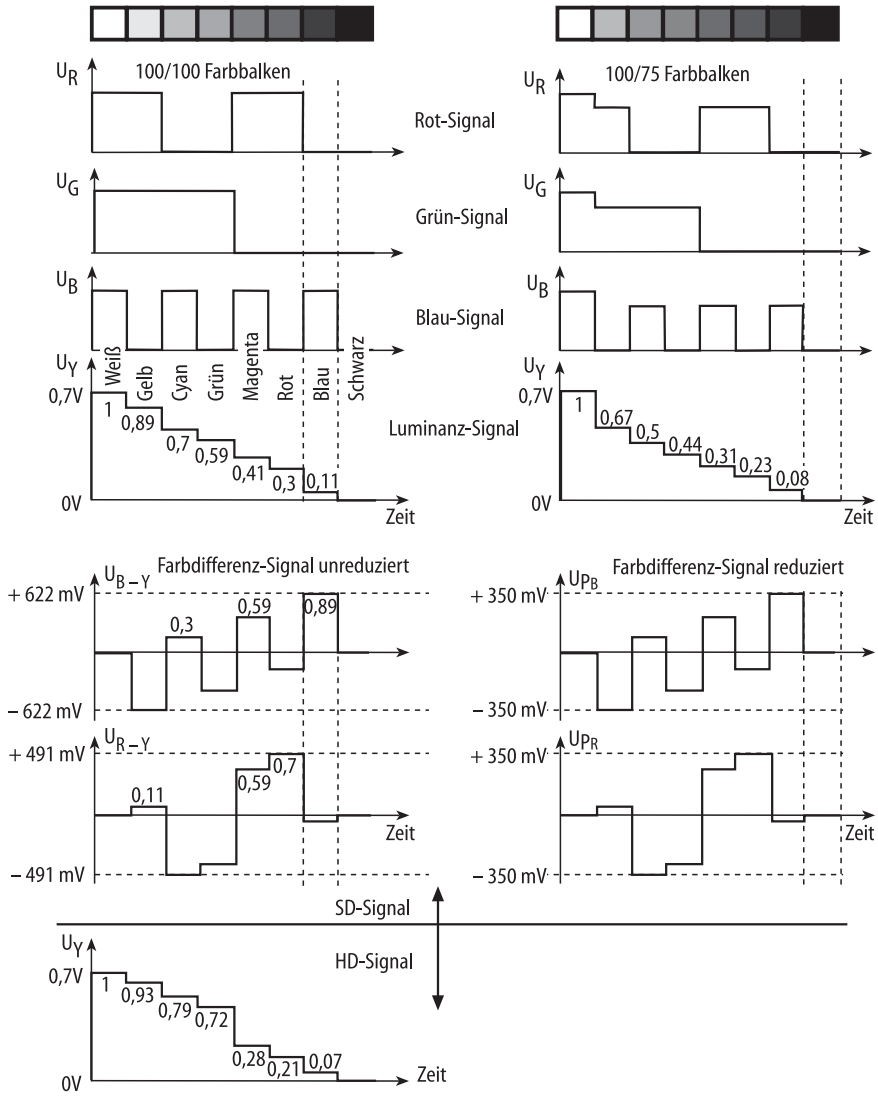


Abb. 2.79. Luminanz- und Farbdifferenzsignale

zierten Signale werden dann mit U und V bezeichnet. Es gilt die Beziehung:

$V = 0{,}877\ (R - Y)$

$U = 0{,}493\ (B - Y)$ (für die PAL-QAM-Modulation).

Die Signale U und V sind an keiner Schnittstelle zu finden. Die Buchstaben U und V werden aber häufig fälschlicherweise an Stelle von P_R und P_B als allgemeine Bezeichnung für Farbdifferenzkomponenten verwendet.

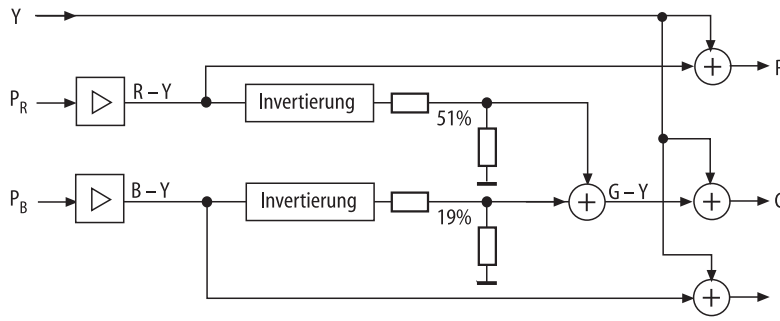


Abb. 2.80. Matrix zur Gewinnung der RGB-Signals aus dem SD-Komponentensignal

Abbildung 2.80 zeigt eine Matrixschaltung zur Rückgewinnung des RGB-Signals. Rechnerisch gilt für die Rückgewinnung der RGB-Komponenten aus den Farbdifferenzkomponenten U und V für SD-Signale:

$$R = V/c_v + Y; \quad B = U/c_u + Y; \quad G = Y - 0{,}299/0{,}587 \cdot V/c_v - 0{,}114/0{,}587 \cdot U/c_u,$$

Bei der Rückgewinnung von RGB aus Y, P_R, P_B statt aus YUV müssen anstelle von c_u und c_v die Werte 0,564 und 0,713 eingesetzt werden. Unabhängig von der jeweiligen Pegelreduktion gilt für SD-Signale die Beziehung:

$$(G - Y) = - 0{,}509\ (R - Y) - 0{,}194\ (B - Y).$$

Bie HD-Signalen müssen in die obigen Gleichungen statt c_u und c_v die Werte 0,5389 und 0,6350 eingesetzt werden und es gilt:

$$R = P_R/c_v + Y; \quad B = P_B/c_u + Y; \quad G = Y - 0{,}213/0{,}715 \cdot P_R/c_v - 0{,}07/0{,}715 \cdot P_B/c_u.$$

2.4.5.2 Bandbreitenreduktion

Y, P_R und P_B bilden zusammen das Komponentensignal (Component Analogue Video, CAV). Zur Übertragung dieses Signals sind weiterhin, wie bei der RGB-Übertragung, drei Leitungen bei voller Bandbreite erforderlich, der einzige Vorteil ist die S/W-Kompatibilität. Zur S/W-Übertragung wird nur das Y-Signal verwendet, die Farbdifferenzsignale werden einfach weggelassen.

Die Komponentenform bietet die Möglichkeit, die Übertragungsbandbreite für die Farbdifferenzsignale zu vermindern, mit dem Argument, dass das menschliche Auge, wegen des geringeren Ortsauflösungsvermögens für feine Farbdetails, die aufgrund der kleineren Bandbreite verringerte Farbauflösung nicht bemerkt (s. Abschn. 2.3.1). Wird die Bandbreite für die Farbdifferenzsignale z. B. jeweils halbiert, so ist im Vergleich zum S/W-System statt der dreifachen insgesamt nur noch die doppelte Übertragungsbandbreite erforderlich. In der Praxis wird die Bandbreite der SD-Farbdifferenzsignale bei einer Y-Bandbreite von 5 MHz auf 1,3 ... 2,5 MHz reduziert. Der Wert 1,3 MHz ist bei visueller Beurteilung des SD-Bildes ausreichend. Er wird benutzt, wenn das Komponentensignal als Basis für eine PAL-Codierung zur Übertragung zu den Zuschauern dient. Im Produktionsbereich wird mit einer Bandbreite von etwa 2 MHz für die SD-Farbdifferenzkomponenten gearbeitet, da die dort vorgenom-

menen Signalbearbeitungen, wie z. B. Chroma Key, eine höhere Farbauflösung erfordern. Komponentensignale sind trotz Bandbreitenreduktion Signale mit hoher Qualität und sind nicht mit FBAS-Artefakten (s. Abschn. 2.4.6) behaftet. Studios mit analoger Technik waren daher oft für Komponentensignale ausgelegt. Das Komponentensignal ist auch das Ausgangsformat zur Gewinnung eines hochwertigen Digitalsignals, sowohl im SD- als auch im HD-Bereich.

Der wesentliche Nachteil gegenüber der FBAS-Technik ist die dreifache Verkabelung. Die Komponentensignalübertragung erfordert je eine Leitung für Y, P_R und P_B. Dem Y-Signal ist das Synchronsignal beigemischt, auf der Y-Leitung beträgt die Spannung insgesamt 1 V_{ss}, auf den Leitungen für die Buntanteile 0,7 V_{ss}. Aufgrund der getrennten Signalführung können Laufzeitdifferenzen, z. B. wegen unterschiedlicher Leitungslängen, entstehen. Wie bei jeder getrennten Video-Signalführung muss dafür gesorgt werden, dass diese Laufzeitdifferenzen kleiner als 10 ns bleiben, da sonst im Bild Farbflächen gegenüber den zugehörigen Konturen verschoben erscheinen.

2.4.6 Das Farbsignal mit Farbhilfsträger

Die Farbdifferenzsignale und das BAS-Signal sollen im Bereich der Standard-Definition-Auflösung zu einem Farb-Bild-Austast-Synchronsignal derart zusammengefasst werden, dass das Gemisch nicht nur S/W-kompatibel ist, sondern auch die o. g. Forderung erfüllt, dass es nicht mehr Übertragungsbandbreite beansprucht als das S/W-Signal allein. Die Erfüllung der Forderung gelingt durch Nutzung der Lücken im Spektrum des BAS-Signals (s. Abb. 2.44). Die Farbdifferenzsignale werden dazu zunächst zum Chrominanzsignal C zusammengefasst, das aufgrund der gleichen Zeilenstruktur auch etwa die gleichen Lücken im Amplitudenspektrum aufweist wie das Luminanz-Signal Y (Abb. 2.45). Das Chrominanzsignal wird dann so auf einen Farbhilfsträger moduliert, dass die Spektren für Y und C gerade ineinander verkämmt sind (Abb. 2.81).

2.4.6.1 Das Chrominanzsignal

Im Chrominanzsignal sind die beiden Farbdifferenzkomponenten zusammengefasst. Es gibt diesbezüglich die drei wichtigen Varianten NTSC, SECAM und PAL. In Deutschland wird das PAL-Verfahren angewandt, die Darstellung sei hier zunächst an diese Form angelehnt.

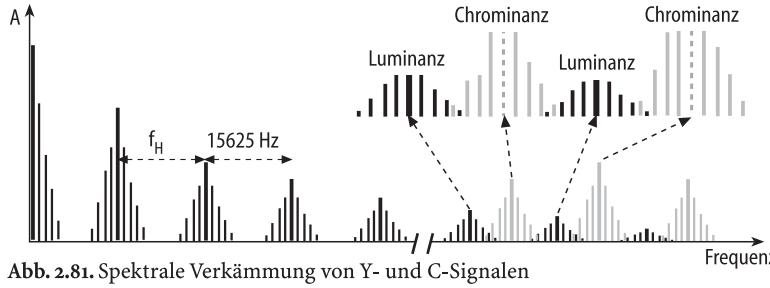


Abb. 2.81. Spektrale Verkämmung von Y- und C-Signalen

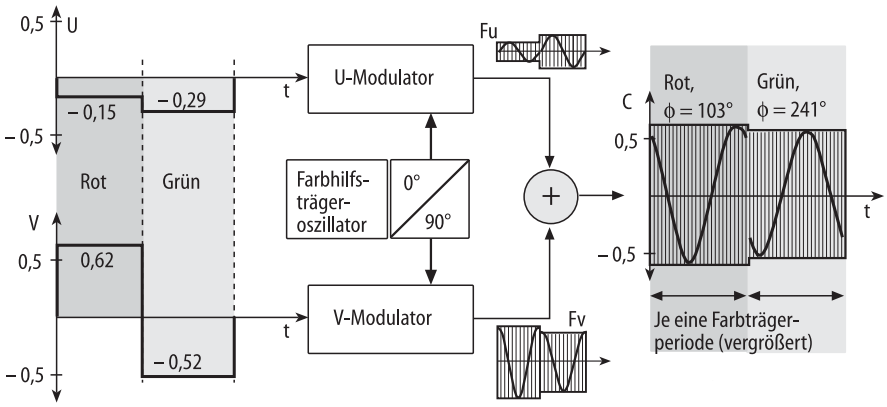


Abb. 2.82. QAM-Signalbildung bei 100% Rot und Grün

Die beiden Differenzsignale werden auf ca. 1,3 MHz Bandbreite begrenzt und so pegelreduziert, dass die Anteile U und V entstehen. Da das Chrominanzsignal und das zugehörige Luminanzsignal im Frequenzmultiplex übertragen werden sollen, muss das Chrominanzsignal auf einen Hilfsträger (engl.: Subcarrier, Abk.: SC) moduliert werden. Die Farbhilfsträgerfrequenz f_{sc} liegt für PAL bei ca. 4,4 MHz. Um zwei Signale auf nur einer Trägerfrequenz unterbringen zu können, wird als Modulationsart die Quadraturamplitudenmodulation (QAM) gewählt [156]. QAM bedeutet zweifache Amplitudenmodulation mit zwei Trägern gleicher Frequenz, die um 90° gegeneinander phasenverschoben sind (Abb. 2.82). Damit ergeben sich aus U und V die modulierten Signale F_u und F_v:

$$F_u = U \sin(\omega_{sc} t) \text{ und } F_v = V \cos(\omega_{sc} t),$$

die anschließend addiert werden. Durch diese Art der Modulation wird nicht nur die Amplitude der Trägerschwingung, sondern auch die Phasenlage verändert. Abbildung 2.82 zeigt am Beispiel des Übergangs zwischen den gesättigten Farben Rot und Grün, wie sich bei QAM Amplituden- und Phasensprünge ergeben. Das Gesamtsignal heißt Chrominanzsignal C, es wird, nicht ganz exakt, auch als Farbartsignal F bezeichnet.

Entsprechend der Verschiebung um 90° können die Differenzsignale U und V als Zeiger in einem rechtwinkligen Koordinatensystem dargestellt werden, U liegt auf der x- und V auf der y-Achse (Abb. 2.84). Es entsteht eine Fläche analog zum Farbkreis, in der jeder Punkt einer Farbart entspricht. Die Chrominanzfläche ist aber nicht gleich der Farbartfläche, da U und V mit verschiedenen Reduktionsfaktoren gebildet wurden und die Farbdifferenzkomponenten vom Luminanzsignal abhängen. Jeder Punkt der Fläche kann durch einen Vektor beschrieben werden. Für die Vektorlänge gilt:

$$C = \sqrt{U^2 + V^2}, \text{ bzw. } C' = \sqrt{(B - Y)^2 + (R - Y)^2}$$

für unreduzierte Komponenten. Die Länge des Chrominanzvektors hängt von der Farbsättigung ab, bei Schwarz, Weiß und Grautönen ist seine Länge gleich

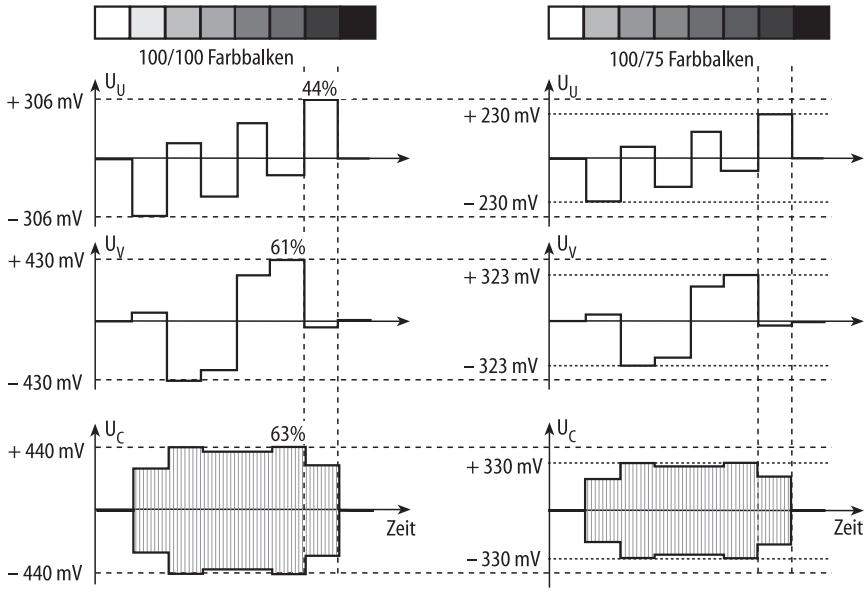


Abb. 2.83. Chrominanzsignale beim 100/100- und 100/75-Farbbalken

Null. Für den Winkel ϕ zwischen Vektor und U-Achse gilt:

$$\tan \phi = V/U.$$

Der Winkel ϕ repräsentiert den jeweiligen Farbton. Zu Rot gehört z. B. der Winkel $\phi = 103°$, zu Grün $\phi = 241°$, zu Blau $\phi = 347°$ etc.

Die Chrominanzsignale, die sich aus den Farbbalkenprüfsignalen ergeben, sind in Abbildung 2.83 dargestellt. Der schnelle Wechsel zwischen hochgesättigten Farben, deren Phasenwinkel sich stark unterscheiden (Komplementärfarben), geht mit einem Amplituden- oder Phasensprung einher. Dabei können kurzzeitig Einschwingvorgänge auftreten, die sich als unscharfe Trennung zwischen den beiden betroffenen farbigen Bereichen äußeren. So erscheint z. B. ein rotes Fußballtrikot auf grünem Rasen mit einer unscharfen Kontur.

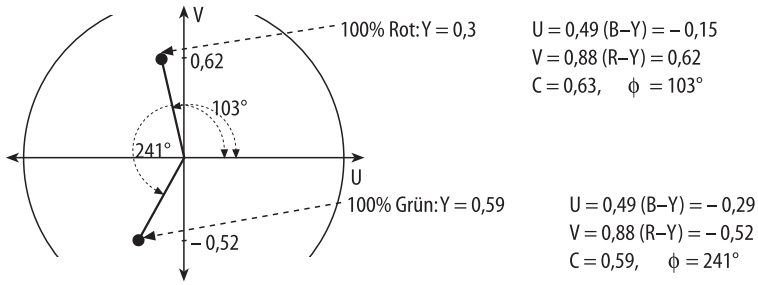
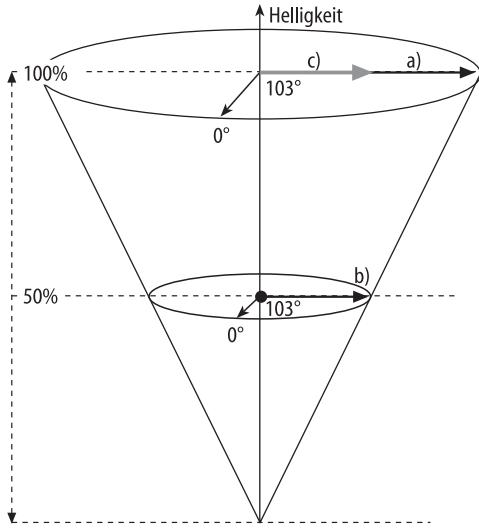


Abb. 2.84. Zeigerdarstellung der Chrominanzkomponenten bei Rot und Grün



a) Rot mit 100% Farbsättigung
und Y = 100% :
R = 1, G = B = 0, Y = 0,3R = 0, 3
U = 0,49 (B – Y) = – 0,15
V = 0,88 (R – Y) = 0,62
C = 0,63 ϕ = arctan (0,62/–0,15) = 103°

b) Rot mit 100% Farbsättigung
und Y = 50% (Dunkelrot):
R = 0,5, G = B = 0, Y = 0,3R = 0,15
U = 0,49 (B – Y) = – 0,07
V = 0,88 (R – Y) = 0,31
C = 0,315 ϕ = 103°

c) Rot mit 50% Farbsättigung
und gleicher Helligkeit wie in a) (Rosa):
R = 0,5 + 0,15, G = B = 0,15, Y = 0,3
U = 0,49 (B – Y) = – 0,07
V = 0,88 (R – Y) = 0,31
C = 0,315 ϕ = 103°

Abb. 2.85. Chrominanzvektor in Abhängigkeit von Helligkeit und Farbsättigung

Die Vektorlänge hängt nicht allein von der Farbsättigung, sondern auch vom Luminanzanteil ab. Wenn der Luminanzanteil konstant ist, so entspricht die Länge allein der Sättigung, wenn die Sättigung konstant ist, repräsentiert die Vektorlänge die Helligkeit. Abbildung 2.85 zeigt eine dreidimensionale Darstellung, die farbmetrisch nicht korrekt ist, die den Zusammenhang an Beispielen für den Farbton Rot aber gut veranschaulicht.

2.4.6.2 Das Y/C-Signal

Das gesamte Farbvideosignal besteht aus den beiden Anteilen Luminanz Y und Chrominanz C. Wenn diese beiden Anteile separat auf eigenen Leitungen übertragen werden, spricht man vom Y/C-Signal. Diese Art der getrennten Signalführung findet bei hochwertigen analogen Heimsystemen und semiprofessionellen Geräten (S-VHS, Hi8, Computerschnittsysteme) Verwendung und wird auch als S-Video bezeichnet. Der Luminanzkanal beinhaltet das BAS-Signal mit einer Spannung von 1 V_{ss}, also auch die Synchronimpulse. Das Chrominanzsignal ist um den Burst ergänzt (s. u.), der mit einer Spannung von 0,3 V_{ss} auf der hinteren Schwarzschulter liegt. Die Signalverläufe für Y und C sind bezüglich eines Farbbalkens in Abbildung 2.86 dargestellt.

Das Y/C-Signal ist qualitativ schlechter als ein Komponentensignal, da die Bandbreite des geträgerten Chromasignals stärker eingeschränkt ist. Gegenüber dem FBAS-Signal ergibt sich dagegen ein Qualitätsvorteil, da bei getrennter Signalführung Filter zur Y/C-Trennung umgangen werden können (s. Abschn. 2.4.8). Um den Einsatz zweier separater Stecker zu vermeiden, werden für die Y/C-Signale besondere Steckverbindungen verwendet. Meistens der 4-polige Hosidenstecker (Abb. 2.86). Teilweise gibt es die Möglichkeit, das S-Videosignal auch über SCART-Stecker zu führen (Abb. 2.52).

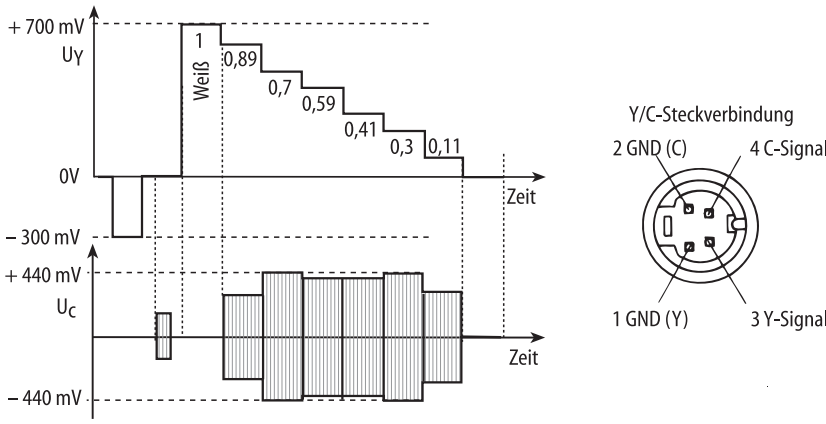


Abb. 2.86. Luminanz- und Chrominanzsignal beim 100/100-Farbbalkensignal

2.4.6.3 Das FBAS-Signal

Durch die Zusammenführung von Luminanz- und Chrominanzsignal entsteht das zusammengesetzte (Composite) Farb-, Bild-, Austast-, Synchronsignal (FBAS). Die frequenzabhängige Darstellung (Abb. 2.81) zeigt, dass bei der Zusammensetzung die Spektren von Luminanz- und Chrominanz-Signalen miteinander verkämmt werden. In der zeitabhängigen Darstellung des Gesamtsignals wird deutlich, dass die Y- und C-Amplitudenwerte addiert werden. Das Chrominanzsignal hat positive und negative Anteile, bei der Addition der Amplitudenwerte zum Y-Signal ergibt sich für das Farbbalkensignal der in Abbildung 2.87 dargestellte Verlauf.

Hier zeigt sich, dass bezüglich des 100%-Farbbalkens eine Übermodulation bis zum Wert 133% bzw. –33% zugelassen wird. Bei geringerer Amplitude für das Chrominanzsignal würde sich ein zu schlechter Signal-Rauschabstand ergeben, denn ohne Zulassung von Übermodulation dürfte die Chromaamplitude z. B. für Gelb nur 11% der Weißwertamplitude betragen. Die Farben Gelb

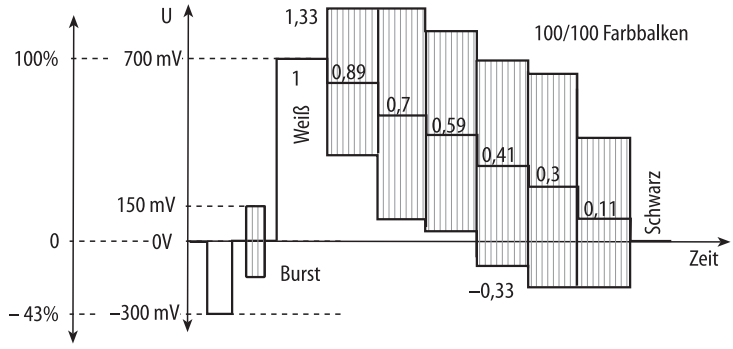


Abb. 2.87. FBAS-Signal beim 100/100-Farbbalkensignal

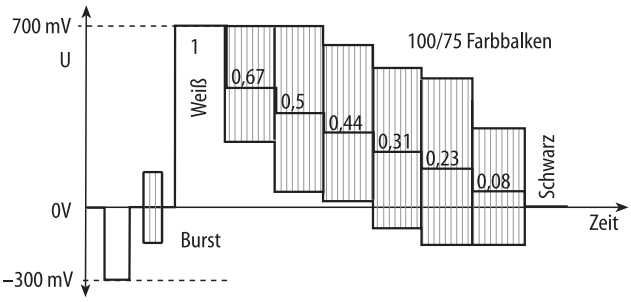


Abb. 2.88. FBAS-Signal beim 100/75-Farbbalken

und Cyan sind besonders kritisch, da die zugehörigen Chrominanzamplituden zu den höchsten Luminanzpegeln addiert werden. Für 100% Gelb gilt:

$$Y = 0{,}3\ R + 0{,}59\ G = 0{,}89$$

$$R - Y = 1 - 0{,}89 = 0{,}11 \text{ und } B - Y = -0.89.$$

Für 100% Cyan gilt:

$$Y = 0{,}59\ G + 0{,}11\ B = 0{,}7$$

$$R - Y = -0{,}7 \text{ und } B - Y = 1 - 0{,}7 = 0{,}3.$$

Bei Zulassung einer Gesamtamplitude von 133% darf die Chrominanzamplitude bei Gelb den Wert 1,33 – 0,89 = 0,44 und bei Cyan den Wert 1,33 – 0,7 = 0,63 aufweisen [156]. Aus den Bestimmungsgleichungen für diese beiden Werte lassen sich die erforderlichen, oben bereits genannten, Pegelreduktionsfaktoren c_u und c_v für die Farbdifferenzsignale ermitteln:

$$\text{Gelb: } (0{,}44)^2 = (c_v\ (R - Y))^2 + (c_u\ (B - Y))^2 = (c_v\ 0{,}11)^2 + (c_u\ (-\ 0{,}89))^2$$

$$\text{Cyan: } (0{,}63)^2 = (c_v\ (R - Y))^2 + (c_u\ (B - Y))^2 = (c_v\ (-\ 0{,}7))^2 + (c_u\ 0{,}3)^2.$$

Die Auflösung ergibt die bereits bekannten Beziehungen

$$c_u = 0{,}493 \text{ und } c_v = 0{,}877.$$

Die sich daraus ergebende Übermodulation kann auch deshalb akzeptiert werden, weil in realen Bildvorlagen im wesentlichen Körperfarben auftreten, deren Farbsättigung so gering ist, dass unter diesen Umständen Pegel von 100% nur selten überschritten werden.

In diesem Zusammenhang wird auch klar, warum die Signalüberprüfung erleichtert wird, wenn anstelle des 100/100- der 100/75-Farbbalken verwendet wird. Bei der Erzeugung des 100/75 Farbbalkens weisen außer Weiß alle RGB-Farbsignale nur 75% Pegel auf. Dadurch wird für die Farben Gelb und Cyan, bei denen die Maximalwerte im FBAS-Signal auftreten, die 33%-Übermodulation gerade kompensiert. Die Chrominanzamplitude kann damit leicht durch Vergleich mit dem Luminanzpegel bei Weiß beurteilt werden (Abb. 2.88).

Das FBAS-Signal ist das Standardvideosignal. Es wird auch als Composite-Signal oder einfach Videosignal bezeichnet. Es enthält alle Signalanteile (Colour, Video, Blanking, Sync, CVBS) und erfordert nur eine einzelne Leitung.

2.4.7 Farbfernsehnormen

Farbfernsehnormen definieren verschiedene Formen der Chrominanzsignalbildung sowie verschiedene Arten der Verbindung mit dem Luminanzsignal. Die Spektren der Luminanz- und Chrominanzsignale weisen aufgrund der Zeilenstruktur gleichartige Lücken auf, sodass die Signale ineinander verzahnt werden können. Für diese Verzahnung reicht es im Prinzip aus, dass für die Farbhilfsträgerfrequenz ein halbzahliges Vielfaches der Zeilenfrequenz gewählt wird (Halbzeilenoffset). Auf der Wiedergabeseite müssen die einzelnen Signalanteile zurückgewonnen werden. Bei der Rückgewinnung von Y und C sind die Spektren nicht wieder vollständig voneinander separierbar, es bleibt auch bei besten Filtertechniken ein Übersprechen zwischen den Kanälen erhalten.

Damit möglichst geringe Übersprechstörungen zwischen den Kanälen auftreten, müssen einige Forderungen erfüllt werden: Erstens sollte der Farbträger für das Chrominanzsignal hochfrequent sein, damit sich durch das Übersprechen im Y-Kanal eine möglichst feine Struktur (Perlschnurmuster) ergibt. Da das Auge für hohe Ortsfrequenzen unempfindlicher ist als für mittlere, wirkt das Muster subjektiv weit weniger störend als eine grobe Struktur, die bei tiefer Trägerfrequenz aufträte. Zweitens ist bei der Wahl der Hilfsträgerfrequenz zu beachten, dass bei der Quadraturamplitudenmodulation der größte Teil beider Seitenbänder übertragen werden müssen. Aus beiden Forderungen ergibt sich für ein Bildübertragungssystem mit 5 MHz Bandbreite als Kompromiss eine Farbhilfsträgerfrequenz von ca. 4,4 MHz. D. h. sie liegt um ca. die halbe Chrominanzbandbreite unterhalb der oberen Grenzfrequenz. Abbildung 2.89 zeigt die Lage der Luminanz- und Chrominanzspektren, zusammen mit den zugehörigen Transmissionskurven einfacher Filter.

Das Übersprechen in den Y-Kanal wird als Cross Luminanz bezeichnet, Farbsignale werden hier als Helligkeitssignale interpretiert. Wegen der hohen Frequenz des Farbsignals erscheinen diese Störungen als feines Muster, welches besonders bei computergenerierten Videosignalen hervortritt. Das Übersprechen in den Chroma-Kanal wird als Cross Colour bezeichnet, feine Strukturen im Helligkeitssignal werden im Chroma-Kanal als Farbveränderungen interpretiert. Dieser Effekt zeigt sich deutlich bei der Wiedergabe feiner unbunter Streifenmuster (Nadelstreifenanzüge). Derartige Muster sind von einem Farbschleier umgeben, dessen Farbe sich bei Bewegung des Musters ständig ändert [149].

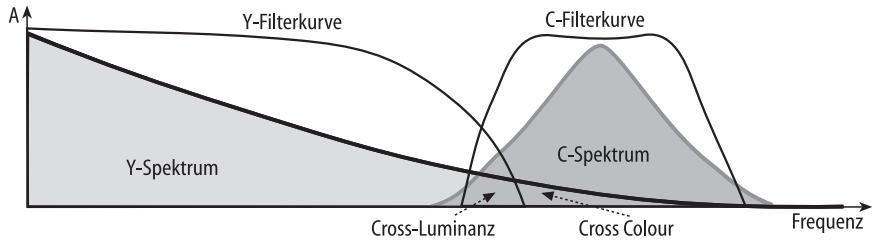


Abb. 2.89. Übersprechen zwischen Y und C beim Einsatz einfacher Filter

2.4.7.1 NTSC

Das vom National Television System Committee (NTSC) in den USA eingeführte Verfahren wurde bereits 1953 standardisiert und war damit das erste verbreitete Farbfernsehsystem. Die bis heute in den USA gültige FCC-M-Norm beruht auf dem Zeilensprungverfahren mit 59,94 Halbbildern pro Sekunde und 525 Zeilen pro Vollbild. Die Videobandbreite beträgt nur 4,2 MHz und der Bild–Tonträgerabstand 4,5 MHz. Das Chrominanzsignal wird aus den oben genannten Gründen gebildet, indem die Farbhilfsträgerfrequenz über einen sogenannten Halbzeilenoffset als halbzahliges Vielfaches der Horizontalfrequenz f_H generiert wird. Für die Farbhilfsträgerfrequenz f_{sc} gilt:

$$f_{sc} = (2n + 1)\, f_H/2.$$

Bei der Grenzfrequenz von 4,2 MHz beträgt der optimale Farbhilfsträgerwert ca. 3,6 MHz. Für den Faktor n wurde die Zahl 227 gewählt, damit folgt:

$$f_{sc} = (2 \cdot 227 + 1) \cdot 525 \cdot 29{,}97\ \text{Hz}/2 = 3579{,}545\ \text{kHz}.$$

Die Gleichung zeigt, dass die Bildfrequenz im Zuge der Definitionen der Farbhilfsträgerfrequenz von ehemals 30 Hz um den Faktor 1000/1001 auf 29,97 Hz gesenkt wurde. Dies geschah mit Rücksicht auf den Abstand zwischen Bild- und Tonträger von 4,5 MHz, der aus Kompatibilitätsgründen zum S/W-Fernsehsystem beibehalten werden sollte. Um dabei Interferenzen zwischen den Ton- und Farbhilfsträgerfrequenzen zu minimieren, wurde festgelegt, dass die Tonträgerfrequenz genau das 286-fache der Zeilenfrequenz betragen sollte, wofür allerdings der Übergang von 30 Hz auf 29,97 Hz Bildwechselfrequenz erforderlich wurde. Die Zeilenfrequenz für NTSC beträgt somit 4,5 MHz /286 = 15 734,265 Hz = 525 · 29,97 Hz = 15 750 Hz · 1000/1001. Diese geringfügige Senkung hat für NTSC-Länder weit reichende Folgen. Da die Konversion von Bildraten mit sehr geringen Unterschieden sehr aufwendig ist, wurden Möglichkeiten zur Anpassung für andere Systeme definiert. So gibt es z. B. für die Filmtechnik neben den üblichen 24 fps die Bildrate 23,976 fps = 24 · 1000/1001 fps. Die Abtastraten für die Digitalisierung der Videosignale wurden ebenso angepasst wie die Abtastraten für digitale Audiodaten (z. B. 47,952 KHz statt 48 KHz) und viele weitere Grundfrequenzen im Videobereich.

Bei NTSC wird das Chrominanzsignal durch Quadraturmodulation aus den Farbdifferenzkomponenten gebildet. Allerdings werden als Farbdifferenzsignale die Komponenten I und Q (Inphase und Quadratur) anstatt U und V verwendet. Die Koordinatenachsen I und Q stehen senkrecht aufeinander, das Koor-

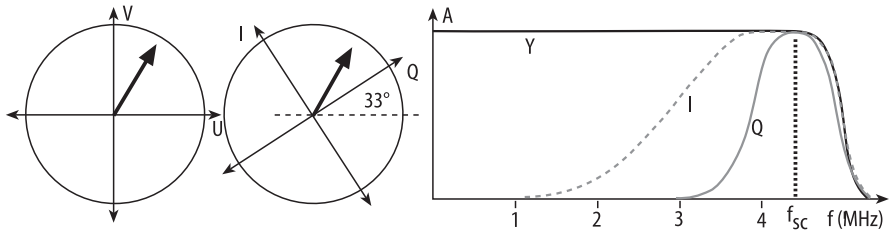


Abb. 2.90. Vergleich der Koordinatensysteme U, V und I, Q und Bandbegrenzung für I, Q

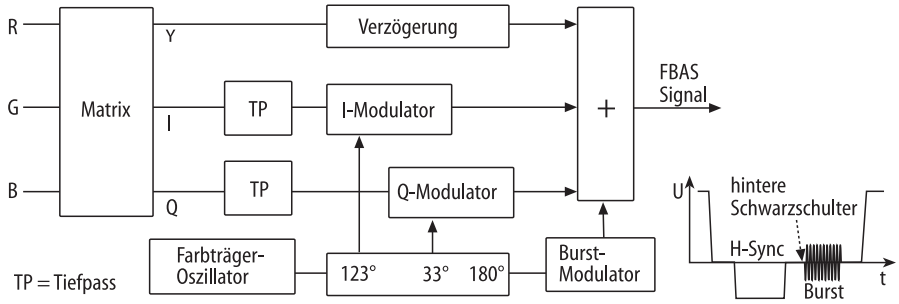


Abb. 2.91. NTSC-Encoder und Eintastung des Burst

dinatensystem ist um 33° gegenüber dem System aus U und V gedreht (Abb. 2.90). Es gilt:

$$I = V\cos 33° - U\sin 33°\,,\quad Q = V\sin 33° + U\cos 33°.$$

Die I, Q-Achsen liegen in Quer- und Längsrichtung der McAdam-Ellipsen, also in Richtungen, die der maximalen (rot – blau/grün, I) bzw. minimalen (gelb – blau, Q) Kontrastempfindlichkeit des menschlichen Auges für diese Farbübergänge entsprechen [149]. Vor der Quadraturmodulation werden dementsprechend für I und Q verschiedene Bandbreitenbegrenzungen vorgenommen (Abb. 2.90). Das Q-Signal wird auf 0,5 MHz begrenzt und kann mit voller Bandbreite für beide Seitenbänder übertragen werden. Das I-Signalspektrum wird auf 1,5 MHz eingegrenzt und das obere Seitenband teilweise beschnitten. Das bedeutet, dass bis 0,5 MHz alle Farben uneingeschränkt übertragbar sind und bis 1,5 MHz quasi eine zweifarbige Übertragung stattfindet.

Die Abbildungen 2.91 und 2.92 zeigen schematisch den Aufbau von NTSC-Coder und Decoder. Bei der Decodierung muss die Amplitude und die Phasenlage bestimmt werden. Das erfordert eine Synchrondemodulation, d. h. der ursprüngliche Träger muss auf der Empfangsseite dem ankommenden Signal phasen- und frequenzrichtig zugesetzt werden. Der Träger steht nicht zur Verfügung, denn die QAM arbeitet mit unterdrücktem Träger. Zur Lösung des

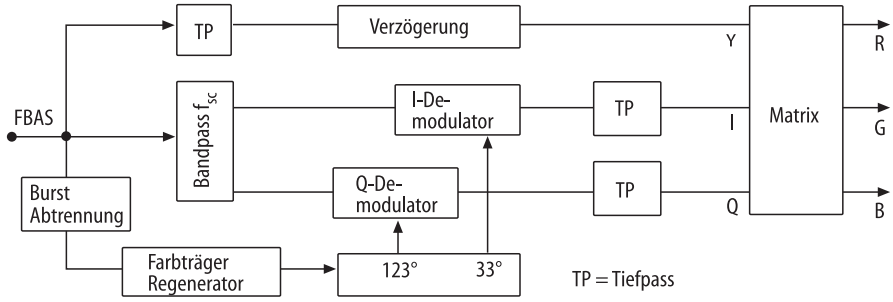


Abb. 2.92. NTSC-Decoder

Problems wird im Empfänger ein Oszillator eingebaut, der auf der festgelegten Farbträgerfrequenz schwingt. Zur phasenrichtigen Verkopplung der Oszillatoren in Sender und Empfänger wird in jeder Zeile im Bereich der hinteren Schwarzschulter der Burst eingetastet (Abb. 2.91). Der Burst besteht aus ca. zehn Schwingungszügen der Farbträgerfrequenz mit definierter Phasenlage. Mit dem Burst ist es möglich, den Empfängeroszillator bezüglich Phase und Frequenz mit dem Sender synchron zu halten.

Bei der QAM-Farbsignalübertragung ist die Farbsättigung in der Amplitude und der Farbton in der Phasenlage des Chromasignals verschlüsselt. Der Farbton wird aus der Relation zwischen Chrominanzsignal- und Oszillatorphase im Empfänger ermittelt. Eine bei der Übertragung unvermeidliche Beeinflussung dieser Phasenrelation führt daher zu einer Änderung des Farbtons. Bei NTSC kann diese Relation zum Ausgleich durch einen Phasenlagensteller (willkürlich) verändert werden. Kurzfristig veränderliche Phasenlagen (differentielle Phasenfehler) sind aber nicht kompensierbar, daher ergibt sich ein Bild mit entsprechend schwankendem Farbton. Dieses Problem hat dem NTSC-Verfahren den Spitznamen „Never The Same Color“ eingebracht und dazu angeregt, verbesserte Farbcodiermethoden zu entwickeln. NTSC wird bis heute in Amerika und Japan verwendet.

2.4.7.2 SECAM

1957 wurde in Frankreich ein Verfahren vorgestellt, das die oben genannten Farbstabilitätsprobleme umgeht, indem auf die kritische QAM verzichtet wird. Statt dessen werden die Farbdifferenzsignale zeilenweise alternierend übertragen (Abb. 2.95). Damit ist die Farbauflösung in vertikaler Richtung zwar halbiert, weil in jeder Zeile nur die Information für eine Differenzkomponente vorliegt, doch stellt dieser Umstand kaum ein Problem dar, da auch in horizontaler Richtung die Farbauflösung begrenzt ist.

Die Abbildungen 2.93 und 2.94 zeigen den Aufbau von SECAM-Coder und SECAM-Decoder. Im Empfänger wird die jeweils aktuelle Zeile gespeichert, damit die entsprechende Farbkomponente auch zur Zeit der Übertragung der nächsten Zeile, bzw. der anderen Komponente, zur Verfügung steht. Die sich ergebende, nicht zeitrichtige Zusammenfassung der Farbdifferenzsignale erzeugt kaum sichtbare Fehler, da sich die Inhalte zweier benachbarter Zeilen i. a. wenig unterscheiden. Die Speicherung kann mit Hilfe einer Ultraschall-

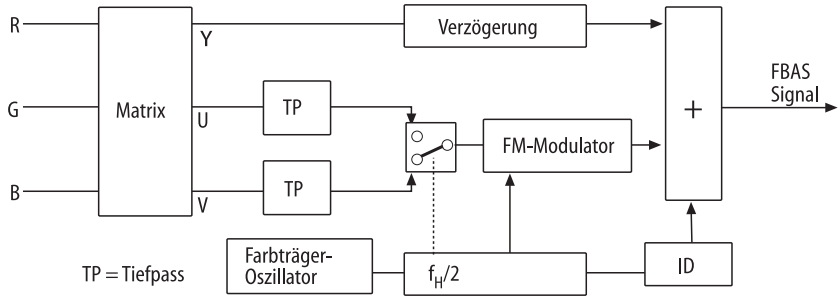


Abb. 2.93. SECAM-Encoder

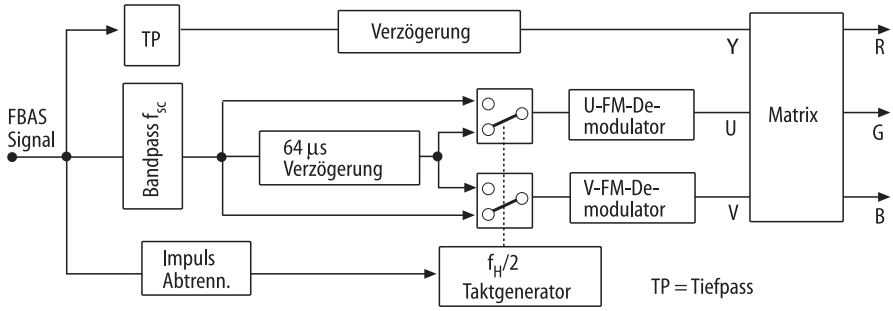


Abb. 2.94. SECAM-Decoder

Laufzeitleitung vorgenommen werden, in der das Signal um die Zeilendauer (64 µs) verzögert wird. Der Name des Verfahrens ist eine Abkürzung für das Funktionsprinzip: **Sé**quentiel **c**ouleur **à m**émoire.

Auch bei SECAM werden zunächst speziell pegelveränderte und auf 1,3 MHz bandbegrenzte Farbdifferenzsignale gebildet, sie werden hier als D_R und D_B bezeichnet. Es gilt: $D_R = -1{,}9\ (R - Y)$ und $D_B = 1{,}5\ (B - Y)$. Die Farbdifferenzsignale werden mit Frequenzmodulation (FM, s. Kap. 4) auf einen Farbhilfsträger moduliert. Dabei darf die maximale Frequenzabweichung zwischen 3,9 MHz und 4,75 MHz betragen. Es werden zwei verschiedene Farbträgerfrequenzen verwendet, die mit der Zeilenfrequenz verkoppelt sind; für die Komponente D_B gilt: $f_{scB} = 272 \cdot f_H = 4{,}25$ MHz und für D_R: $f_{scR} = 282 \cdot f_H = 4{,}4$ MHz.

Da das Farbträgermuster auf dem Bildschirm zu deutlich in Erscheinung trat wurde SECAM mit den Varianten I–IIIb oft verändert, wobei versucht wurde, das Übersprechen in den Luminanzkanal durch Pegelreduktionen und Preemphasis-Schaltungen zu reduzieren. Das Übersprechen ist hier stärker als im Fall der QAM, denn aufgrund der variierenden Frequenz kann die spektrale Verkämmung nicht gewährleistet werden. Der Vorteil des Verfahrens gegenüber NTSC ist die Farbstabilität. Als Nachteil ist zu nennen, dass es für die Studiotechnik nicht geeignet ist, da aufgrund der FM keine additive Mischung möglich ist. SECAM wird für die analoge Fernsehausstrahlung in Frankreich, Nordafrika und osteuropäischen Ländern verwendet.

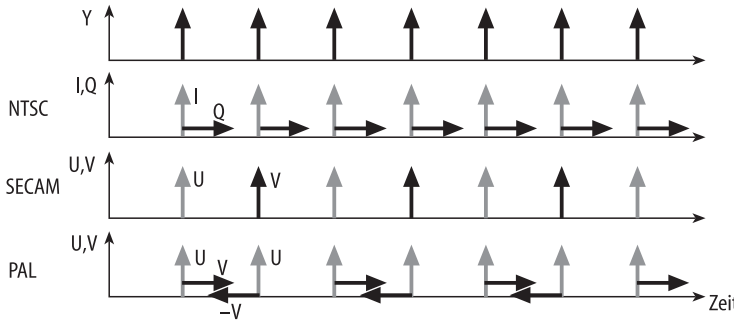


Abb. 2.95. Vergleich von NTSC, SECAM und PAL

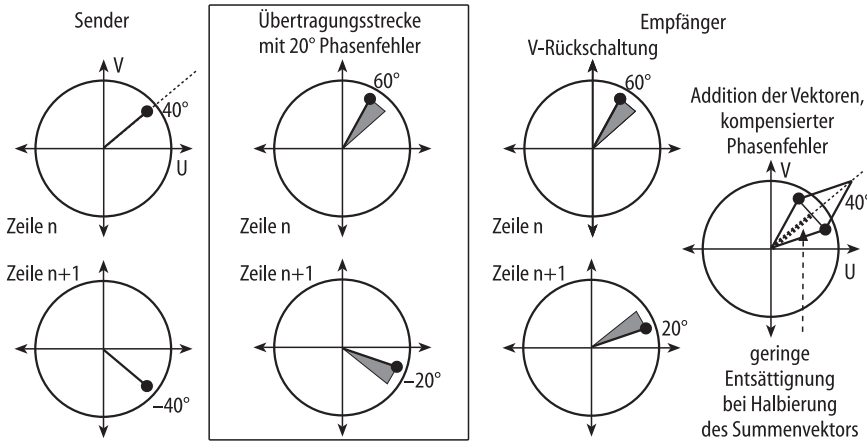


Abb. 2.96. Prinzip der PAL-Phasenfehlerkompensation

2.4.7.3 PAL

Das 1963 von W. Bruch konzipierte PAL-Verfahren (Phase Alternation Line) beruht auf dem NTSC-Verfahren, umgeht aber die Empfindlichkeit gegenüber Phasenschwankungen. Im Gegensatz zu NTSC werden bei PAL direkt die Farbdifferenzkomponenten U und V, und nicht I und Q verwandt. Beide Komponenten werden auf ca. 1,3 MHz bandbegrenzt. Auch bei PAL wird die QAM benutzt, allerdings mit dem Unterschied, dass die V-Komponente zeilenweise alternierend invertiert wird (Abb. 2.95 und Abb. 2.96). Damit wechseln sich immer Zeilen mit nichtinvertierter (NTSC-Zeilen) und invertierter V-Komponente (PAL-Zeilen) ab, während die U-Komponente unbeeinflusst bleibt.

Abbildung 2.96 zeigt, warum der Aufwand mit der invers geschalteten V-Komponente betrieben wird: Ein auftretender Phasenfehler wirkt sich bei geschalteter und ungeschalteter Zeile in gleicher Richtung aus. Wird die geschaltete V-Komponente im Empfänger wieder invertiert, so erscheint der Fehler negativ, sodass er sich bei der Summation mit dem in einer Ultraschall-Leitung gespeicherten Signal der vorhergehenden Zeile aufhebt. Auch PAL nutzt

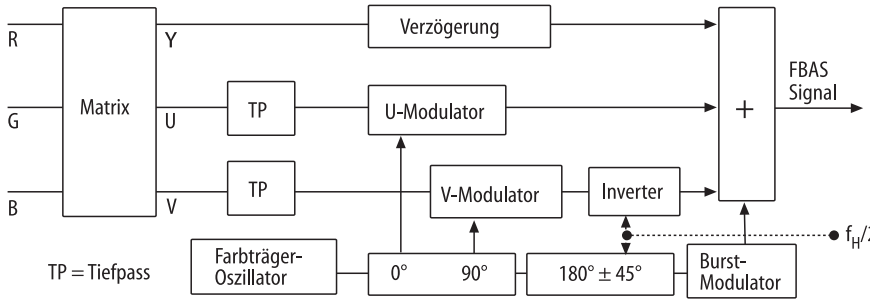


Abb. 2.97. PAL-Encoder

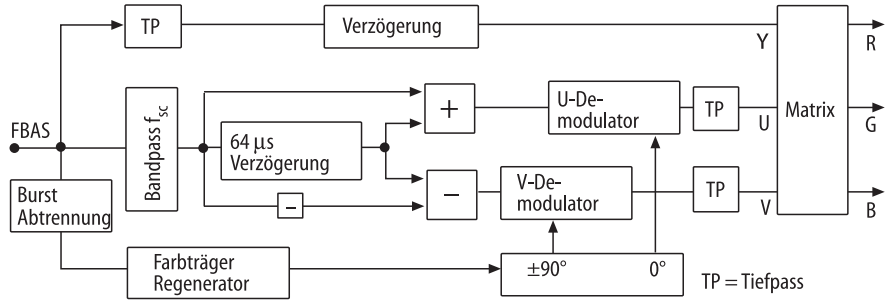


Abb. 2.98. PAL-Decoder

dabei den Umstand, dass sich bei gewöhnlich auftretenden Bildern die Inhalte und die Phasenfehler der Nachbarzeilen nur wenig voneinander unterscheiden. In Abbildung 2.96 wird auch deutlich, dass aufgrund der Vektoraddition der Summenzeiger kürzer ist als die Summe der Einzelzeiger. Das bedeutet, dass die Farbsättigung verringert wird. Im PAL-System wird also ein Farbtonfehler in einen Sättigungsfehler umgewandelt, der aber subjektiv wesentlich weniger stört. Weiterhin ist wie bei SECAM auch bei PAL die vertikale Farbauflösung gegenüber NTSC verringert. Die Abbildungen 2.97 und 2.98 zeigen den Aufbau von PAL-Coder und -Decoder.

Die Umschaltung der V-Komponente wird auch bei der Kontrolle eines PAL-Signals mit dem Vektorskop (s. Abschn. 2.5.1) deutlich. Beim PAL-Vektorskop sind zwei verschiedene Einstellungen möglich: es kann die V-Komponente zurückschalten oder nicht. Bei zurückgeschalteter Komponente sollten bei Vorliegen eines Farbbalkens die sechs entstehenden Leuchtpunkte innerhalb der Toleranzfelder auf dem Vektorskopschirm liegen. Bei nichtgeschalteter Komponente gilt dies für zwölf Komponenten, denn alle Werte sind an der Horizontalachse gespiegelt (Abb. 2.99). Der Burst liegt im geschalteten Fall in der Diagonalen bei 135° (s. u.), auch er wird im ungeschalteten Fall verdoppelt.

Die dauernde V-Umschaltung wirkt wie eine zusätzliche Modulation mit halber Zeilenfrequenz, d. h. die Chromaspektren für U und V unterscheiden sich um $f_H/2$. Wenn der Farbhilfsträger, wie bei NTSC, einfach mit dem Halb-

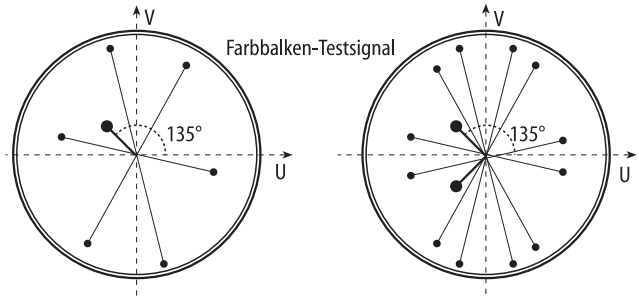


Abb. 2.99. PAL-Signal mit nicht-invertierter und invertierter V-Komponente

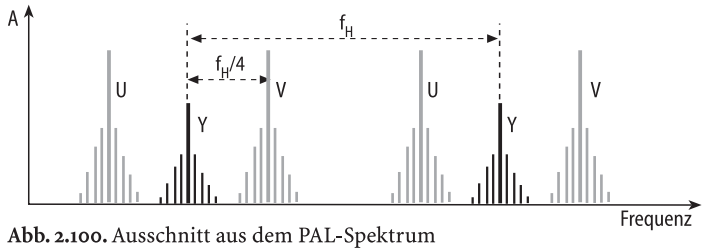


Abb. 2.100. Ausschnitt aus dem PAL-Spektrum

zeilenoffset gebildet würde, ergäbe sich eine Überlagerung der Komponenten V und Y. Die PAL-Hilfsträgerfrequenz wird daher mit einem Viertelzeilen-Offset berechnet: $f_{sc} = (n - 1/4) \cdot f_H$, wodurch die Anteile U und V untereinander einen Abstand von $f_H/2$ und zu den benachbarten Y-Linien einen Abstand von $f_H/4$ aufweisen (Abb. 2.100). Durch den Viertelzeilen-Offset entsteht eine unangenehme Störmusterwiederholung über vier Halbbilder. Dieses Muster erscheint bewegt und damit weniger störend, wenn die Farbträgerfrequenz mit einem zusätzlichen Versatz von 25 Hz gebildet wird. Mit n = 284 folgt daraus

$$f_{sc} = 283{,}75 \cdot f_H + 25\ \text{Hz} = 4{,}43361875\ \text{MHz}$$

für die Berechnung der PAL-Hilfsträgerfrequenz f_{sc} bezüglich der Norm B/G.

Ein Coder für FBAS-Signale muss die Farbträgerfrequenz mit den Synchronimpulsen verkoppeln. Dazu werden alle Impulse aus einem hochstabilen Farbträgeroszillator abgeleitet. Neben den Austast- (A) und den H- (S) und V-Synchronimpulsen wird zusätzlich ein Coder-Kennimpuls (K) generiert, der bestimmt, wie der Burst in die Austastlücke eingetastet wird. Weiterhin wird ein PAL-Kennimpuls (P) erzeugt, der die Zeilen mit invertierter V-Komponente kennzeichnet. Abbildung 2.101 zeigt den zeitlichen Zusammenhang zwischen den S-, K-, P-und A-Impulsen.

Der PAL-Burst wird in jede Zeile eingetastet und dient wie bei NTSC dazu, einen mit der Farbhilfsträgerfrequenz schwingenden Oszillator im Empfänger phasen- und frequenzsynchron zum Sendesignal zu halten, denn auch bei PAL ist aufgrund der QAM eine Synchrondemodulation erforderlich. Außerdem dient der Burst als Referenz für die Regelung der Chrominanzamplitude. Der Burst hat eine Dauer von zehn Schwingungszügen, bzw. 2,25 µs, einen Spannungswert von 0,3 V_{ss} und liegt auf der hinteren Schwarzschulter (Abb. 2.101).

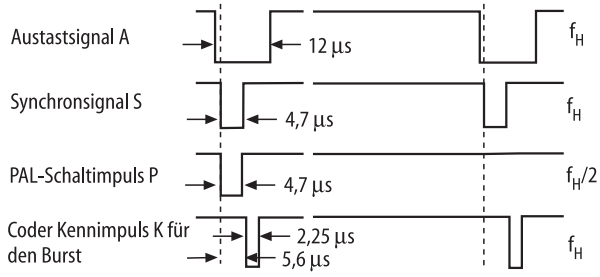


Abb. 2.101. Zusammenhang der Impulse A, S, P, K

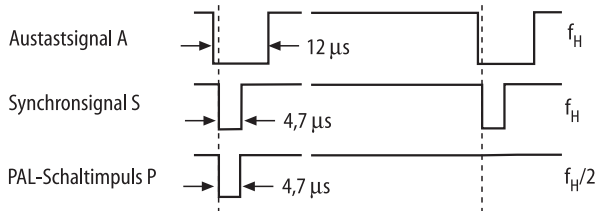


Abb. 2.102. PAL-Burst und Burst-Phasenlage

Mit dem Burst wird zusätzlich die Information über den Schaltzustand der V-Komponente übertragen, damit die Invertierung im Empfänger für die richtige Zeile rückgängig gemacht wird (Abb. 2.102). Der Schaltzustand wird durch eine Veränderung der Phasenlage des Burst auf 180° + 45° oder 180° – 45° repräsentiert, die mittlere Phase ergibt sich, wie bei NTSC, zu 180°. Der Burst wird in einigen Zeilen der vertikalen Austastlücke nicht übertragen. Bei der Burst-Austastung (Burst-Blanking) wird die PAL-4 V-Sequenz (s. u.) berücksichtigt.

Bei der Demodulation des PAL-Signals muss vor allem die Invertierung der V-Komponente rückgängig gemacht werden (Abb. 2.103). Wie bei SECAM wird auch bei PAL eine 64 µs-Verzögerungsleitung verwandt, sodass die V-Komponente immer in beiden Schaltzuständen vorliegt. Wird gerade eine NTSC-Zeile empfangen, so liegt am Ausgang der Verzögerungsleitung die vorhergehende PAL-Zeile. In den geradzahligen Zeilen liegen die geträgerten Komponenten F_U und $+F_V$, in den ungeradzahligen Zeilen die Komponenten F_U und $-F_V$ vor. Durch Summenbildung der aktuellen und der verzögerten Zeile entstehen die Anteile $2\,F_U$ und $\pm\,2\,F_V$ (Abb. 2.103). Die F_V-Komponente wird schließlich in jeder zweiten Zeile zurückgeschaltet und die Signale werden den Synchrondemodulatoren zugeführt, wo durch phasenrichtige Zusetzung der Farbhilfsträgerfrequenz die Differenzsignale U und V entstehen.

2.4.7.4 PAL-Sequenzen

Das PAL-Verfahren bietet, verglichen mit SECAM und NTSC, die beste Bildqualität. Es wird in Deutschland und vielen weiteren Ländern angewendet. Es unterliegt jedoch verschiedenen Periodizitäten, die bei anderen Codierungsarten nicht auftreten und die in einigen Fällen besonderer Beachtung bedürfen:

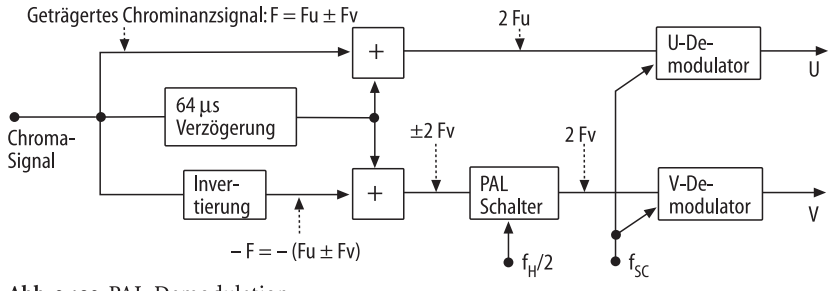


Abb. 2.103. PAL-Demodulation

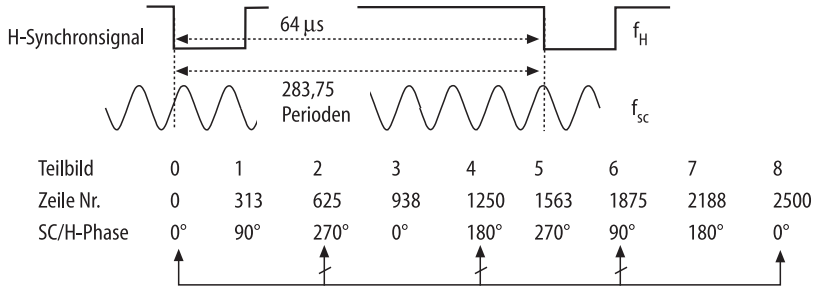


Abb. 2.104. PAL-8V-Sequenz

Ein Vollbild besteht aus zwei Halbbildern, damit ergibt sich zunächst eine Periodizität über zwei Halbbilder, die sogenannte 2V-Sequenz. Dieselbe Phasenlage zwischen Bild-Synchronpuls und PAL-V-Schaltsignal erfordert eine gerade Zeilenzahl, die frühestens nach Ablauf von 1250 Zeilen erreicht ist. Damit entsteht die PAL-4V-Sequenz mit einer Periodizität von zwei Voll- bzw. vier Halbbildern.

Aus dem Viertelzeilen-Offset beim PAL-Farbhilfsträger folgt weiterhin eine als PAL-8V-Sequenz bezeichnete Periodizität der F/H-Phase (SC-H-Phase), also der Phasenbeziehung zwischen Farbträger- und Horizontalsignal. Hier wird auch die V-Phase mit einbezogen, der Begriff F/H-Phase bezieht sich auf eine definierte Zeile des Vollbildes. Betrachtet man die 1. Zeile eines Bildes, bei dem diese Phasendifferenz gleich Null ist, so ergibt sich nach Ablauf des ersten Vollbildes eine Phasenverschiebung um den Faktor 283,75 · 625. Erst nach vier Vollbildern, acht Halbbildern oder 2500 Zeilen wird dieser Faktor ganzzahlig und damit die Phasenverschiebung wieder Null (Abb. 2.104).

Die 4V- und 8V-Sequenzen sind PAL-spezifisch und mussten vor allem bei der Umordnung von gespeicherten Bildsequenzen (elektronischer Schnitt) beachtet werden. Bei Nichtbeachtung der Periodizitäten konnte es vorkommen, dass auf das erste Bild einer Sequenz das sechste Bild der nächsten Sequenz folgt, sodass am Schnittpunkt die 8V-Sequenz gestört war. Es konnte in diesem Fall an der Schnittstelle zu Farbverfälschungen oder einem horizontalen Bildruck kommen (s. Kap. 9.5). Die Beachtung der PAL-Sequenzen bei der Nachbearbeitung wird mit dem Begriff Colour-Framing bezeichnet.

Bei der Produktion stellen die PAL-Sequenzen kein Problem dar, denn die beteiligten Signalquellen (Kameras, MAZ-Geräte) werden fest an einen zentralen Studiotakt gekoppelt. Durch Justierung der H- und SC-Phase an den Geräten kann eine sehr genaue Anpassung an den Takt vorgenommen werden.

2.4.7.5 Das Frequenzspektrum des PAL-Signals

Die frequenzabhängige Darstellung des PAL-Signals zeigt, dass die Luminanz- und Chrominanzsignale ein Linienspektrum aufweisen und dass die Spektren ineinander verzahnt sind. Die Y-Amplituden nehmen zu hohen Frequenzen hin ab, sind aber durchaus auch im Bereich über 5 MHz noch von Bedeutung. Die Chrominanzsignalamplituden spielen vor allem in der Nähe der Farbträgerfrequenz eine Rolle. In Abbildung 2.105 rechts sind nur die Chromakom-

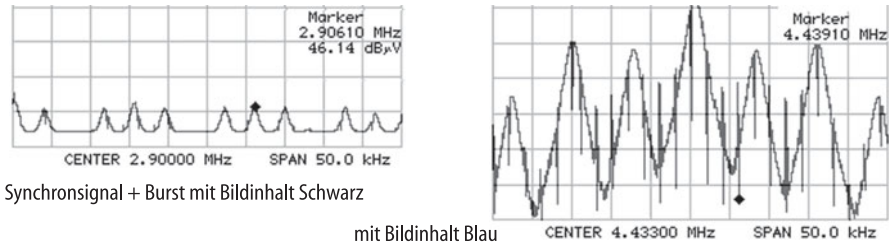


Abb. 2.105. Gemessene PAL-Amplitudenspektren

ponenten zu sehen, die einen Abstand von $f_H/2$ aufweisen. Im Bereich um 3 MHz weisen die Amplituden für Y und C etwa gleiche Werte auf. Anders als bei NTSC erscheinen aber bei PAL aufgrund des Viertelzeilenoffsets und der V-Umschaltung die Komponenten U und V separat neben den Y-Linien (Abb. 2.105 links). Neben der zeitlichen weist das Videosignal auch eine örtliche Periodizität auf. Das Videosignal kann damit auch in drei Dimensionen bezüglich der Ortsfrequenzen sowie der (Zeit-) Frequenz dargestellt werden.

2.4.7.6 Verbreitungsgebiete der Farbfernsehnormen

Abbildung 2.106 und Tabelle 2.9 zeigen die Verbreitungsgebiete der verschiedenen Farbfernsehnormen. Oftmals werden die Begriffe PAL und NTSC auch im Umfeld der Digitalsignale benutzt. Da dort jedoch fast ausschließlich Komponentensignale zum Einsatz kommen, und keine Farbsignale mit Farbhilfsträger, ist die Bezeichnung falsch. Sie wird aus Bequemlichkeit benutzt, um die Bildaufbaustandards 525/60 und 625/50 zu unterscheiden. In vielen Fällen korrespondieren die Farbfernsehstandards mit dem Bidaufbauparametern, doch gilt dies nicht immer. So arbeitet ein großes Land wie Brasilien mit 525/60 Hz, benutzt jedoch PAL als Farbcodierung für das analoge Fernsehen.

Abb. 2.106. Nutzungsgebiete der verschiedenen Farbfernsehnormen

Tabelle 2.9. Farbfernsehsysteme verschiedener Länder

NTSC		SECAM		PAL	
Bolivien	Chile	Frankreich		Ägypten	Australien
Equador	Haiti	Marokko	Niger	Brasilien	BRD
Hawaii	Japan	Mauretanien	Monaco	China	Großbritannien
Kanada	Kolumbien	Mongolei	Senegal	Indien	Indonesien
Korea	Kuba	Russland	Ukraine	Italien	Kenia
Mexiko	Phillippinen	Tunesien		Neuseeland	Niederlande
Peru	USA			Südafrika	Thailand

2.4.8 Farbsignalhierarchie

Als Überblick sollen hier noch einmal die verschiedenen Farbsignalcodierungen, abgestuft nach ihrer Qualität, dargestellt werden (Abb. 2.107).

Die höchste Farbsignalqualität bietet die RGB-Übertragung. In allen drei Kanälen ergeben sich Spannungswerte von 0,7 V_{ss}. Die zweitbeste Farbsignalqualität ergibt das Komponentensignal Y P_R P_B. Das Luminanz- und das Synchronsignal werden gemeinsam mit einem Gesamtspitzenspannungswert von 1 V_{ss} übertragen, die Farbdifferenzkomponenten liegen zwischen ± 0,35 V. Die Bandbreite der Farbdifferenzsignale kann auf Werte um ca. 2 MHz eingeschränkt sein, dabei ist visuell kein Unterschied zum RGB-Signal feststellbar.

Bei der Zusammenfassung der Komponenten zum PAL-Signal werden die Farbdifferenzsignale gemeinsam auf einen Träger moduliert, durch Bandbreitenbegrenzung auf ca. 1,3 MHz wird die Qualität dabei weiter verschlechtert. Als Zwischenschritt ermöglicht die Y/C-Übertragung eine getrennte Führung von Luminanz- und Chrominanzanteilen, sodass sich keine Übersprechprob-

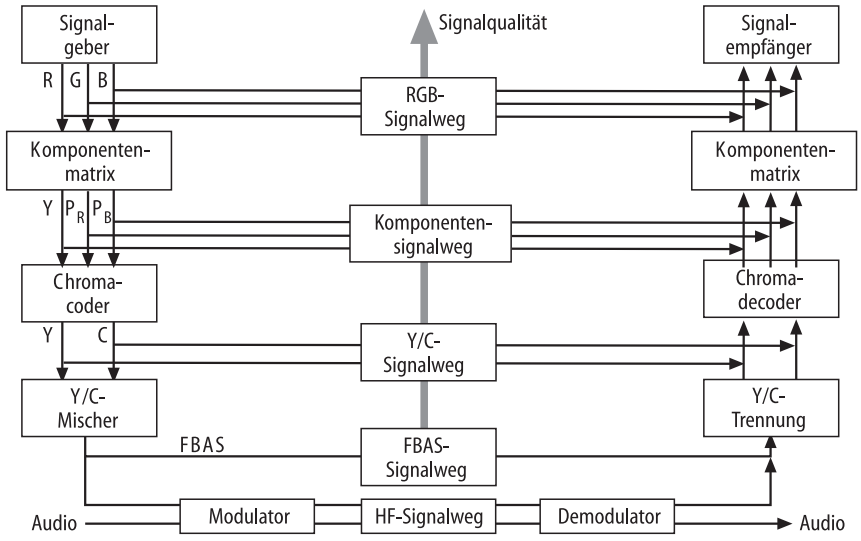


Abb. 2.107. Farbsignalhierarchie

leme ergeben. Luminanz- und Synchronsignal werden beim Y/C-Signal (S-Video) gemeinsam auf einer Leitung mit einem Gesamtspannungsspitzenwert von 1 V_{ss} übertragen, die zweite Leitung führt das Chrominanzsignal und den Burst. Beim FBAS-Signal sind schließlich alle Videosignalanteile zusammengefasst, die maximale Spannung beträgt bei Verwendung des 100/75 Farbbalkens 1 V_{ss}. Die Qualität ist gegenüber dem Y/C-Signal durch Cross Luminanz und Cross Colour verringert, da Y und C mittels Filtern getrennt werden müssen.

Auf der untersten Stufe der Hierarchie in Abbildung 2.107 steht schließlich das auf einen Träger modulierte Signal (Antennensignal, HF-Signal), bei dem noch das Begleittonsignal hinzukommt. Hier muss vor der Decodierung noch HF-demoduliert werden, womit ein weiterer Qualitätsverlust einhergeht.

2.4.8.1 Übergang zwischen den Signalformen

Der Codierungsweg von RGB bis zu FBAS ist technisch prinzipiell unproblematisch. Allerdings wirkt sich ein Fehler, der z. B. im RGB-Signal auftritt, auch auf die Signale aus, die auf dieser Basis beruhen und umgekehrt. Ein Fehler ist z. B. die Über- oder Unterschreitung der zulässigen Spannungen von 0 V und 700 mV. Diese Grenzen werden auch als Gamut Limit bezeichnet, wobei der Begriff Gamut für die Palette aller reproduzierbaren Farben in einem Farbraum steht. Die Verletzung des Gamut Limits macht das Signal illegal. Ein Signal, das bei der Transformation in ein anderes legal bleibt, ist auch ein gültiges Signal. Ein gültiges Signal ist immer legal, aber ein legales Signal ist nicht immer gültig. Das zeigt Abbildung 2.108, wo im oberen Teil ein legales Komponentensignal mit allen drei Anteilen dargestellt ist. Die Y-Komponente ist nun oben rechts verzerrt, was nach der Transformation zum RGB-Signal dazu führt, dass dieses illegal wird. Das verzerrte Komponenten-Signal ist damit ungültig. Gültige Signale verletzen in keiner Signalform die Limits [3].

Bei der Rückgewinnung des RGB-Signals aus dem FBAS-Signal ergibt sich als große Schwierigkeit die Trennung von Chrominanz- und Luminanzanteil, die nur bei unbewegten Bildvorlagen theoretisch völlig rückwirkungsfrei möglich ist. In dem in Abbildung 2.98 dargestellten Blockschaltbild des PAL-Farbdecoders ist die Y/C-Trennung durch Chromabandpass- und Tiefpassfilter realisiert. Das Luminanzsignal wird durch das Tiefpassfilter auf 3,5 MHz begrenzt,

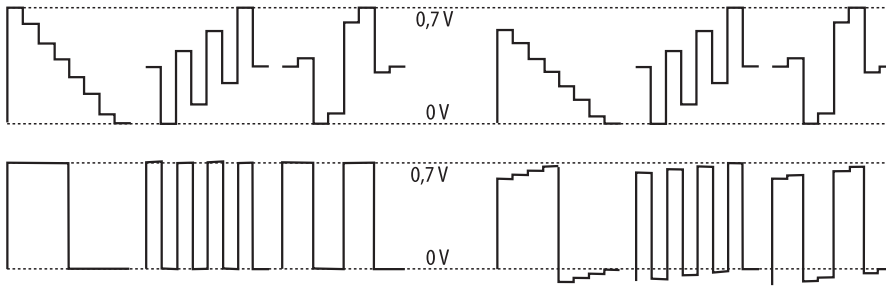


Abb. 2.108. Legales und illegales Signal

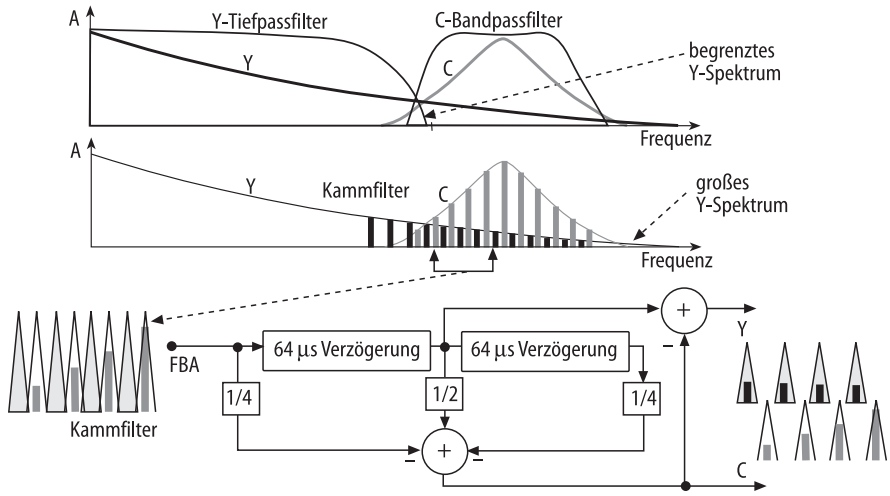


Abb. 2.109. Bandpass- und Kammfilter zur Trennung von Y und C

das Bandpassfilter hat die geringste Dämpfung im Bereich der Farbträgerfrequenz und arbeitet mit einer Bandbreite von ca. 1,5 MHz (Abb. 2.109). Hier entsteht ein erheblicher Auflösungsverlust für das Y-Signal und ein erhebliches Maß an Y/C-Übersprechen. Diese Methode zur Y/C-Trennung wird daher nur in einfachen Geräten angewendet. Bei Geräten der Studiotechnik wird eine effektivere Cross-Colour- und Cross-Luminanzunterdrückung verlangt. Daher werden höherwertige Filter eingesetzt, die die Periodizitäten im Luminanz- und Chrominanzspektrum ausnutzen. Die Nullstellen des Filterverlaufs für das Luminanzsignal liegen dann genau auf den Spektrallinien des Chrominanzsignals und umgekehrt, man spricht von einem Kammfilterverlauf.

Ein Kammfilter für ein NTSC-Signal mit 1/2-Zeilen-Offset kann mit einer 2-Zeilen-Verzögerungsleitung nach Abbildung 2.109 aufgebaut werden. Der Mittelabgriff zwischen den Teilen der Verzögerungsleitung bietet die Möglichkeit, die Phasenbeziehung zwischen drei aufeinander folgenden Zeilen auszunutzen. Bezogen auf das Signal am Mittelabgriff liegen vor und hinter den beiden Verzögerungsgliedern die vorhergehende und die nachfolgende Videozeile vor, in denen das Chrominanzsignal aufgrund des 1/2-Zeilen-Offsets eine Phasenverschiebung von 180° aufweist, sodass sie bei einer Subtraktion vom Signal am Mittelabgriff dieses periodisch verstärken, bzw. nach Ablauf einer halben Periode mit ihm zu Null interferieren (Abb. 2.109). Das Luminanzsignal entsteht einfach durch Subtraktion des Chrominanzsignalanteils vom Gesamtsignal.

Bei der Anwendung von Kammfiltern für PAL-Signale muss bedacht werden, dass diese spektrale Anteile für U und V aufweisen, die wegen des 1/4-Zeilen-Offsets zwischen den Spektrallinien des Luminanzsignals platziert sind. Daher ist eine Verdopplung der Nullstellenanzahl des Filters erforderlich, d. h. dass die Verzögerung der Laufzeitglieder verdoppelt werden muss, sodass ein Vierzeilen-Kammfilter zum Einsatz kommt. Für eine hochqualitative Y/C-Trennung werden horizontale, vertikale und zeitliche Kammfiltertechniken kombiniert.

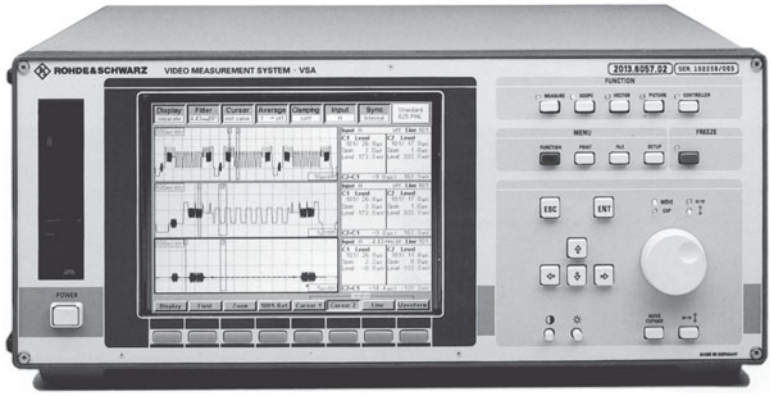


Abb. 2.110. Video-Messsystem [163]

2.5 Signalkontrolle

Der Signalkontrolle kommt in Videosystemen besondere Bedeutung zu. Einerseits müssen die technischen Parameter überprüft werden können, andererseits ist die Signalkontrolle auch zur korrekten Aussteuerung des Bildsignals während des Programmablaufs erforderlich. Zu den technischen Parametern gehören die Signalpegel, die Signallaufzeiten, die Grenzfrequenzen des Systems, das Übersprechen zwischen Teilsignalen etc.

Die generelle technische Prüfung wird in den messtechnischen Abteilungen der Sendeanstalten vorgenommen. Hier stehen besonders hochwertige Messsysteme zur Verfügung und es werden spezielle Prüfsignale verwendet. Ein Beispiel ist das Videomesssystem VSA von Rohde & Schwarz (Abb. 2.110), das geeignet ist, das Signal auch mehrkanalig darzustellen und die Messung vieler Parameter anhand definierter Prüfsignale vorzunehmen. Ein solches Gerät übernimmt auch die Funktionen von Waveformmonitor und Vektorskop (s. u.). Es ist in der Lage, ein Signalspektrum darzustellen und erlaubt die Beurteilung anhand eingeblendeter Schablonen. Darüber hinaus können viele weitere Parameter, wie z. B. Zeitbasisschwankungen (Line und Field Jitter) dargestellt werden, was zur Beurteilung von MAZ-Geräten sehr hilfreich ist (Abb. 2.111).

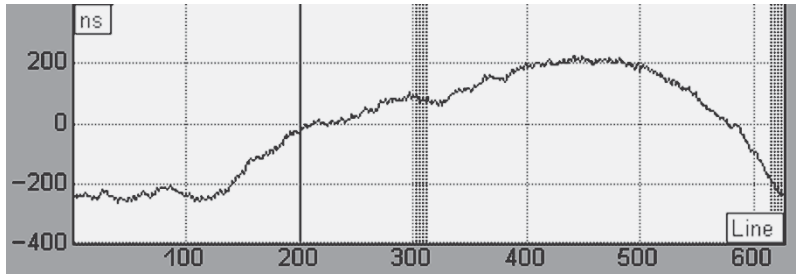


Abb. 2.111. Darstellung von Zeitbasisschwankungen eines MAZ-Systems

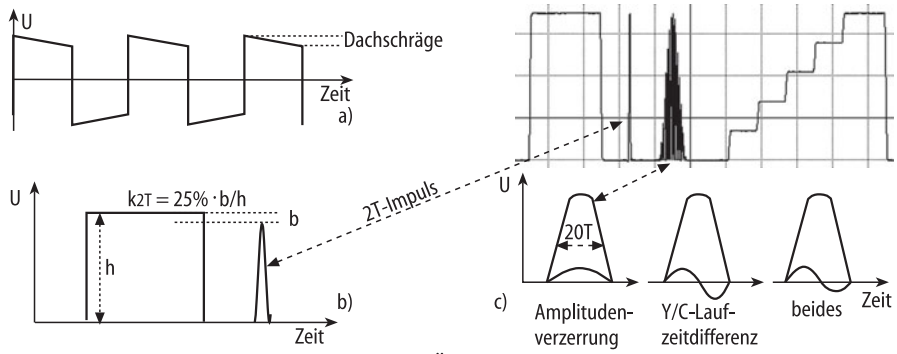


Abb. 2.112. Prüfsignale zur Untersuchung des Übertragungsverhaltens

Als Prüfsignale zur Untersuchung des Übertragungsverhaltens im Bereich niedriger Frequenzen werden die Dachschräge und Verrundung von Rechtecksignalen herangezogen (Abb. 2.112a). Je höher die untere Grenzfrequenz, desto stärker die Schräge des Rechtecks. Das 50-Hz-Signal erlaubt die Erfassung von Fehlern unterhalb 15 kHz, das 15-kHz-Signal im Bereich bis 100 kHz.

Zur Beurteilung der Übertragung höherer Frequenzen dient der 2T-Impuls, dessen Name sich von der Einschwingzeit des Übertragungskanals ableitet. Bei einer Grenzfrequenz von 5 MHz beträgt diese T = 100 ns. Der 2T-Impuls hat einen $\sin^2$-förmigen Verlauf und die Eigenschaft, dass keine Signalanteile oberhalb 5 MHz auftreten. Der Übertragungskanal wird mit Hilfe des k-Faktors für die Impulshöhe beurteilt, dessen Berechnung in Abbildung 2.112b angegeben ist. Ein großer k2T-Faktor bedeutet starke Qualitätsminderung, bei Studioanlagen sollte der k-Faktor weniger als 1% betragen.

Weitere Prüfsignale sind z. B. der 20T-Impuls mit überlagerter Farbträgerschwingung (Abb. 2.112c), mit dem Amplituden- und Laufzeitfehler im Bereich des Farbträgers beurteilt werden können. Hinzu kommen Signale zur Prüfung der statischen und dynamischen Linearität, der Intermodulation etc.

2.5.1 Betriebsmesstechnik

Eine technische Prüfung ist auch im betrieblichen Alltag unerlässlich, allerdings kann sie weniger umfangreich sein. Im Produktionsbereich müssen z. B. Kameras eingerichtet und mit den Regieeinrichtungen verbunden werden. Dabei treten Pegel- und Laufzeitdifferenzen auf. Aufgrund verschiedener Leitungslängen ergeben sich Phasendifferenzen zwischen den Signalen zweier Kameras und müssen ausgeglichen werden. Weiterhin müssen die Kameras nach der Einrichtung der Studiobeleuchtung mit einem Referenzweiß zu Unbunt abgeglichen werden, indem die Verstärker in den Farbartkanälen der Kamera so eingestellt werden, dass sie bei Vorliegen des Referenzweiß gleiche Pegel aufweisen. Für dieses Kamera-Matching müssen Prüfgeräte zur Verfügung stehen. Die Standardprüfgeräte der Betriebsmesstechnik sind der Waveformmonitor und das Vektorskop, als Prüfsignal wird das Farbbalkensignal verwendet.

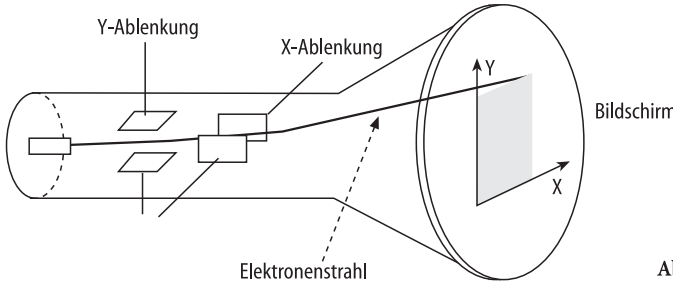


Abb. 2.113. Oszilloskop

2.5.1.1 Waveformmonitore

Diese Kontrollgeräte sind spezielle Oszilloskope für die Betriebsmesstechnik im Videobereich. Oszilloskope basieren, ebenso wie alte Fernsehgeräte auf der Braunschen Röhre (Abb. 2.113), in der ein Elektronenstrahl erzeugt wird, der einen Bildpunkt auf der den Bildschirm bedeckenden Leuchtschicht zur Licht–emission anregt. Der Strahl wird hier in Abhängigkeit von einer Messspannung abgelenkt. Beim XY-Betrieb werden zwei Spannungen für die Ablenkung in X- und Y-Richtung verwendet. Moderne Geräte arbeiten intern digital, simulieren nach außen jedoch das Verhalten der bekannten Oszilloskopanzeigen.

Die übliche zeitabhängige Darstellung eines Signalverlaufs wird erreicht, indem nur eine Messspannung zur Ablenkung in Y-Richtung benutzt wird, während die Horizontale automatisch in einer einstellbaren Zeit durchlaufen wird. Um dabei ein periodisches Signal als stehendes Bild darzustellen, wird mit einer Triggerstufe die Zeitablenkung gezwungen, immer bei einem bestimmten Pegelwert und damit immer zur selben Zeit mit derselben Phasenlage zu beginnen. So können alle Zeilen übereinander abgelesen werden.

Ein Waveformmonitor (Abb. 2.114) ist ein Oszilloskop mit einer speziellen Ausstattung für Videosignale. Z. B. ermöglicht ein Tiefpassfilter die Abschaltung der Farbinformation bei einem FBAS-Signal, sodass nur der Luminanzanteil bis ca. 2 MHz dargestellt wird (Lum) oder die Darstellung der Farban-

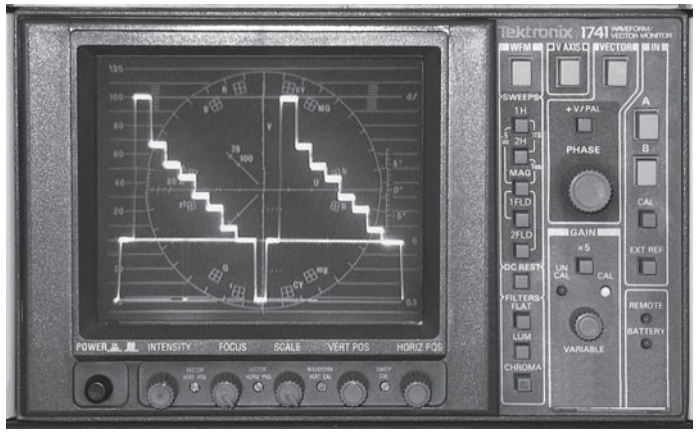


Abb. 2.114. Waveformmonitor

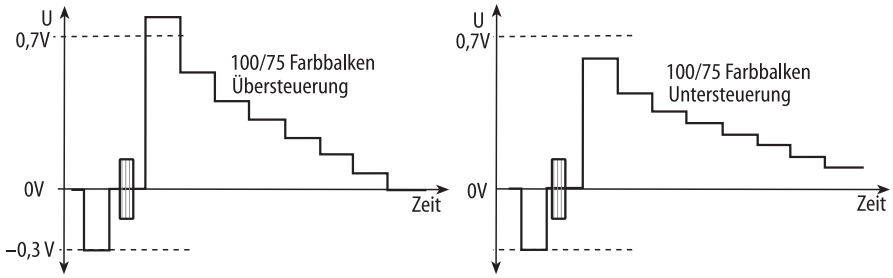


Abb. 2.115. Über- und Untersteuerung des Luminanzsignals

teile oberhalb dieses Wertes (Chrom). Die Gesamtdarstellung wird bei abgeschaltetem Filter (Flat) erreicht. Statt des üblichen Zeitbasiseinstellers ist ein Wahlschalter für H, 2H, V und 2V vorhanden, damit wird der Elektronenstrahl horizontal so abgelenkt, dass er während je einer oder zweier Zeilen- bzw. Bildperioden einmal über den Schirm läuft. In den meisten Fällen werden alle Zeilen übereinander geschrieben. Bei einigen Geräten kann in der 2H-Darstellung der Chrominanzanteil in einer der H-Perioden unterdrückt werden (Abb. 2.116). Auf dem Bildschirm ist eine Skala aufgetragen, mit deren Hilfe die Periodendauer und der Pegel einfach überprüft werden können. Die Skalierung bezieht sich entweder nur auf die Bildamplitude, d. h. 0 V und 0,7 V entsprechen 0 und 100%, oder auch auf das Synchronsignal, die Spannungen 0 V und 1 V entsprechen dann –43% und 100%.

Die Kontrolle der Amplituden eines FBAS-Signals kann leicht anhand dieser Skala erfolgen. Das Gerät wird so eingestellt, dass der Austastwert bei 0% liegt. Das Luminanzsignal ist normgerecht, wenn der Pegel des Weißbalkens 100% erreicht (Abb. 2.115 und 2.116). Der richtige Pegel für das Chrominanzsignal ist gerade dann erreicht, wenn die Amplituden der dem Grauwert überlagerten Farbhilfsträgerschwingung in den Bereichen für Gelb und Cyan des 100/75-Farbbalkens auch den Wert 100% erreichen (Abb. 2.117).

Bei der Überprüfung der Phasendifferenz zwischen verschiedenen Signalen muss beachtet werden, dass sich immer die gleiche Phasenlage zeigt, wenn sich die Triggerung des Gerätes auf das Eingangssignal einstellt (Reference

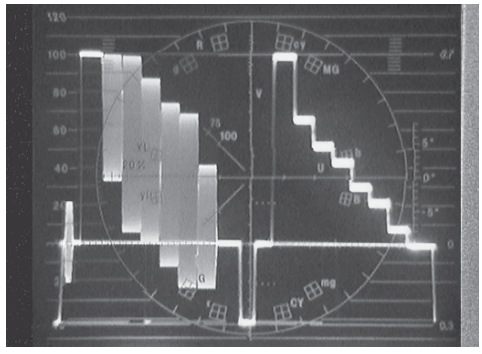


Abb. 2.116. 2H-Darstellung des normgerechten 75%-Farbbalkens auf dem Waveformmonitor

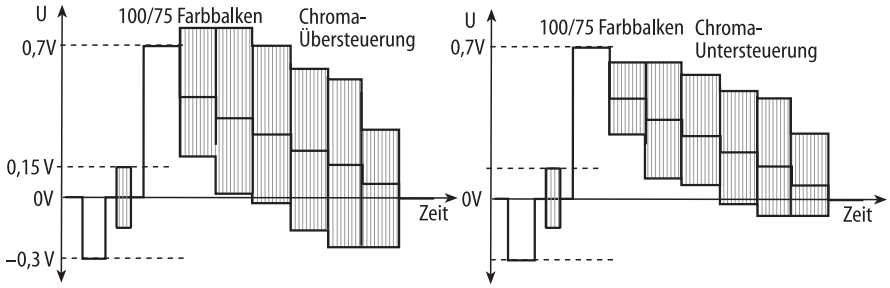


Abb. 2.117. Über- und Untersteuerung des Chrominanzsignals

Internal). Differenzen werden nur sichtbar, wenn alle Signale bezüglich einer festen Referenz, meist dem Studiotakt (Black Burst), dargestellt werden. Das bedeutet, dass der Waveformmonitor auf das extern zugeführte Referenzsignal getriggert werden muss (Reference External).

Waveformmonitore für Komponentensignale bieten besondere Funktionen und Darstellungsarten, um die drei Signalanteile Y, P_R und P_B gleichzeitig betrachten zu können. Zur Beurteilung der Pegelverhältnisse werden die drei Komponenten mit der Parade-Darstellung in zeitlich komprimierter Form nebeneinander dargestellt (Abb. 2.118). Durch einen besonderen Gleichspannungsoffset wird hier die Pegelkontrolle erleichtert. Die Nulllinie wird dabei für die Farbdifferenzkomponenten auf 50% des Videopegels (bzw. 37,5% beim 75% Farbbalken) angehoben, sodass alle Signale im positiven Bereich liegen.

Um den Zeitbezug der Signale untereinander beurteilen zu können, bzw. Zeitbasisfehler aufzudecken, werden die drei Anteile ohne zeitliche Kompression im Overlay übereinander dargestellt (Abb. 2.118). Für die gleichzeitige Pegel- und Zeitfehlerprüfung wurde schließlich die Ligthning-Darstellung entwickelt, die ihren Namen von der Zackenform erhält, die ein normgerechtes Farbbalkensignal erzeugt. Diese Darstellung wird erreicht, wenn die vertikale Ablenkung durch das Luminanzsignal gesteuert wird, wobei es zeilenweise wechselnde Polarität aufweisen muss. Die Horizontalablenkung wird zeilenweise wechselnd von je einem Farbdifferenzsignal gesteuert, sodass in den Zeilen mit nichtinvertiertem Luminanzanteil die Komponente (B – Y) anliegt und in der nächsten Zeile (R – Y). Die Beurteilung der Normrichtigkeit geschieht mit

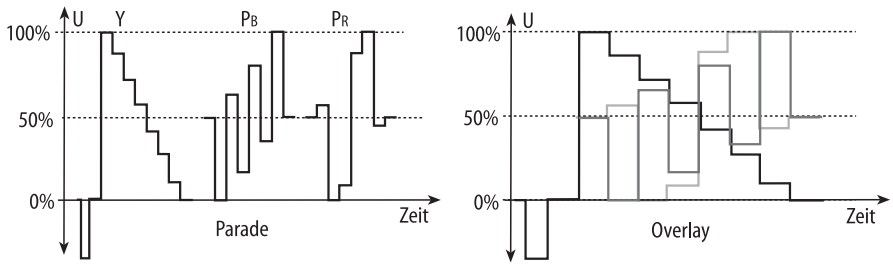


Abb. 2.118. Komponentensignale in Parade- und Overlay-Darstellung

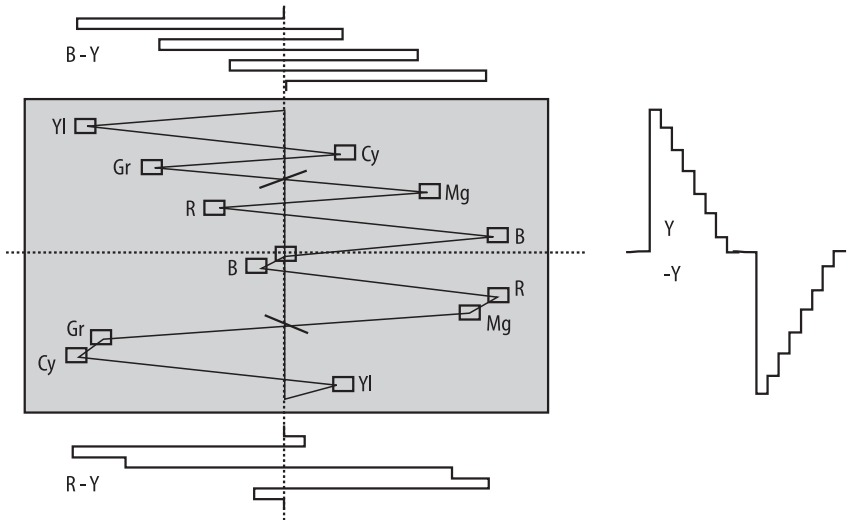


Abb. 2.119. Lightning-Darstellung des normgerechten Farbbalkensignals

Hilfe von Toleranzfeldern (Abb. 2.119), in denen die Spitzen des zackigen Signals bei korrekten Pegeln liegen müssen [73].

2.5.1.2 Vektorskop

Zur Überprüfung des Chrominanzsignals ist ein Waveformmonitor schlecht geeignet, denn der den Farbton bestimmende Phasenwinkel ist im hochfrequenten Farbträgersignal schlecht bestimmbar (Abb. 2.120). Daher wird ein Vektorskop eingesetzt, ein Gerät, das speziell zur Überprüfung des Chrominanzsignals bestimmt ist. Das Vektorskop ist ein Oszilloskop, ggf. als ein umschaltbarer Waveformmonitor, und arbeitet im XY-Betrieb. Das Gerät enthält Filter und Demodulatoren, mit denen aus dem FBAS-Signal ein Komponentensignal gewonnen wird, denn zur Ablenkung in x-Richtung wird die Komponente U und zur Ablenkung in y-Richtung die Komponente V verwandt (Abb. 2.121). Die Bildschirmfläche repräsentiert alle Farben. Im Mittelpunkt liegt der Unbuntpunkt.

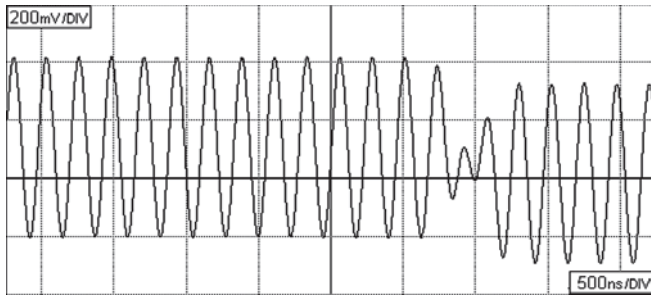


Abb. 2.120. Übergang im Farbbalkensignal in zeitlicher Darstellung

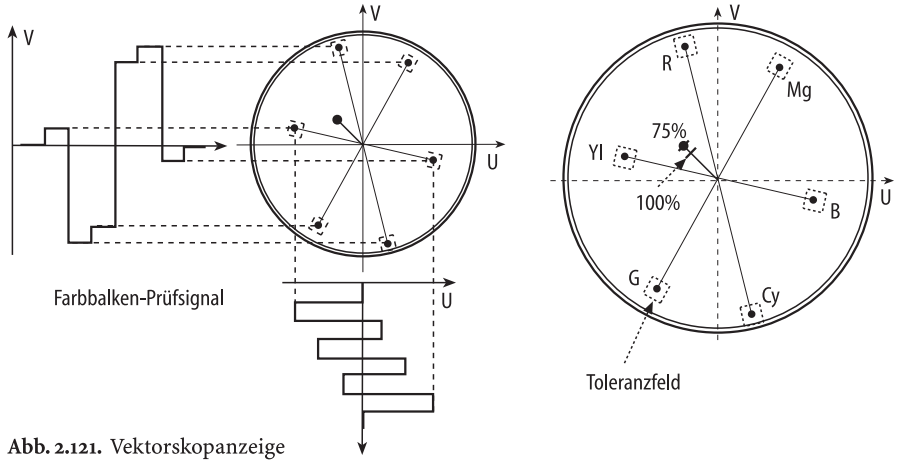


Abb. 2.121. Vektorskopanzeige

Für die Überprüfung des Chrominanzsignals mit Hilfe eines Farbbalkens sind auf der Skala des Vektorskops Punkte und Toleranzfelder für die sechs Farbwerte des Testsignals aufgebracht. Die Farben werden bei der Abtastung einer Zeile ständig durchlaufen, sodass dauernd sechs Punkte sichtbar sind, die bei optimaler Einstellung innerhalb der Toleranzfelder liegen sollten (Abb. 2.121 sowie 2.122 mit nicht invertierter V-Phase). Eine Winkelabweichung entspricht einem Phasenfehler, der sich als Farbtonfehler äußert, eine Amplitudenabweichung einem Sättigungsfehler. Die Toleranzfelder weisen einen Toleranzbereich von 5% für die Chrominanzamplitude und von 3% für die Phase auf.

Die Kalibrierung des Vektorskops geschieht bezüglich der kurzen Linie, die vom Burst herrührt. Zunächst wird der Anfangspunkt dieser Linie auf den Skalennullpunkt zentriert. Dann wird die Phase so eingestellt, dass nur eine Burstlinie unter einem Winkel von 135° erscheint. Schließlich wird die Verstärkung so gewählt, dass der Endpunkt der Burstlinie in Abhängigkeit vom

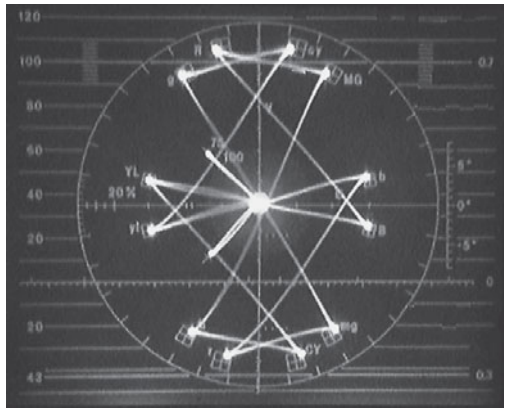


Abb. 2.122. Darstellung des normgerechten 75%-Farbbalkensignals auf dem Vektorskopschirm

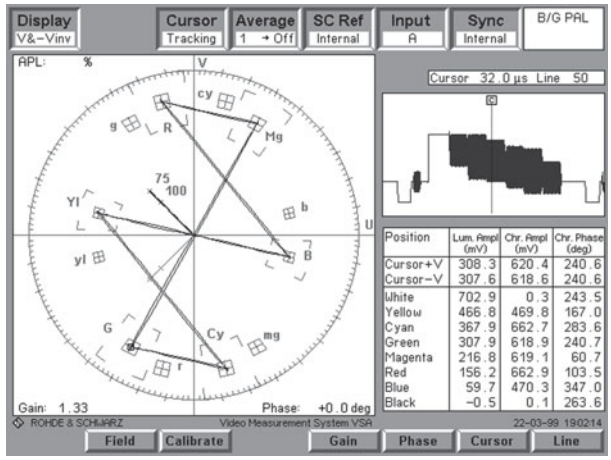


Abb. 2.123. Digitale Vektorskopdarstellung des normgerechten Farbbalkensignals

verwendeten Farbbalken auf der 75% oder 100% Marke liegt (Abb. 2.121, Abb. 2.123). Diese etwas ungewöhnliche Methode zur Beurteilung der Länge der Chrominanzvektoren wurde gewählt, weil bei dem logisch richtigeren Bezug zu einem festen Wert für den Burst zwei Gruppen von Toleranzfeldern für die beiden Typen von Farbbalkensignalen erforderlich wären und damit die Anzeige zu unübersichtlich würde.

Abbildung 2.123 zeigt die Vektorskopanzeige bei einem digital arbeitenden Messgerät das daneben noch die Waveform-Darstellung und die gemessenen Pegel angibt. Hier zeigt sich, dass die Anzeige von den analogen Oszilloskopen übernommen wird. Auch bei der Messung digitaler Videosignale selbst wird die Waveform- und Vektorskopdarstellung beibehalten.

Die Betriebsmesstechnik dient also insgesamt vornehmlich zur Überprüfung der Signalwege, wobei die Einstellungen mit Hilfe des Farbbalkensignals in der dargestellten Form vorgenommen werden. Generell muss dafür gesorgt werden, dass die Signale aller Geräte bezüglich eines Referenzsignals amplituden- und phasenrichtig sind. Diese und ähnliche Überprüfungen müssen je nach Produktionsbereich oft wiederholt werden, im Bereich der Postproduktion muss das Signal z. B. nach jedem Wechsel eines analogen Magnetbandes kontrolliert werden. Zwar sollte das Signal mit optimalem Pegel gespeichert vorliegen, aufgrund verschiedener Geräteeinstellungen ergeben sich aber Abweichungen. Professionelle Magnetbandgeräte erlauben es daher, Luminanz- und Chrominanzpegel des Ausgangssignals einzustellen (meist am TBC, s. Kap. 8). Zur Orientierung dient der aufgezeichnete technische Vorspann mit dem Farbbalkensignal. Vor der Produktion werden die Kameras zudem zu Unbunt abgeglichen. Hier ist das Vektorskop hilfreich, denn es zeigt die Abweichung vom Nullpunkt.

Neben der rein technischen Überprüfung ist die Signalkontrolle auch für das Programm erforderlich. Während des Programmablaufs werden, bei paralleler Betrachtung auf einem guten Monitor, mit Hilfe des Waveformmonitors

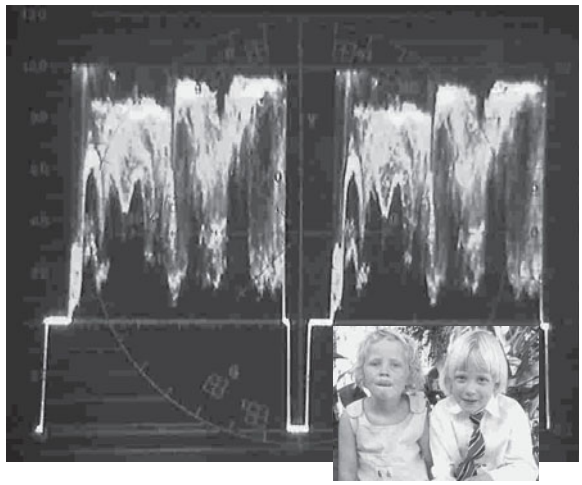


Abb. 2.124. Überlagerung aller Zeilen auf dem Waveformmonitor (mit zugehöriger Bildvorlage)

vor allem die Kamerablende und der Schwarzwert ständig so eingestellt, dass die Maximalwerte des Signals nur sehr geringfügig über 100% (0,7 V) hinausgehen und die dunkelsten Bildpartien den Austastwert nicht unterschreiten. Falls diese Bedingungen nicht eingehalten werden, ergibt sich in den Lichtern bzw. in den dunklen Bildpartien keine Zeichnung mehr, d. h. dort können keine Grauwerte mehr differenziert werden. Auch der umgekehrte Fall, die Unterschreitung des Maximalpegels und eine wesentliche Überschreitung der Schwarzabhebung, sollte vermieden werden, da dann der Kontrastbereich nicht ausgenutzt wird und das Bild flau erscheint (s. Abschn. 6.2.3). Für diesen Einsatzbereich ist es unerlässlich, dass die Geräte alle Bildzeilen gleichzeitig anzeigen (Abb. 2.124).

Es gibt allerdings Fälle, in denen die Darstellung aller Zeilen in ihrer Fülle auch verwirrend ist. Ein Beispiel ist die Nachrichtenproduktion, bei der es für die Bildtechnik wesentlich darauf ankommt, mit Hilfe der Blende vor allem für das Gesicht der Moderatoren den richtigen Pegel einzustellen. Der Helligkeitswert bei hellhäutigen Menschen sollte bei 63% liegen. Gerade wenn sich die Person nicht bewegt, ist es aber oft schwer, festzustellen welcher Signalanteil auf diesen Wert ausgesteuert werden soll. In diesem Fall ist ein Gerät hilfreich, das den Farbton des Gesichts als Kriterium nimmt, um einen Schaltvorgang auszulösen, der bewirkt, dass alle Signale außerhalb eines Toleranzbereichs um diesen Farbton herum ausgeblendet werden [239]. Abbildung 2.125 zeigt ein Beispiel für diese Funktion. Obwohl hier jeweils nur eine Zeile und nicht die Überlagerung gezeigt ist, wird deutlich, wie hilfreich es sein kann, wenn im gefilterten Signal nur noch die Gesichtsbereiche dargestellt werden.

Ein bekannter Messgerätehersteller für den Videobereich ist die Firma Tektronix. Ein universell einsetzbares Gerät aus diesem Hause ist der Typ WFM 700, das zur Messung von SD und HD-Signalen in analoger oder digitaler Form verwendet werden kann. Er arbeitet intern digital und ist in der Lage,

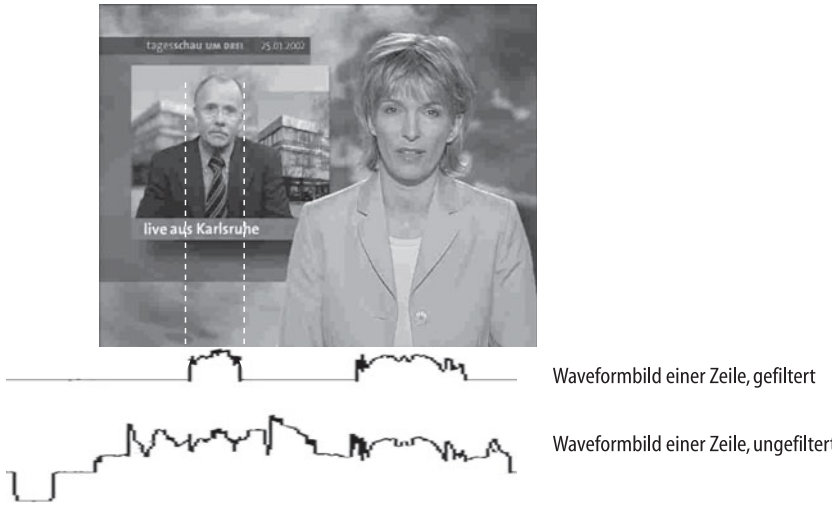


Abb. 2.125. Waveformsignale mit und ohne Hauttonfilter [239]

die Ergebnisse als Zahlwerte, aber auch als Abbildung darzustellen. Die Darstellungen sind von den Displays der analogen Messgeräte abgeleitet.

Der WFM 700 übernimmt Funktionen von Waveformmonitor und Vektorskop und stellt auf Wunsch zusätzlich das Bild selbst dar (Abb. 2.126 links). Im Waveformmode können die Zeilen gemeinsam oder zeilenselektiv dargestellt

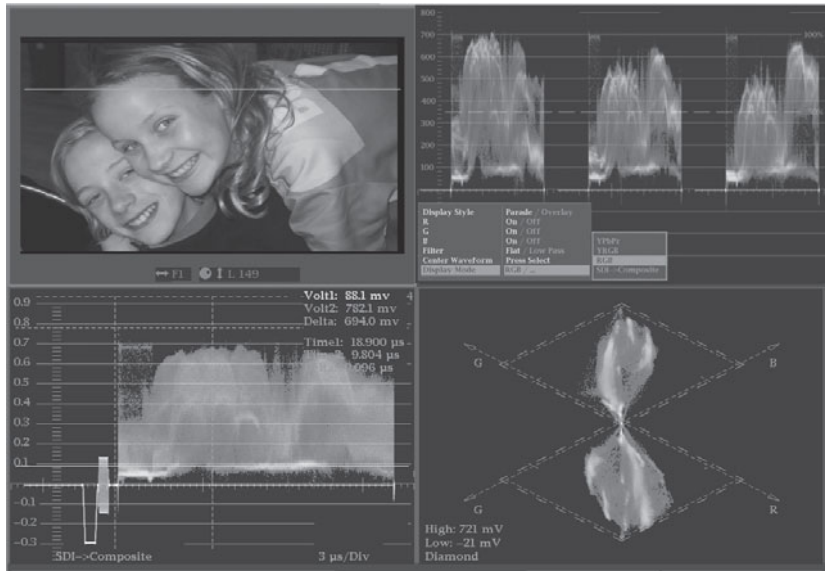


Abb. 2.126. Bild-, RGB-Parade-, Waveform- und Diamond-Darstellung

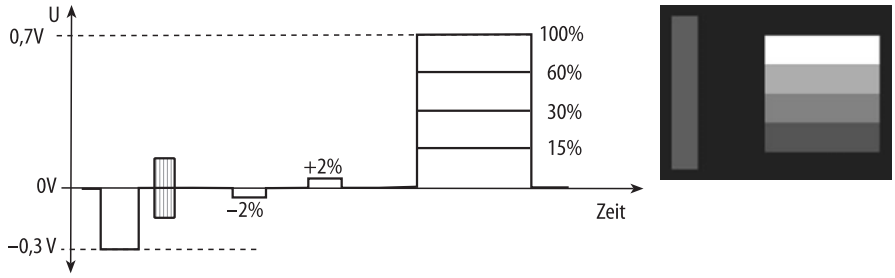


Abb. 2.127. Pluge-Signal zur Helligkeits- und Kontrasteinstellung

werden. Dabei kann im Gerät eine Umrechnung der Signale vorgenommen werden, sodass z. B. ein Komponentensignal auch in seinen RGB-Anteilen oder als FBAS-Simulation angezeigt werden kann.

Die Programmdarstellung als RGB-Parade (Abb. 2.126 rechts oben) ist für Farbkorrekturen sehr hilfreich. Um weitere Hilfen für den Bereich der Farbkorrektur zu schaffen, hat Tektronix eine besondere Farbraumanzeige entwickelt, die auch auf RGB-Signalanteilen beruht und als Diamond-Display bezeichnet wird (Abb. 2.126 rechts unten). Der obere Teil der Fläche entsteht, indem das Signalgemisch B+G auf die y-Achse und das Gemisch B–G auf die x-Ablenkung eines Oszilloskops gegeben wird. Der untere Teil der Fläche wird aus –(R+G) für die y- und R–G für die x-Richtung gebildet. Wenn alle gemessenen Signalanteile innerhalb der Fläche liegen, ist sichergestellt, dass für den RGB-Farbraum ein legales Signal vorliegt.

Schließlich muss festgestellt werden, dass nicht nur Messgeräte, sondern auch die verwendeten Bildmonitore sehr wichtige Kontrollgeräte sind. Damit sie diese Funktion erfüllen können, müssen sie natürlich korrekt kalibriert sein und mit der richtigen Umgebungsleuchtdichte betrieben werden. Als Hilfe für die Einstellung von Schwarz- und Weißwert dient das PLUGE-Signal (Picture Lineup Generating Equipment, Abb. 2.127). Die Kalibrierung wird folgendermaßen vorgenommen: Der Monitor sollte sich in einer Umgebungshelligkeit mit einer Leuchtdichte befinden, die ca. 10% der des Weißwertes beträgt. Zunächst wird der Schwarzwert festgelegt, dafür ist der Bildhelligkeitssteller maßgeblich. Die Bildhelligkeit muss so eingestellt werden, dass die – 2%-Untersteuerung des PLUGE-Signals im Bild verschwindet und die + 2%-Abhebung gerade zu erkennen ist. Mit dem Kontrasteinsteller (Gain) wird dann das Bildschirmweiß so eingestellt, dass sie bei Studiomonitoren 80 cd/m² erreicht (s. auch Abschn. 7.4.1.2). Dabei sei zum Abschluss dieses Kapitels noch bemerkt, dass hier und bei allen anderen Prüfvorgängen der normrichtige Leitungsabschluss nicht vergessen werden darf.

3 Das digitale Videosignal

3.1 Digitale Signalformen

Das Wort Digital bezieht sich auf zählbare Quantitäten. Die Nachrichtenübertragung mit Hilfe einer begrenzten Menge ganzer Zahlen anstelle von unendlich vielen Analogwerten hat den Vorteil, dass die wenigen Digitalwerte im Prinzip eindeutig rekonstruierbar sind. Zwar treten die bei der Signalübertragung unvermeidlichen Störungen und das Rauschen bei der Digitaltechnik ebenso auf wie im analogen Fall, die Digitalwerte können aber so gewählt werden, dass sie sich eindeutig von Störungen und Rauschen abheben.

Ein zweiter Vorteil ist, dass Digitaldaten vor allem aufgrund der zeitlichen Diskretisierung flexibler als ein Analogsignal verarbeitet werden können. Einfache Kurzzeitspeicher (RAM) mit sehr schneller Zugriffszeit sind realisierbar. Digitale Daten können einfach umgeordnet und optimal an Übertragungskanäle angepasst werden, zusätzlich gibt es die Möglichkeit der Fehlerkorrektur. Digitaldaten können mit Rechnern auf vielfältige Weise manipuliert werden, dabei werden extreme Bildveränderungen möglich, die analog nicht realisierbar sind. Hinzu kommt heute, dass die Digitalisierung ein Zusammenwachsen des Videobereichs mit den Bereichen Telekommunikation und Informatik hervorruft, sodass aufgrund des dort vorhandenen Massenmarktes digital arbeitende Systeme auch preisgünstiger werden als analoge.

Bezüglich der elektronischen Schaltungen ergibt sich der Vorteil, dass digitale Schaltkreise wesentlich stabiler arbeiten als analoge und sehr viel weniger Abgleich erfordern. Digitale Schaltungen sind oft auch störfester und brauchen daher nicht so aufwendig abgeschirmt zu werden.

Bereits in den 70er Jahren wurde die Digitaltechnik im Videobereich eingesetzt, zunächst für Grafiksysteme und zur Normwandlung (z. B. NTSC in PAL). Mit digitalen Normwandlern konnte der bis dahin notwendige und erheblich qualitätsmindernde Weg über die optische Wandlung, d. h. die Wiedergabe des Bildes in einer Norm und die Abtastung in der anderen, vermieden werden. Eine zweite frühe Anwendung der digitalen Signalverarbeitung betraf den Einsatz in Zeitfehlerausgleichern (TBC) für MAZ-Maschinen. Das aufgrund des nichtkonstanten Magnetbandlaufs mit Zeitfehlern behaftete Signal wird im TBC digitalisiert, gespeichert, mit einem stabilen Studiotakt wieder ausgelesen und zurück in ein analoges Signal gewandelt.

Da das digitale Videosignal gut vom Rauschen separiert werden kann, hat es besondere Bedeutung für die Magnetaufzeichnungsverfahren (MAZ). Es erlaubt sehr viele Kopiergenerationen, während bei analoger Aufzeichnung die

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4_3, © Springer-Verlag Berlin Heidelberg 2013

Schwelle des tolerierbaren Rauschens schon nach wenigen Generationen überschritten ist. Schon Mitte der 80er Jahre wurde ein digital arbeitendes MAZ-Gerät (D1) entwickelt. Anfang der 90er Jahre folgten weitere Typen (D2 – D5). Digitale MAZ-Geräte zur Aufzeichnung von HDTV-Signalen stehen seit Mitte der 90er Jahre ebenso zur Verfügung wie digitale Formate für Heimanwendungen (DV, HDV).

Der Studioeinsatz der digitalen Videotechnik begann mit kleinen digitalen Inseln, die jeweils von A/D- und D/A-Umsetzern umgeben waren. Im Laufe der Zeit wurden die Inseln immer größer, sodass heute eher analoge Inseln zwischen digitalen Komplexen vorliegen. Mit der Einführung digitaler Broadcast-Verfahren wuchsen die Inseln immer weiter zusammen, sodass ein vollständig digitales Videosystem entsteht.

3.2 Grundlagen der Digitaltechnik

Die Digitaltechnik nutzt das Dualzahlensystem, eine Zahlenstelle (Binary Digit, Bit) kann nur zwei Werte annehmen (anstatt 10 im Dezimalsystem). Die beiden Werte werden üblicherweise als »0« und »1« bezeichnet und elektronisch durch Schaltzustände »ein« (z. B. 5 V, high, H) und »aus« (z. B. 0 V, low, L) repräsentiert. Wird statt der Spannung 0 V durch eine Störspannungsinduktion z. B. der Wert 0,31 V übertragen, so ist dieser Wert immer noch viel kleiner als 5 V und kann im Empfänger als 0 V interpretiert werden, sodass der Fehler beseitigt ist. Ungenauigkeiten bezüglich der zeitlichen Dimension (Jitter), die z. B. in allen Aufzeichnungssystemen auftreten, werden in ähnlicher Weise durch kurzzeitige Abtastung ausgeblendet (Abb. 3.1).

Das Dualzahlensystem beruht auf der Basiszahl Zwei. Die Dualzahl 1 0 1 1 entspricht $2^3 + 2^1 + 2^0$ und damit der Dezimalzahl 11. Vier Bit können durch eine Hexadezimalzahl repräsentiert werden, dabei werden Werte, die über 9 hinausgehen, als Buchstaben dargestellt (z. B. 1 0 1 1 = B). Darüber hinaus werden 8 Bit zu einem Byte (By Eight) zusammengefasst (Abb. 3.2). Ein Byte kann auch einen Buchstaben repräsentieren, da die Buchstaben des Alphabets eindeutig bestimmten Zahlenwerten zugeordnet wurden (ASCII-Code).

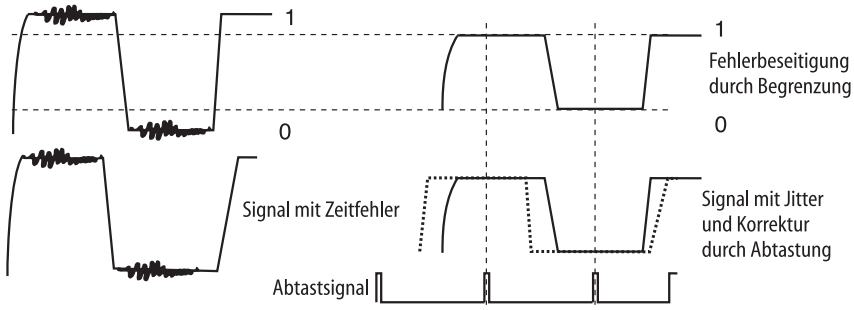


Abb. 3.1. Rekonstruktion gestörter Digitalsignale

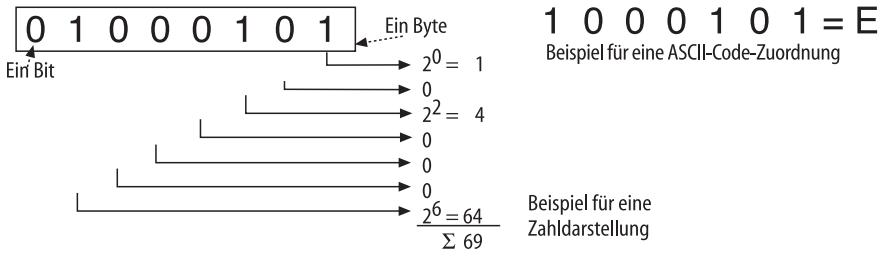


Abb. 3.2. Zuordnung von Binärzahlen zu Dezimalzahlen und ASCII-Code

3.2.1 Digitalisierung

Das digitale Videosignal wird aus dem analogen gewonnen. Das Analogsignal ist zwar bereits in vertikaler und zeitlicher Hinsicht in Zeilen und Bilder eingeteilt, innerhalb dieses Rahmens liegt es aber zeitkontinuierlich und auch wertkontinuierlich vor. In beiden Bereichen sollen diskrete Werte erzeugt werden, denn in der Digitaltechnik werden die Signalwerte in Klassen eingeteilt (quantisiert) und die diskreten Klassennummern anstelle der analogen Werte zu definierten Zeitpunkten übertragen. Zunächst wird der Zeitablauf durch Abtastung diskretisiert, anschließend werden die Werte im A/D-Umsetzer quantisiert und codiert. Dieser Vorgang wird auch als »Digitalisierung« bezeichnet. Das Digitalsignal kann durch D/A-Umsetzer wieder in die analoge Form überführt werden.

3.2.1.1 Zeitdiskretisierung

Die zeitliche Diskretisierung wird durch Abtastung realisiert. In regelmäßigen Abständen werden dem zu wandelnden Signal Proben (Samples) entnommen und bis zur nächsten Probenentnahme gespeichert (Hold). Eine Sample & Hold-Schaltung ist prinzipiell einfach durch Aufladung eines Kondensators C realisierbar (Abb. 3.3). Nach der Zeitdiskretisierung liegt dann ein Puls-Amplitudenmoduliertes (PAM) Signal vor (Abb. 3.4), dessen Werte noch analoger Art sind.

Je höher die Abtastfrequenz ist, d. h. je mehr Abtastwerte pro Zeiteinheit vorliegen, um so besser wird der ursprüngliche Signalverlauf angenähert, desto mehr Daten müssen aber auch übertragen werden. Shannon hat beschrieben,

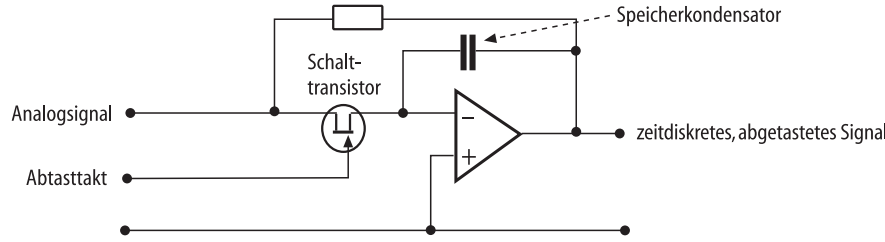


Abb. 3.3. Abtast- und Halte-Schaltung

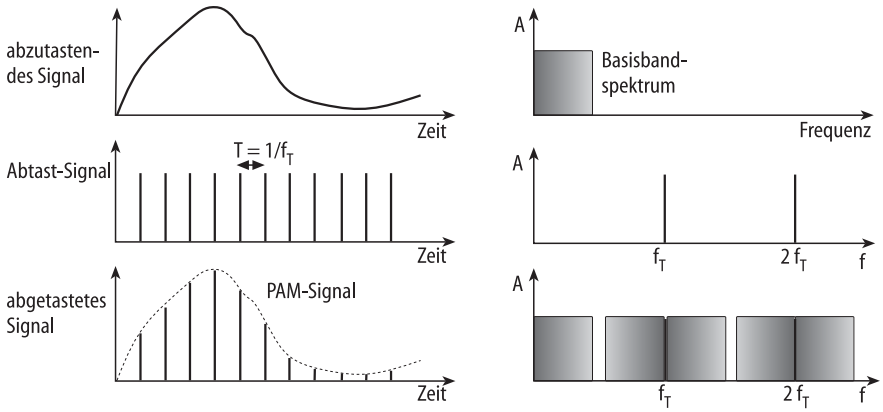


Abb. 3.4. Signal, Abtastsignal, abgetastetes Signal und zugehörige Amplitudenspektren

welche Abtastfrequenz f_T (Samplingrate) mindestens erforderlich ist, um das Signal störungsfrei aus dem PAM-Signal rekonstruieren zu können. Demnach reicht es aus, die Abtastrate mehr als doppelt so groß wie die höchste auftretende Signalfrequenz f_{go} zu wählen, das zugehörige Abtasttheorem lautet:

$$f_T > 2 \cdot f_{go}.$$

Dabei kommt es nicht nur darauf an, dass die höchste zu berücksichtigende Signalfrequenz (z. B. 20 kHz im Audiobereich) erfasst wird, es muss auch dafür Sorge getragen werden, dass oberhalb f_{go} keine Spektralanteile im Signal auftreten. Vor der A/D-Umsetzung ist also eine Tiefpassfilterung erforderlich.

Zur Veranschaulichung des Abtasttheorems kann die Abtastung des Signalverlaufs als Modulation und damit als Produktbildung zwischen dem Signal s(t) und den Abtastimpulsen $f_T(t)$ (Träger) betrachtet werden:

$$s_p(t) = s(t) \cdot f_T(t).$$

Zur Ermittlung des Produktspektrums sei angenommen, dass das zu übertragende Signal s(t) die Fouriertransformierte S(ω) hat, sodass gilt:

$$s(t) = \int S(\omega)\, e^{j\omega t}\, d\omega,$$

mit den Integralgrenzen $-\infty$ und $+\infty$ (s. Abschn. 2.3.8). Das periodische Abtastsignal $f_T(t)$ mit der Periodendauer $T = 1/f_T$ habe die Fourierkoeffizienten c_k mit k zwischen $-\infty$ und $+\infty$, sodass gilt:

$$f_T(t) = \sum c_k\, e^{jk\Omega t}, \text{ mit Zählindex k und } \Omega = 2\pi/T.$$

Für das Produkt $s_p(t) = s(t) \cdot f_T(t)$ gilt dann [20]:

$$s_p(t) = \sum c_k \int S(\omega)\, e^{j(\omega + k\Omega)t}\, d\omega.$$

Mit der Substitution $\omega' = \omega + k\Omega$ folgt

$$s_p(t) = \int \left(\sum c_k\, S(\omega' - k\Omega)\right) e^{j\omega' t}\, d\omega'.$$

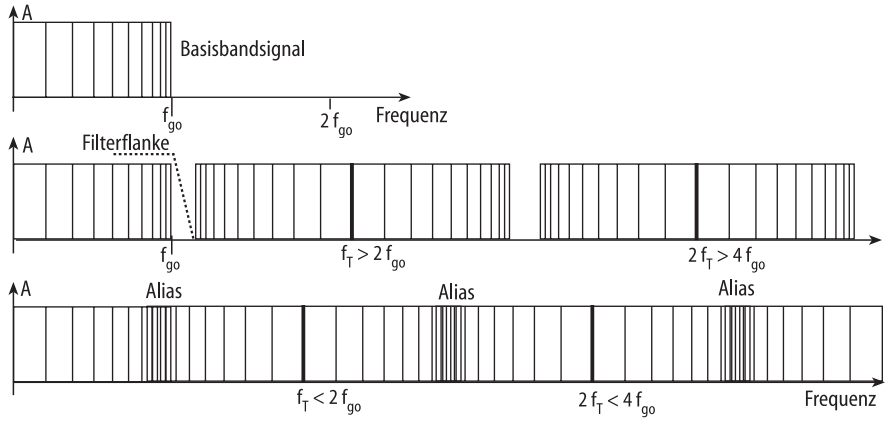


Abb. 3.5. Entstehung von Alias-Anteilen

Die Fouriertransformierte der Zeitfunktion $s_p(t)$ hat damit also folgende Form:

$$S_p(\omega) = \sum c_k S(\omega - k\Omega).$$

Die grafische Darstellung der Spektren (Abb. 3.5) zeigt, dass das Spektrum $S(\omega)$ des Signals $s(t)$ mit seinen beiden Seitenbändern mehrfach auftritt, wobei es um den Wert $k\Omega$ verschoben ist. Die Größe der Amplituden um $k\Omega$ hängt von c_k ab. Das Abtasttheorem sagt nun nichts anderes aus, als dass sich zur fehlerfreien Signalrekonstruktion die Teilspektren eindeutig trennen lassen müssen, was wiederum heißt, dass sie sich nicht überschneiden dürfen. Da das obere Seitenband des einen Teilspektrums an das untere Seitenband des nächsten grenzt, muss gelten:

$$\Omega > 2 \cdot \omega_{go}, \text{ bzw. } f_T > 2 \cdot f_{go} \text{ (Abtasttheorem).}$$

Für jede Halbwelle der höchsten im Signal auftretenden Frequenz muss also mindestens ein Abtastwert zur Verfügung stehen. Falls das nicht gilt und sich die Teilspektren überlappen, entstehen als Fehler scheinbare Informationen, die als Alias bezeichnet werden (Abb. 3.5). Eine Verletzung des Abtasttheorems ist z. B. dann gegeben, wenn ein Videosignal mit 5 MHz Bandbreite von einem 7 MHz-Störsignal begleitet wird und die Abtastfrequenz bei 11 MHz liegt, was eigentlich für die störfreie Übertragung des 5 MHz-Signals ausreicht. In diesem Beispiel führt die bei der Modulation auftretende Summen- und Differenzbildung zwischen Stör- und Abtastfrequenz zu den Frequenzen 18 MHz und 4 MHz. Letztere tritt also im Basisband in Erscheinung, so als ob sie bereits im Originalsignal vorgelegen hätte (Alias). Wird als zweites Beispiel ein Audiosignal mit 32 kHz abgetastet und eine Signalfrequenz von 18 kHz zugelassen, so ergibt die Differenz zur Abtastfrequenz ein 14 kHz-Signal und erscheint im Hörbereich als Alias-Frequenz, die im Original nicht vorhanden ist.

Zur Alias-Vermeidung ist der Einsatz hochwertiger Filter erforderlich. Da aber auch gute Filter nie beliebig steilflankig arbeiten, darf die Samplingfrequenz nicht zu nah an $2 \cdot f_{go}$ liegen, je kleiner die Lücke, desto aufwendiger der

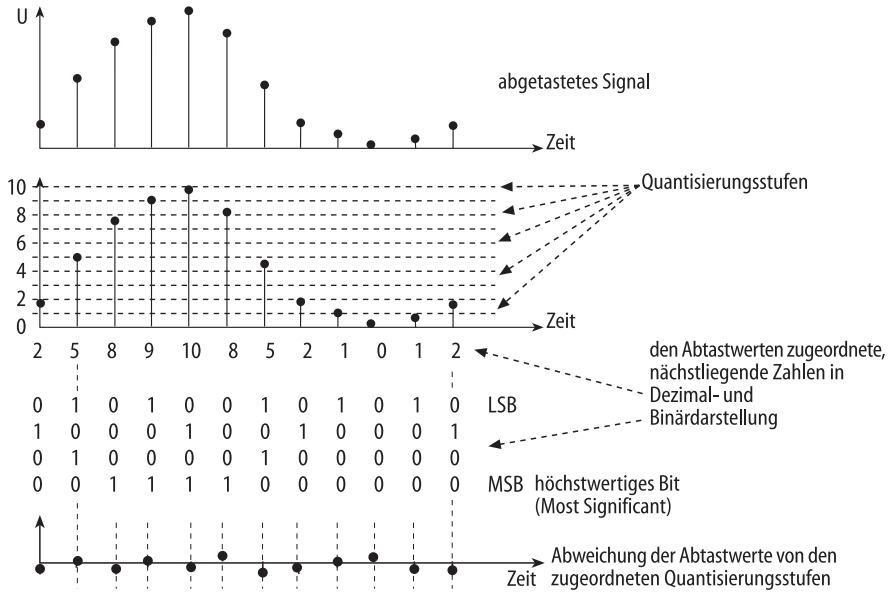


Abb. 3.6. Wertquantisierung und Abweichung vom Originalwert

Filter (Abb. 3.5). Im Audiobereich mit $2 \cdot f_{go} = 40$ kHz werden z. B. Abtastraten von 44,1 kHz oder 48 kHz verwandt. Im Videobereich wird das SD-Luminanz-Signal meist mit 13,5 MHz abgetastet. Bei mehrfacher Filterung mit der gleichen Grenzfrequenz ergibt sich eine verringerte Gesamt-Grenzfrequenz, mehrfache A/D- bzw. D/A-Umsetzungen (Wandlungen) sollten daher vermieden werden.

3.2.1.2 Amplitudenquantisierung

Das nächste Stadium bei der Signaldiskretisierung ist die Quantisierung der Abtastwerte. Hier gibt es keine einfache Grenze für fehlerfreie Rekonstruktion, hier gilt: je mehr Quantisierungsstufen zugelassen werden, desto besser ist die Signalqualität. Die sinnvolle Auflösung richtet sich nach den Qualitätsansprüchen, die durch die menschlichen Sinne bestimmt sind. Unter der Annahme, dass der bei der Quantisierung auftretende Fehler wie ein Rauschen wirkt, lässt sich der tolerierbare Fehler über den Signal-Rauschabstand (S/N) bestimmen. Für den Audiobereich wird S/N > 90 dB und für den Videobereich S/N > 45 dB gefordert. Im Folgenden wird gezeigt, dass bei einer Amplitudenauflösung in 65000 Stufen (darstellbar mit 16 Bit) keine Verfälschung gegenüber dem Original hörbar ist. Für den Gesichtssinn reicht es meist aus, das Videosignal in 256 Stufen (8 Bit) aufzulösen. Wenn höhere Auflösungen verwendet werden, liegt das vor allem daran, dass bei der Signalbearbeitung Spielraum für Rechenprozesse und die dabei unvermeidlichen Rundungsfehler benötigt wird.

Die zwangsläufig auftretenden Abweichungen, die entstehen, wenn unendlich viele mögliche Werte auf eine endliche Anzahl reduziert werden (Abb. 3.6), führen zum Quantisierungsfehler, der bei genügend hoher Aussteuerung

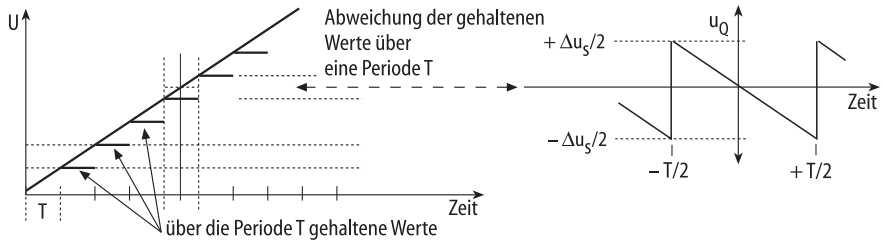


Abb. 3.7. Zur Bestimmung der Quantisierungsrauschleistung

als über den Frequenzbereich statistisch gleich verteilt angenommen werden kann. Unter dieser Voraussetzung spricht man auch von Quantisierungsrauschen. Die Voraussetzung ist jedoch bei kleinen Signalpegeln nicht gegeben. Der Quantisierungsfehler tritt dann mit starken periodischen Anteilen auf, sodass im Audiobereich ein unangenehmes Geräusch zu vernehmen ist, das mit Hilfe von zugesetzten unregelmäßigen Signalen (Dither) verdeckt wird. Der Fehler wird bei Videosignalen als stufiger Grauwertübergang sichtbar und kann ebenfalls durch Dither verdeckt werden.

Unter der einfachen Annahme eines statistisch gleichverteilten Fehlers der maximal um eine halbe Quantisierungsstufenhöhe Δu_s nach oben oder unten abweicht, gilt mit der Anzahl $s = 2^N$ gleichverteilter Quantisierungsstufen:

$$u_{ss} = (s - 1) \cdot \Delta u_s = (2^N - 1) \cdot \Delta u_s.$$

Bei gleichmäßiger Verteilung der Signalwerte wird der Verlauf der Quantisierungsfehlerspannung u_Q über dem Abtastintervall T sägezahnförmig (Abb. 3.7), und es gilt [148]:

$$u_Q = - \Delta u_s \cdot t/T.$$

Damit folgt für eine große Zahl s die maximale Quantisierungsrauschleistung P_Q aus dem Integral über dem Quadrat des Signalverlaufs in der Periode zwischen +T/2 und –T/2, bezogen auf den Widerstand R und die Periodendauer:

$$P_Q = 1/T \cdot \int_t u_Q^2/R \, dt = 1/T \cdot \int_t (\Delta u_s \cdot t/T)^2/R \, dt = (\Delta u_s)^2/R \cdot 1/12.$$

Wegen der verschiedenen Signalformen und der damit zusammenhängenden Bestimmung der Signalleistungen ergeben sich für Audio- und Videosignale unterschiedliche Formeln zur Bestimmung des Signal-Rauschabstands S/N_Q. Bei einem Sinussignal folgt die maximale Signalleistung P_S aus dem Effektivwert:

$$P_S = (u_s/\sqrt{2})^2/R = u_{ss}^2 /8R = (2^N \cdot \Delta u_s)^2 /8R,$$

und für den Signal-Quantisierungsrauschabstand gilt beim Audiosignal:

$$S/N_Q = 10 \log P_S/P_Q = 10 \log 3/2 \cdot 2^{2N} = (N \cdot 6{,}02 + 1{,}76) \text{ dB} \approx (N \cdot 6 + 2) \text{ dB}.$$

Bei einem Videosignal folgt die höchste Leistung aus dem Spitze-Spitze-Wert:

$$P_S = u_{ss}^2/R = (2^N \cdot \Delta u_s)^2 /R$$

und als Signal-Quantisierungsrauschabstand für das Videosignal:

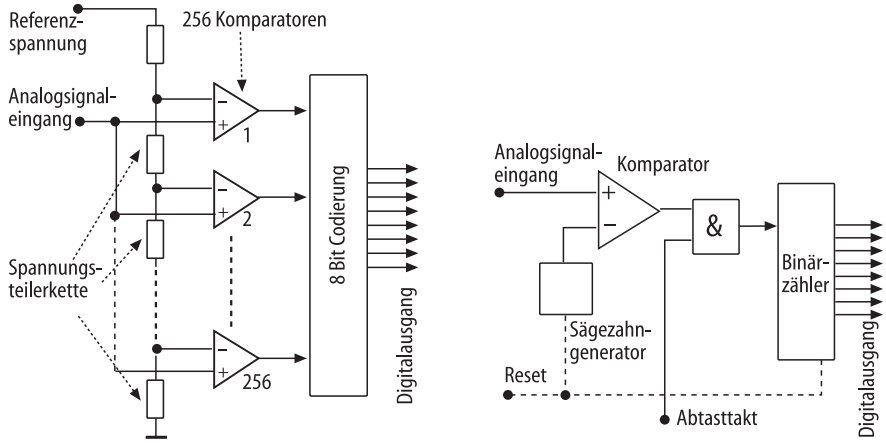


Abb. 3.8. Prinzipschaltbilder von Parallel-A/D-Wandler und Zählwandler

$$S/N_Q = 10 \log P_S/P_Q = 10 \log 12 \cdot 2^{2N} = (N \cdot 6{,}02 + 10{,}8)\ \text{dB} \approx (N \cdot 6 + 11)\ \text{dB}.$$

Bei Verwendung der Stufenzahl s anstelle der Bitanzahl N gilt die Beziehung:

$$S/N_Q = 20 \log s \cdot \sqrt{12}.$$

Der Signal-Rauschabstand bezüglich des Quantisierungsrauschens steigt mit 6 dB pro Quantisierungsbit, also pro Faktor 2, denn 20 log 2 ≈ 6 dB. Ein Audiosignal hat bei einer Quantisierung mit 16 Bit eine hohe Signalqualität, bei 8 Bit ist dagegen das Quantisierungsgeräusch deutlich hörbar. Ein mit 8 Bit digitalisiertes Videosignal hat dagegen einen Signal-Rauschabstand von 59 dB, wenn alle Stufen genutzt werden. Dies ist im Videobereich ein guter Wert, da ab ca. 42 dB Rauschstörungen visuell nicht mehr wahrnehmbar. sind

Die angegebenen Werte für das Signal-Rauschverhältnis beziehen sich auf die Vollaussteuerung unter Nutzung aller Stufen, also der vollen Systemdynamik. Meist können aber nicht alle Stufen genutzt werden, denn bei der Signalaussteuerung vor der Wandlung wird darauf geachtet, dass ein Sicherheitsbereich zur Verfügung steht. Die Vermeidung von Übersteuerungen hat für das Audiosignal sehr große Bedeutung, denn der A/D-Umsetzer bewirkt eine harte Beschneidung des Signals, was nicht lineare Verzerrungen mit hochfrequenten Störsignalanteilen hervorruft, für die das Gehör sehr empfindlich ist. Im Audiobereich wird mit Übersteuerungsreserven (Headroom) vom Faktor 2 bis 4 gearbeitet (6 bis 12 dB). Beim Videosignal werden dagegen fast alle Stufen genutzt, da eine Übersteuerung keine derart gravierenden Störungen hervorruft. Die mit der Übersteuerung verbundene Beschneidung führt nur zum Verlust einiger Grauabstufungen, der Zeichnung, im Weißbereich.

Zur praktischen Realisierung der Amplitudenquantisierung (A/D-Umsetzung) gibt es mehrere Verfahren. In der Videotechnik wird die schnellste Variante, der Parallelumsetzer verwendet, bei dem die umzuwandelnde Spannung gleichzeitig mit allen Quantisierungsstufen verglichen wird (Abb. 3.8). Die Spannungswerte der Stufen werden aus einem Spannungsteiler gewonnen. Be-

reits nach einem Vergleichsschritt liegt das quantisierte Signal vor. Das Verfahren ist allerdings technisch anspruchsvoll, denn bereits für eine Auflösung mit 8 Bit müssen gleichzeitig 256 Vergleichswerte zur Verfügung stehen. Mit diesem Prinzip können 10-Bit-Abtastungen mit mehr als 75 MHz realisiert werden. Einfacher ist die Wägewandlung. Hier wird zuerst das höchstwertige Bit des zu konstruierenden Datenwortes gleich eins gesetzt, sodass es den halben Analogwert repräsentiert. Der Digitalwert wird in einen Analogwert gewandelt, und je nachdem ob dieser größer oder kleiner ist, wird er dann auf den Analogwert 3/4 oder 1/4 eingestellt. Dann wird wieder verglichen, bis sich nach N Schritten der Digitalwert dem analogen bestmöglich genähert hat.

Beim dritten Typ, dem Zählwandler, läuft eine Sägezahnspannung vom Minimum bis zum Maximum des zu wandelnden Signals. Während der Anstiegszeit läuft auch ein Digitalzähler, der gestoppt wird, sobald die Sägezahnspannung den zu wandelnden Wert erreicht hat (Abb. 3.8). Dieser Wandlertyp kann im Audiobereich verwendet werden.

3.2.2 D/A-Umsetzung

D/A-Umsetzter (D/A-Wandler) dienen zur Rückwandlung der Digital- in Analogwerte. Es entsteht dabei natürlich nicht das Originalsignal, die Quantisierungsfehler bleiben erhalten. Ein D/A-Umsetzer kann im einfachsten Fall mit gewichteten Leitwerten aufgebaut werden, deren Werte im Verhältnis 1:2:4:... stehen. Addiert man die Ströme, die bei konstanter Spannung durch diese Leitwerte bestimmt sind, so entsteht als Summe ein Signal, das dem PAM-Signal auf der A/D-Umsetzerseite ähnlich ist. Die übertragenen Digitalwerte dienen dazu, die richtigen Leitwertkombinationen einzuschalten (Abb. 3.9).

Die Schaltvorgänge sind mit kurzen Pulsen verbunden die hochfrequente Spektralanteile aufweisen. Daher wird am Umsetzerausgang zunächst wieder eine Sample & Hold-Schaltung eingesetzt, die den gewandelten Wert zu einem Zeitpunkt abtastet, an dem die Störungen weitgehend abgeklungen sind. Schließlich wird zur Glättung der sich aus dem Hold-Vorgang ergebenden Stufenfunktion ein Rekonstruktionsfilter eingesetzt, das als Tiefpass Frequenzen oberhalb der höchsten Grenzfrequenz des Originalsignals (z. B. 5 MHz bei einem SD-Videosignal) unterdrückt und damit das Originalsignal herausfiltert.

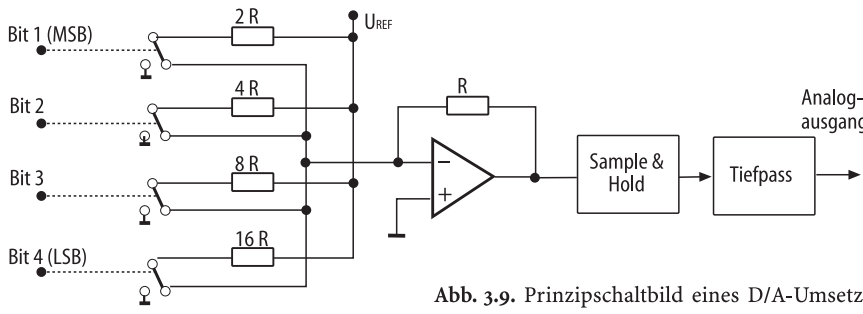


Abb. 3.9. Prinzipschaltbild eines D/A-Umsetzers

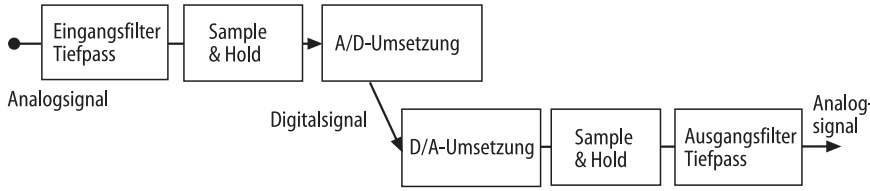


Abb. 3.10. A/D – D/A-Übertragungskette

Der gesamte digitale Übertragungsprozess ist als Blockschaltbild in Abbildung 3.10 dargestellt. Er beinhaltet die Eingangstiefpassfilterung, die Zeitdiskretisierung, die Quantisierung (A/D) und den Übertragungsweg. Die Rückwandlung geschieht mit dem D/A-Umsetzer und dem Rekonstruktionsfilter.

3.2.3 Digitale Signalübertragung im Basisband

Die Digitaldatenübertragung unterliegt den gleichen Bedingungen wie die Analogübertragung, Störungen und Rauschen treten hier wie dort auf. Auch Digitaldaten können kabelgebunden im Basisband oder in Trägerfrequenzsystemen mit elektromagnetischen Wellen übertragen werden. Zunächst wird hier die Übertragung im Basisband betrachtet, das bedeutet, dass die Signale in bestimmter Weise codiert sind, was auch als Modulation bezeichnet wird, dass sie aber nicht auf einen HF-Träger moduliert werden.

Im Basisband liegt zunächst das PCM-Signal (Puls Code Modulation) vor, dessen Code aber oft verändert wird. Codierung bedeutet in diesem Zusammenhang Veränderung des Datenstroms durch Umordnung und Einfügen von Bits derart, dass sich der Originaldatenstrom eindeutig rekonstruieren lässt. Dies geschieht zum einen, um eine Datenreduktion zu erhalten. Die dazu erforderliche Basis ist das PCM-Signal an der Quelle, daher spricht man bezüglich dieser Codierung auch von Quellencodierung (source coding). Ein zweites Motiv für die Veränderung der Daten ist die Anpassung an die besonderen Eigenschaften des Übertragungskanals. Diese können sich erheblich unterscheiden, so sind z. B. die Störabstände und Fehlertypen bei MAZ-Aufzeichnungssystemen ganz andere als bei digitaler Kabelübertragung. Bei der Kanalcodierung ist neben der Bandbreite zu berücksichtigen, welche spektralen Anteile das Signal aufweist, ob Gleichwertanteile auftreten, welche Modulationsarten verwendet werden oder ob eine Taktgewinnung aus dem Signal selbst erforderlich ist. Die Kanalcodierung findet unter Berücksichtigung des Fehlerschutzes statt.

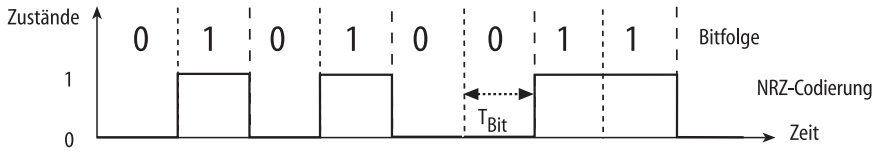


Abb. 3.11. NRZ-Signal

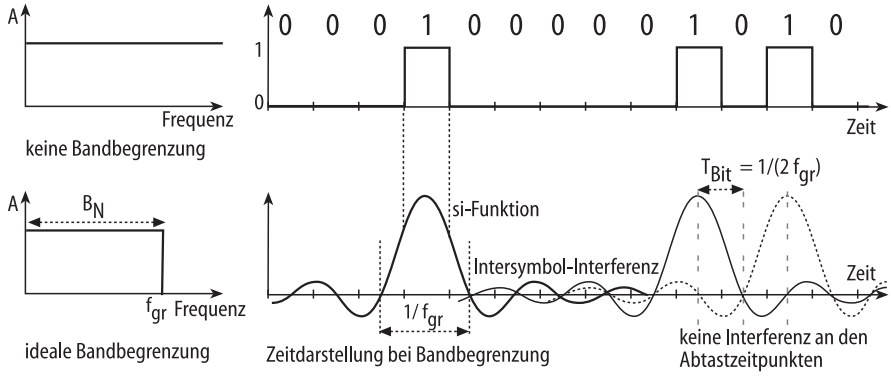


Abb. 3.12. Verformung des NRZ-Signals und Intersymbol-Interferenz durch Bandbegrenzung

Auch der Fehlerschutz selbst ist schließlich ein Motiv für die Umordnung der Daten.

Die Repräsentation der Bits als aufeinanderfolgende Zustände von ein- und ausgeschalteter Spannung wird am einfachsten, wenn das Signal bei aufeinanderfolgenden High-Zuständen nicht zurückgesetzt wird (Abb. 3.11), der Code wird dann als NRZ-Code (Non Return to Zero) bezeichnet. Das NRZ-Signal ist sehr unregelmäßig. Wenn lange Folgen gleicher Bits auftreten, ändert sich das Signal nicht und damit ist es nicht gleichwertfrei. Die höchste Frequenz ergibt sich aus einem ständigen Wechsel der Zustände, der kürzeste Signalzustand entspricht der Bitdauer T_{Bit}. Daraus ergibt sich die Bitrate

$$v_{Bit} = 1 \text{ Bit}/T_{Bit},$$

bzw. die Bittaktfrequenz $f_{Bit} = 1/T_{Bit}$.

Das entstehende Signal muss nicht formgetreu übertragen werden, dafür wäre bei einem Rechteck-Signal auch eine sehr große Bandbreite erforderlich. Zur Rekonstruktion des Signals wird es abgetastet, daher genügt es, dass es an den Abtastzeitpunkten eindeutig erkannt wird, d. h. das Übertragungsband kann begrenzt werden. Eine strenge ideale Bandbegrenzung mit rechteckförmigem Verlauf hängt nach der Fouriertransformation aber wiederum mit einem Signal zusammen, das im Zeitbereich einen Verlauf nach der si-Funktion aufweist (Abb. 3.12). Die zeitliche Verbreiterung führt zu einem Übersprechen zwischen den Nachbarimpulsen, der Intersymbol Interferenz (ISI). Die si-Funktion weist Nullstellen auf, die im Zentrum den Abstand $2 \cdot T_{Bit}$ haben. Wenn die Maxima der Nachbarfunktionen mit diesen Nullstellen zusammenfallen, werden die Symbole gerade nicht verfälscht, sodass hier ein idealer Abtastzeitpunkt für die Rekonstruktion des Signals vorliegt. Dieser Umstand wird durch die Küpfmüller-Nyquist-Bedingung mit der Bandbreite B_N und

$$B_N = 1/2 \cdot 1/T_{Bit} = 1/2 \cdot f_{Bit}$$

ausgedrückt (Abb. 3.12). Das bedeutet: Von dem entstehenden Impulssignal muss mindestens die Grundfrequenz des kürzesten Wechsels übertragen wer-

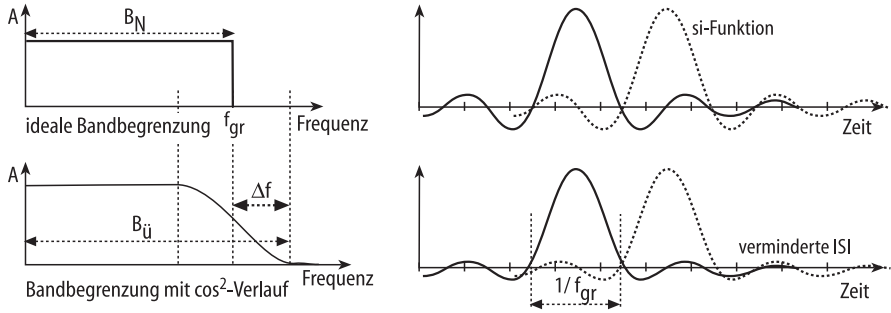


Abb. 3.13. Vergleich der Auswirkung von idealer und realer Bandbegrenzung

den. Für ein digitales Komponentensignal mit 270 Mbit/s (s. Abschn. 3.3.2) folgt daraus, dass eine Minimalbandbreite von 135 MHz erforderlich ist.

Im Realfall ist der Übertragungskanal aber weder streng bandbegrenzt, noch unendlich breit, sondern durch Tiefpassfilter mit flachen Flanken bestimmt. Eine Bandbegrenzung mit $\cos^2$-förmigem Verlauf bewirkt eine Dämpfung der Überschwinger der si-Funktion im Vergleich zur idealen Bandbegrenzung. Die Flankensteilheit der Filter bestimmt damit das Ausmaß der ISI. Die Steilheit des Kurvenverlaufs wird mit dem Roll-off-Faktor r beschrieben:

$$r = \Delta f/B,$$

wobei B der Nyquist-Bandbreite B_N entspricht, wenn als Maß der 50%-Abfall der Flanke herangezogen wird (Abb. 3.13).

Je größer der Roll-off-Faktor, desto größer wird die erforderliche Übertragungsbandbreite $B_Ü$ in Relation zur Nutzbandbreite B, es gilt:

$$B_Ü = (1 + r)\ B.$$

Mit r= 0,5, folgt also z. B. $B_Ü = 0{,}75 \cdot v_{Bit}/Bit$.

Der Roll-off-Faktor bestimmt damit unter anderem das Verhältnis zwischen der Bitrate und der Übertragungsbandbreite. Diese Relation wird auch als Bandbreiteneffizienz oder spektrale Effizienz bezeichnet und als $\varepsilon = v/B_Ü$ mit der Einheit Bit/s pro 1 Hz angegeben.

Je kleiner r, desto steiler ist die Bandbegrenzung und die erforderliche Übertragungsbandbreite sinkt. Daraus resultiert wiederum ein stärkeres Übersprechen. Das Übersprechen bewirkt eine verringerte Fehlertoleranz bei der Rekonstruktion der Signalzustände auf der Empfangsseite und führt zusammen mit den Störungen des Kanals zu Fehlern, die in der Form Bitfehlerzahl pro Sekunde als Bit Error Rate (BER) angegeben werden. Welche Fehlerrate toleriert werden kann, hängt vom Einsatzzweck ab. Wenn bei einem Datenstrom von 28 Mbit/s ein Fehler pro Stunde erlaubt ist, hat die BER den Wert 10^{-11}.

Die Qualität des Übertragungssystems kann visuell mit Hilfe des Augendiagramms beurteilt werden. Zur Gewinnung des Augendiagramms wird ein Oszilloskop auf den Schritttakt der Digitaldaten getriggert. Alle Signalübergänge werden überlagert, und es entsteht ein augenförmiges Bild wie es in Abb. 3.14 dargestellt ist. Je größer die Augenöffnung, desto kleiner sind die Amplituden-

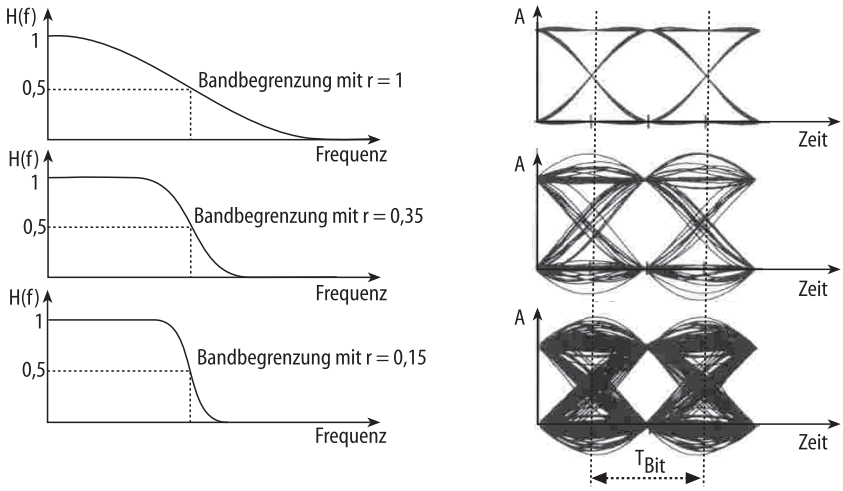


Abb. 3.14. Auswirkung verschiedener Bandbegrenzungen auf das Augendiagramm

und Zeittaktschwankungen (Jitter), desto einfacher sind die Signalzustände rekonstruierbar und desto weniger Fehler treten auf. Solange das Auge nicht geschlossen ist, kann das Signal regeneriert werden.

Abbildung 3.14 zeigt auch die Veränderung des Augendiagramms in Abhängigkeit von r [174]. Es wird deutlich, dass bei steiler Bandbegrenzung das Symbolübersprechen steigt, daher kann mit kleinem r nur dort gearbeitet werden wo der Kanal weitgehend frei von zusätzlichen Störungen ist, wie z. B. bei der kabelgebundenen Übertragung.

Der Einfluss des auf dem Übertragungsweg auftretenden Rauschens kann sehr effektiv gemindert werden, wenn am Empfängereingang ein Filter eingesetzt wird, dessen Impulsantwort mit dem Verlauf des übertragenen Impulses übereinstimmt. Mit dieser signalangepassten Filterung (Matched Filter) wird erreicht, dass der Rauscheinfluss vom mittleren Energiegehalt pro Bit E_b abhängt und nicht mehr von der Signalform über der Zeit. Bei einem redundanz-

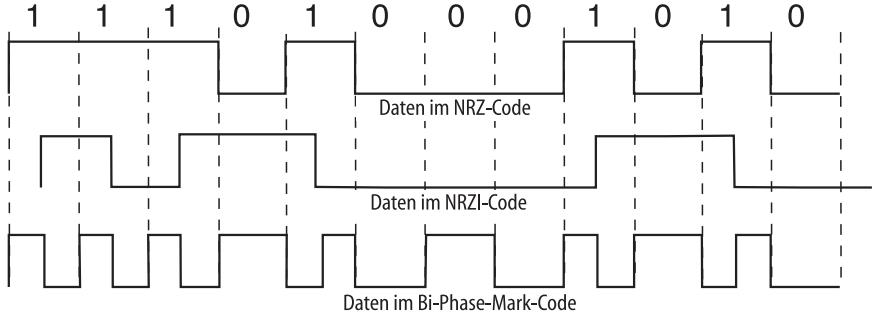


Abb. 3.15 Vergleich von NRZ-, NRZI- und Bi-Phase-Mark-Code

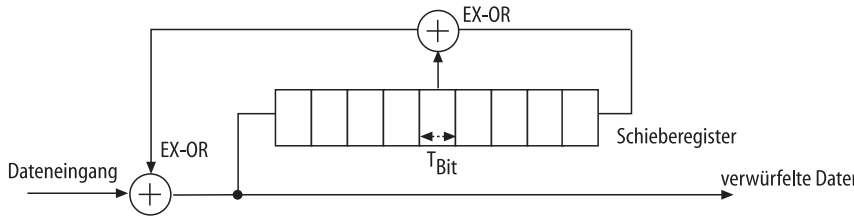


Abb. 3.16. Datenverwürfelung, Scrambling

freien Signal, das zum idealen Zeitpunkt abgetastet wird, gilt die Beziehung

$$S/N = E_b/N.$$

NRZ ist ein sehr simpler Code, aber weder selbsttaktend noch gleichspannungsfrei. Der Code kann zu NRZI (Non Return to Zero Inverse) geändert werden, wenn bei jeder 1 der Signalzustand zwischen 0 und 1 bzw. 1 und 0 umgeschaltet wird (Abb. 3.15). Jede 1 entspricht einem Pegelwechsel und damit wird die Codierung polaritätsunabhängig. Das NRZI-Spektrum ähnelt dem NRZ-Spektrum.

Ein NRZI-Code wird in modifizierter Form bei der grundlegenden digitalen Signalform der Videotechnik, beim seriell-digitalen Videosignal verwendet. Die Schnittstelle für dieses Signal ist so definiert, dass die Taktrate aus dem Datenstrom selbst rekonstruiert werden kann. Weder NRZ noch NRZI sind aber selbsttaktend, denn im Spektrum sind niederfrequente Anteile stark vertreten. NRZI wird daher durch Verwürfelung (Scrambling) zu SNRZI verändert. Die Verwürfelung wird mit einem n-stufigen Schieberegister durchgeführt, durch das der serielle Datenstrom hindurchgeschickt wird (Abb. 3.16). An wenigstens zwei Stellen hat das Register Anschlüsse für eine Signalrückkopplung. Durch eine Exclusiv-Oder-Verknüpfung der Daten an dieser Stelle wird eine Pseudo-Zufallsdatenfolge erzeugt, die ein Signalleistungsspektrum mit ausgeglichener oder gezielt gewichteter Verteilung bewirkt. Die Verteilung wirkt zufällig, beruht aber auf einem festen Bildungsgesetz, das Signal kann damit auf der Empfangsseite eindeutig rekonstruiert werden.

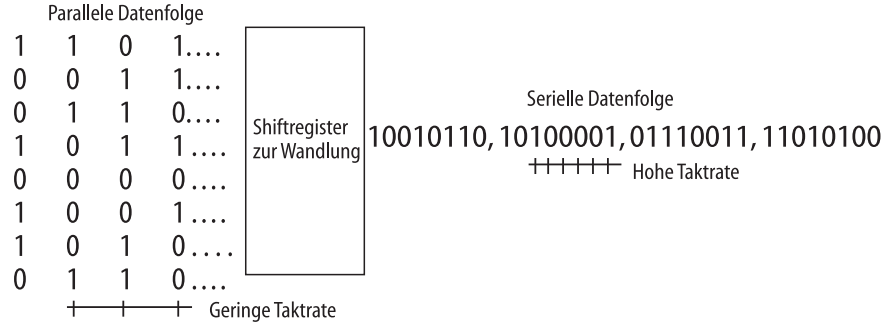


Abb. 3.17 Parallele und serielle Datenübertragung

Ein weiterer selbsttaktender Code ist der Bi-Phase-Mark-Code. Hier findet an jeder Bitzellengrenze ein Pegelwechsel statt, wobei die »1« so codiert wird, dass ein zusätzlicher Pegelwechsel dazwischen liegt (Abb. 3.15). Dieser Code wird benutzt, wenn eher eine einfache Codierung als eine hohe Effizienz erforderlich ist, was in der Videotechnik z. B. bei der Timecode-Übertragung der Fall ist.

Die Datenübertragung geschieht parallel oder seriell. Wurde z. B. mit 8 Bit quantisiert, so werden bei der Parallelübertragung acht Leitungen eingesetzt, sodass im Rhythmus der Abtastpulse der gesamte codierte Wert vorliegt (Abb. 3.17). Für die serielle Übertragung wird dagegen nur eine Leitung benötigt, die Bits werden nacheinander im Zeitmultiplex übertragen. Die serielle Übertragung ist in der Praxis deutlich weniger aufwendig, andererseits ist aber eine viel höhere Übertragungsgeschwindigkeit erforderlich. Da mit wachsendem technischen Fortschritt die Bewältigung hoher Übertragungsraten zunehmend geringere Schwierigkeiten bereitet, wird wegen der praktischen Vorteile meistens die serielle Datenübertragung bevorzugt.

Wie bei Analogsystemen ist die maximale Übertragungskapazität außer von der Bandbreite B auch vom Signal-Rauschabstand S/N abhängig. Die Verwendung von nur zwei Zuständen, wie bei NRZI, ist sehr störunempfindlich aber nicht effektiv. Die bessere Nutzung des großen Signal-Rauschabstands ist durch mehrwertige Modulation, d. h. durch die Zulassung von mehr als zwei Signalzuständen möglich (Abb. 3.18). Jeder Zustand wird durch eine Gruppe von Bits gebildet, die zu einem Abtastzeitpunkt gehören und durch den Begriff Symbol gekennzeichnet werden. Ein Symbol umfasst den Informationsgehalt der einem Bit oder mehreren Bits entspricht, damit ergibt sich die maximale Datenrate aus der Symbolrate multipliziert mit dem Informationsgehalt eines Symbols.

Die Nutzung mehrwertiger Modulation macht das Signal störanfälliger, meist wird damit ein Fehlerschutz erforderlich. Bei der Basisbandübertragung steht bei Nutzung von Koaxialkabeln genügend Bandbreite zur Verfügung, daher wird bei der digitalen Videoübertragung im Basisband auf mehrwertige Modulation und auf einen Fehlerschutz verzichtet, im Gegensatz zur digitalen HF-Übertragung bei DVB (s. Kap. 4.5), bei der die mehrwertige Codierung explizit genutzt wird, um Bandbreite zu sparen.

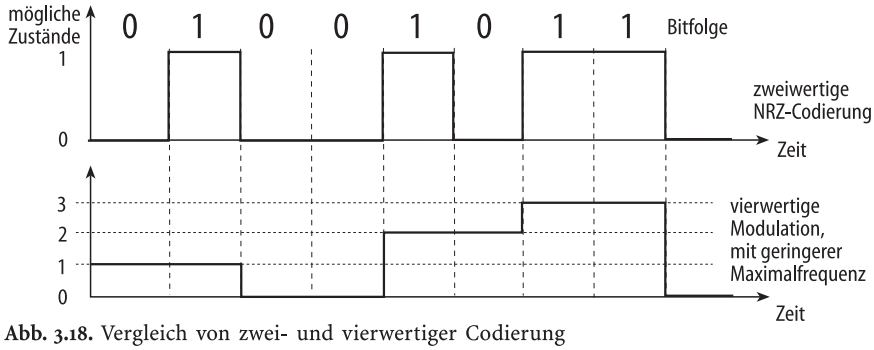


Abb. 3.18. Vergleich von zwei- und vierwertiger Codierung

3.3 Digitale SD-Videosignale

Das analoge Videosignal ist durch die Bildwandlung bereits weitgehend quantisiert. Es ist zeitlich in Einzelbilder und räumlich in Zeilen eingeteilt. Beim Einsatz eines CCD- oder CMOS-Bildwandlers (s. Kap. 6.1) werden darüber hinaus auch diskrete Bildpunkte gebildet. Die darin gebildeten Ladungen liegen aber nicht in einer wertdiskreten Form vor. Das Videosignal muss daher den gewöhnlichen Digitalisierungsprozess durchlaufen. Dabei wird das analoge Signal mit konstanter Rate abgetastet, die vorliegenden Signalwerte werden quantisiert und dann als binär codierte Zahlenwerte übertragen. Entsprechend dem Abtasttheorem ist dabei eine Abtastfrequenz erforderlich, die mindestens doppelt so groß ist wie die höchste Signalfrequenz. Ausgehend von einem SD-Analogsignal mit 5 MHz Bandbreite ergibt sich damit eine Abtastrate von 10 MHz und eine minimale Datenrate von 80 Mbit/s, denn im einfachen Fall ist eine Quantisierungsstufenzahl von 256 ausreichend, die mit 8 Bit dargestellt werden kann.

Mit NRZ-Codierung ist zur Übertragung dieser Datenrate eine Bandbreite von mindestens 40 MHz erforderlich (s. Abschn. 3.2.3). Die Digitalübertragung erfordert schon in diesem Beispiel, verglichen mit dem Analogsignal, die achtfache Bandbreite. Die genannte Datenrate ist aber ein Minimalwert, der sich in der Praxis wesentlich erhöht. Im Produktionsbereich ist das vertretbar, die Digitalübertragung zum Heimempfänger wurde aber erst praktikabel, als effektive Algorithmen zu Reduktion dieser Datenmenge zur Verfügung standen.

Digitale Videosignale beruhen auf den bekannten analogen Signalformen (s. Kap. 2). Das RGB-Signal wird selten verwendet, dies gilt sowohl im analogen als auch im digitalen Umfeld. Kaum größerer Bedeutung hat das auf dem analogen FBAS-Signal beruhende Digital Composite Signal. Dagegen ist das Digital Component Signal die zentrale Form sowohl im Produktionsbereich als auch als Ausgangsformat für die Datenreduktion.

3.3.1 Digital Composite Signal

Das Digital Composite Signal ist das digitalisierte FBAS-Signal. Hier steigt die Datenrate erheblich über 80 Mbit/s, denn die Abtastrate muss ein Vielfaches der Farbhilfsträgerfrequenz betragen, um Aliasfehler möglichst gering zu halten (Abb. 3.19). Es kommen die Faktoren 3 und 4 in Frage. Es wurde der Faktor vier festgelegt, denn insbesondere wegen der PAL-Sequenzen (s. Kap. 2.3) werden die Störungen dann minimal, wenn die Abtastfrequenz $f_T = 4 \cdot f_{sc} =$

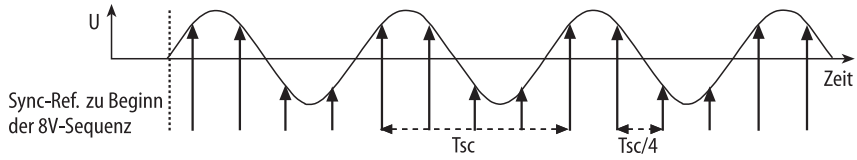


Abb. 3.19. Abtastung der Farbhilfsträgerfrequenz mit 4 f_{sc}

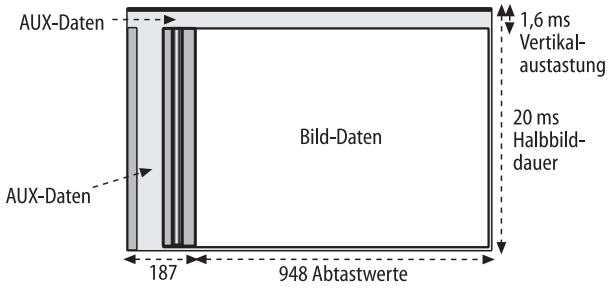


Abb. 3.20. Lücken für Zusatzdaten beim Digital Composite-Format

17,73 MHz (PAL) beträgt. Für NTSC gilt: $f_T = 4 \cdot f_{sc} = 14{,}32$ MH. Die Abtastpunkte haben bei PAL die in Abbildung 3.19 gezeigte Relation zur Farbhilfsträgerschwingung. Auch der Synchronimpuls und der Burst werden digitalisiert, aus jeder Zeile werden bei PAL-Signalen 1135 Abtastwerte gewonnen, davon gehören 948 zur aktiven Zeile. Teile der H- und V-Austastbereiche können für Zusatzdaten (AUX) genutzt werden (Abb. 3.20). Das Digital Composite Signal ist für die Übertragung mit 8 oder 10 Bit definiert, maximal ergibt sich für ein PAL-Signal eine Datenrate von 177 Mbit/s.

Das Digital Composite Signal ist heute im Wesentlichen nur noch im Zusammenhang mit den digitalen MAZ-Formaten D2 und D3 (s. Kap. 8.8) von Bedeutung. D2- und D3-Geräte arbeiten mit 8 Bit-Auflösung, von den 256 möglichen Quantisierungsstufen entfallen 64 auf das Synchronsignal, sodass für die Bildamplitude nur 191 Stufen zur Verfügung stehen (Abb. 3.21). Für das Luminanzsignal allein verringert sich dieser Wert auf 148, woraus ein Signal/Rauschabstand von 54 dB folgt. Mit D2- und D3-MAZ-Geräten wird eine gegenüber der analogen Aufzeichnungstechnik deutlich verbesserte Anzahl möglicher Kopier-Generationen erreicht. Das Digital Composite-Signal ist aber aufgrund der Dynamikreduktion durch Einbeziehung des Synchronsignals als genereller Studiostandard schlecht geeignet, außerdem bleiben auch die Nachteile durch PAL-Artefakte und die FBAS-typische schmalbandige Farbsignalübertragung erhalten.

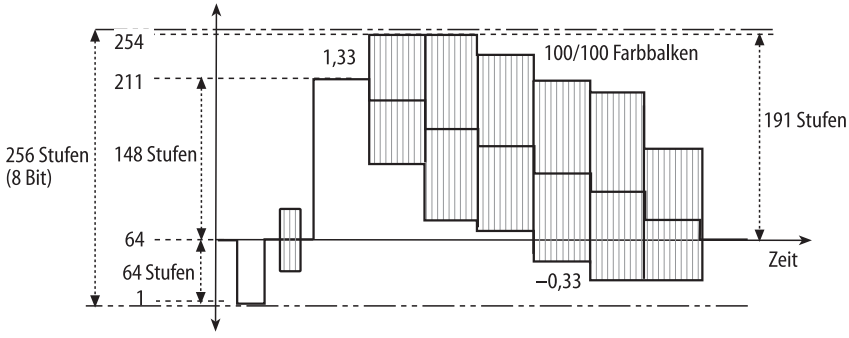


Abb. 3.21. Quantisierungsstufen beim Digital Composite Signal

3.3.2 Digital Component Signal

Die Aufteilung des Farbvideosignals in die drei Komponenten für die Helligkeit und die Farbdifferenzen ermöglicht gegenüber FBAS-Signalen eine wesentlich bessere Signalverarbeitungs- und Übertragungsqualität. Dieses gilt auf analoger Ebene ebenso wie beim Digitalsignal. Die Komponentencodierung bietet den Vorteil, unabhängiger von Fernsehnormen zu sein und ermöglicht eine einfachere Datenreduktion. Weiterhin kann die Abtastrate ohne Rücksicht auf die Farbhilfsträgerfrequenz gewählt werden. Der einzige Nachteil, die dreifache Leitungsführung, wird bei digitalisierten Komponenten irrelevant, da die drei Signalteile einfach per Multiplex zusammengefasst werden können.

Digitale Komponentensignale basieren auf einem analogen Signal, bei dem die Farbdifferenzkomponenten (P_R und P_B) derart pegelreduziert sind, dass sich ohne Berücksichtigung der Synchronimpulse für alle drei Signale Spannungswerte von 0,7 V_{ss} ergeben (s. Kap. 2.3). Wichtigster Punkt für die Normung war die Wahl der Abtastrate. Um zu erreichen, dass die Bildpunkte in jeder Zeile dieselbe Position haben (orthogonale Abtaststruktur) muss die Abtastrate ein Vielfaches der Zeilenfrequenz betragen. 1979 gab es einen Standardisierungsvorschlag für SD-Signale mit der Bezeichnung 12:4:4, d. h. das Luminanzsignal sollte mit 12 MHz und die Farbdifferenzkomponenten mit 4 MHz abgetastet werden. Dieser wurde nicht akzeptiert, weil die in USA und Europa verwendeten Systeme mit 525- bzw. 625-Zeilen auf digitaler Ebene nur noch möglichst wenige Unterschiede aufweisen sollten. Bei 12 MHz Abtastrate für das 625-Zeilensystem ist eine Anpassung an 525 Zeilen schlecht möglich.

3.3.2.1 ITU-R BT. 601

Das orthogonale Abtastraster sollte international einheitlich mit nur einer Abtastrate erreicht werden, d. h., dass sie sowohl im 625- als auch im 525-Zeilensystem, mit den Halbbildfrequenzen 50 Hz und 59,94 Hz, ein ganzzahliges Vielfaches der Zeilenfrequenz betragen muss. Das kleinste gemeinsame, ganzzahlige Vielfache für beide Systeme ist der Wert 2,25 MHz. Ausgehend von einem SD-Analogsignal von 5 MHz Bandbreite würde das Fünffache dieses Wertes, also 11,25 MHz, als Abtastrate genügen. Um den Filteraufwand zur Rückgewinnung

Tabelle 3.1. Spezifikationen bei ITU-R BT. 601

System	525/29,97	625/25
Komponenten	Y / C_R / C_B	Y / C_R / C_B
Abtaststruktur	orthogonal, C_R, C_B zusammen mit den ungeraden Y-Werten	
Abtastfrequenz	13,5 / 6,75 / 6,75 MHz	
Abtastwerte/Zeile	858 / 429 / 429	864 / 432 / 432
Wertanzahl der digitalen aktiven Zeile	720 / 360 / 360	
Wertanzahl der analogen aktiven Zeile	714 / 355 / 355	702 / 350 / 350
Anzahl aktiver Zeilen pro Bild	486	576
Anzahl Bits pro Sample	8 / 10 Bit	
Nutzbare Stufenanzahl bei 8 Bit	220 / 225 / 225	
Nutzbare Stufenanzahl bei 10 Bit	877 / 897 / 897	

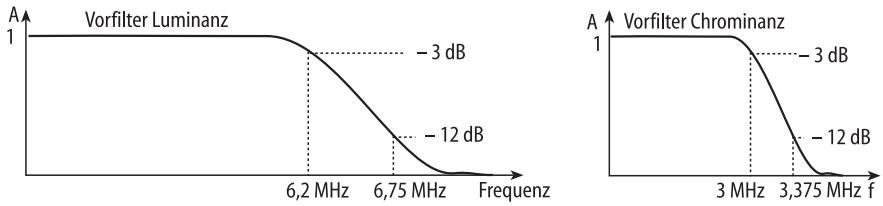


Abb. 3.22. Anti-Aliasing-Filter für das Digital Component Signal

des Basisbandes nicht zu groß zu machen, und um auch Bandbreiten über 5,5 MHz zu ermöglichen, wurde mit dem Standard ITU-R BT.601 international das Sechsfache von 2,25 MHz, also eine Abtastrate von 13,5 MHz, für das Luminanzsignal festgelegt (Tabelle 3.1). Die Recommendation 601 gilt für SD-Signale, die Variante für HD-Komponenten hat die Bezeichnung 709.

Aus der Abtastfrequenz 13,5 MHz ergeben sich im 625-Zeilensystem (50 Hz) 864 Abtastwerte und im 525-Zeilensystem (59,94 Hz) 858 Abtastwerte pro Zeile für das Luminanzsignal. Für Europa gilt: 864 · 15,625 kHz = 13,5 MHz und für Amerika: 858 · 15,734265 kHz = 13,5 MHz.

Für die Farbdifferenzsignale C_R und C_B sind die Werte halbiert, die Abtastfrequenz hat hier den Wert 6,75 MHz, da die Chrominanzanteile wie beim Analogsignal mit geringerer Bandbreite übertragen werden. Die erlaubte Bandbreite für die analogen Eingangssignale beträgt beim Luminanzsignal 5,75 MHz (±0,1 dB), bzw. 6,2 MHz bei –3 dB und für C_R und C_B 2,75 MHz, bzw. 3 MHz bei –3 dB (Abb. 3.22). Bei der halben Abtastfrequenz (6,75 bzw. 3,375 MHz) müssen die Tiefpassfilter mehr als 12 dB dämpfen.

In allen Fernsehsystemen ist der für das Videosignal nutzbare Zeilenbereich, die aktive Zeile, durch die horizontale Austastlücke begrenzt. Mit ITU-R BT.601 wurde für die europäische und die US-Norm gleichermaßen festgelegt, dass die aktive Zeile 720 Luminanz-Abtastwerte enthält, sowie je 360 für C_R und C_B. Diese Zahlen sind durch 8 teilbar und damit günstig im Hinblick auf die Datenreduktion, die oft eine Blockbildung mit 8 x 8 Pixeln erfordert. Insgesamt befinden sich in der aktiven Videozeile 1440 Datenworte (Abb. 3.23). Aus 720 dividiert durch 13,5 MHz ergeben sich 53,33 µs, die digitale aktive Zeile ist somit länger als die analoge aktive Zeile, deren Dauer im 625-Zeilensystem 52 µs beträgt. Da die aktive Zeile im 525-Zeilensystem wiederum etwas länger ist als 52 µs, sind die in die Austastlücke fallenden Abtastwerte in beiden Normen verschieden. Das bedeutet, dass bei der Digitalisierung der analogen aktiven Zeile in Europa nur 702 und in den USA nur 718 Bildpunkte von 720 nutzbar sind. Der Rest gehört zur analogen H-Austastlücke. Diese Bildpunkte erscheinen oft als schmale schwarze Ränder bei der Darstellung von digitalen Videobildern auf dem PC. Die vertikale Austastlücke ist für das 625-Zeilensystem so festgelegt, dass pro Halbbild 288 aktive Zeilen zur Verfügung stehen.

Die 720 x 576 aktiven Bildpunkte werden gelegentlich auch als Non Square Pixels bezeichnet. In Europa sind sie breiter als hoch, denn für ein Seitenverhältnis 1 : 1 d. h. gleiche Horizontal- wie Vertikalauflösung, müssten in der aktiven Zeile 576 x 4/3 = 768 Bildpunkte zur Verfügung stehen. Das Seitenver-

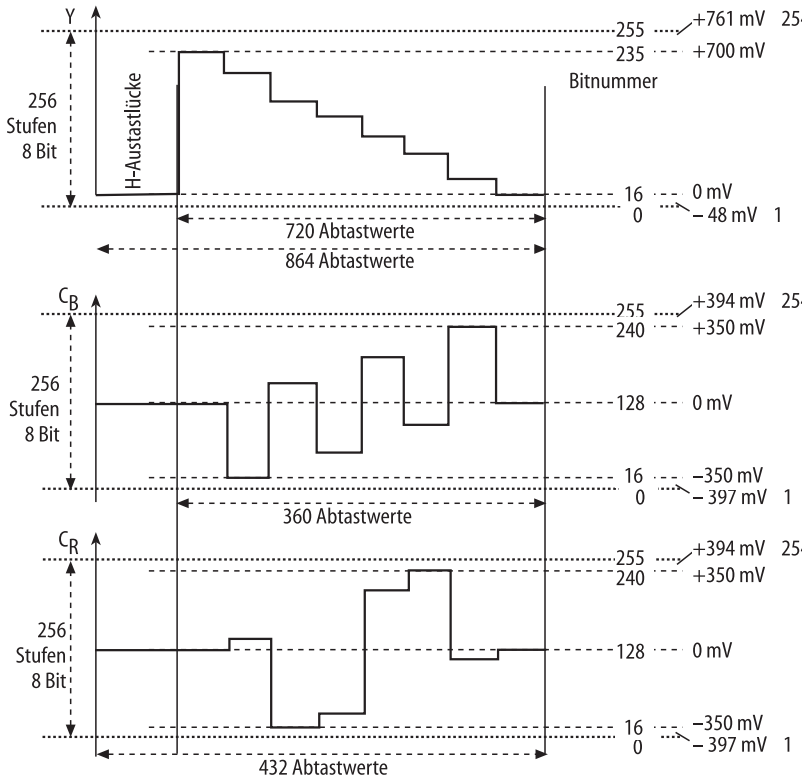


Abb. 3.23. Abtast- und Quantisierungswerte nach ITU-R BT. 601

hältnis des Bildpunktes (Pixel Aspect Ratio) beträgt x : y = 1 : 720/768 = 1,0667. Dieser Umstand führt manchmal zu Problemen, wenn Grafiken für Videosysteme mit PC-Programmen, wie z. B. Adobe Photoshop, erstellt werden, die mit Square Pixels, also dem Aspect Ratio 1:1 arbeiten. In diesem Falle benutzt man entweder eine Dokumentvorlage von 768 x 576 Pixeln und lässt das Bild vor der Ausgabe auf eine Breite 720 Pixel umrechnen, oder man wählt für die Vorlage 720 x 576 Pixel und berücksichtigt, dass die Darstellung auf einem Videomonitor horizontal gedehnter erscheinen wird als am PC-Monitor.

Die horizontale Dehnung auf dem Videomonitor wird noch um den Faktor 4/3 extremer, wenn statt des Bildseitenverhältnisses 4/3 das Verhältnis 16/9 benutzt wird. Eigentlich müsste in diesem Fall die Abtastfrequenz erhöht werden, auf acht mal 2,25 MHz = 18 MHz, einige technische Systeme erlauben das auch, ebenso wie ITU-R BT.601. Es ist jedoch ausreichend, auch für 16/9-Systeme die 720 Bildpunkte aus der Abtastung mit 13,5 MHz beizubehalten und die Bildpunktverbreiterung auf x : y = 4/3 : 720/768 = 1,42 akzeptieren, mit dem Argument, dass diese auf dem Niveau des analogen Videosignals liegt. Leitet man aus der Grenzfrequenz von 5 MHz ein Pixel Aspect Ratio für analoge Signale ab, so ergeben sich aus den 520 aktiven Bildpunkten: x : y = 1 : 520/768 = 1,48.

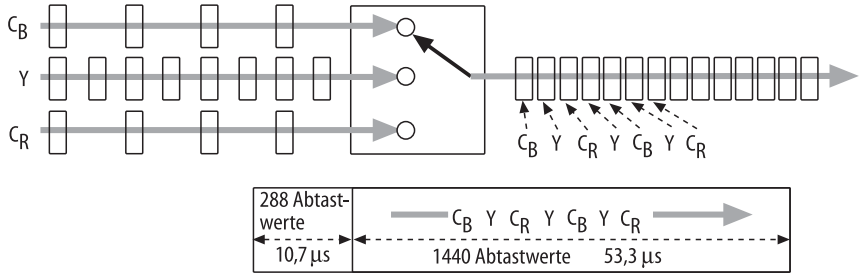


Abb. 3.24. Multiplexbildung der Luminanz- und Chrominanzkomponenten

Um die Digitaldaten der Komponenten in einen Gesamtdatenstrom zu integrieren, werden sie wortweise im Zeitmultiplex in der Folge [C_B Y_1 C_R] Y_2 [C_B Y_3 C_R] Y_4 ... übertragen (Abb. 3.24). Die Werte für die Chrominanz schließen dabei jeweils die zugehörigen ungeradzahligen Luminanzwerte ein. Der Datenstrom umfasst beim 625-Zeilensystem insgesamt 1728 Abtastwerte pro Zeile.

Die Quantisierung der Signalabtastwerte wurde im Jahre 1986 in der ursprünglichen Fassung von CCIR 601 für alle drei Komponenten mit 8 Bit festgelegt und später auf 10 Bit erweitert. Die 8 oder 10 Bit sind jedoch nicht vollständig als Graustufenanzahl nutzbar, denn es sind Über- und Untersteuerungsreserven (Head- und Feetroom) vorgesehen, um auch die Pegelüberschreitungen erfassen, die beim analoge Signal erlaubt sind und dort ca. 7 % betragen. Beim Luminanzsignal mit 8 Bit-Quantisierung ist daher der Digitalwert 16 für Schwarz und der Wert 235 für Weiß definiert, dazwischen sind 220 Graustufen verfügbar. Die zugehörigen Farbdifferenzwerte können Analogwerte zwischen + 350 mV und – 350 mV annehmen, im Digitalbereich sind 225 Quantisierungsstufen (16 ... 240) reserviert. Der Analogwert 0 V entspricht dem Digitalwert 128 (Abb. 3.23).

Im 10-Bit-System sind die genannten Stufenzahlen vervierfacht, sodass ein Übergang von 10 auf 8 Bit einfach durch Weglassen der beiden minderwertigsten Bits (LSB) erfolgen kann. Das simple Abschneiden der Bits kann allerdings zum Contouring führen, d. h., dass bei kleinen Grauwerten die Abstufungen sichtbar werden. Dieses kann durch sogenanntes Dynamic Rounding oder mit Hilfe des Zusatzes eines statistischen Signals (Dither) vermindert werden.

Im 10-Bit-System liegt das Luminanzsignal zwischen den Werten 64 (0 V) und 940 (700 mV), die Farbdifferenzwerte liegen zwischen 512 (0 V) und 64 (– 350 mV) bzw. 960 (+ 350 mV). Zur Berechnung der Luminanz-Digitalwerte bezüglich 8 Bit (D_{Y8}) oder 10 Bit (D_{Y10}) aus den Analogwerten gelten die Beziehungen:

$$D_{Y8} = \{U_Y/0{,}7V \cdot 219\} + 16$$

$$D_{Y10} = \{U_Y/0{,}7V \cdot 876\} + 64$$

Die Klammern stehen dabei für die Rundung auf ganze Zahlen. Für die Digitalwerte der Farbdifferenzkomponenten C_R und C_B gilt:

$$D_{C8} = \{U_C/0{,}7V \cdot 224\} + 128$$

$$D_{C10} = \{U_C/0{,}7V \cdot 896\} + 512$$

Tabelle 3.2. Datenbeispiel für das Digital Component Signal

Übergang von Weiß zu Gelb im Farbbalken (Zeile 58, 10 Bit)							
Nr. ab SAV		Hex	Dec	Nr. ab SAV		Hex	Dec
185	Y	3AC	940	186	R-Y	200	512
187	Y	3AC	940	188	B-Y	1FA	506
189	Y	3AC	940	190	R-Y	201	513
191	Y	3AC	940	192	B-Y	1C4	452
193	Y	3A7	935	194	R-Y	20A	522
195	Y	378	888	196	B-Y	158	344
197	Y	319	793	198	R-Y	21B	539
199	Y	2BA	698	200	B-Y	0EC	236
201	Y	28C	652	202	R-Y	22D	557
203	Y	286	646	204	B-Y	0B6	182
205	Y	286	646	206	R-Y	236	566
207	Y	286	646	208	B-Y	0B0	176

Tabelle 3.2 zeigt als Beispiel die nach einer 10-Bit-Digitalisierung gemessene Zahlenfolge anhand eines 75%-Farbbalkens an der Grenze des Weißbalkens in Zeile 58. Der Farbbalken enthält acht Bereiche, die sich auf die 702 Y-Abtastwerte der aktiven analogen Zeile aufteilen, sodass jeder Balken ca. 88 Bildpunkte bzw. 176 Datenworte umfasst. Im Bereich der Zählwerte 193 bis 203 liegt der Übergang vom Weiß- zum Gelbbalken. Der Luminanzwert für Gelb (Y-Anteil $\approx 0{,}886 \cdot 0{,}75$) hat nach dem Übergang den Wert

$$D_{Y10} = 646 \approx 0{,}886 \cdot 0{,}75 \cdot 876 + 64.$$

Für C_R gilt:

$$D_{C10R} = 566 \approx 0{,}713 \cdot (1 - 0{,}886) \cdot 0{,}75 \cdot 896 + 512$$

und für C_B:

$$D_{C10B} = 176 \approx 0{,}564 \cdot (0 - 0{,}886) \cdot 0{,}75 \cdot 896 + 512.$$

Ein Vollbild mit 720 H x 576 V aktiven Bildpunkten erfordert einen Speicherbedarf von 1440 x 576 = 829 440 Byte bzw. 6,635 MBit bei Quantisierung mit 8 Bit, oder 8,29 MBit bei 10 Bit-Quantisierung. Das aktive Bild belegt also einen Speicherplatz in der Größenordnung 1 MByte. Da im Datenstrom nicht nur die aktiven Bildwerte, sondern auch die Werte der Austastlücke übertragen werden, ergibt sich bei 8 Bit ein Datenstrom von 8 Bit $\cdot$ (13,5 + 2 $\cdot$ 6,75) MHz = 216 Mbit/s, bei 10 Bit hat er den Wert 270 Mbit/s.

Wegen des 2:1-Verhältnisses der Abtastraten zwischen den Komponenten Y und C_R, C_B wird für ITU-R BT.601 häufig die Bezeichnung 4:2:2 verwandt. Mit 4:2:2 sind beim SD-Signal also die Abtastraten als Vielfaches der Frequenz 3,375 MHz gekennzeichnet. Abweichend werden auch andere Abtaststrukturen benutzt: 4:4:4 beschreibt die Abtastung der RGB-Komponenten mit jeweils 13,5 MHz (erhöhte Qualität). Die Bezeichnung 4:2:2:4 wird für ein Standardkomponentenformat verwandt, das um einen vierten Kanal erweitert ist, der für Stanzsignale benutzt wird und die Transparenz des Bildes angibt (Alphakanal).

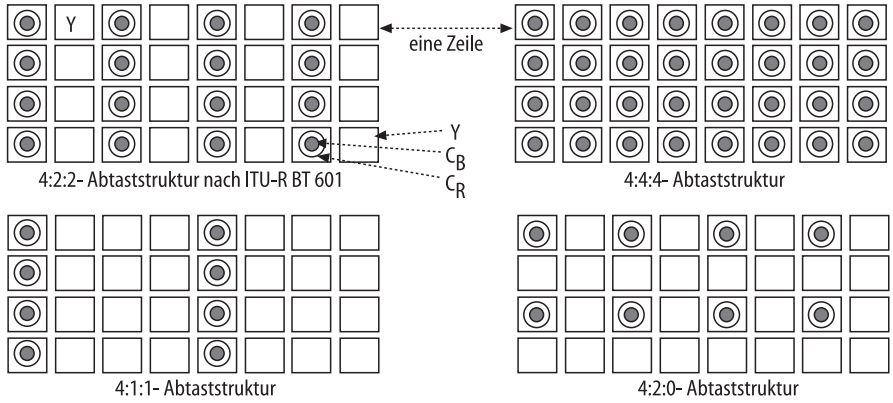


Abb. 3.25. Vergleich der Abtaststrukturen

Die Varianten 4:1:1 und 4:2:0 werden im Zusammenhang mit Datenreduktionsverfahren eingesetzt. Bei 4:1:1 werden die SD-Chrominanzkomponenten nur mit 3,375 MHz (unter-)abgetastet, die mögliche Chrominanzsignal-Bandbreite und damit die Horizontalauflösung ist gegenüber ITU-R BT.601 halbiert.

Die Bezeichnung 4:2:0 bezieht sich auf ein Format, bei dem die vertikale Farbauflösung reduziert ist, da die Farbdifferenzwerte nur in jeder zweiten Zeile vorliegen, während die Horizontalauflösung der Chrominanz die gleiche bleibt wie bei 4:2:2 (Abb 3.25). Beide Einschränkungen der Chrominanzauflösung sind für die Wiedergabe beim Konsumenten kein Problem, denn die Chrominanzbandbreite liegt bei 4:1:1 mit ca. 1,5 MHz immer noch über der des analogen PAL-Systems. Problematisch ist die Unterabtastung aber im Produktionsbereich. Insbesondere für das Chroma Key-Verfahren (s. Kap. 9.2) ist die Chrominanzbandbreite zu gering und die Anordnung der Chrominanzwerte ist für Rechenprozesse bei der Bildbearbeitung ungünstig.

3.3.2.2 Synchronsignale und Zeitreferenz

Nach ITU-R BT.601 wird das vollständige Komponentensignal einschließlich der Austastlücken abgetastet, allerdings wird der Synchronimpuls nicht in die Quantisierung einbezogen, da sonst der Rauschabstand unnötig verschlechtert würde. Der Synchronimpuls trägt auch nur eine simple Information: nämlich die über das Ende der Zeile. Diese Information kann mit nur einem Bit codiert werden. Ein zweites Bit ist erforderlich, um die horizontale von der vertikalen Synchroninformation zu unterscheiden. Diese beiden Bits sind, ebenso wie die analogen Sync-Signale von allergrößter Bedeutung, da von ihnen der gesamte Bildaufbau abhängt. Deshalb werden sie fehlergeschützt und umgeben von besonders gut erkennbaren Bitkombinationen als Timing Reference Signal TRS übertragen. Das TRS tritt zweimal auf, der Anfang und das Ende der aktiven Zeilen werden mit je vier Datenworten (Bytes) gekennzeichnet und in der H-Austastlücke untergebracht. Diese Zeitreferenzen werden als SAV und EAV, Start und End of Active Video, bezeichnet. Die Zeitreferenz stellt den Bezug zum analogen H-Sync und damit für die Signalmischung und -umschaltung her.

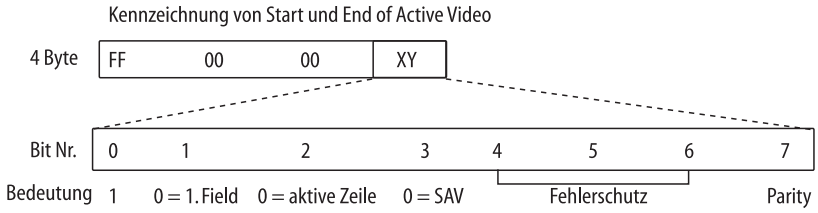


Abb. 3.26. Bedeutung des Synchronwortes

Die gute Erkennbarkeit des Timing Reference Signals wird erreicht, indem im ersten Datenwort alle 8 oder 10 Bit gleich 1 gesetzt werden und in den nächsten beiden Worten gleich Null. So ergeben sich die festen Hexadezimalwerte FF, 00, 00, bzw. 3FF, 000, 000 bei 10 Bit-breiten Datenworten. Diese kommen nicht im normalen Datenstrom vor, da nur eine Wertebereich von 1 bis 254 bzw. 1023 für das Signal genutzt wird. Die Synchron-Information steckt im vierten, mit XY bezeichneten Rahmensynchronwort. In diesem Byte ist die Kennzeichnung der vertikalen Austastlücke enthalten, weiterhin eine Halbbildkennung, die EAV/SAV-Kennung und Fehlerschutzbits. Der Inhalt ist aus Abb. 3.26 ersichtlich. Das erste Bit ist immer gleich 1, das zweite gleich 0 im ersten und gleich 1 im zweiten Halbbild. Das dritte Bit kennzeichnet die aktiven Bildzeile mit einer Null und die Vertikalaustastung mit einer 1 und das vierte Bit ist 0 bei SAV und 1 bei EAV. Die vier weiteren Bit bilden eine Prüfsumme zum Fehlerschutz. Das Rahmensynchronwort wird bei der 10-Bit-Übertragung um 2 Nullen erweitert. Tabelle 3.3 zeigt als Beispiel die Zahlwerte aus einer 10-Bit-Digitalisierung im Bereich EAV und SAV in Zeile 3.

Die Zählung der Datenworte umfasst für digitale SD-Signale abwechselnd immer einen Chrominanz- und eine Luminanzwert. Die Zählung geschieht Wortweise (0 ... 1727) und beginnt nach der SAV-Information (Abb. 3.27). Der Wert C_0 gehört zur Blaudifferenzkomponente, der Wert Y_0 zur Luminanz des ersten Bildpunktes der Zeile. Die Zählung geht dann über die 1440 aktiven Bildpunkte bis zu Y_{1439}. Bei C_{1440} beginnt dann die EAV-Kennzeichnung. Die

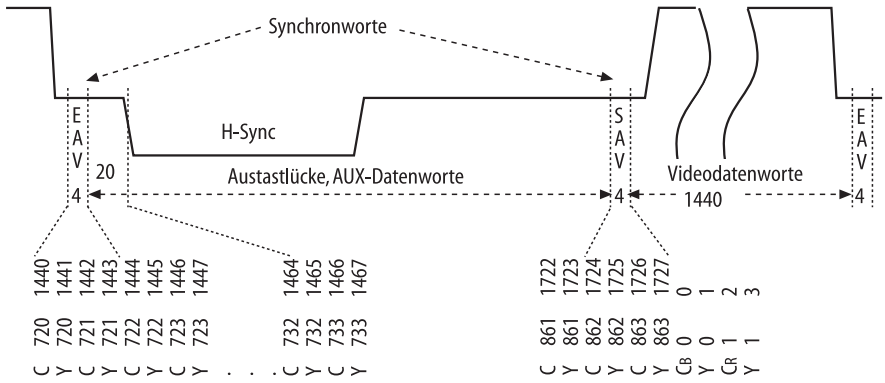


Abb. 3.27. Zeitbezug zwischen analogem und digitalem Videosignal

Tabelle 3.3. EAV- und SAV-Beispieldaten

Field 1, 10 Bit, Line 3				1723	Y	40	
1440	(B–Y)	3FF	EAV	1724	(B–Y)	3FF	SAV
1441	(Y)	000		1725	(Y)	000	
1441	(R–Y)	000		1726	(R–Y)	000	
1443	(Y)	2D8	10 1101 1000	1727	(Y)	2AC	10 1010 1100
1444	B–Y	200	512	0	B–Y	200	512
1445	Y	40	64	1	Y	40	64

SAV-Kennung umfasst die Werte C_{1724} ... Y_{1727} und anschließend beginnt die nächste Zeile. Der Bezug zum analogen H-Sync-Signal ist so gewählt, dass das erste Wort des EAV-Signals (C_{1440}) 24 Worte bzw. 12 Abtastperioden vor der Mitte der fallenden Flanke des H-Sync-Impulses liegt. Der V-Sync ergibt sich aus der Kennzeichnung der Zeilen der V-Austastung.

3.3.3 Schnittstellen

Ein digitales SD-Komponentensignal kann mit einer Datenwortrate von 27 MHz bitparallel übertragen werden, oder die Bits jedes Datenwortes werden seriell, also zeitlich nacheinander mit 270 Mbit/s auf nur einer Leitung gesendet. Für dieses und das digitale Compositesignal gibt es eigens definierte Schnittstellen.

3.3.3.1 Die Parallelschnittstelle

Zur Parallelübertragung wird jedem Bit des Datenwortes ein Leitungspaar zugeordnet, zusätzlich ist eine Taktleitung erforderlich (Abb. 3.28). Die Taktrate ist gleich der Datenwortrate (27 MHz bzw. 17,7 MHz beim PAL-Compositesignal). Es wird der NRZ-Code benutzt, eine Umordnung oder Erweiterung der Daten wird nicht vorgenommen. Die Datenworte können 8 Bits oder 10 Bits umfassen, für eine 8-Bit-Übertragung werden nur die acht höchstwertigen Bits (Most Significant Bits, MSB) verwendet. An die Übertragungsleitungen werden hohe Anforderungen gestellt. Zur Störimpulsreduzierung sind die Leitungen symmetrisch mit einer Impedanz von 110 Ω ausgeführt (Twisted Pair), Störimpulse addieren sich gegenphasig und löschen sich damit weitgehend aus. Für die Parallelübertragung von 10 Bit sind damit 11 Leitungspaare und eine Masseleitung, insgesamt also 23 Leitungen, erforderlich. Als Steckverbinder dient der häufig in der Computertechnik verwandte 25-polige Sub-D-Stecker. Alle Geräte tragen Buchsen, die Kabel sind mit Steckern versehen, Abb. 3.29 zeigt die Belegung.

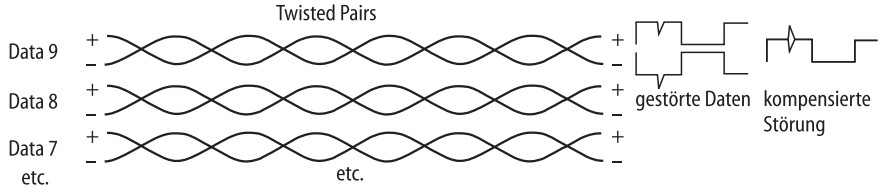
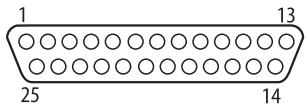


Abb. 3.28. Symmetrische Leitungsführung bei der Parallelschnittstelle



1	Takt +	14	Takt –
2	System Masse	15,13	System Masse
3	Daten 9 +	16	Daten 9 –
.	.	.	.
.	.	.	.
11	Daten 1 +	24	Daten 1 –
12	Daten 0 +	25	Daten 0 –

Abb. 3.29. Steckerbelegung bei der Parallelschnittstelle

Die zur Parallelübertragung verwendeten vieladrigen Kabel (Multicore) sind unflexibel und teuer. Trotz symmetrischer Signalführungen sind die nutzbaren Leitungslängen beschränkt. Ein weiterer Nachteil ist, dass nur einfache Verbindungen, wie z. B. zwischen MAZ-Player und Recorder, vorgenommen werden können. Komplexe Verteilungen, Kreuzschienen etc. sind ökonomisch nicht realisierbar. Heute werden daher serielle Schnittstellen eingesetzt.

3.3.3.2 Die serielle Schnittstelle

Mit dem seriellen digitalen Interface (SDI), das nach ITU-R BT.656 standardisiert ist, ist eine digitale Datenübertragung über herkömmliche Koaxialleitungen möglich, was den Vorteil hat, dass eine bestehende Studioverkabelung, die für Analogsignale vorhanden ist, weiter genutzt werden kann. Es eignet sich gleichermaßen für Digital Composite- und Digital Component-Signale.

Die Umsetzung eines parallel vorliegenden Digitalsignals in ein serielles Signal ist technisch einfach mit Hilfe von Schieberegistern realisierbar (Abb. 3.30). Wenn z. B. Komponentensignale vorliegen, so werden sie mit einer Rate von 27 MHz in das Schieberegister eingelesen. Die Ausgabe aus dem Register erfolgt mit der 10-fachen Taktrate, das LSB kommt dabei zuerst. Auf diese Weise entsteht das seriell digitale Komponentensignal (DSK oder DSC), welches das Standardsignal bei der digitalen Produktion ist.

Bei SDI beträgt die Datenrate in der Praxis stets 270 Mbit/s, bei der Übertragung von 8-Bit-Signalen werden zwei Nullen angehängt und der Datenstrom auf 10 Bit erweitert. Aufgrund der NRZI-Codierung ist für die Übertragung eine Bandbreite von mindestens 135 MHz erforderlich. Die Spannung hat an der Quelle den Wert 0,8 V_{ss}. Zur Übertragung wird die Standardkoaxialleitung mit 75 Ω Wellenwiderstand verwendet, derartige Leitungen haben bei 135 MHz

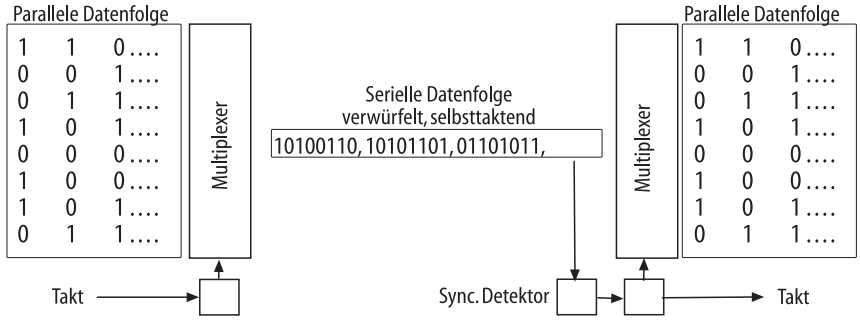


Abb. 3.30. Parallel-Seriell-Wandlung

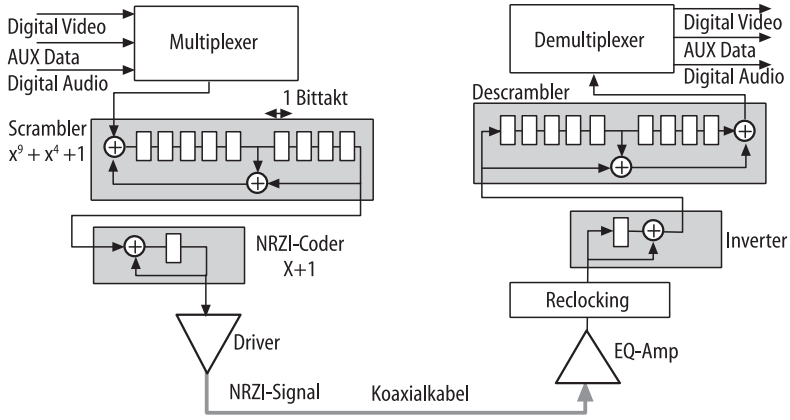


Abb. 3.31. Serial Digital Interface, SDI

eine Dämpfung von weniger als 20 dB pro 100 m Leitungslänge. Ausgehend von einem möglichen Dämpfungsausgleich von ca. 40 dB ergibt sich ohne weitere Einrichtungen eine Leitungslänge von mehr als 250 m, ein Wert, der für den größten Teil der Studioverkabelung ausreichend ist. Größere Entfernungen können mit Taktrückgewinnungseinrichtungen überbrückt werden.

Entscheidend für das Schnittstellenformat ist die Codierung. Diese soll derart vorgenommen werden, dass das Signal selbsttaktend und gleichspannungsfrei wird. Eine frühe Schnittstellendefinition sah dazu eine Erweiterung der 8 Bit-Datenworte auf 9 Bit vor. Beim SDI-Datenformat ist eine derartige Einführung von Zusatzbits unnötig, denn als Kanalcode wird SNRZI, ein verwürfelter NRZI-Code (Scrambled Non Return to Zero, s. Abschn. 4.2.3) benutzt. Der Datenstrom x wird mit Hilfe eines Schieberegisters codiert (Abb. 3.31). Mit der Codiervorschrift $x^9 + x^4 + 1$ wird zunächst das verwürfelte NRZ-Signal erzeugt, anschließend mit x +1 die Polaritätsfreiheit von NRZI generiert. So wird erreicht, dass die Frequenzen relativ klein bleiben und keine separate Taktleitung erforderlich ist. Der Takt kann aus dem Datenstrom abgeleitet werden, denn er ist vom Betrag her bekannt und ein Oszillator im Empfänger kann einfach synchronisiert werden, da das Srambling lange Null-Folgen verhindert. Die SDI-Schnittstelle ist eigentlich sehr spezialisiert, sie hat aber im Videobereich eine weite Verbreitung gefunden und wird heute, z. B. bei SDTI und ASI (s. Kap. 10) auch zum Transport von komprimierten Videodaten benutzt.

3.3.3.3 Auxiliary Data

Bei der Einführung der analogen Videonorm waren ein großer Horizontalsynchronimpuls und eine große Austastlücke erforderlich, um eine sichere Synchronisation zu gewährleisten. In der 625-Zeilennorm umfasst die Horizontalaustastlücke nach ITU-R BT.601 für das Luminanzsignal 864–720 = 144 Datenworte. Da alle drei Komponenten nacheinander übertragen werden, ergeben sich hier zusammen 288 Datenworte, von denen nur wenige erforderlich sind, um als Synchronworte den Zeilenbeginn zu kennzeichnen. Anstatt die

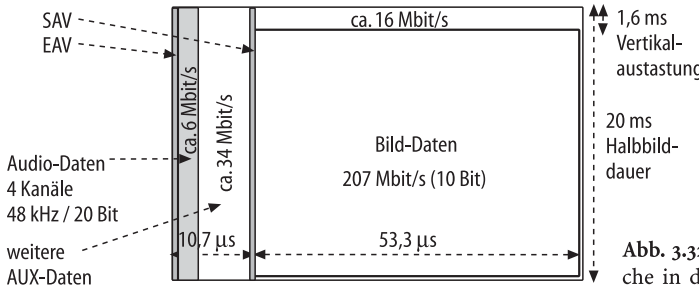


Abb. 3.32. Zusatzdatenbereiche in den Austastlücken

Daten in den Austastlücken einfach wegzulassen und damit die Datenrate zu reduzieren, werden, um eine möglichst einfache A/D-Umsetzung zu erzielen, auch die Austastlücken abgetastet. Die darin befindlichen Datenworte werden als Ancillary Data (ANC) bezeichnet und können zur Übertragung von Zusatzinformationen, wie z. B. digitalen Audiosignalen, benutzt werden (Abb. 3.32). Für diese Zusatzdaten (Auxiliary Data) stehen beim Komponentensignal in der horizontalen Austastlücke 256 Datenworte pro Zeile zur Verfügung, mit 625 Zeilen pro Bild folgt daraus eine Datenrate von ca. 40 Mbit/s. Durch Nutzung der vertikalen Austastlücke ergibt sich eine Datenrate von mehr als 16 Mbit/s.

Die Nutzung der ANC-Datenbereiche wird mit einer Präambel, bestehend aus den Datenworten 000, 3FF, 3FF gekennzeichnet, die, ähnlich wie das TRS, im Datenstrom leicht erkennbar ist (Abb. 3.33). Im darauf folgenden Data ID-Byte definieren die untersten acht Bit 256 verschiedene Datenarten, z. B. 00 für frei definierte Daten und FF für Audiodaten. Die Data Block-Nr. dient zur

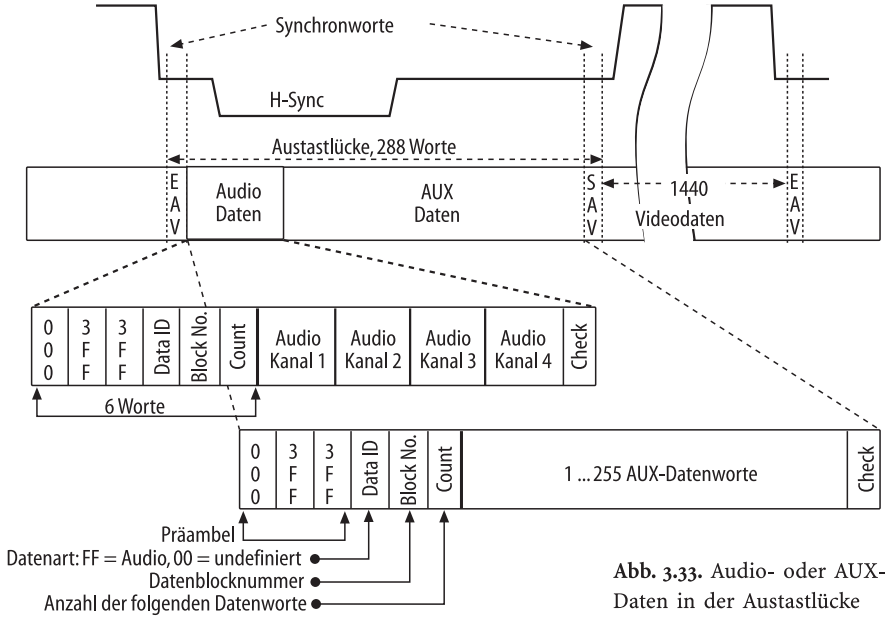


Abb. 3.33. Audio- oder AUX-Daten in der Austastlücke

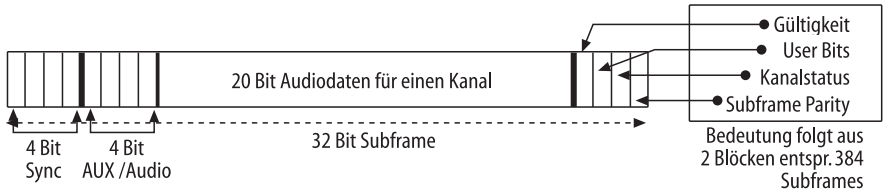


Abb. 3.34. Audio-Datenformat nach AES/EBU

Zählung zusammenhängender Daten, und im Data Count-Wort wird die Anzahl der im laufenden Datenblock enthaltenen Daten angegeben. Im Nutzdatenbereich liegen die Audio- oder Aux-Daten, dieser User-Data-Bereich umfasst maximal 255 Worte. Abschließend wird in der Check Sum die Kontrollsumme aus allen Worten ab Data-ID angegeben [80].

Die gesamte Hardware für die serielle Übertragung beinhaltet somit zunächst ggf. einen Multiplexer, der in der Austastlücke Aux- und Audio-Daten in den Videodatenstrom integriert. Es folgen die Stufen zur Parallel/Seriell-Wandlung, und zur SNRZI-Codierung. Die Parallel/Seriell-Wandlung geschieht mit einfachen Schieberegistern, das Least Significant Bit (LSB) wird zuerst übertragen. Auf der Empfangsseite ist zunächst ein Entzerrer vorgesehen, der die frequenzabhängige Leitungsdämpfung ausgleicht. Es folgen eine Reclocking-Stufe zur Stabilisierung der Datenrate, die Decodierung und Umformung in ein NRZ-Signal und schließlich der Demultiplexer (Abb. 3.31).

3.3.3.4 Embedded Audio

Embedded Audio bezeichnet die Übertragung digitaler Audiodaten im Ancillary-Bereich des Signals an der SDI-Schnittstelle. Die Übertragung von Bild- und Tonsignalen ist damit auf einer einzigen Leitung möglich, die Audiodaten müssen nur zur richtigen Zeit in den Videodatenstrom eingetastet werden. Bei der Übertragung über große Strecken lohnt sich dieser Aufwand, innerhalb des Studios wird das Audiosignal auch separat übertragen. Embedded Audio beinhaltet vier oder mehr Audiokanäle, die paarweise nach AES/EBU codiert sind und pro Paar eine Datenrate von ca. 3 Mbit/s erfordern. Das AES/EBU-Format beschreibt das Standardsignal in der professionellen digitalen Audiotechnik. Danach wird pro Abtastzeitperiode T=1/48 kHz ein 64 Bit umfassender Datenrahmen (Frame) übertragen, der in zwei gleiche Subframes unterteilt ist. Jedes Subframe enthält die Daten für einen Audiokanal. Die Daten bestehen zunächst aus vier Synchronisationsbits, dann folgen vier Bits für Zusatzdaten, die auch als Audiodaten genutzt werden können, weiter 20 Bits für Audiodaten und schließlich vier weitere für allgemeine Kennzeichnungen [225]. Deren Bedeutung folgt aus zwei Blöcken, die 384 Subframes umfassen. Die genaue Bitzuordnung im Subframe ist in Abbildung 3.34 dargestellt. Insgesamt ergibt sich eine Datenrate von 1,536 Mbit/s bzw. 192 kByte/s für jeden Audiokanal.

3.3.3.5 SDI-Signalkontrolle

Die Kontrolle digitaler Videosignale ist nicht so einfach wie bei analogen, denn aus dem Oszilloskopbild eines Digitalsignals ergibt sich kein Hinweis auf feh-

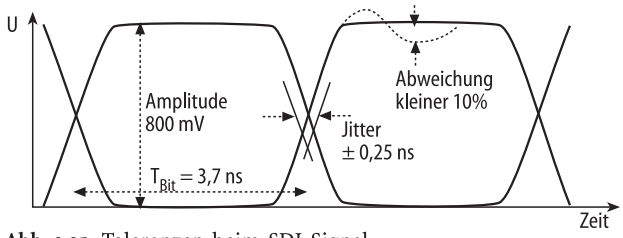


Abb. 3.35. Toleranzen beim SDI-Signal

lerhafte Bits. Logische Bitfehler lassen sich über Messungen der Bitfehlerrate (BER) und mittels Error Detection and Handling (EDH) ermitteln, wobei bestimmte Bitmuster, die eine Fehleranalyse erlauben, im Ancillary-Bereich des Datenstroms mitgeführt werden. Für die Prüfung der Kabelentzerrung und der Taktrückgewinnung werden auch sogenannte pathologische Signale eingesetzt, die zwar legal sind, aber das System an seine Grenzen bringen.

Zur systematischen Kontrolle kann die Augenmusteranalyse verwendet werden, um physikalische Signalverzerrungen zu bestimmen. In Messgeräten können dafür Toleranzfelder für den maximalen Zeit- bzw. Amplitudenfehler angegeben sein. Abbildung 3.35 zeigt ein normgerechtes SDI-Signal und erlaubte Abweichungen. Generell gilt, dass das Signal noch decodierbar ist, solange das Auge noch nicht geschlossen ist [218]. Insgesamt ist diese Art der Fehleranalyse aber aufwendig und vor allem wenig anschaulich. Daher werden zur Kontrolle der Digitalsignale die Daten in den meisten Fällen interpretiert und in einer Weise dargestellt wie es aus der Analogtechnik bekannt ist (s. Kap.2).

Das digitale Signal ist sehr störsicher, zeigt aber an der Grenze tolerierbarer Störungen das typische »Alles oder Nichts«-Verhalten der Digitaltechnik (Brick Wall Effect). Als Beispiel ist in Abbildung 3.36 die Auswirkung einer über die Toleranzgrenze hinaus gesteigerten Leitungslänge bei einem bestimmten Kabeltyp aufgezeigt. Es wird deutlich, dass durch eine kleine Leitungsverlängerung von 270 m auf 290 m das bis dahin einwandfreie Signal plötzlich stark fehlerbehaftet wird [143]. Datengeber für das digitale Komponentensignal besitzen häufig Fehlergeneratoren die dieses Verhalten simulieren, sodass künstlich z. B. 50 m Kabellänge zugeschaltet werden können um zu prüfen, wann die Toleranzgrenze für Fehler erreicht ist.

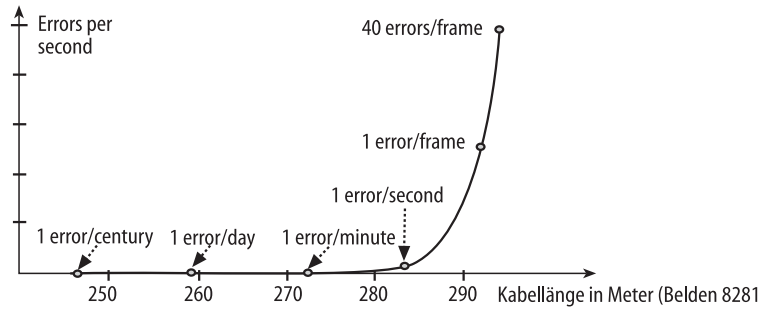


Abb. 3.36. Anstieg der Bit Error Rate beim SDI-Signal in Abhängigkeit von der Kabellänge

3.4 Digitale HD-Signale

Nachdem die in den 90er-Jahren begonnene Entwicklung von HDTV, also der Nutzung hoch aufgelöster Bilder im Fernsehen, aus verschiedenen Gründen scheiterte (s. Kap. 4.4), waren in Europa hoch aufgelöste Bilder für die Fernsehausstrahlung vorerst von geringer Bedeutung. Erst im Jahre 2004 wurde EURO 1080, der erste HDTV-Sender, in Betrieb genommen und zwar nicht als nationaler Sender, sondern für ganz Europa. Im Zuge der Entwicklung in den USA, die die Digitalisierung des TV-Bereichs mit der Einführung höherer Auflösungen verknüpfte, wurde HDTV aber für die Produktionsseite doch immer interessanter, denn HD-Produktionen haben den Vorteil international verwertbar zu sein, bis hin zur Auswertung im Kino. Zudem werden sie zunehmend auch unabhängig vom Fernsehen genutzt. Die HD-Produktion sollte dabei möglichst formatunabhängig sein, wie z. B. von 50-Hz- und 60-Hz-Bildwechselfrequenzen. Daher wurden vermehrt Multiformatgeräte entwickelt, die zusätzlich mit 24 Hz Bildwechselfrequenz bei progressiver Abtastung (24p) arbeiten und damit in diesen Aspekten weitgehend dem Filmformat entsprechen, das bis heute das einzige internationale Austauschformat geblieben ist.

Die analogen HD-Signalformen wurden bereits im vorigen Kapitel vorgestellt, hier werden die HD-Signale in digitaler Form beschrieben. Die digitalen HD-Signalformen sind in enger Anlehnung an die SD-Signalparameter definiert worden und basieren von den Grundkonzepten auf den analogen HD-Signalen. So gilt, dass als Farbsignalgrundform meist das Komponentenformat verwendet wird und weiterhin, dass das Bildseitenverhältnis 16/9 beträgt und die Grenzfrequenz ohne Kellfaktor berechnet wird (Tabelle 3.4). Im Gegensatz zu den SD-Signalen wird allerdings mit quadratischen Bildpunkten (square pixels) und einem den aktuellen Leuchtstoffen entsprechenden Farbraum gearbeitet. Für das HD-Luminanzsignal gilt:

$$Y = 0{,}213\ R + 0{,}715\ G + 0{,}072\ B.$$

Damit ändern sich auch die Pegelreduktionsfaktoren zur Bildung von C_R und C_B, denn die Signalspannung in allen drei Kanälen soll weiterhin 700 mV bzw. ± 350 mV betragen. Es gilt:

$$C_R = 0{,}635\ (R - Y) \text{ und } C_B = 0{,}539\ (B - Y).$$

Auf dieser Basis werden die Amplitudenwerte der γ-vorentzerrten HD-Signale mit 8 oder 10 Bit digitalisiert, wobei die Wertzuordnung vollständig von den SD-Signalen übernommen wurde (Abb. 3.23).

Nachdem zunächst mit zwei national verschiedenen HD-Zeilenzahlen gearbeitet wurde, die auf einer ungefähren Verdopplung der beiden dominanten SD-Zeilenzahlen beruhten, gab es schließlich einen international einheitlichen Standard. Die zugehörige Empfehlung ITU-R BT.709 sah nur noch eine Zeilenzahl von 1125 brutto bei 1080 aktiven Zeilen vor [114]. Diese ergeben sich aufgrund der geforderten square pixels aus 1920 aktiven Pixels in der Horizontalen. Diese errechnen sich wiederum aus der Verdopplung der aktiven Bildpunktzahl des SD-Systems und dem Faktor 1,33 aus dem Übergang von B/H = 4/3 zu B/H = 16/9, also: $2 \cdot 720 \cdot 1{,}33 = 1920$. Wie bei den Standard-Definition-

Tabelle 3.4. Spezifikationen von ITU-R BT 709-5

Bildseitenverhältnis		16:9	
Komponentenform		4:2:2	
Zeilen/Bild		1125	
Aktive Bildpunkte/Zeile		1920	
Aktive Zeilen/Bild		1080	
Abtastrate		74,25 MHz	
Bildfrequenzen	30p/30psF / 60i	25p/25psF / 50i	24p/24psF
Bildpunkte/Zeile	2200	2640	2750

Formaten wird dabei meist mit dem Zeilensprungverfahren gearbeitet. Dieses HD-Format, das abgekürzt oft nach der aktiven Zeilenzahl benannt wird, war damit von der Bildauflösung her das einzige und wurde national nur den verschiedenen Bildfrequenzen angepasst.

Diese Situation änderte sich zu dem Zeitpunkt, als die US-TV-Sender zur Einführung von digitalem Fernsehen und zum gleichzeitigen Umstieg auf HDTV verpflichtet wurden. Um den Übergang zu erleichtern und die Investitionskosten dafür abzumildern, wurde ein geringer aufgelöstes HD-Format zugelassen. Dieses hat zwei Drittel der aktiven Zeilen des „großen" HD-Formats, also 720 Zeilen. Da auch hier square Pixels entstehen sollen, errechnet sich aus B/H = 16/9 eine aktive Bildpunktzahl von 1280 pro Zeile. Die Bruttozeilenzahl wurde auf 750 Zeilen festgelegt. Der erste HD-Standard hat damit eine Bildauflösung die bei ungefähr 2 Millionen Bildpunkten liegt (1920 · 1080 = 2,0736 Millionen), während der zweite ca. eine Million Bildpunkte (1280 · 720 = 0,9216 Millionen) aufweist.

3.4.1 Bildaufbau digitaler HD-Signale

Bei der Gewinnung von Digitalsignalen hat die Wahl der Abtastfrequenz eine entscheidende Bedeutung. Wie bei SD-Signalen beruht sie hier auf einem ganzzahligen Vielfachen von 2,25 MHz. Während bei SD-Signalen das sechsfache davon, also 13,5 MHz, als Basisabtastfrequenz gewählt wurde, ergibt sich für HD-Signale, ausgehend von einer Verdopplung der Bildauflösung in beiden Bilddimensionen und dem Übergang von 4/3 auf 16/9, das 5,33-fache dieses Wertes, also 5,33 · 13,5 MHz = 32 · 2,25 MHz = 72 MHz. Dieses war die Abtastfrequenz des alten europäischen HD-Systems. Bei der internationalen Angleichung wurde dann das 33-fache von 2,25 MHz, also 74,25 MHz festgelegt.

Der Wert 74,25 MHz ist aktuell die Grundabtastfrequenz für die HD-Systeme. Das gilt für das „große" HD-Format, das standardmäßig im interlaced-Mode arbeitet. Es gilt aber auch für das „kleine" HD-Format mit 720 aktiven Zeilen, denn dieses hat die Besonderheit, dass es ausschließlich mit progressiver Abtastung verknüpft ist und nicht mit dem Zeilensprungverfahren betrieben werden darf. Bei 750 (Brutto-) Zeilen und 50 Hz Vollbildbildwechselfrequenz ergibt sich eine ähnliche Bildpunktfrequenz wie bei 1125 Zeilen und 50 Hz-Halb- bzw. 25 Hz-Vollbildwechselfrequenz: 1280 · 750 · 50 Hz = 48 MHz und 1920 · 1080 · 25 Hz = 51,84 MHz.

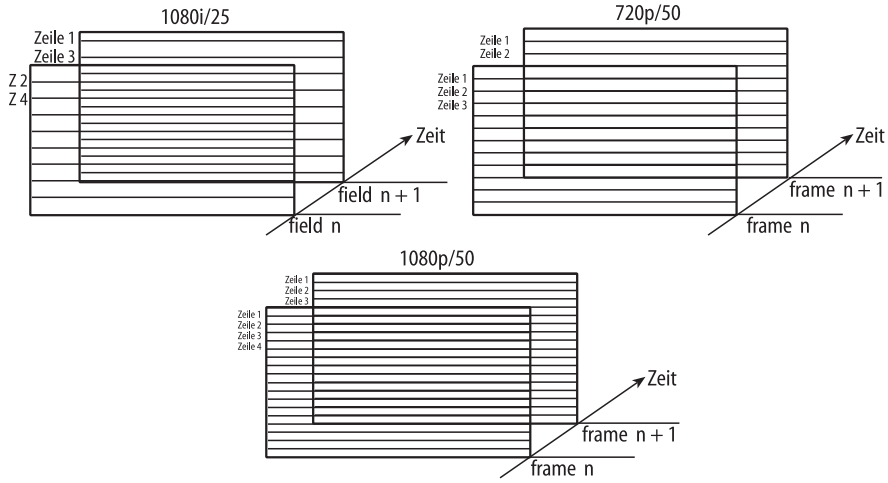


Abb. 3.37. Vergleich der Orts- und Zeitauflösungen der HD-Signale

Die Festlegung der Abtastrate und aktiven Pixel führt bei den verschiedenen, zugelassenen Bildfrequenzen zu unterschiedlichen Zeilenfrequenzen und damit variieren die Abtastwerte pro Zeile (Tabelle 3.4). Es gilt: 1125 · 30 Hz = 33, 75 kHz, 1125 · 25 Hz = 28,125 kHz und 1125 · 24 Hz = 27 kHz. Die Variation wird durch die Austastlücken aufgefangen, kleinere Bildfrequenzen führen zu längeren Austastlücken. Bei einer Bildfrequenz von 25 Hz und 1125 Zeilen ergeben sich z. B. insgesamt 74,25 MHz/28,125 kHz = 2640 Abtastwerte. Aufgrund der 1920 aktiven Bildpunkte umfasst die Austastlücke 720 Werte. Das zweite HD-Format nutzt brutto 750 Zeilen. Die Zeilenfrequenz beträgt bei 50 Hz Bildfrequenz 37,5 kHz die 1980 Abtastwerten entsprechen, sodass die Austastlücke bei 720p/50 genau 1980 – 1280 = 700 Abtastwerte umfasst.

Die entstehenden Datenraten und Grenzfrequenzen sind bei der 720p/50- und der 1080i/25-Formatvariante auch dann ähnlich, wenn man sie nur auf die aktiven Pixel bezieht. Die Variante 720p/50 zeigt 50-mal in der Sekunde ein voll aufgelöstes Bild mit 1280 x 720 sichtbaren Bildpunkten, während das 1080i/25-Signal 50-mal 540 Zeilen mit je 1920 Pixels zeigt (Abb. 3.37). Das bedeutet, dass in einer 50stel Sekunde 0,92 MPixel (720p) bzw. 1,04 MPixel (1080i) erzeugt werden. Die Bildwirkung ist allerdings verschieden. Während bei Bildsequenzen mit wenig Bewegung das 1080i-Format Bildqualitätsvorteile bietet (denn das voll aufgelöste Bild aus 1/25 Sekunde hat 2,07 MPixel), ist bei schnellen Bewegungen das 720p-Format vorteilhaft, da es keine Zeilensprungartefakte aufweist. Für das 720p-Format spricht besonders, dass PC-Systeme zur Bildbearbeitung und die mit stark steigender Tendenz verwendeten Flachbildschirme progressive Bilder erfordern. Diese Geräte müssen Interlaced-Signale meist progressiv wandeln (Deinterlacing), was nur bei hohem technischen Aufwand mit guter Qualität gelingt und so die Geräte teuer macht.

Optimal für die Bildqualität wäre hier das 1080p/50-Format, das pro Sekunde 50 voll aufgelöste Bilder mit 1920 x 1080 Bildpunkten zeigt. Es vereint

Tabelle 3.5 HD-Definitionen der EBU

System	Bezeichnung/ Abkürzung	Framerate (Hz)	aktive Pixel pro Zeile	Abtastrate (MHz)	Gesamt-zeilenzahl	Bitrate (netto) (4:2:2, 10 bit)
S1	1280 x 720/p/50 720p/50	50	1280	74,25	750	921,6 Mbit/s
S2	1920 x 1080/i/25 1080i/25	25 50 (Field)	1920	74,25	1125	1036,8 Mbit/s
S3	1920 x 1080/p/25 1080p/25	25	1920	74,25	1125	1036,8 Mbit/s
S4	1920 x 1080/p/50 1080p/50	50	1920	148,50	1125	2073,6 Mbit/s

die bessere Bildauflösung und den progressiven Aufbau. Das 1080p-Format hat gegenüber den anderen erwähnten HD-Formaten eine verdoppelte Auflösung, muss aber entsprechend mit 148,5 MHz abgetastet werden (Abb. 3.37).

Das 1080p-Format ist bisher wenig verbreitet. Das liegt einerseits daran, dass es wegen der damit verbundenen Datenrate von ca. 3 Gbit/s kaum noch in der üblichen Studioinfrastruktur mit Kabeln zu übertragen ist (s. Abschn. 3.4.2) und andererseits die Heimdisplays meist nicht über diese Auflösung verfügen. Hinzu kommt, dass bei den üblichen Betrachtungsabständen, die in Europas Wohnzimmern nach EBU-Ermittlungen ca. 2,7 m betragen, 1080 Zeilen auch gar nicht erforderlich sind. Die deutschen Rundfunkanstalten haben den HDTV-Betrieb daher mit 720p/50 aufgenommen [189]. Für die Produktion sind aber auch 1080-Zeilen-HD-Formate von Bedeutung.

Die European Broadcasting Union (EBU) definiert für Europa die vier in Tabelle 3.5 dargestellten HD-Formate. Zunächst die beiden wichtigsten: S1 mit der abgekürzten Benennung 720p/50 und S2 mit der Bezeichnung 1080i/25. Das Format S3 wurde bisher noch nicht erwähnt. Es arbeitet mit progressivem Bildaufbau bei 25 Hz und bietet eine schlechte Bewegungsauflösung, die der der Filmtechnik ähnelt. S4 ist schließlich das 1080p/50-Format mit progressivem Bildaufbau bei 50 Hz. Auch die EBU empfiehlt aus den oben genannten Gründen, und auch, um die Kompressionsalgorithmen nicht mit einer zu hohen Eingangsdatenrate zu überfordern, die Verwendung des Formates 720p/50.

Außer von der EBU und der ITU wird die Standardisierung auch von der SMPTE (Society of Motion Picture and Television Engineers) vorangetrieben. Die Empfehlung SMPTE 274M definiert die digitale Darstellung und das Timing des Formats mit 1080 und SMPTE 296M des Formats mit 720 Zeilen.

Bei den HD-Formaten ist für die meisten praktischen Fälle eine Verwendung von Komponentensignalen vorgesehen, deren Abtastratenverhältnis wie bei SD mit 4:2:2 bezeichnet wird, da die Chrominanzkomponenten mit der Hälfte von 74,25 MHz, also 37,125 MHz abgetastet werden. Bei der Verwendung von Datenworten mit 10 Bit Breite ergibt sich daraus ein serieller Datenstrom mit einer Bitrate von 1,485 Gbit/s. Bei hohen Qualitätsansprüchen kann auch mit 4:4:4 gearbeitet werden, allerdings erhöht sich dann die Datenrate. In diesem Fall, also ohne Unterabtastung, wird als Basis allerdings in den meisten Fällen das RGB-Signal verwendet.

In Verbindung mit Datenreduktionsverfahren gibt es bei HD-Signalen auch eine Unterabtastung als generelle Vorverarbeitung. Im Gegensatz zu SD-Signalen, wo die Unterabtastung nur die Farbdifferenzsignale betrifft (4:2:0 oder 4:1:1), greifen einige Hersteller, insbesondere im Zusammenhang mit Magnetbandaufzeichnungsverfahren von 1080i-Signalen, auch zur Reduktion der Y-Auflösung. Verbreitet sind hier Abtastmuster von 3:1:1 (HDCAM) und 3:1,5:1,5 (DVCProHD) oder 3:1,5:0 (HDV), d.h. anstelle von 1920 wird mit 1440 Bildpunkten gearbeitet. Die Grundabtastfrequenz ist dabei von 74,25 MHz auf 3/4 davon, also 55,6875 MHz verringert und die Chrominanzauflösung entsprechend der Zahlen angepasst. So arbeitet z. B. HDCAM intern mit einer H-Auflösung von nur 480 Pixels für die Chrominanzkanäle.

3.4.1.1 HD und gesteigerte Bildauflösung

Die dargestellten Signalformate gelten für die Anwendung im Bereich HDTV bzw. HD-Video. Neuerdings wird viel über die Verwendung dieser Formate für die Filmproduktion diskutiert. In diesem Zusammenhang sei darauf hingewiesen, dass die Digitalbearbeitung von Filmmaterial eine Abtastung des Filmbildes erfordert, solange keine elektronische Kamera verwendet wird. Diese Abtastvorgänge werden meist mit 2k, teilweise auch mit 4k vorgenommen, was bedeutet, dass die Bildhorizontale in 2048 oder 4096 Bildpunkte aufgelöst wird. Für die Vertikale ergeben sich beim Vollformat des Filmbildes daraus 1536 bzw. bei 4k-Abtastung 3112 Bildpunkte (s. Kap. 5 und Abb. 3.38).

Die HD-Auflösung mit 1920 x 1080 Pixels liegt also nahe an der 2k-Auflösung. Das gilt auch für die Zeitauflösung, da eine Abtastung mit 24p im Standard vorgesehen ist. HD und 2k sind aber nicht identisch. Um 4k-Auflösungen zu erzielen, müssen eigene, hoch auflösende Videoformate definiert werden. Mit den neuen Definitionen wird auch die Verwendung digitaler RGB-Signale

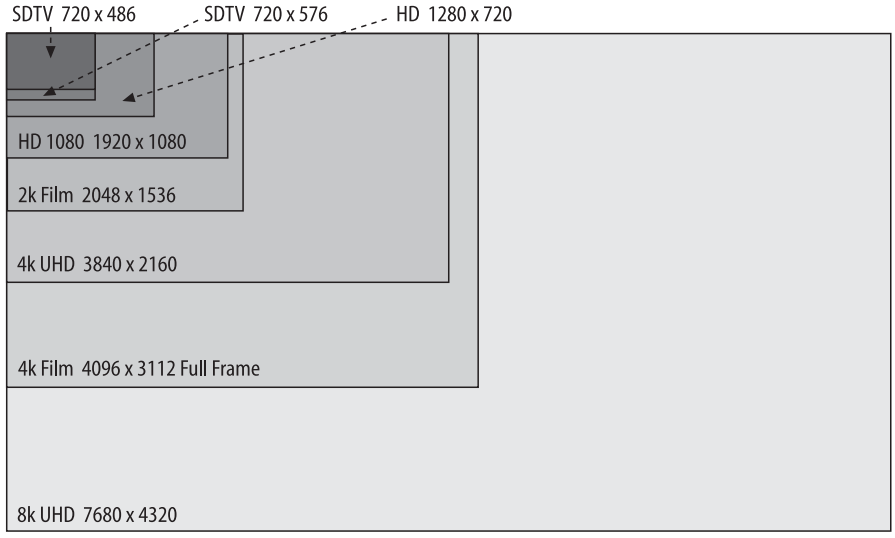


Abb. 3.38. Maßstabsgerechter Vergleich von Video- und Filmauflösungen

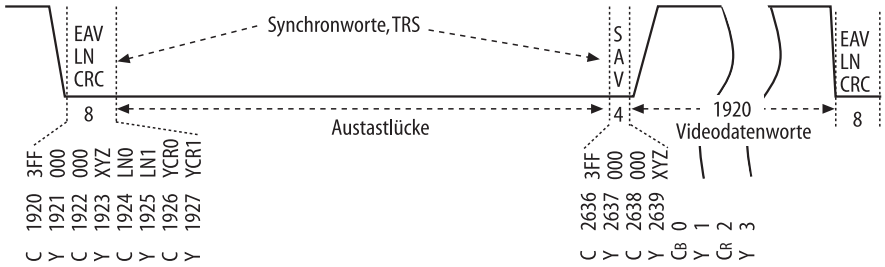


Abb. 3.39. Darstellung eines von zwei Datenströmen beim 1080i/25-Signal

(4:4:4) erfasst, was der Bearbeitung im Filmbereich wesentlich näher kommt als die Verwendung der im HDTV-Umfeld gebräuchlichen Komponenten mit ihrer gegenüber RGB verringerten Farbauflösung.

Ohne Rücksicht auf die Filmseite wird die Steigerung der Bildauflösung auf der Videoseite mit den Formaten 4k-UHD (Ultra-HD1) und 8k-UHD (UHD2) vorangetrieben [86]. Sie sehen Auflösungen vor, die in jeder Dimension das Zwei- bzw. Vierfache des HD-Formats mit 1080 Zeilen hat, also 3840 x 2160 und 7680 x 4320 Bildpunkte bei 8k-UHD (Abb. 3.38).

3.4.1.2 HD-Synchronsignale

Das digitale Synchronsignal (Timing Reference Signal, TRS) ist für HD-Signale wieder in weitgehender Übereinstimmung mit dem digitalen SD-Komponentensignal definiert worden. Wie bei SD wird es mit den Datenworten 3FF, 000, 000 eingeleitet und enthält die Information im Datenwort XYZ, oder eigentlich in den 8 Bit XY. Diese werden bei 10 Bit-Signalen um zwei Nullen erweitert, damit beim Übergang von 10 Bit auf 8 Bit trotz Weglassens der beiden letzten Bits der Informationsgehalt gleich bleibt.

Während das SAV-Signal sich gegenüber SD-Signalen nicht verändert, werden nach einer Empfehlung in SMPTE 292 im Unterschied dazu beim 1080i-System den EAV-Daten noch je zwei Datenworte mit Informationen über die Zeilennummerierung (Line Number) und einen CRC-Fehlerschutz angefügt (Abb. 3.39). Der Fehlerschutz gilt für die Informationen der zugehörigen Zeile.

Außerdem ist bei 1080i definiert, dass zwei Datenströme zu bilden sind, die im ersten die Y-Datenworte und im zweiten Strom die Datenworte von C_R und C_B im Multiplex enthalten. Die TRS-Signale mit dem EAV-Zusatz werden beiden Datenströmen zugefügt, sodass nach einem nachfolgenden Multiplexing der beiden Datenströme die TRS-Daten doppelt erscheinen (Abb. 3.40).

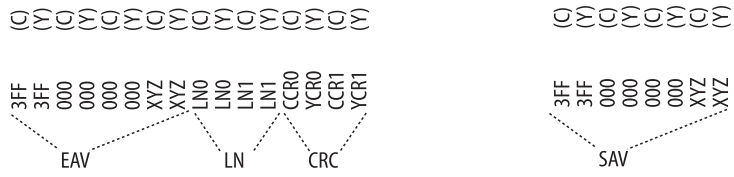


Abb. 3.40. Multiplexbildung und Datenwortbildung beim 1080i-Signal

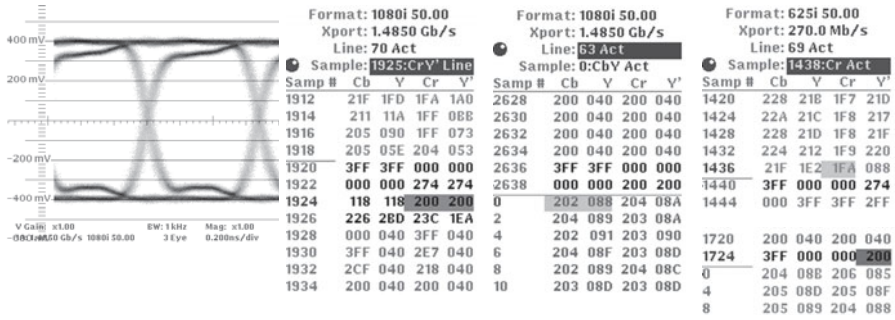


Abb. 3.41. Augendiagramm eines SDI-Signals und Datenworte im bei SD- und HD-Signalen

Entsprechend bezieht sich auch die Datenwortzählung auf einen Datenstrom allein (Datenworte bei 50 Hz: 0 .... 2639 in einer Zeile), während sie sich bei SD-Signalen direkt auf alle drei Anteile des vorher vorgenommenen Multiplexes beziehen (Datenworte bei 50 Hz: 0 .... 1727) (Abb. 3.41 und 3.27) [224].

3.4.2 Digitale HD-Schnittstellen

Auch für HD-Signale stehen eigene Schnittstellen zur Verfügung. Diese sind nach ITU-BT R.1120 und SMPTE 292M definiert. Es gibt es hier wieder eine weitgehende Übereinstimmung mit der Schnittstellendefinition für das Standardsignal, dem Serial digital Interface, SDI. Die serielle HD-Schnittstelle wird in Anlehnung an das Standardsignal als HD-SDI bezeichnet und ist fast identisch definiert. Als wesentlicher Unterschied ist die Erweiterung um LN und CRC und die abweichende Multiplexbildung bei 1080i-Signalen zu nennen (s. Abschn. 3.4.1.2) und natürlich die höhere Datenrate. Sie beträgt insgesamt 1,485 Gbit/s, da die Schnittstelle zwei Datenströme trägt, die jeweils mit einem Takt von 74,25 MHz bitparallel vorliegen und dann serialisiert werden. Der Einfachheit halber wird dabei immer mit 10-bit-Datenworten gearbeitet. Signale mit 8 Bit werden um 2 Nullen erweitert. Die Spannungswerte und die SNRZI-Codierung des Datenstroms werden von SDI übernommen.

Als Leitungen kommen wieder 75-Ω-Koaxialkabel zum Einsatz. Aufgrund der hohen Frequenz und der damit verbundenen Dämpfung können hier nur Längen erreicht werden, die deutlich unter denen der SDI-Schnittstelle liegen. Abhängig von der Kabelqualität kann mit Leitungslängen um die 80 m gerechnet werden.

Bezüglich der Messtechnik bzw. Signalkontrolle gilt das im Abschnitt 3.3.3.5 über SDI Gesagte. Auch hier können die Bits als Augendiagramm dargestellt werden (Abb. 3.41 links) oder als Zahlenwerte. Abbildung 3.42 rechts zeigt diese für ein 1080i-HD-im Vergleich zu einem 576i-SD-Signal im Bereich der Austastlücke. Hier wird die oben angesprochene unterschiedliche Zählweise der Datenworte deutlich. Die Darstellung stammt von dem bekannten Messgerät WFM 700 der Firma Tektronix. Dieses kann die Daten zur einfacheren Beurteilung in Waveform-, Vektorskop- und andere Darstellungsformen bringen.

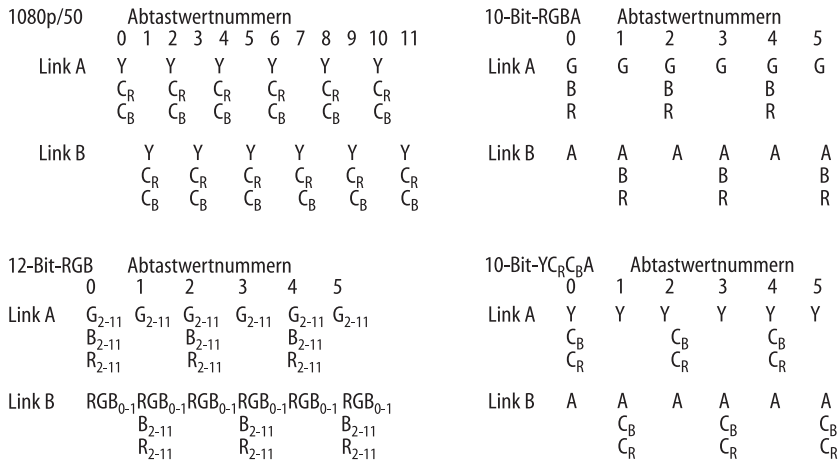


Abb. 3.42. Dual-Link-Datenzuordnungen für verschiedene HD-SDI-Signalformen

3.4.2.1 Dual Link und 3G-SDI

Während die HD-SDI Schnittstelle aufgrund ihrer Kapazität von ca. 1,5 Gbit/s für 4:2:2-HD-Signale nach 1080i oder 720p und damit für die meisten Standardanwendungsfälle geeignet ist, gibt es im Produktionsbereich doch Einsatzgebiete, die eine höhere Leistung erfordern. Beispiele sind Produktionen mit voller Farbauflösung RGB (4:4:4) bzw. auch zusätzlichem Alphakanal für Stanzmasken RGBA (4:4:4:4) oder Anwendungen, die eine Quantisierung der Daten erfordern, die 10 bit übersteigen

Besonderes Interesse gibt es in letzter Zeit an Produktionen, die zwar mit 4:2:2 bei 10 Bit, aber mit 1080p arbeiten und zwar mit 50 oder 60 Hz Vollbildwechselfrequenz. Dieses Format verbindet die Vorteile des 720p- und 1080i-Standards. Der Übergang von 1080i/25 auf 1080p/50 erfordert in diesem Zusammenhang jedoch die Verdopplung der Datenrate.

Für all diese Fälle ist ein Interface gefragt, das $2 \cdot 1{,}485$ Gbit/s, also 2,97 Gbit/s transportieren kann. Um das Ziel zu erreichen, wurde unter der Bezeichnung Dual link zunächst eine Parallelübertragung mit zwei HD-SDI-Leitungen definiert (SMPTE 372M) und später unter der Bezeichnung 3G-SDI eine einkanalige Übertragung (SMPTE 424M).

Die Dual-Link-Verbindung basiert auf zwei gewöhnlichen HD-SDI-Schnittstellen. Die Verteilung der Daten auf die beiden Ströme, die Link A und B bzw. Kanal 1 und 2 genannt werden, wird so vorgenommen, dass allein aus Link A noch weitgehend ein gewöhnliches HD-Signal gewonnen werden kann. Im Falle, dass ein 1080p/50-Signal, also ein progressiv abgetastetes Signal mit 2,97 Gbit/s vorliegt, wird es zeilenweise abwechselnd auf die beiden Datenströme verteilt (Abb. 3.42). Falls ein 10-Bit-RGBA-Signal vorliegt, werden in Link A alle G-Abtastwerte und die geradzahligen R- und B-Werte untergebracht und in Link B der Rest. Im Falle, dass ein RGB-Signal mit 12 Bit vorliegt, werden in Link A die 10 most significant Bits der G- und der geradzahligen R- und B-Abtastwerte untergebracht. Link B trägt die 10 most significant Bits der ungeradzahligen

R- und B-Abtastwerte und die least significant Bits von RGB. Die Festlegungen für die letzten beiden Fälle gelten auch dann, wenn statt RGB ein Komponentensignal Y, C_R, C_B mit voller Auflösung (4:4:4) vorliegt.

Bei den Definitionen des Interfaces 3G-SDI nach SMPTE 424 werden die HD-SDI-Festlegungen für eine Koaxialkabel-basierte Schnittstelle übernommen, die in der Lage ist, die Datenrate von 2,97 Gbit/s zu transportieren. Dabei wird entweder die Bitrate verdoppelt (Level A für 1080p/50-Anwendungen) oder die Daten von zwei parallelen 1,5 Gbit/s-Datenströmen werden im Multiplex übertragen (Level B für Dual Link- oder zwei unabhängige 1,5-GBit/s-Ströme, z. B. bei 3D-Anwendungen) [38], EAV- und SAV erscheinen hier vervierfacht.

Die hohe Datenrate beschränkt die praktisch verwendbare Leitungslänge, abhängig vom Kabeltyp, auf ca. 40 m. Dieser Wert ist für viele Anwendungen zu gering, sodass alternativ auch auf Glasfaserverbindungen zurück gegriffen wird.

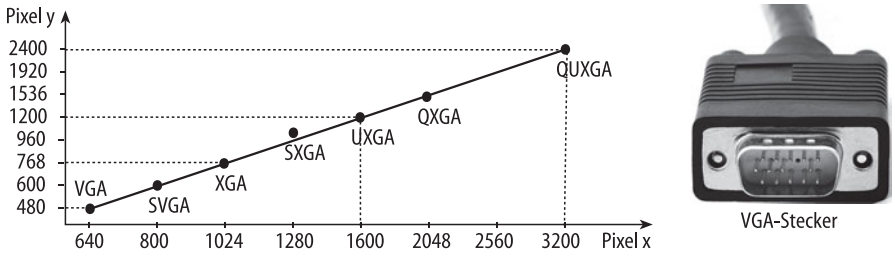
3.4.2.2 Digitalschnittstellen im Heimanwenderbereich

Im Heimanwenderbereich werden Videosignale noch z. T. in analoger Form übertragen, auch dann, wenn Digitalgeräte verwendet werden. Als Signalformen werden FBAS, Y/C und RGB verwendet, mit Cinch-, Hosiden- oder SCART-Steckern.

Daneben gibt es analoge Signale aus dem Computerbereich, die immer als RGB-Signale übertragen werden. Die zugehörige Schnittstelle ist der Videographics Adapter, VGA, der RGB- und Sync-Signale parallel mit Hilfe eines 15-poligen Steckers überträgt (Abb. 3.43). Mit der Bezeichnung VGA war anfangs eine Bildauflösung von 640 x 480 Bildpunkten verknüpft. Unter Beibehaltung der physikalischen Schnittstelle wurde die Bildauflösung dann sukzessive erweitert. Die Signale erhielten dabei die in Abbildung 3.43 angegebenen Bezeichnungen, wie SVGA, XGA etc. Das zugehörige Bildseitenverhältnis war 4:3. Der Vorsatz „W“ beschreibt Formen mit B/H = 16/10 bei gleicher Vertikalauflösung.

Digitale Schnittstellen für den Heimanwenderbereich sind so konzipiert, dass Geräte, wie z. B. Monitore oder Projektoren, mit der gleichen Schnittstelle und unter Vermeidung mehrfacher D/A und A/D-Umsetzungen sowohl an PC als auch an Geräte der Unterhaltungselektronik geschlossen werden können. Aufgrund ihrer Leistungsfähigkeit und des günstigen Preises kommen sie teilweise auch im professionellen Sektor zum Einsatz.

Als erstes wurde das Digital Visual Interface DVI definiert. Es ist so gestaltet, dass es sowohl die analogen VGA-Signale als auch die Digitalsignale



VGA-Stecker

Abb. 3.43. Standard-Auflösungen im Computerbereich

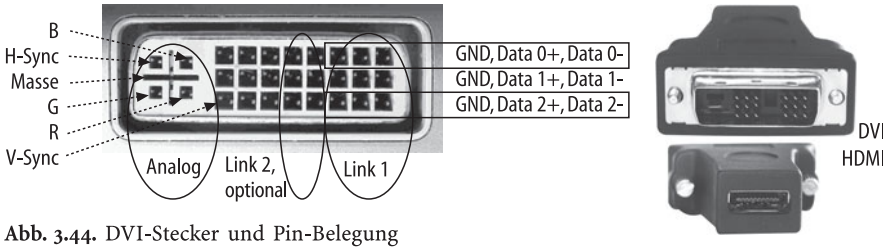


Abb. 3.44. DVI-Stecker und Pin-Belegung

mit einer Leitung übertragen kann. DVI-A beschreibt die analoge, DVI-D die rein digitale und DVI-I die kombinierte Übertragung. Abbildung 3.44 zeigt einen Stecker und die Pinbelegung. Im linken Teil sind die Pins für das analoge RGB-Signal zu sehen. Mittels dieser kann eine einfache Adaptierung von DVI auf VGA vorgenommen werden. Die Digitaldaten werden rechts mittels TDMS (Transition Minimised Differential Signalling) über 4 Datenleitungen übertragen (3 symmetrische Datenleitungen, 1 x Sync). Dabei werden Bitraten bis zu 4 Gbit/s unterstützt, d. h. bei 60 Hz Vollbildwechselfrequenz und 8 Bit-RGB-Daten können 165 Millionen Bildpunkte pro Sekunde (z. B. UXGA oder 1080p) übertragen werden. Höhere Datenraten werden durch Duallink-Verbindungen erreicht, die die Geschwindigkeit durch einen zweiten Satz von Pins verdoppeln.

Als Nachfolger von DVI ist das High Definition Multimedia Interface, HDMI konzipiert. Es ist das digitale Pendant zum SCART-Stecker, denn es erlaubt die Übertragung digitaler Audio-, Video- und Steuerdaten über eine einzige kleine Steckverbindung (Abb. 3.44). Die Audioübertragung ist für bis zu 8 Kanäle mit 24 Bit und bis zu 192 kHz Abtastrate definiert. Die Videoübertragung ist DVI-kompatibel, aber nicht mehr analog möglich. Es kommt weiterhin TDMS zum Einsatz. Die Daten sind mit einem Kopierschutz mit der Bezeichnung HDCP (High-bandwidth Digital Content Protection) versehen.

3.4.2.3 Zusammenhang der Videofrequenzen

Zum Abschluss des Abschnitts über die digitalen unkomprimierten Videosignale sei hier in Abbildung 3.45 verdeutlicht, dass die Taktfrequenzen für die Abtastung mit den Zeilen- und Bildfrequenzen aller Videosignale so zusammenhängen, dass sie stets einen festen Zusammenhang bilden, der in den meisten Fällen auf der Multiplikation von Primzahlen beruht, bzw. Frequenzen aus einer ganz- oder halbzahligen Teilung gewonnen werden können.

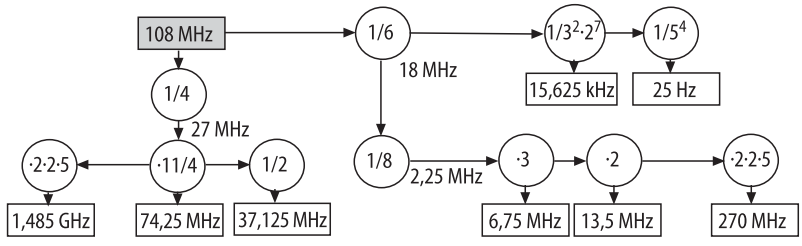


Abb. 3.45. Zusammenhang der Video- und Abtastfrequenzen

3.5 Videodatenreduktion

Die Übertragung des digitalen seriellen Komponentensignal erfordert ein Mindestbandbreite von 135 MHz, also mehr als das 13-fache des analogen Komponentensignals, dessen Bandbreite mit ca. 9 MHz angegeben werden kann (5 MHz + 2 · 2 MHz). Für die Basisbandübertragung stellt die hohe Bandbreite kein Problem dar, sie ermöglicht aber weder eine ökonomische Speicherung der Daten, noch eine Digitalübertragung für die Verteilung an die Fernsehzuschauer. Aus diesem Grunde wurde das digitale Videosignal zunächst nur in sogenannten digitalen Inseln im Produktionsbereich verwendet. Im Laufe der Zeit wurden dann effektive Methoden zur Datenreduktion entwickelt, die so leistungsfähig sind, dass bei gleicher visueller Bildqualität statt einer gegenüber dem Analogsignal erhöhten, nur noch eine verminderte Bandbreite erforderlich ist.

Auch bei hoher visueller Qualität ist die Datenreduktion meist mit Verlusten verbunden. Bei der Produktion soll das digitale Videosignal mehrfach verschiedene Bearbeitungsstufen zur Mischung, für Chroma Key etc. durchlaufen können. Etwaige Verluste können sich in diesen Stufen akkumulieren, sodass im Produktionsbereich keine Datenreduktion oder maximal eine Reduktion um den Faktor 2...5 verwendet wird. Man sagt, dass das Digitalsignal transparent bleiben soll. Bei der Distribution, wie auch bei der Programmzulieferung zu den Sendeanstalten (Contribution), liegen die Programme dagegen meist in endgültiger Form vor. Nachbearbeitungsschritte entfallen, sodass für die Contribution eine stärkere Datenreduktion erlaubt ist. Bei der Verteilung an die Zuschauer mittels Digitalem Fernsehen, DVD etc. werden Reduktionsfaktoren zwischen 20 und 100 verwendet.

Bei der Datenreduktion werden die Begriffe Redundanz-, Relevanz- und Irrelevanzreduktion verwendet. Redundanzreduktion im strengeren Sinne bezeichnet die Beseitigung von Daten, die im Original mehrfach vorliegen. Redundante Information wird z. B. eliminiert, wenn ein Standbild nur einmal übertragen wird, anstatt es 25-mal in der Sekunde zu wiederholen. Das Originalsignal ist aus einem redundanzreduzierten Signal ohne Informationsverluste rekonstruierbar. Da die Redundanz mit dem Maß Entropie bestimmt werden kann, wird die Redundanzreduktion auch als Entropiecodierung bezeichnet. Die Entropie ist ein Maß für den mittleren Informationsgehalt oder auch der Informationsdichte eines Zeichens. Für Videosignale ist, je nach Bildinhalt, etwa bei einer Halbierung der Datenrate, also bei einem Kompressionsfaktor von 2:1, die Grenze zwischen Redundanz- und Irrelevanzreduktion erreicht. Die Redundanzreduktion wird bei allen Datenreduktionsverfahren angewandt.

Der Begriff Irrelevanzreduktion bezeichnet eine Datenreduktion, die nicht oder kaum wahrnehmbare Fehler erzeugt. Bereits die verminderte Grenzfrequenz bzw. verminderte Auflösung der Farbdifferenzkomponenten gegenüber dem Luminanzsignal im Bereich der Analogsignale ist ein Beispiel für Irrelevanzreduktion. Irrelevanzreduktion liegt bei Verwendung heute üblicher Algorithmen etwa bis zu einem Reduktionsfaktor von 50:1 vor, darüber hinaus werden die Fehler immer deutlicher sichtbar, d. h. dann tritt Relevanzreduktion auf, die ggf. aus ökonomischen Gründen in Kauf genommen wird.

Tabelle 3.6. Datenreduktion durch Vorverarbeitung

Signal	625i/25	720p/50	1080i/p/25	1080p/50
Schnittstelle	SDI	HD-SDI	HD-SDI	Dual link/3G
Datenrate der Schnittstelle	270 Mbit/s	1,485 Gbit/s	1,485 Gbit/s	2,97 Gbit/s
Datenrate bei 4:2:2 YC_RC_B 10 Bit	270 Mbit/s	1,485 Gbit/s	1,485 Gbit/s	2,97 Gbit/s
	Weglassen der Austastwerte			
Nettodatenrate	207, 36 Mbit/s	921,6 Mbit/s	1,036 Gbit/s	2,072 Gbit/s
	Reduktion von 10 auf 8 Bit			
Datenrate bei 8 Bit	165, 89 Mbit/s	737,28 Mbit/s	0,8288 Gbit/s	1,6576 Gbit/s
	Unterabtastung			
Datenrate bei 4:2:0 oder 4:1:1	124,4 Mbit/s	552,96 Mbit/s	621,6 Mbit/s	
Datenrate bei 3 : 1,5 : 1,5		552,96 Mbit/s	621,6 Mbit/s	
Datenrate bei 3 : 1 : 1			518 Mbit/s	
Datenrate bei 3 : 1,5 : 0			466,2 Mbit/s	

3.5.1 Grundlagen

Die Datenreduktion kann zunächst in die Stufen Vorverarbeitung und Algorithmus unterteilt werden. Die Vorverarbeitung beschreibt das schlichte Weglassen von Daten ohne aufwendige Berechnung.

3.5.1.1 Vorverarbeitung

Die Vorverarbeitung ist nicht einheitlich und wird je nach Anwendung oder Hersteller variiert. Als gewöhnliche Schritte wird auf jeden Fall aber nur das aktive Bild benutzt, d. h. die Daten der Austastlücken werden weggelassen. Als zweiter Schritt wird in sehr vielen Fällen ein 10-Bit-Signal auf 8 Bit reduziert. Nach diesen beiden Stufen ist die Bildqualität visuell meist nicht gemindert, trotzdem ist ein 4:2:2-SD-Komponentensignal bereits von 270 Mbit/s auf ca. 166 Mbit/s reduziert, ein 1080i/25-HD-Signal von 1,485 Gbit/s auf 829 Mbit/s.

Darüber hinaus kann als Vorverarbeitung eine Unterabtastung eingesetzt werden, die über 4:2:2 hinausgeht. Bei SD-Signalen wird nur die Chrominanzauflösung weiter reduziert und zwar auf 4:2:0 oder 4:1:1, was gegenüber 4:2:2 nur noch 3/4 der Daten erzeugt (ca. 125 Mbit/s bei SD und 622 Mbit/s bei HD-1080i). Bei HD-Signalen wird, abhängig vom Gerätehersteller, auch oft mit einer Unterabtastung der Y-Daten gearbeitet. Typische Muster sind 3:1,5:1,5, 3:1:1 oder 3:1,5:0. So arbeitet z. B. das verbreitete HD-Magnetbandaufzeichnungssystem HDCAM mit 3:1:1, also 1440 Bildpunkten für Y und je 480 Pixels für die beiden Chrominanzkomponenten (Tabelle 3.6).

3.5.1.2 Bildcodierung

Dieser Begriff beschreibt die Darstellung von Bildern mit Hilfe von Zahlen. Das Bild wird in einzelne Bildpunkte, in Pixel zerlegt. Die Grau- oder Farbwerte jedes Bildpunktes werden durch eine Zahl beschrieben. Ein digitalisiertes Bild ist also bereits codiert, zusätzlich kann eine, im Hinblick auf die

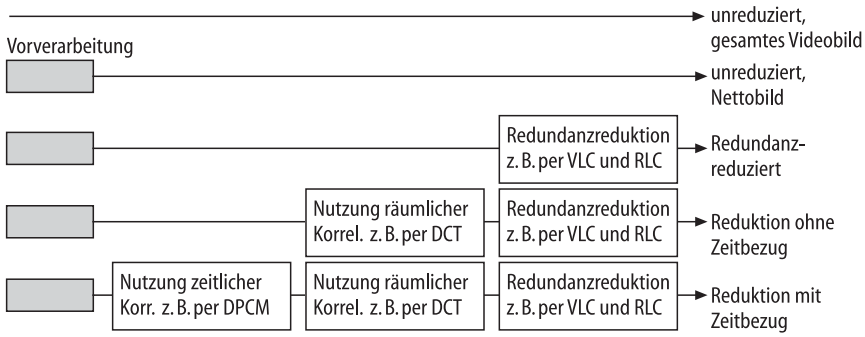


Abb. 3.46. Grundmechanismen der Datenreduktion

angestrebte Datenreduktion effiziente Transformation der Daten und eine Codewortzuordnung vorgenommen werden, die auf der Wiedergabeseite wieder rückgängig gemacht wird. Das Codierungs- /Decodierungspaar wird als Codec bezeichnet.

Die Grundlage der Verarbeitung für die Datenreduktion von Farbbildern ist in den meisten Fällen das digitale Komponentenformat (4:2:2). Das Luminanzsignal und die Farbdifferenzsignale werden hier separat und parallel nach den gleichen Verfahren behandelt.

Nach der Vorverarbeitung wird die weitere Verkleinerung der Datenmenge durch Rechenprozesse mit Hilfe spezieller Algorithmen erreicht. Der wesentliche Grund dafür, dass eine Datenreduktion überhaupt möglich ist, ist das Vorliegen von Ähnlichkeiten im Bild. Bei der Betrachtung zweier zeitlich aufeinanderfolgender Bilder wird deutlich, dass sich der Bildinhalt bei gewöhnlichen Bildfolgen von Bild zu Bild nur wenig ändert. Meist ist die Änderung auf die Bewegung von Objekten zurückzuführen. Bei der Betrachtung eines Einzelbildes fällt auf, dass oft große Flächen mit ähnlichen Farb- und Grauwerten auftreten, d. h. dass benachbarte Pixel einander häufig ähnlich sind. Außerdem sind mittlere Graustufen häufiger zu finden als die Werte Schwarz und Weiß. Diese Ähnlichkeiten des Bildes mit sich selbst oder folgenden Bildern – man spricht auch von hoher örtlicher und zeitlicher Korrelation – stellen Redundanzen dar, die zur Datenreduktion ausgenutzt werden können [78]. Die Datenreduktion bewirkt eine geringere Datenrate, die Bits rücken scheinbar enger zusammen. Daher wird anstelle des Begriffs Datenreduktion auch die Bezeichnung Kompression verwendet.

Abbildung 3.46 zeigt eine Übersicht über übliche Kombinationen der Bearbeitungsstufen für die Nutzung der genannten Korrelationen. Außer bei unkomprimierten Daten wird immer die Redundanzreduktion eingesetzt. In vielen Fällen werden dann zusätzlich örtliche Ähnlichkeiten ausgewertet. Die höchsten Datenreduktionsfaktoren werden durch die Kombination zeitlicher und örtlicher Korrelationen erzielt. Die Stufen sind so geordnet, dass die am häufigsten eingesetzten hinten liegen. Dementsprechend werden hier die Stufen auch in der Reihenfolge von hinten nach vorn, also beginnend mit der Redundanzreduktion vorgestellt.

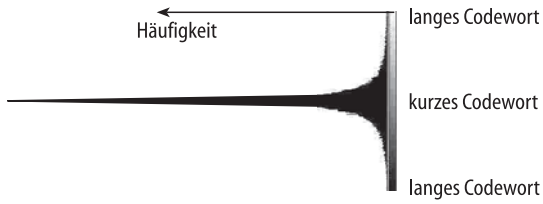


Abb. 3.47. Prinzip des Variable Length Coding

3.5.2 Redundanzreduktion

3.5.2.1 Variable Längencodierung

Durch geschickte Codewortzuordnung kann bereits bei einem nichttransformierten Signal ein verringerter Datenstrom erzeugt werden (Redundanzreduktion durch Entropiecodierung). Nach einer gewissen Übertragungszeit lässt sich nämlich feststellen, dass einige Werte häufiger übertragen werden als andere. Anstatt für alle Daten, d. h. Grauwerte der Pixel eines Bildes, die gleiche Wortlänge mit z. B. 8 Bit zu verwenden, können die Wortlängen für häufig auftretende Werte kurz und für selten auftretende Werte lang gewählt werden (Abb. 3.47). Je größer die Wahrscheinlichkeitsunterschiede sind, die im zu komprimierenden Signal auftreten, desto effektiver arbeitet diese variable Längencodierung (VLC). Ein einfaches Beispiel für die Funktion ist das Morsealphabet. Der häufig auftretende Buchstabe »e« wird hier nur durch einen Punkt (·) repräsentiert, während der seltener auftretende Buchstabe »f« durch zwei Punkte, einen Strich und einen weiteren Punkt dargestellt wird (· · – ·).

Damit die variable Längencodierung effektiv arbeiten kann, muss dafür gesorgt werden, dass die Unterschiede in der Auftrittswahrscheinlichkeit der Werte groß werden, d. h. dass viele Werte möglichst häufig und nur wenige Werte selten auftreten, denn dann können sehr viele Werte mit kurzen Symbolen codiert werden.

Eine verbreitete Methode zur variablen Längencodierung ist die Anwendung des Huffman Codes (Abb. 3.48). Dabei werden zunächst für alle auftretenden Zeichen die Auftrittswahrscheinlichkeiten ermittelt und nach Häufigkeit geordnet. Die Einzelwahrscheinlichkeiten der am seltensten auftretenden Zeichen werden zu einer gemeinsamen Wahrscheinlichkeit addiert und dann die Sum-

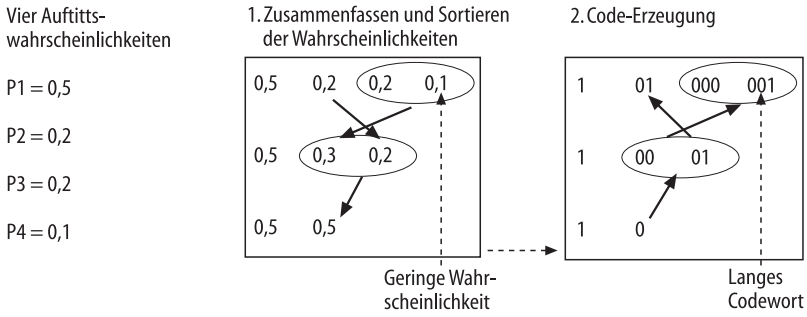


Abb. 3.48. Beispiel für eine Huffman-Codierung

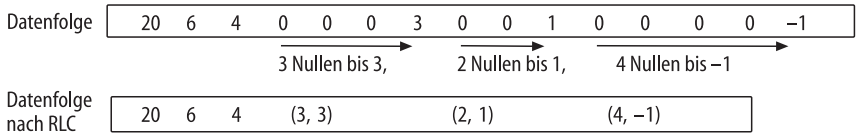


Abb. 3.49. Prinzip des Run Length Coding

me mit den restlichen Zeichen erneut nach Häufigkeit geordnet. Der Prozess wird solange wiederholt, bis nur noch zwei übrig bleiben, denen die Codeworte 1 und 0 zugeordnet werden können. Anschließend werden die Stufen zurückverfolgt, und den Stufen entsprechend zunächst die kurzen und dann die längeren Codeworte vergeben. So resultiert eine Bitzuweisung, bei der den größten Wahrscheinlichkeiten die kürzesten Codeworte zugeordnet sind.

Auf der Empfängerseite muss dieser Prozess zur Decodierung bekannt sein. Dazu kann die Information in Form einer Huffmann-Tabelle dem Decodierer als festgelegt bekannt sein, oder sie wird mit im Datenstrom übertragen.

Die Effektivität der Redundanzreduktion kann weiter durch die Lauflängencodierung (Run Length Coding, RLC) gesteigert werden. Dabei wird mehreren gleichen Werten, die in Folge auftreten, ein eigenes Symbol zugeordnet, indem zuerst die Anzahl der Wiederholungen und dann der wiederholte Wert gespeichert wird. Der Beginn einer Wiederholung wird mit Marker-Bytes gekennzeichnet, die nicht im Datenstrom vorkommen. Mittels eines Offset kann die Mindestwiederholrate angegeben werden, ab der kodiert wird (Abb. 3.49).

Ein Problem, dass bei der variablen Längencodierung auftritt, ist die sich ergebende veränderliche Datenrate (Variable Bitrate, VBR), die in einigen Fällen, z. B. bei der Aufzeichnung von digitalen Videosignalen, nicht verarbeitet werden kann. Eine konstante Bitrate (CBR) kann mit Hilfe von Pufferspeichern (Buffer) erzielt werden. Der Puffer fängt die ankommenden Daten auf und gibt sie in gleicher Reihenfolge wieder aus, wobei die Datenrate konstant gehalten wird. Im optimalen Fall ist der Puffer halb gefüllt. Falls er über einen längeren Zeitraum leerer wird, muss die eintreffende Datenrate erhöht werden und umgekehrt. Die vom Füllstand abhängige Regelung der Datenrate kann über eine Rückkopplung (Abb. 3.50) mit Hilfe des Quantisierers geschehen, dem mitgeteilt wird, dass er z. B. gröber quantisieren soll, wenn der Puffer zu voll wird. Die Quantisierungsstufe, die eine Rundung und die Beschränkung der Anzahl erlaubter Werte bewirkt, wird der VLC und RLC vorgeschaltet. Sie ist die Stufe in der Bearbeitungskette, die für die eigentliche Relevanz- oder Irrelevanzreduktion, also eine Qualitätsverschlechterung im Signal verantwortlich ist.

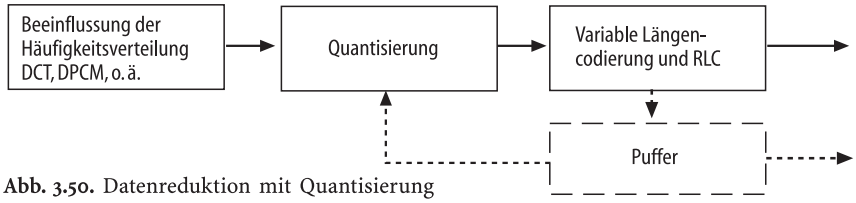


Abb. 3.50. Datenreduktion mit Quantisierung

3.5.3 DCT

Die Diskrete Cosinustransformation, DCT, ist neben der Redundanzreduktionsstufe die bedeutungsvollste unter den Datenreduktionsalgorithmen, denn sie wird bei fast allen Datenreduktionsstandards verwendet. Sie dient dazu, die Korrelationen innerhalb eines Einzelbildes (intraframe) auszunutzen. Als Sonderform der allgemeinen Transformationscodierung nutzt die DCT den Umstand, dass sich, ähnlich wie bei einem fouriertransformierten Analogsignal (s. Abschn. 2.3.2), aus dem Frequenzspektrum andere Bearbeitungsmöglichkeiten erschließen als bei der Zeitdarstellung. Das Frequenzspektrum eines Videosignals kann dreidimensional dargestellt werden. Neben der Frequenz bezüglich der Zeit ergeben sich die Ortsfrequenzen bezüglich der Koordinaten der Bildfläche. Tiefe Ortsfrequenzen repräsentieren große Bildstrukturen und langsame Helligkeitsübergänge, hohe Frequenzen werden dagegen durch kleine Strukturen und abrupte Übergänge verursacht. Da in natürlichen Bildern grobe Strukturen weit häufiger vorkommen als feine, findet eine Konzentration auf niederfrequente Komponenten statt und die nachfolgende Redundanzreduktion mit VLC und RLC kann effektiv arbeiten. Zudem können hochfrequente Komponenten zur Datenreduktion ungenauer dargestellt oder ganz weggelassen werden, da der dabei entstehende Fehler meist irrelevant ist.

Vor der Anwendung der DCT wird das Bild in Blöcke aufgeteilt, die meist quadratisch angeordnet sind, ein Block besteht dann aus N x N benachbarten Pixeln (Abb. 3.51). Die Transformation wird jeweils auf alle Pixel eines Blockes angewendet. Bei zu kleinen Blöcken werden die Ähnlichkeiten benachbarter Pixel nicht gut ausgenutzt, bei großen Blöcken ist die erforderliche Rechenleistung sehr hoch. Häufig wird eine Blockgröße mit N = 8 gewählt.

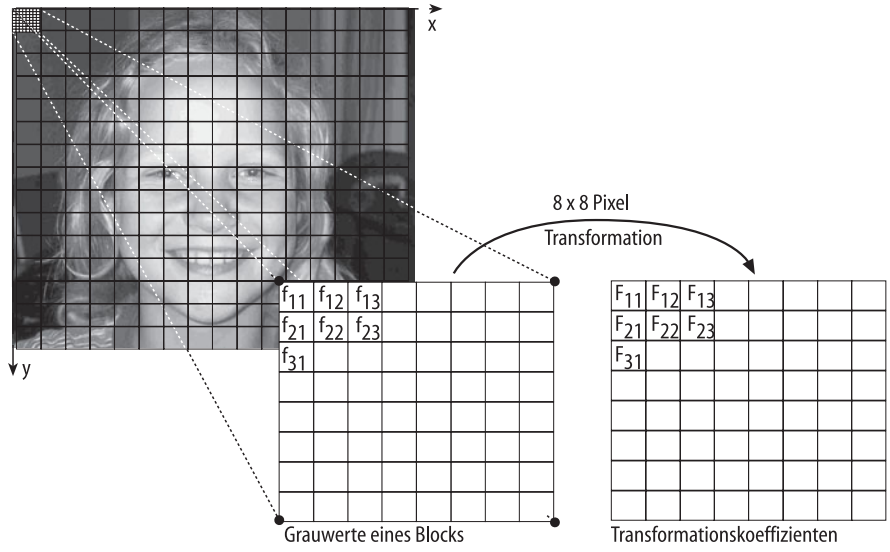


Abb. 3.51. DCT-Blockbildung

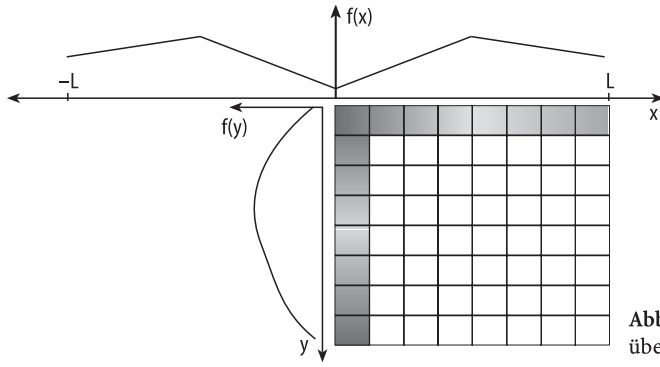


Abb. 3.52. Funktionsverlauf über den Bilddimensionen

Die Transformation eines Blockes von Bildpunkten führt zu einem gleich großen Block von Transformationskoeffizienten. Die Transformation ist prinzipiell reversibel, ein N x N Block von Bildpunktwerten wird in N x N Koeffizienten überführt, die statt der Werte selbst übertragen werden. Die diskreten Pixelpositionen im Bild können mit x und y bezeichnet werden, die Positionen der Transformationswerte mit u und v. Die Werte der Bildpunkte sind dann durch f(x,y) beschrieben, die Transformationswerte durch F(u,v). Damit entsteht die DC-Transformation folgendermaßen aus der Fouriertransformation: Eine eindimensionale Funktion f(x), die stetig über dem Intervall –L bis +L definiert ist, lässt sich als Fourierreihe mit den Summen über die mit den Fourierkoeffizienten a und b gewichteten cos- und sin-Funktionen darstellen, die als Argumente die Vielfachen der Grundfrequenz enthalten (s. Abschn. 2.3.2). Betrachtet man nun ein Intervall von o bis +L, wie es z. B dem Funktionsverlauf über einen Teil der Videozeile entspricht, so lässt sich die Funktion über diesen Bereich auf das Intervall –L bis +L erweitern. So entsteht eine gerade Funktion, die die Eigenschaft hat, dass alle zur sin-Funktion gehörigen Fourierkoeffizienten zu Null werden und nur noch die cos-Anteile übrigbleiben (Abb. 3.52). Da die erweiterte Funktion über dem halben Intervall aber der Ursprungsfunktion entspricht, bleibt es beim Wegfall der sin-Anteile. Es gilt:

$$f(x) = \sum_u a_u \cos(u\pi x/L).$$

Für die Transformation aller Werte, die auf einer zweidimensionalen Fläche vorliegen, kann zunächst von der Multiplikation zweier Funktion ausgegangen werden, die von x und y abhängen [20]:

$$f(x,y) = f(x) \cdot f(y).$$

Damit folgt die zweifache Reihenentwicklung

$$f(x,y) = \sum_u \sum_v c_{uv} \cos(u\pi x/L_1) \cos(v\pi y/L_2).$$

Mit dem Übergang von einer kontinuierlichen auf eine diskrete Intensitätsverteilung, entsprechend der Werte einzelner Bildpunkte innerhalb eines Blockes der Größe N x N folgen schließlich die Beziehungen:

$$f(x,y) = 2/N \sum_u \sum_v C(u)\, C(v)\, F(u,v) \cos((2x+1)u\pi/2N) \cos((2y+1)v\pi/2N)$$

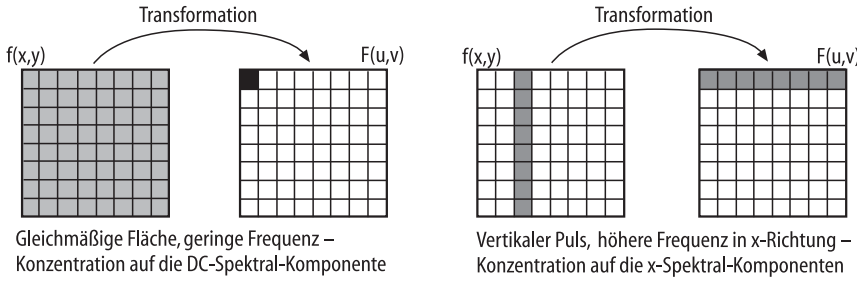


Abb. 3.53. DCT-Beispiele

Die genannte Gleichung gilt für die inverse DCT (IDCT). Die Umkehrung, d. h. die eigentliche DC-Transformation, erfolgt nach der Beziehung:

$$F(u,v) = 2\ C(u)\ C(v)/N\ \sum_x\ \sum_y f(x,y)\ \cos((2x+1)u\pi/2N)\ \cos((2y+1)v\pi/2N)$$

mit der Summe von 0 bis N–1 und $C(w) = 1/\sqrt{2}$ für w = 0 und C(w) = 1 sonst.

Es wird deutlich, dass in die Berechnung eines einzelnen Transformationskoeffizenten alle Werte des zu transformierenden Blockes einfließen. Die DC-Komponente für den mittleren Grauwert des Blocks ist von großer Bedeutung, sie befindet sich in der linken oberen Ecke der DCT-Matrix (Abb. 3.53). Wenn für jeden Block nur dieser Wert übertragen würde, ergäbe sich bereits ein Bild, das jedoch eine horizontale und vertikale Auflösungsverminderung um den Faktor N aufweisen würde. Die Koeffizienten für die niederfrequenten AC-Spektralanteile befinden sich in direkter Nachbarschaft darunter bzw. rechts vom DC-Koeffizienten. Die hochfrequenten AC-Anteile sind unbedeutender. Sie liegen bei der unteren rechten Ecke der Matrix und weisen i. d. R. wesentlich kleinere Werte auf als die Koeffizienten in der Nähe der DC-Komponente, zudem sind die zugehörigen hohen Ortsfrequenzen visuell weniger gut wahrnehmbar.

Mit der Transformation wird, abhängig vom Bildgehalt meist eine Konzentration der Signalleistung auf wenige Teilbänder des Frequenzspektrums vorgenommen. Abbildung 3.53 zeigt zwei Beispiele für einfache Transformationen. Bei einer gleichmäßig hellen Fläche ergibt sich nach der Transformation nur der DC-Koeffizient, während ein vertikaler Impuls einer Sprungfunktion in der Horizontalen entspricht, woraus entsprechende Koeffizienten folgen.

Nachdem hier vor allem die Bildanalyse mit der DCT betrachtet wurde, also der Übergang von den Grauwerten zu den Koeffizienten, soll das Wesen der DCT noch einmal anhand der Rücktransformation, also quasi der Synthese verdeutlicht werden. Dazu kann man sich, wie bei der Fouriertransformation, vorstellen, dass die Grauwerte des betrachteten Bildblocks durch Zusammensetzung von bekannten Mustern entstehen. Die Muster sind die Basisfunktionen der DCT, d. h. die ganzzahligen Vielfachen von Hell-Dunkel-Wechseln, hier in zwei Dimensionen. Die Frequenzen sind bekannt: Die geringste Frequenz ist Null und entspricht einer einheitlich grauen Fläche, die höchste Frequenz ist der S/W-Wechsel von Bildpunkt zu Bildpunkt in beiden Dimensionen, also ein Schachbrettmuster mit N x N Feldern.

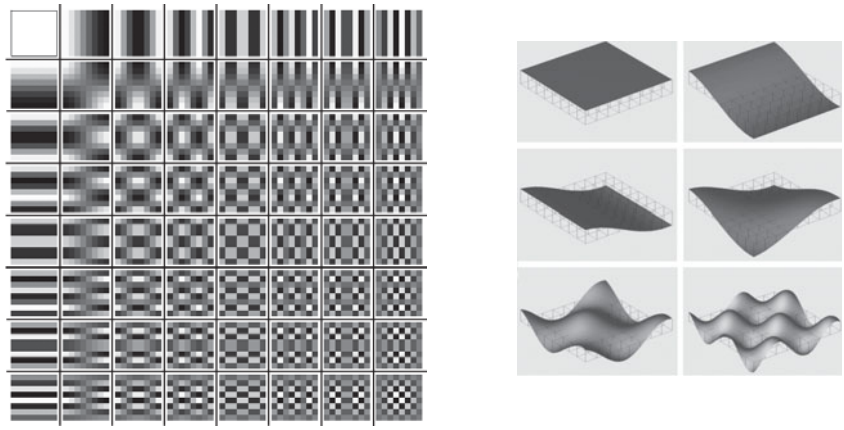


Abb. 3.54. DCT-Basisfunktionen bei 8 x 8 Pixeln

Abbildung 3.54 zeigt links alle DCT-Basisfunktionen, bezogen auf diskrete Werte aus 8 x 8-Blöcken und rechts einige davon als 3D-Veranschaulichung. Für die Rücktransformation in Grauwerte kann man sich nun vorstellen, dass alle Muster überlagert werden wobei von der Koeffizienten bestimmt wird, mit welcher Helligkeit die Muster an der Synthese teilnehmen.

Als konkretes Beispiel zeigt Abbildung 3.55 die Blockbildung bei einem Graustufenbild, die Grauwerte des ausgewählten 4 x 4-Blocks sind als Zahlen dargestellt. Nach einer DCT ergibt sich der in der Abbildung dargestellte Block von Koeffizienten. Es wird deutlich, dass der Gleichanteil den größten Wert aufweist und sich hier die Zahlen erheblich mehr voneinander unterscheiden als im Bildblock. Damit sind die Voraussetzungen für eine effektive variable Längencodierung geschaffen.

3.5.3.1 Quantisierung

Die bisher beschriebene Form der DCT ist hinsichtlich der Datenreduktion nicht effektiv, dafür findet eine verlustlose Reduktion statt, die DCT ist vollständig reversibel. Zunächst können für die Darstellung der DCT-Koeffizienten (insbesondere für die den DC-Anteil) sogar mehr als 8 Bit erforderlich sein. Ausgehend von einem Block mit 8 x 8 Bildpunkten, die mit je 8 Bit codiert sind, ergibt sich nämlich durch die Summierung über die 64 Werte, die zur Gewinnung jedes Koeffizienten erforderlich ist, eine Erhöhung um 6 Bit. Aufgrund der Normierung mit dem Divisor 8 bzw. 4 in der DCT-Bestimmungsformel reduzieren sich diese wiederum, sodass alle Koeffizienten mit 11 Bit dargestellt werden könnten.

Um die Daten effektiv zu reduzieren, werden die Koeffizienten quantisiert und dabei unterschiedlich bewertet, d. h. sie werden geteilt und dann gerundet (Abb. 3.55). Die hochfrequenten Signalanteile werden durch größere Werte geteilt (gröber quantisiert) als die niederfrequenten, oder sie werden ganz weggelassen, denn das menschliche Auge nimmt die zu den hochfrequenten Anteilen gehörigen feinen Bildstrukturen weniger gut wahr als die groben. Ab-

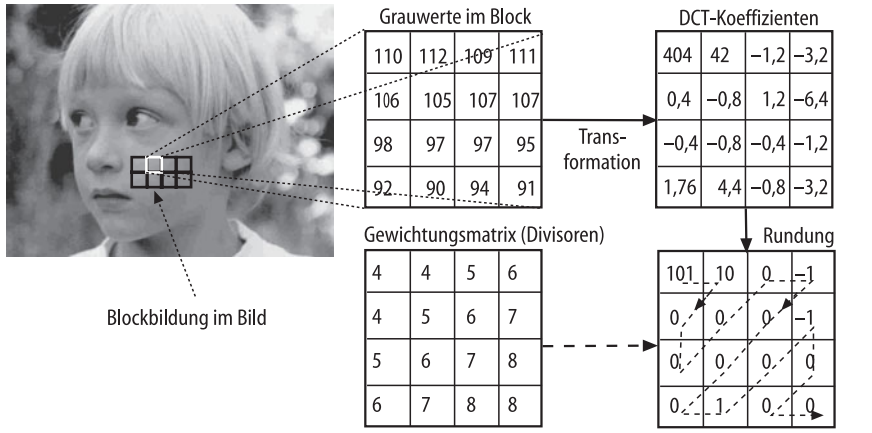


Abb. 3.55. DCT mit Gewichtung und Rundung

bildung 3.56 zeigt zwei anhand eines psycho-physiologischen Experiments optimierte Quantisierungstabellen für Luminanz und Chrominanz, in denen für jeden Koeffizienten eines 8 x 8-DCT-Blocks ein Divisor aufgeführt ist. Es wird deutlich, dass die zu hochfrequenten Bildinhalten gehörigen Koeffizienten ein geringeres Gewicht bekommen. Der DC-Divisor der Tabelle ist 16, sodass der DC-Koeffizient, der gewöhnlich den größten Wert aufweist, bei dieser Bewertung mit 7 Bit statt 11 Bit dargestellt werden kann. Die Teilung der höchsten Koeffizientenwerte durch ca. 100 reduziert die erforderliche Bitzahl auf 5.

Nachdem die Werte auf ganze Zahlen gerundet wurden sind nur noch sehr wenige Koeffizienten von Null verschieden. Dieser Umstand wird besonders deutlich, wenn die Original-Abtastwerte nur sehr wenige Grauwertabstufungen umfassen (Abb. 3.55). Fein strukturierte Bildblöcke können eine größere Zahl von Null verschiedener Koeffizienten enthalten. Daran zeigt sich, dass die Datenrate bei der Übertragung in Abhängigkeit vom Bildinhalt schwankt. Die Darstellung der Menschenmenge in einem Fußballstadion erfordert z. B. eine höhere Datenrate als das Sprecherbild einer Nachrichtensendung. Wenn die DCT in Systemen eingesetzt wird, die diese Schwankung nicht verarbeiten können, wie z. B. Bandaufzeichnungsverfahren, muss der Datenstrom geglättet werden.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Quantisierungstabelle für die Luminanz

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Quantisierungstabelle für die Chrominanz

Abb. 3.56. Quantisierungstabellen [114]

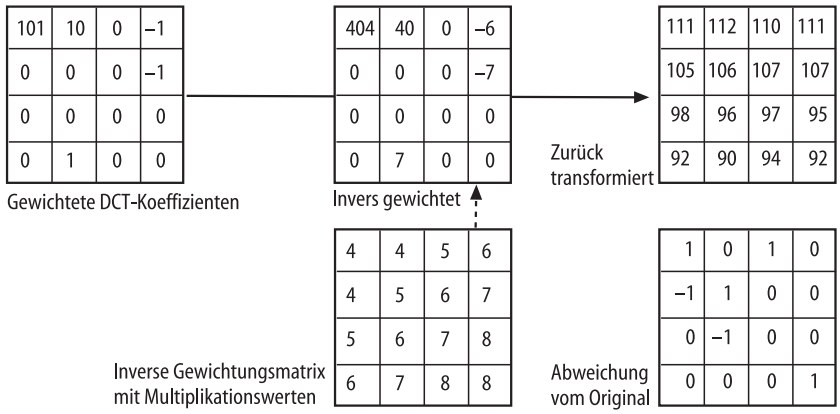


Abb. 3.57. DCT-Rücktransformation

In Abbildung 3.57 ist die Rückgewinnung der Grauwerte aus dem einfachen Beispiel von Abbildung 3.55 dargestellt. Sie wird nach der Übertragung der gerundeten DCT-Koeffizienten im Empfänger vorgenommen. Zu Beginn wird invers gewichtet, aus dem quantisierten Koeffizientenfeld ergibt sich danach durch Rücktransformation das decodierte Bildsegment. Schließlich wird deutlich, dass sich nur wenige der zurückgewonnenen Bildpunktwerte von den Ausgangsbildpunkten unterscheiden.

Aber auch wenn die Fehler klein sind, so wird das Signal doch durch den Quantisierungsprozess verfälscht. Ein typischer Fehler, der bei der DCT auftritt, ist der Blocking-Effekt, d. h. dass die Blockgrenzen sichtbar werden. Es hängt von der Wahl der Bewertungs-Matrizen ab, in welchem Maße die Fehler in Erscheinung treten. Bei starker Quantisierung, wenn also viel hochfrequente Anteile verschwinden, dominieren in der Bildblöcken immer mehr die DC-Koeffizienten, sodass die Unterschiede in den mittleren Grauwerten der Nachbarblöcke als Blocking-Artefakte deutlich in Erscheinung treten (Abb. 3.58). Die Wahl der Bewertungs-Matrizen kann einen wesentlichen Unterschied bei der Implementierung der DCT hervorrufen. Daher können Datenreduktionsfehler bei einem Gerät deutlicher in Erscheinung treten als bei anderen.

Abb. 3.58. Vergleich verschieden stark DCT-komprimierter Bilder

3.5.3.2 Hybride DCT

Eine deutliche Erhöhung der Effektivität der Datenreduktion ergibt sich, wenn die Nachbarschaftsbeziehungen zwischen den Pixelwerten gleichzeitig innerhalb eines Bildes (Intraframe) und auch zwischen verschiedenen Bildern einer Bildfolge (Interframe) ausgenutzt werden. Die Erweiterung der DCT unter Einbeziehung der zeitlichen Dimension ist sehr rechen- und speicheraufwendig, deswegen wird in der Praxis meist eine Kombination aus einer DPCM (s. Abschn. 3.5.5) und einer DCT, die sogenannte hybride DCT, angewendet, bei der die DCT die Ähnlichkeitsbeziehungen innerhalb eines Bildes und die DPCM die Ähnlichkeiten zwischen zeitlich benachbarten Bildern ausnutzt (Abb. 3.59).

Bei der hybriden DCT wird die Differenz zweier aufeinander folgender Bilder ermittelt und anschließend wird das Differenzbild transformationscodiert. So ergibt sich eine sehr effektive Datenreduktion. Bilder in guter Qualität können mit einer Datenrate von ca. 5 Mbit/s übertragen werden. Bei akzeptabler Qualität sind Kompressionsfaktoren bis zu 100, bei geringen Qualitätsansprüchen auch Faktoren über 500 erreichbar. Die hybride DCT ist die Basis der Video-Quellencodierung beim MPEG-Standard und damit auch die Basis für das digitale Fernsehen, Digital Video Broadcasting (DVB).

Dort, wo das Videosignal bearbeitet wird, muss der Zugriff auf jedes einzelne Bild gewährleistet sein, die Interframe-Codierung ist hier nicht geeignet. Falls die Bilder aber nur noch übertragen werden sollen, ist die hybride Codierung gut einsetzbar. Beim MPEG-Standard werden die meisten Bilder hybrid codiert, in festen Abständen werden aber einige unabhängige, intraframecodierte Bilder in den Datenstrom eingefügt. Damit wird der Decodereinstieg und statt des Zugriffs auf Einzelbilder der Zugriff auf Bildgruppen möglich.

Die DCT hat eine große Bedeutung für die Videodatenreduktion. Sie wird heute bei den Datenreduktionsstandards MPEG, JPEG und DV (s. u.) ebenso eingesetzt wie bei den speziellen Datenreduktionsverfahren für digitale Videorecorder (z. B. Digital Betacam) die nicht allgemeingültig standardisiert und veröffentlicht sind.

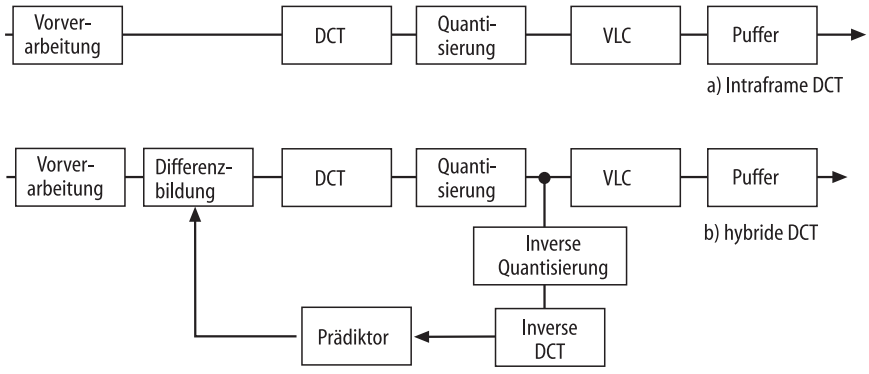


Abb. 3.59. Vergleich von Intraframe-DCT (a) und Hybrid-DCT (b)

3.5.4 Wavelet-Transformation

Ein Nachteil der Videocodierung mittels DCT ist die Tatsache, dass hierbei eine Blockbildung mit zweidimensionaler Transformation zum Einsatz kommt, die eine gezielte Rekonstruktion einzelner Pixel, Pixelgruppen oder des gesamten Bildes ohne vollständige Rücktransformation nicht unmittelbar zulässt. Bei einer groben Quantisierung des Bildinhaltes und einer daraus resultierenden geringen Datenrate sind zudem deutliche Blockartefakte zu erkennen. Man kann bezüglich der Blöcke die Datenreduktion mit der Faltung eines Blatt Papiers vergleichen: Nach der Entfaltung bleiben die Spuren sichtbar.

Um dieses Problem zu mindern, kann als Alternative zur Transformation in den Frequenzbereich auch eine Intraframe-Codierungsmethode gewählt werden, die auf einer Filterung des gesamten Bildes beruht. Dieses leistet die Diskrete Wavelet-Transformation (DWT), bei der die Zeilen und Spalten des Originalbildes jeweils hoch- und tiefpassgefiltert (HP und TP) werden. Mathematisch geschieht dies über eine Faltung. Wavelets sind Funktionen, die nur über einen definierten Bereich von Null verschieden sind und deren Integral gleich Null ist (Abb. 3.60). Sie können verschoben und in ihrer Ausdehnung gedehnt oder gestaucht werden und so zur Untersuchung eines Signals auf seine Frequenzanteile benutzt werden. Bei der Anwendung als Videofilter wird das Wavelet so verschoben, dass eine ganze Zeile nach und nach abgetastet wird. Dabei wird jedem Bildpunkt ein Wavelet-Koeffizient zugeordnet.

Als Filter kommen verschiedene Wavelet-Funktionen zum Einsatz. Am Beispiel der einfachsten Wavelet-Transformation, der Haar-Transformation, kann das Verfahren verdeutlicht werden. Dabei wird ein Filter der Länge Zwei verwendet, d. h. von den Grauwerten zweier Nachbarpixel wird als TP-Koeffizient der Mittelwert übertragen und als HP-Koeffizient die Differenz: $TP(n) = 1/2 \cdot x(n) + x(n+1)$ und $HP(n) = 1/2 \cdot x(n) - x(n+1)$. Trotz der Unterdrückung jedes zweiten Koeffizienten lassen sich bei der Rücktransformation über die Summen- und Differenzbildung wieder die Ursprungswerte der Pixel berechnen.

In praktischen Einsatzfällen, z. B. bei Videostandard JPEG2000, werden komplexe Filter eingesetzt: für verlustbehaftete Codierung ein 9/7- und für verlustlose Kompression ein 5/3-Filterpaar (Abb. 3.60). Die Zahlen beschreiben die Filterlängen jeweils für die Hoch- und Tiefpassfilter bei der Analyse [16].

Da das Filter über die gesamte Bildzeile oder -spalte geführt wird, wird jeder Pixelwert mehrfach gefaltet. Beginnend am Anfang der Zeile werden die ersten Pixel, je nach Länge des Filters, weniger oft gefaltet, als diejenigen, die nicht am Rand liegen. Um nun unerwünschte Artefakte an den Bildrändern zu vermeiden, werden die Ränder des Bildes künstlich erweitert.

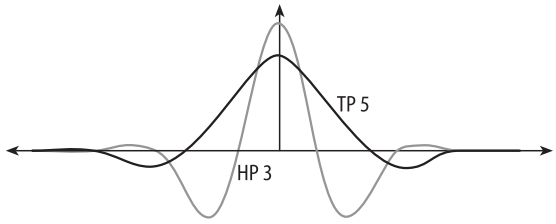


Abb. 3.60. Beispiel für 5/3-Wavelet-Filterfunktionen

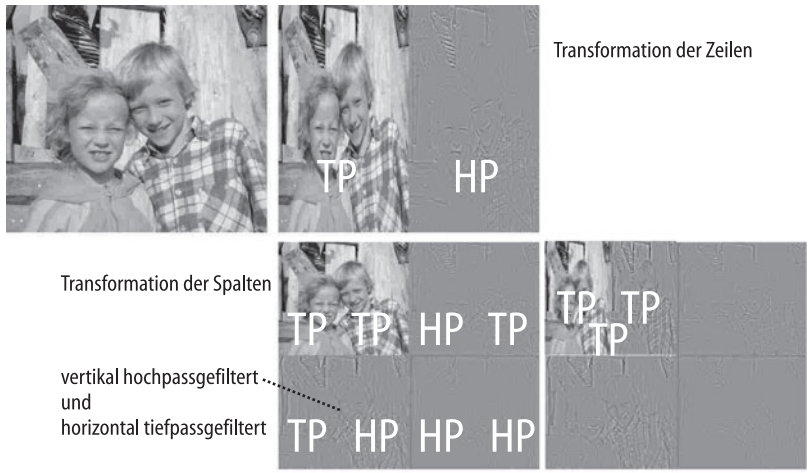


Abb. 3.61. Funktionsprinzip der Wavelet-Transformation

Nachdem das Originalbild zeilenweise transformiert wurde, wird entsprechend die Vertikale bearbeitet. So entstehen vier Teilbilder, die zusammen die Fläche des Originalbildes bedecken und die verlustfreie Rekonstruktion des Originalbildes erlauben. Dasjenige Teilbild, welches sowohl horizontal als auch vertikal tiefpassgefiltert wurde, wird nun als Ausgangsbild für die weitere Transformation verwendet (Abb. 3.61). Der Prozess kann mehrfach wiederholt werden. Ab einer Tiefe von 5 bis 10 Stufen ist jedoch gewöhnlich kaum noch eine Effizienzsteigerung der Codierung zu erwarten [88].

Die eigentliche Datenreduktion beruht wie bei der DCT auf der nachgeschalteten Quantisierung der Werte in den Teilbändern und wie gewöhnlich schließt sich dann an die Quantisierung eine Redundanzreduktion an.

Schließlich folgt die Organisation des entstehenden Bitstroms. Dabei werden die Koeffizienten in Paketen zusammengefasst, wobei die Bildinformationen entsprechend der Bedeutung für den Bildinhalt angeordnet werden können. Die so entstehende Hierarchie richtet sich nach der Art der Verwendung. So kann der Bildwiederaufbau mit wachsender Größe, mit steigender Qualität oder sortiert nach Regions of Interest (ROI) durchgeführt werden. In letzterem Falle werden hervorzuhebende Bildanteile zuerst in voller Qualität aufgebaut, während die übrigen Bildteile noch schlecht aufgelöst bleiben. Dies ist z. B. bei Bildern von Interesse, die möglichst schnell übertragen werden sollen, in denen es aber im Wesentlichen nur auf bestimmte Bereiche ankommt. Der Bitstrom kann jederzeit, auch vom Benutzer, bzw. Betrachter des sich aufbauenden Bildes unterbrochen werden, der Bildaufbau bleibt dann in der entsprechenden Phase stehen. Dadurch sind Einsparungen in der Übertragung möglich. Außerdem besteht die Möglichkeit, ein Bild einmal zu transformieren und dann in unterschiedlichen Qualitätsstufen abzurufen, ohne für jede Stufe ein neues Bild in neuer Quantisierung erzeugen zu müssen [96]. Die DWT hat heute ihre größte praktische Bedeutung bei der Reduktion hochaufgelöster Bilder, z. B. der Datenreduktion für das Digital Cinema.

3.5.5 DPCM

Eine naheliegende Möglichkeit, Ähnlichkeiten in Videosequenzen zu nutzen, ist die Bildung von Differenzen. Dies geschieht mit der Differential Puls Code Modulation (DPCM). Hier werden die Regelmäßigkeiten ausgenutzt, die sich über ausgedehnte Teile des Bildes oder Bildfolgen erstrecken, indem anstelle der Werte selbst nur die Grauwertunterschiede zwischen Nachbarpunkten oder aufeinander folgenden Bildern übertragen werden.

Bei der Differenzbildung innerhalb der Bilder (intraframe) wird die Tatsache genutzt, dass die Bilder oft große Flächen enthalten, sodass die Differenz benachbarter Pixel nur gering ist. Nur die Konturen erzeugen große Differenzen und führen zu den wenigen hohen Werten. Als Beispiel für eine Intraframe-DPCM zeigt Abbildung 3.62 ein Graustufenbild und das zugehörige Differenzbild, sowie im Histogramm jeweils die zugehörigen Häufigkeitsverteilungen der Grauwerte. Zur Gewinnung des Differenzbildes wurde das Bild um ein Pixel horizontal verschoben, die Differenz zwischen verschobenem und unverschobenem Bild gebildet und der Wert 0 dem Wert 128 als mittlerem Grauwert zugeordnet. Positive Differenzen erscheinen heller und negative dunkler. Mittlere Grauwerte überwiegen deutlich, auch im Histogramm wird wieder die Konzentration um den Wert 128 sichtbar. Die Statistik des Differenzbildes hat sich gegenüber dem Ausgangsbild verändert, der mittlere Grauwert tritt viel häufiger auf als die anderen Differenzwerte. Die Häufigkeitsverteilung ist damit so umgestaltet, dass die variable Längenzuordnung effektiv arbeiten kann.

Die Intraframe Prädiktion wird nur bei modernen Codecs verwendet, im Gegensatz zur Interframe Prädiktion, bei der die Ähnlichkeit zwischen aufeinander folgenden Bildern ausgenutzt wird. Zwei zeitlich benachbarte Bilder

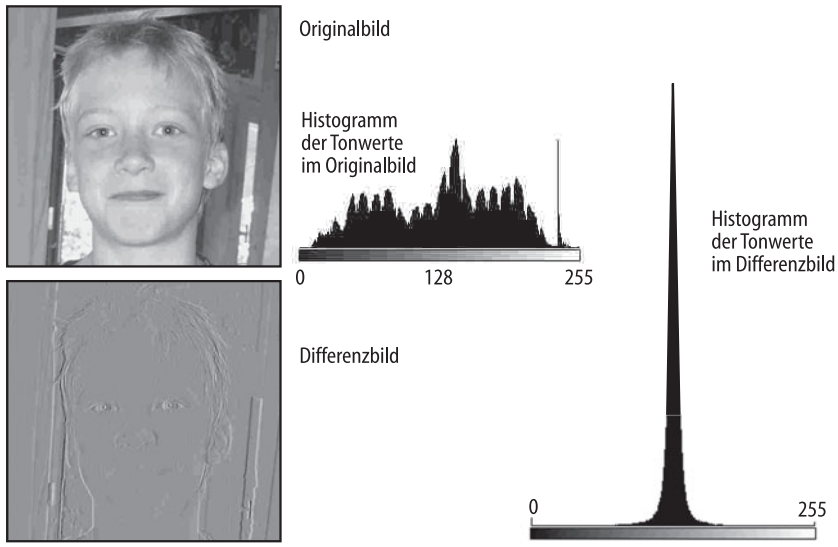


Abb. 3.62. Original- und Differenzbild mit Grauwertverteilungen

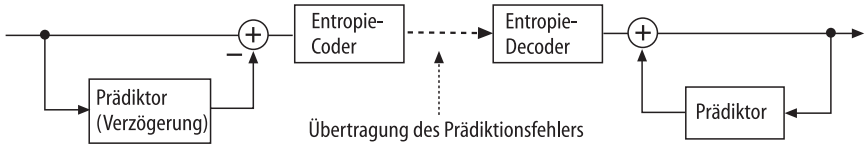


Abb. 3.63. DPCM-Prinzip

zeigen bei gewöhnlichen Bildsequenzen noch weit mehr Ähnlichkeiten als sie innerhalb des Einzelbildes auftreten. Die DPCM wird gewöhnlich der DCT vorgeschaltet, die DC-Transformation also auf Differenzwerte angewandt.

Sowohl bei der Intra- als auch bei der Interframe-DPCM kann die Differenzbildung so betrachtet werden, dass aus den übertragenen Bildpunkten eine Vorhersage (Prädiktion) für den nächsten gebildet wird. Vom Signalwert wird die Prädiktion subtrahiert, schließlich wird nur die Abweichung vom wirklichen Wert, der Prädiktionsfehler, übertragen. Der Vorgang entspricht einer digitalen Filterung. Schon eine einfache Verzögerung (Abb. 3.63) eignet sich als Prädiktionsfilter, und kann auch beim Integrierglied verwendet werden, das die Originalwerte wieder rekonstruiert.

Bei ungestörter Übertragung ist eine fehlerfreie Rekonstruktion möglich, es ergibt sich aber keine hohe Codiereffizienz. Um den Datenstrom weiter zu reduzieren, werden die Daten häufig quantisiert. Die dabei unvermeidlichen Quantisierungsfehler werden nun im Decoder aufaddiert und diese Fehlerverschleppung macht das Signal innerhalb kurzer Zeit untauglich. Die Fehlerfortpflanzung lässt sich aber vermeiden. Dazu wird bereits im Sender eine Decodierung vorgenommen, die genau der Decodierung im Empfänger entspricht. Das decodierte Signal wird wieder der Additionsstufe zugeführt (Abb. 3.64), sodass die Eingangsbedingungen am Decoder vorweggenommen sind [154].

Abhängig vom Bewegungsgehalt innerhalb einer Bildsequenz sind die Ähnlichkeiten benachbarter Bilder unterschiedlich groß. Standbildfolgen erzeugen minimale, Kameraschwenks dagegen große Unterschiede. In sehr vielen Fällen unterscheiden sich aufeinander folgende Bilder vor allem durch veränderte Bewegungsphasen bewegter Objekte. Die Bewegung ändert sich jedoch nicht von Bild zu Bild, sondern bleibt in ihrer Richtung über eine längere Bildsequenz erhalten, sodass z. B. bezüglich eines Objekts, das sich in einer Sekunde über zwei Drittel der Bildbreite bewegt hat, mit sehr großer Wahrscheinlichkeit ab-

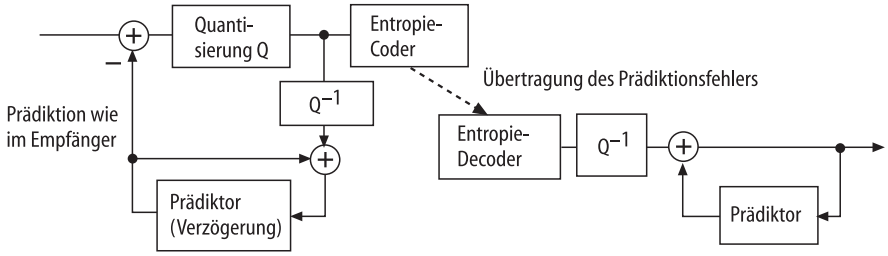


Abb. 3.64. DPCM ohne Fehlerverschleppung

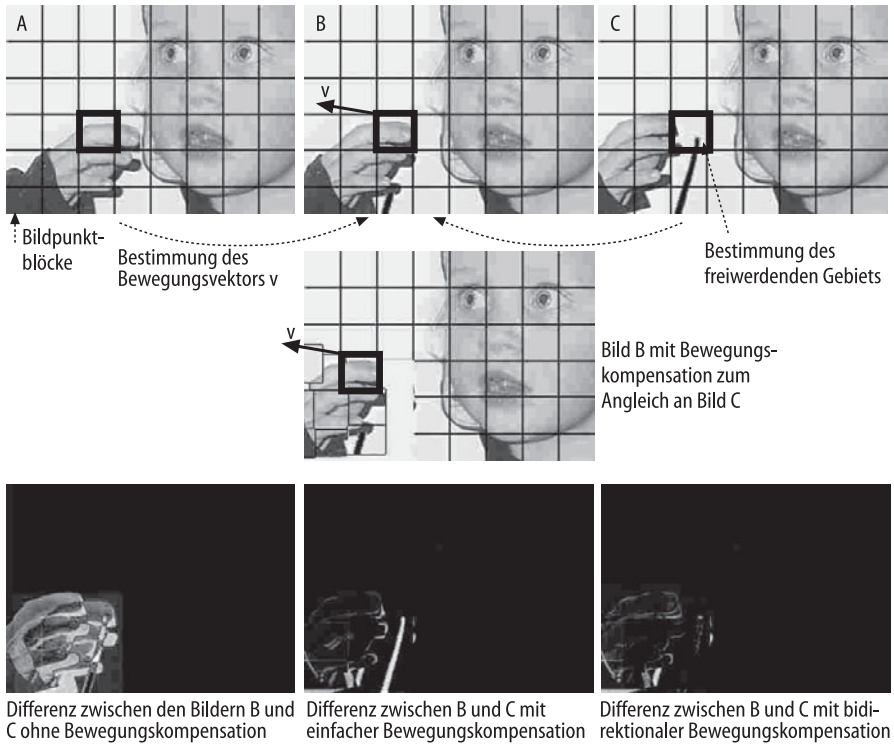


Abb. 3.65. Bidirektionale Prädiktion

geschätzt werden kann, an welcher Position es sich im nächsten Bild befinden wird. Mit dieser Abschätzung verschobener Bildteile aus dem Vergleich von aktuellem und vohergehendem Bild kann die Effizienz der DPCM erheblich gesteigert werden, nämlich indem eine Bewegungskompensation vorgenommen wird. Dabei wird das aktuelle Bild der Bewegungsvorhersage entsprechend verändert und ist damit dem folgenden Bild wesentlich ähnlicher als ohne Bewegungskompensation, die Differenz wird also kleiner.

Abbildung 3.65 zeigt ein Beispiel für eine Bildsequenz mit statischem Hintergrund aber bewegter Hand. Die Bilder A, B und C zeigen die Bewegung der Hand, die immer weiter nach links verschoben ist. Für die Prädiktion, z. B. des Bildes C aus B, kann nun im einfachsten Fall das Bild B unverändert benutzt werden. Als Differenzbild erscheint dann das linke Bild der unteren Reihe in dem eine Differenz von Null schwarz dargestellt ist. Es wird deutlich, dass sich noch relativ große Differenzen ergeben. Diese können minimiert werden, indem aus den Bildern A und B die Bewegung der Hand abgeschätzt wird und als Zusatzinformation in die Prädiktion einbezogen wird. Das Problem dabei ist die Erkennung des bewegten Gegenstands. Objektorientierte Methoden dazu befinden sich noch in der Entwicklung, sodass gegenwärtig meist das einfache Blockmatching-Verfahren zum Aufsuchen ähnlicher Bildteile verwen-

det wird. Dabei werden aus Gründen des Rechenaufwands nicht jeder einzelne Bildpunkt, sondern Bildpunktblöcke (oft 16 x 16 Pixel) betrachtet, von denen angenommen wird, dass sie jeweils zu einem translatorisch bewegten Objekt gehören. Beim Bildvergleich wird im Nachbarbild der Block mit der größten Ähnlichkeit gesucht und dessen relative Verschiebung zum Ausgangsblock als Vektor gespeichert (Abb. 3.65). Die Bewegungskompensation erfolgt schließlich so, dass mit Hilfe dieser Information die Blockverschiebung bei der Ausgabe des vorhergesagten Bildes berücksichtigt wird.

Das Bild in der Mitte von Abbildung 3.65 zeigt Bild B mit entsprechend verschobenen Blöcken und am Differenzbild in dem Mitte der unteren Reihe wird deutlich, dass die Differenz zwischen dem veränderten Bild B und Bild C mit Bewegungskompensation geringer ist als ohne.

Bei der Verfolgung der Bewegung erscheint ein weiteres Problem, nämlich dass das Objekt zwar im, dem aktuellen Bild folgenden, Bild weitgehend richtig lokalisiert wird, aber bei der Bewegung an der Stelle, von wo aus es sich fortbewegt, den Hintergrund verdeckt hat und damit in diesem Bereich die Prädiktion versagt. Der Hintergrund wird erst im folgenden Bild freigegeben. Wenn diese Information für die Vorhersage des aktuellen Bildes benutzt werden könnte, wäre die Prädiktion noch besser. Eine optimale Prädiktion ist also bidirektional, sie stützt sich auf das dem aktuellen Bild vorhergehende und auf das nachfolgende Bild (Abb. 3.65). Obwohl das folgende Bild in der Zukunft liegt und eigentlich gar nicht bekannt sein kann, gibt es die Möglichkeit zur bidirektionalen Prädiktion durch zeitliche Verzögerung auf der Coderseite, d. h. das nächste Bild wird vor der Übertragung berücksichtigt.

Aus diesem Grunde ist es günstig, die Bildfolge vor der Übertragung zu ändern, sodass zunächst beide Nachbardifferenzbilder des gewünschten Bildes übertragen werden und dann am Decoder vorliegen, wo es durch Integration aus den Differenzen zurückgewonnen wird. Zum Schluss müssen alle Bilder im Decoder dann wieder in die richtige Reihenfolge gebracht werden.

Bilder, die auf einfacher, unidirektionaler Prädiktion beruhen, werden mit P gekennzeichnet, die aus bidirektionaler Prädiktion mit B. Intraframe-codierte Bilder ohne Prädiktion werden als I-Frames bezeichnet.

Codierungsmethoden, die auf Bildfolgen beruhen, arbeiten effektiv, haben aber auch einige Nachteile. Insbesondere bei der Schnittbearbeitung (Editing), also der Umstellung von Bildfolgen, ergeben sich hier Probleme, da die Abhängigkeit der Bilder, die keinen einfachen, direkten Zugriff auf jedes Einzelbild erlaubt. Hinzu kommt das Problem, dass die Decodierung zeitlich abhängiger Bilder (Group of Pictures, GOP) mit erheblich mehr Zeitverzögerung verbunden ist als bei Folgen aus unabhängigen Einzelbildern (I-Frame only).

DPCM, DCT und DWT sind die wichtigsten Algorithmen für die Videocodierung. Es gibt Weiterentwicklungen, wie z. B. die 3D-DWT bei der die zeitliche Ebene in die Codierung mit einbezogen wird und noch andere Verfahren, doch sind die drei erstgenannten heute die Grundlage aller wesentlichen Datenreduktionsstandards wie JPEG MPEG, H.26x, DV etc., die in den folgenden Kapiteln behandelt werden. Zur Festlegung der Standards wurden Expertengruppen gebildet, u. a. die Joint Photographics Expert Group (JPEG) und die Moving Pictures Expert Group (MPEG), die den Standards die Namen gaben.

3.6 JPEG

Die Abkürzung JPEG steht für Joint Photographics Expert Group und ist einer der ersten Standards für die Bilddatenkompression. JPEG beschreibt einen universellen Algorithmus zur Einzelbildcodierung und wurde ohne besondere Rücksicht auf Videosignale entwickelt. Das Verfahren beruht auf der Intraframe-DCT, weitgehend in der Form wie sie im vorigen Abschnitt beschrieben wurde. Zunächst wird das Bild in Blöcke von 8 x 8 Bildpunkten zerlegt, die dann mit DCT in den Frequenzbereich transformiert werden. Als nächstes folgt die Gewichtung der DCT-Koeffizienten nach psycho-visuellen Gesichtspunkten. Als Besonderheit bei JPEG werden die DC-Koeffizienten einer DPCM unterzogen, d. h. ihre Werte werden verringert, indem nur die Differenz zu den DC-Werten der Nachbarblöcke im Bild codiert wird.

Über die Quantisierung und Rundung wird der Reduktionsfaktor festgelegt, wobei viele der hohen Ortsfrequenzen anschließend den Wert Null aufweisen. Die endgültige Datenreduktion wird mit der sich anschließenden VLC und RLC per Huffman-Codierung erreicht (Entropiecodierung). Damit diese effektiv arbeiten, werden die Werte der zweidimensionalen Matrix in eine eindimensionale Folge gebracht. Dies geschieht mit Hilfe des sogenannten Zick-Zack-Scanning, das die Abtastung bei dem DC-Koeffizienten beginnt und die AC-Koeffizienten von tiefen zu hohen Ortsfrequenzen mit ihren immer kleineren Werten hin anordnet. Auf diese Weise wird erreicht, dass zum Ende der Kette sehr viele Nullen auftauchen, die mit der RLC sehr effektiv zusammengefasst werden können (Abb. 3.66). An dem Punkt, ab dem bis zum Blockende nur noch Nullen auftreten, wird das Zeichen End of Block (EOB) übertragen. Der Decoder kann die restlichen Zeichen durch Nullkoeffizienten auffüllen, da die Blockgröße ja festliegt.

JPEG steht für eine Vorschrift zur Einzelbildcodierung, der Rechenprozess kann als Software auf Standardrechnern implementiert werden und braucht im Prinzip nicht in Echtzeit abzulaufen. Wenn die Datenreduktion nach JPEG aber für Videobilder eingesetzt werden soll, was z. B. bei einigen nichtlinearen Schnittsystemen (s. Kap. 9.4) der Fall war, so müssen die Bilder nacheinander

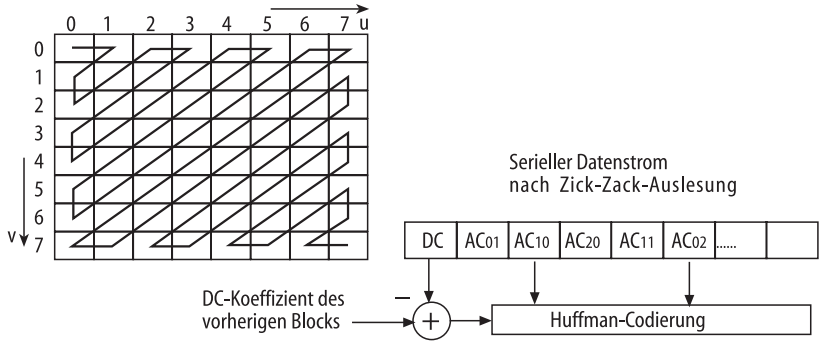


Abb. 3.66. Serialisierung der Koeffizienten durch Zick-Zack-Auslesung

in Echtzeit komprimiert werden, was als Motion JPEG (M-JPEG) bezeichnet wird.

JPEG arbeitet mit Videodaten im Komponentenformat und sieht vor, dass die Unterabtastung der Farbdifferenzkomponenten gegenüber den Luminanzwerten horizontal und vertikal gleich ist und die Chrominanz-Abtastwerte mittig zwischen den Luminanzwerten liegen. Dieses Format widerspricht den Empfehlungen in ITU-R BT.601, was zur Folge hat, dass bei Videoanwendungen meist die 601-Empfehlungen beachtet und das JPEG-Format verletzt wird.

Das JPEG-Nachfolgeformat heißt JPEG2000 (ISO/IEC 15444). Hier wird statt der DCT die Wavelet-Transformation eingesetzt, womit ein flexibler und skalierbarer Bildzugriff und eine gute Codiereffizienz erreicht werden. Abbildung 3.67 zeigt den Vergleich eines unkomprimierten Bitmap-Bildausschnitts mit einem nach JPEG- und einem JPEG2000-reduzierten Bild bei einem hohen Reduktionsfaktor von ca. 100. Es wird deutlich, dass bei JPEG2000 die Fehler wesentlich geringer sind als bei der Codierung nach JPEG.

Nach JPEG2000 darf ein Bild bis zu 16384 Komponenten enthalten, mit jeweils bis zu 38 Bit Farbtiefe. Darüber hinaus sind Mechanismen zur Synchronisation eines Motion-JPEG2000-Datenstroms mit einem Audiodatenstrom definiert. Die Auflösung der Bilder beträgt bis zu ca. 4,3 Mrd. Pixel. Die wichtigste Anwendung ist der Einsatz im Bereich der hochaufgelösten Bilder. Es ist das einzige Datenredukionsverfahren, das für das digitale Kino (Digital Cinema) standardisiert ist. Die empfohlene Dateiendung für JPEG2000 lautet ».jp2«.

Mit JPEG2000 es ist auch möglich, verlustfrei (lossless) zu codieren. Dafür wird eine Integer-Wavelet-Codierung benutzt, die mit symmetrischen Filtern der Längen 5 (TP) und 3 (HP) nur ganzzahlige Koeffizienten produziert. Für eine verlustbehaftete (lossy) Kompression werden Filter der Längen 9 und 7 eingesetzt. Im Gegensatz zur verlustfreien Transformation, bei der die Filterkoeffizienten einfache Brüche darstellen, handelt es sich hierbei um Gleitkommazahlen. Für die Rücktransformation werden die inversen Filter verwendet.

Die Koeffizienten werden im Anschluss ihrer Wichtigkeit für den Bildaufbau entsprechend sortiert. Dafür werden die Subbänder in rechteckige Bereiche (Precincts) unterteilt. Zu einem Precinct gehören meist die drei HP-Subbänder einer Auflösungsebene. Diese werden wiederum in Codeblöcke untergliedert, die typischerweise eine Größe von 64 x 64 Koeffizienten haben. Die Precincts erhalten jeweils einen eigenen Paket-Header. Angefangen mit dem MSB (Most Significant Bit) werden die Codeblöcke nun einem arithmetischen Entropie-Codierer zugeführt. Der Codestrom erhält somit zuerst die für den Bildeindruck relevanten Informationen, Zur Decodierung ausgewählter Bildbereiche (ROI) werden nur die benötigten Codeblöcke aus dem Datenstrom ausgelesen.

Abb. 3.67. Vergleich von Originalbild mit gewöhnlicher JPEG- und JPEG 2000-Kompression

3.7 Der DV-Algorithmus

Dieser Algorithmus ist als Datenreduktionsverfahren für das DV-Magnetbandaufzeichnungsverfahren im Heimanwenderbereich entwickelt worden. Da sich das DV-Format in modifizierter Form im Laufe der Zeit auch im Produktionsbereich etabliert hat, hat der DV-Algorithmus inzwischen als wichtiges Datenreduktionsverfahren neben MPEG und H.264 übergeordnete Bedeutung erhalten. Wie JPEG beruht DV auf einer Intraframe-Codierung ohne sich auf Nachbarbilder zu beziehen. Die Grundlage ist bei beiden Verfahren die DCT mit anschließender Quantisierung und Entropiecodierung per VLC, wobei eine Bildblockgröße von 8 x 8 Bildpunkten festgelegt ist. Im Gegensatz zu JPEG ist DV aber für Videoanwendungen mit konstanter Datenrate optimiert.

DV arbeitet mit einer 8-Bit-Auflösung für jeden Bildpunkt. Das Verfahren ist im Gegensatz zu MPEG symmetrisch, d. h. Encoder und Decoder haben etwa den gleichen Funktionsumfang. Die Basis ist ein digitales Komponentensignal, aus dem zunächst mit Hilfe eines Vertical Chroma Filters ein nach 4:2:0 (in 625/50-Hz-Systemen) oder 4:1:1 (bei 525/59,94 Hz) unterabgetastetes Signal gewonnen wird. Für DV in Europa erzeugt die Y-Abtastung somit 720 x 576 und die C-Abtastung 360 x 288 Bildpunkte. Die Aufteilung in Blöcke mit 8 x 8 Pixeln erzeugt 90 x 72 Luminanz- und 45 x 36 Chrominanzblöcke.

Zur Anpassung an Videosignale und zur effizienten Verarbeitung der Halbbilder wird der Voll- vom Halbbildmodus unterschieden. Der Vollbildmodus (Frame Processing) verwendet die ineinander geschobenen Halbbilder bei der Blockbildung, während im Halbbild-Modus (Field Processing) der 8 x 8-Block in zwei 4 x 8-Blöcke aufgeteilt wird, die jeweils nur die Information aus einem Halbbild enthalten und separat der DCT unterworfen werden. Da sich bei ruhenden Bildbereichen die Halbbilder wenig unterscheiden, ist die Korrelation örtlich benachbarter Zeilen groß, und es wird der Vollbildmodus verwendet, während bei Blöcken mit schnell bewegten Inhalten auf den Halbbildmodus

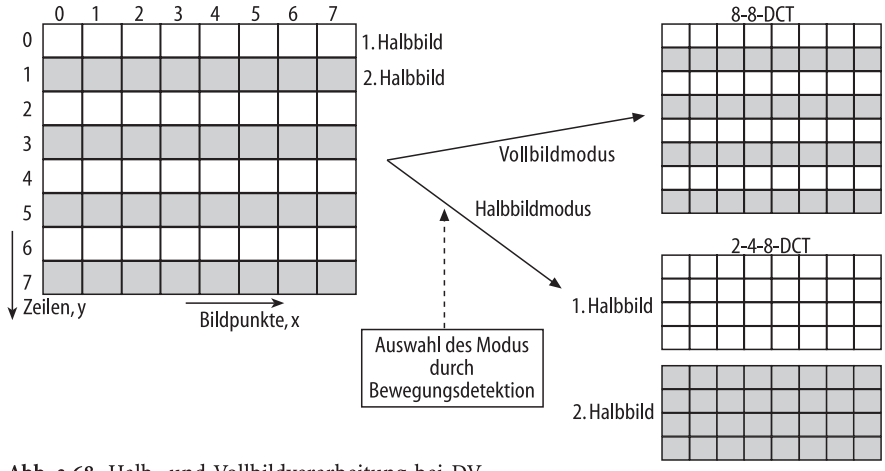


Abb. 3.68. Halb- und Vollbildverarbeitung bei DV

umgeschaltet wird, da sich die größten Ähnlichkeiten innerhalb des Halbbildes ergeben. Die Umschaltung wird von einem Bewegungsdetektor gesteuert (Abb. 3.68) und für die einzelnen Bildbereiche separat vorgenommen, sodass auch Bildvorlagen optimiert werden, die innerhalb des Bildes ruhende und bewegte Teile enthalten.

DV erzielt die eigentliche Datenreduktion mit der einfachen Stufenfolge aus DCT, Quantisierung und anschließender Redundanzreduktion per VLC und RLC. Das Verfahren enthält aber zwei Optimierungen, die darauf zielen, dass eine konstante Datenrate bei möglichst optimaler Bildqualität erzielt wird. Die erste betrifft die Verteilung von Bildbereichen mit feinen und groben Strukturen. Dazu werden je 4 Luminanzblöcke und je ein C_R- und C_B-Block zu einem Makroblock zusammengefasst. Die 45 x 36 Makroblöcke werden dann im sogenannnten Intra-Frame-Shuffling-Prozess systematisch verwürfelt. Dazu werden immer 5 Makroblöcke mit jeweils 384 Bytes (6 x 8 x 8 Bytes) zu einem Superblock der Größe 1920 Bytes zusammengefasst und zwar so, dass die Makroblöcke eines Segments aus weit entfernten Bildbereichen stammen (Abb. 3.69). Damit wird schon vor dem Einsatz des eigentlichen Algorithmus eine Angleichung der durchschnittlichen Redundanz, d. h. eine annähernde Gleichverteilung feiner und grober Bildstrukturen pro Superblock, erzielt und eine effektive Weiterverarbeitung gewährleistet.

Die zweite Optimierung betrifft die Steuerung der Datenrate direkt. Die Datenrate ist zunächst wie üblich vom Bildinhalt abhängig, da bei fein strukturierten Bildblöcken weniger DCT-Koeffizienten zu Null werden als bei Blöcken die nur wenig Unterschiede bei den Grauwerten aufweisen. Der wesentliche Steuerparameter für die Datenrate ist die Quantisierungstabelle mit der bei großer Datenrate auf grobe und bei kleiner Rate auf feine Quantisierung geschaltet werden kann. Während MPEG und andere Verfahren die konstante Datenrate gewöhnlich quasi rückwärtsgerichtet erreichen, indem anhand eines zu voll laufenden Pufferspeichers festgestellt wird, dass die folgenden Daten stärker reduziert bzw. quantisiert werden müssen, wird bei DV zu diesem Zweck mit der sogenannten Feed Forward-Steuerung gearbeitet [55].

Bei diesem Verfahren wird vor der eigentlichen Datenreduktion mit Hilfe einer Vorausberechnung anhand vieler vordefinierter Quantisierungstabellen bestimmt, welche von ihnen die optimale ist. Für jeden Makroblock kann eine

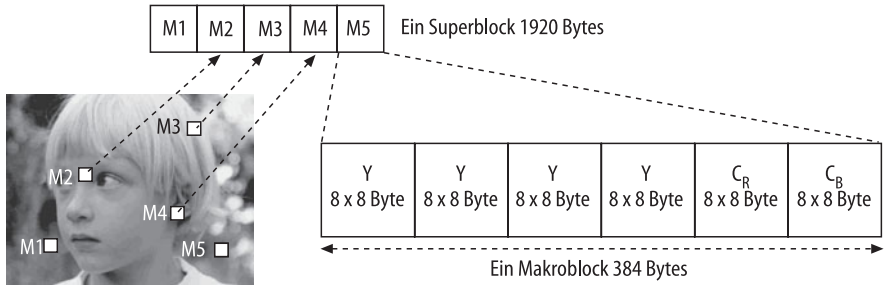


Abb. 3.69. Intraframe Makroblock-Shuffling bei DV

von 64 Tabellen gewählt werden, wobei durch Analyse der Energie der AC-Koeffizienten zunächst eine von vier Tabellenklassen bestimmt wird. Mit den 16 Tabellen der gewählten Klasse werden als virtuelle Kompression nacheinander die verschiedenen Quantisierungen durchgeführt und dann wird diejenige ausgewählt, die eine Byteanzahl erzeugt, das am dichtesten an dem durch das Aufzeichnungsverfahren gesetzten Limit liegt.

Für den DV-Grundalgorithmus beträgt das Limit 77 Bytes pro Makroblock, denn es wurde festgelegt, dass die Ausgangsdatenrate, die sowohl für 4:1:1 als auch für 4:2:0 den Wert 124 416 kbit/s hat, auf ca. 25 Mbit/s reduziert werden soll, um ein komprimiertes Signal hoher Qualität zu erhalten. Der Reduktionsfaktor beträgt etwa 5 und kann nicht variiert werden. Wie oben erwähnt, enthält ein Einzelbild vor der Reduktion 45 x 36 Makroblöcke mit je 384 Bytes bzw. 9 x 36 Superblöcke mit je 1920 Bytes. Mit der Festlegung der Reduktion auf 77 Bytes pro Makro- bzw. 385 Bytes pro Superblock beträgt der Reduktionsfaktor exakt 1920/385 = 4,987. Man spricht bei den reduzierten Daten von 385 Bytes pro Frame. In der reduzierten Ebene gibt es statt der Superblöcke also 9 x 36 Frames pro Bild, die Datenrate ist damit auf 24,948 Mbit/s festgelegt.

Der DV-Algorithmus erzielt mit den beiden genannten Optimierungen sehr gute Ergebnisse, die bei einem Vergleich mit MPEG bei gleicher Datenrate eine geringere Schwankungsbreite der Qualität aufweist, wenn verschiedene Bildsequenzen untersucht werden (s. Abschn. 10.1). Ein weiterer Vorteil gegenüber MPEG ist die einfache Struktur und die feste Datenrate.

Aufgrund des großen Erfolgs im Heimanwenderbereich wurde der DV-Standard fortentwickelt und auch im professionellen Bereich eingesetzt. Unter der Bezeichnung DVCPro, DVCPro25 oder D7 arbeitet er im professionellen Bereich auch beim 625/50 Hz-Standard mit einer Chromaunterabtastung von 4:1:1, während bei der als semiprofessionell bezeichneten Variante DVCam der Algorithmus in seiner Grundform beibehalten wird und diesbezüglich keinen Unterschied zum Heimanwenderbereich aufweist. Bei dem professionellen Format DVCPro50 wird durch die parallele Verwendung zweier Codecs die Datenrate auf 50 Mbit/s verdoppelt. Die zweiten 25 Mbit/s werden so eingesetzt, dass die im professionellen Bereich erforderliche 4:2:2-Struktur erreicht wird. Das erfordert ein Drittel der zweiten 25 Mbit/s. Die restlichen Daten erlauben, den Kompressionsfaktor auf 3,3:1 zu senken. Eine weitere Verdopplung auf 100 Mbit/s ist für HD-Anwendungen ausgelegt (s. Abschn. 8.8).

3.7.1 DV-DIF

Der DV-Algorithmus bezieht sich, ebenso wie JPEG, nur auf das Bild und ist damit nicht wie MPEG für komplette Programmbeiträge ausgelegt, wo neben der Video- auch die Audiocodierung und die Synchronisation erfasst sind. Da in sehr vielen Fällen, auch im Heimanwenderbereich, aber ein Austausch von zusammengehörigen Video- und Audiodaten gewünscht ist, wurde für DV-Daten das spezifische Digital Interface Format DV-DIF (IEC 61883) geschaffen, das z. B. in Verbindung mit der Datenübertragung über die Firewire-Schnittstelle (IEEE 1394, s. Abschn. 10.2 und SMPTE 396) von Bedeutung ist.

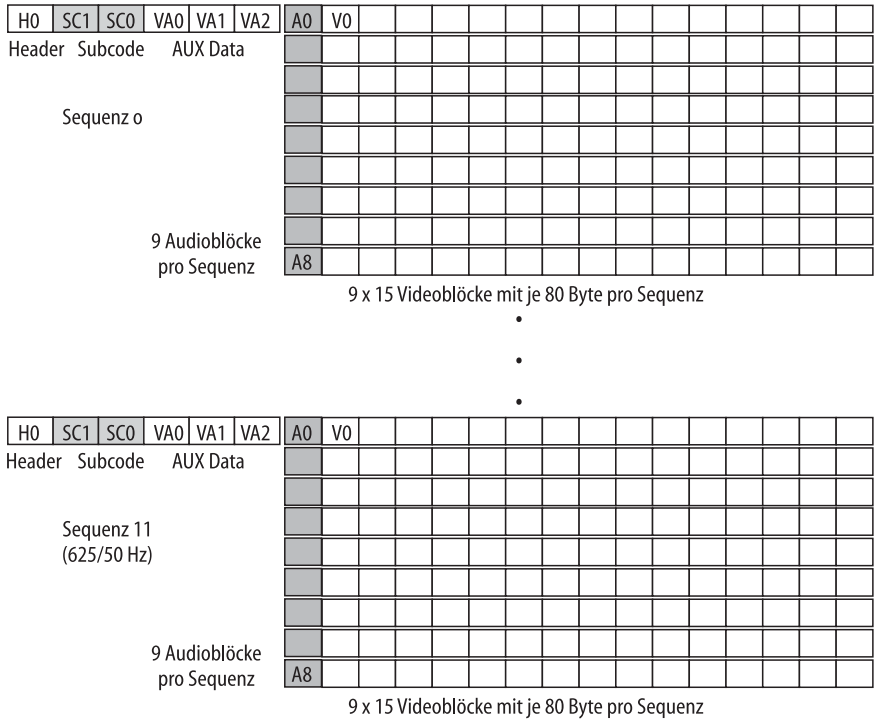


Abb. 3.70. Aufbau eines Bildes nach DV-DIF

Beim DV-DIF-Datenstrom ist das Bild in Blöcke der Größe 80 Bytes eingeteilt, sodass ein DIF-Block die reduzierten 77 Bytes eines Makroblocks aufnehmen kann. Damit das DIF problemlos Videodaten aus 625/50 Hz- und 525/60 Hz-Systemen verarbeiten kann, werden die DIF-Blöcke Sequenzen zugeordnet, die so gebildet werden, dass ein Bild bei 50 Hz in 12 und bei 60 Hz in 10 Sequenzen eingeteilt ist. In Europa enthält eine Sequenz damit 9 x 36 Frames/12, was 9 x 3 Superblöcken bzw. 9 x 15 Blöcken entspricht (Abb. 3.70).

Digitale Audiodaten werden unkomprimiert in den Datenstrom integriert. Es werden dafür bei 40 ms Bilddauer 7680 Bytes pro Bild bzw. 640 Bytes oder 8 Blöcke pro Sequenz gebraucht, denn es wurde festgelegt, dass die Audiodaten entweder zweikanalig mit 48 kHz/16 Bit oder vierkanalig mit 32 kHz/12 Bit geführt werden können. Die DV-DIF-Struktur in Abbildung 3.70 zeigt, dass dafür pro Segment ein zusätzlicher Block für Audiodaten erforderlich ist und eine Sequenz damit aus 15 x 9 Video- und 1 x 9 Audioblöcken besteht. Hinzu kommt eine Erweiterung um 6 Blöcke die als drei zusätzliche Videoblöcke und für Header- und Subcode-Daten zur Verfügung stehen, wobei letztere meistens für Timecode-Daten genutzt werden. Die insgesamt 12 x 150 = 1800 DIF-Audio- und Videodatenblöcke pro Bild führen schließlich zu einer A/V-Datenrate von insgesamt 3,6 MByte/s. Die Speicherung von DV-DIF-Datenmaterial von einer Stunde Dauer erfordert damit eine Kapazität von ca. 13 GByte.

3.8 MPEG

Die Abkürzung MPEG bedeutet Moving Pictures Experts Group und steht für eine Gruppe internationaler Experten, die seit 1988 einen universellen Standard zur Datenreduktion und zum Austausch von Bewegtbildern unter Einbeziehung von Audio- und Zusatzdaten erarbeiten. Die Entwicklung zielte auf eine möglichst umfassende Anwendbarkeit zur Datenübertragung, Speicherung und zum Einsatz in Multimediasystemen ab. Dabei stellt MPEG einen Rahmen dar, den verschiedene Entwickler unterschiedlich füllen können und sorgt vor allem dafür, dass das Format von allen MPEG-konformen Empfängern verstanden wird [217]. Definitionen hierzu und zur Multiplexbildung befinden sich im MPEG-Teil »Systems«. Zwei weitere Teile, nämlich »Video« und »Audio« beschreiben die Video- und Audiocodierung. Bei MPEG-2 wird für die drei Teile auch die Standardbezeichnung ISO/IEC 13818-1 bis -3 verwendet.

Die MPEG-Standards sind in mehreren Stufen festgelegt worden (Abb. 3.71). MPEG-1 wurde zunächst im Hinblick auf kleine Datenraten bis 1,5 Mbit/s (CD-ROM-Anwendung) entwickelt, unter Inkaufnahme einer relativ schlechten Bildauflösung und -qualität. Die Definition von MPEG-2 berücksichtigt die spezifischen Eigenschaften von Videosignalen und lässt bei Standardauflösung Datenraten bis zu 15 Mbit/s zu, womit eine sehr gute Bildqualität erreicht wird. Mit MPEG-3 sollten anschließend die Definitionen für hochauflösendes Fernsehen HDTV festgelegt werden. Die HD-Parameter wurden im Laufe der Entwicklung aber bereits in MPEG-2 integriert, sodass MPEG-3 nicht existiert. Wenn im Zuge der Audio-Datenreduktion von einer MP3-Codierung gesprochen wird, so ist hier der Layer 3 der MPEG-2-Audiocodierung gemeint.

MPEG-4 zeichnet sich dadurch aus, dass neben einer effizienteren Codierung mit Hilfe der Beschreibung von Objekten ein interaktiver Zugriff auf Bildteile ermöglicht wird. MPEG-7 und MPEG-21 beziehen sich nicht mehr auf die Bildcodierung, sondern auf die Verwaltung der audiovisuellen Daten. Die MPEG-Definitionen sind abwärtskompatibel ausgelegt, ein MPEG-4-Decoder soll also z. B. in der Lage sein, auch einen MPEG-2-Datenstrom zu decodieren. Die Variante MPEG-2 hat heute eine sehr große Bedeutung, z. B. als Codierverfahren für das digitale Fernsehen DVB und für die DVD. Wenn hier im Folgenden von MPEG die Rede ist, ist daher zunächst MPEG-1/2 gemeint. MPEG-4 ist deutlich unterschieden und, bezogen auf die Interaktivität, sehr komplex.

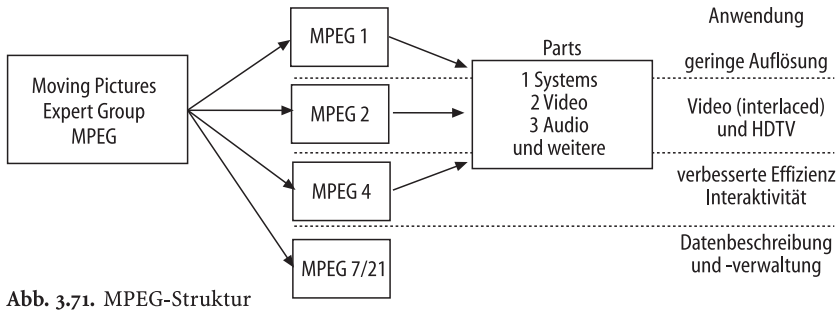


Abb. 3.71. MPEG-Struktur

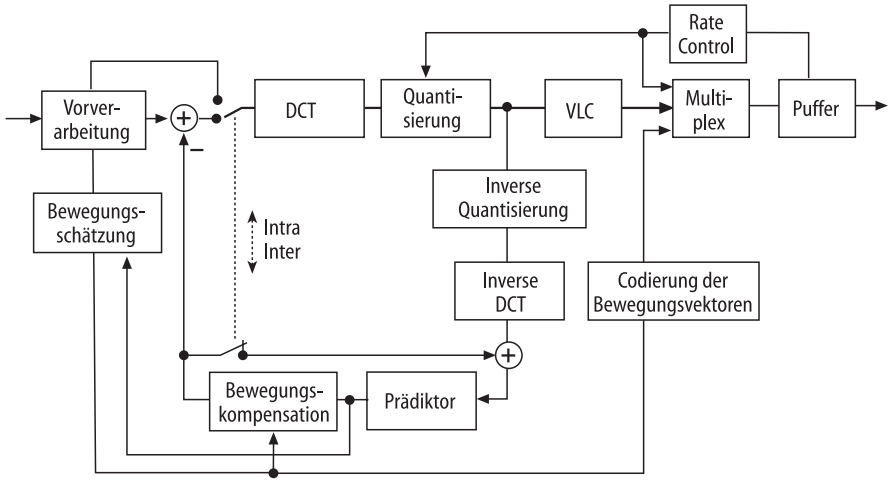


Abb. 3.72. Blockschaltbild des MPEG-Video-Encoders

3.8.1 MPEG-Videocodierung

Die MPEG-Videocodierung beruht auf der hybriden DCT, bei der im Vergleich zu JPEG und DV eine höhere Datenreduktionseffizienz erreicht wird, da sowohl die örtliche Korrelation zwischen den Bildpunkten durch die DCT ausgenutzt wird, als auch die Ähnlichkeit der Bilder in der Zeitdimension durch die DPCM. Abbildung 3.72 zeigt das Blockschaltbild des MPEG-Video-Encoders. Er wird in dieser Form implementiert, ist aber so nicht bei MPEG festgelegt, da dieser Standard nur den Decoder genau beschreibt. Wie bei JPEG ist im Kern eine DCT mit folgender Quantisierungsstufe enthalten, an die sich die Redundanzreduktionsstufe mit VLC und RLC anschließt. Damit die zur Quantisierung benutzten Tabellen nicht immer im Datenstrom zum Decoder übertragen werden müssen, sind für MPEG die in Abbildung 3.73 dargestellten Standardtabellen festgelegt [174] worden, es besteht aber kein Zwang, sie zu benutzen.

Nach Durchlaufen der Quantisierungsstufe wird das Signal verzweigt, und mit reziproker Gewichtung und inverser DCT werden die Bildpunktwerte zu-

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Quantisierungstabelle für die Luminanz

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

Quantisierungstabelle für die Chrominanz

Abb. 3.73. Standard-Quantisierungstabellen bei MPEG

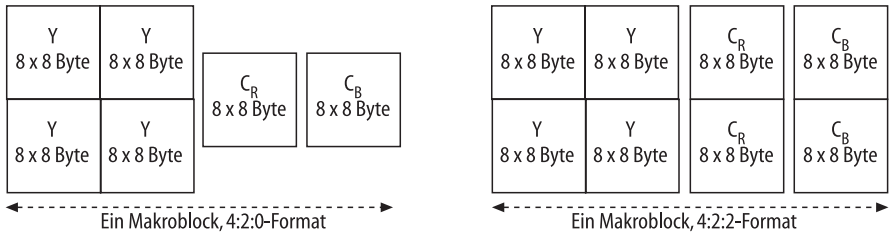


Abb. 3.74. Makroblock-Strukturen bei MPEG

rückgewonnen. Wenn keine I-Framecodierung gefordert ist, wird im Encoder aus dieser Information ein Prädiktionsbild ermittelt, woraus, zusammen mit dem nächsten Bildblock am Codereingang, die DPCM-Differenz gebildet wird. Wie bereits im Abschnitt über DPCM erläutert, kann die Prädiktion durch eine Bewegungskompensation optimiert werden. MPEG unterstützt diesbezüglich auch die dort beschriebene bidirektionale Prädiktion, die auf den Nachbarbildern vor und hinter dem aktuellen Bild beruht. Hierdurch steigt die Datenreduktionseffizienz, da ein bidirektional prädizierter Block mit etwa der Hälfte eines unidirektional prädizierten codiert werden kann.

Die Informationen zur Bewegungskompensation stammen aus einer Verarbeitungsstufe zur Bewegungsschätzung, die meist mit dem Blockmatching-Verfahren arbeitet. Das Verfahren selbst ist aber bei MPEG nicht festgelegt. Die Informationen gelangen in Form eines Bewegungsvektors nicht nur zum Prädiktor, sondern auch an die Daten-Multiplexstufe am Ausgang, da sie auch für den Decoder zur Verfügung stehen müssen und daher mit im Datenstrom übertragen werden (Abb. 3.72).

Das Ausgangsformat vor der eigentlichen Datenreduktion ist ein Komponentensignal, in den meisten Fällen mit 4:2:0-Unterabtastung. Um nicht für Luminanz- und Chrominanzblöcke separate Bewegungsvektoren ermitteln zu müssen, wird wie bei DV aus vier Luminanzblöcken und je einem für C_R und C_B ein Makroblock gebildet, der die Basis für die Bewegungsschätzung bildet. Für die speziell für den professionellen Bereich entwickelte MPEG-2-Variante 4:2:2, wird die Anzahl der Chrominanzblöcke verdoppelt (Abb. 3.74).

Neben den Ausgangsdaten und den Bewegungsvektoren wird für jeden Block eine Information über die Steuerung des Quantisierungsfaktors mittels der Multiplexstufe am Ausgang in den Datenstrom integriert. Die Steuerdaten werden aus einer Pufferstufe am Encoder-Ausgang gewonnen, die die Aufgabe hat, den bildinhaltsabhängig variierenden Datenstrom zu glätten. Falls nun eine lange Folge komplexer Bilder codiert werden muss, wird der Datenstrom so groß, dass der Pufferspeicher überzulaufen droht. Um diesen Zustand zu verhindern, wird die Quantisierungsstufe vom Pufferfüllstand gesteuert. Dabei werden für jeden Block alle Einträge in der Quantisierungstabelle, mit Ausnahme des Wertes für die DC-Komponente bei intraframe-codierten Bildern, mit einem gemeinsamen Faktor variiert. Zur Senkung der Datenrate werden also die Werte in der Tabelle erhöht, sodass die Division der Koeffizienten durch diese Werte eine gröbere Quantisierung bewirkt. Falls auch die Teilung durch den Wert 31 nicht ausreicht, werden die Makroblöcke übersprungen.

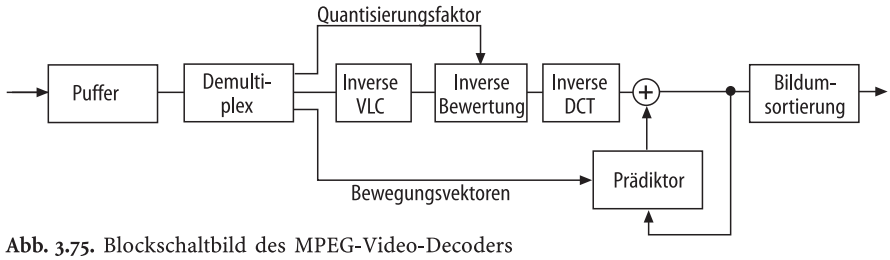


Abb. 3.75. Blockschaltbild des MPEG-Video-Decoders

3.8.1.1 MPEG-Decoder

Abbildung 3.75 zeigt das Blockschaltbild des MPEG-Videodecoders. Auch hier werden die Daten zunächst in einen Pufferspeicher übernommen. Im Demultiplexer werden dann die Bewegungsvektoren und die Quantisierungsinformationen abgetrennt und dem Prädiktor, bzw. der Verarbeitungsstufe zur inversen Quantisierung zugeführt. Nachdem die Koeffizienten der DCT zurücktransformiert sind und die Vorhersagewerte zur Invertierung der DPCM aufaddiert wurden, steht die Bildinformation wieder zur Verfügung. Der Prozess der Decodierung ist damit erheblich einfacher als der der Encodierung, vor allem weil der Rechenaufwand zur Bewegungsschätzung entfällt. MPEG definiert damit ein unsymmetrisches Codierungsverfahren, das den Vorteil hat, dass die Decoder in den vielen Endgeräten erheblich weniger aufwendig und teuer sind als die Encoder in den wenigen Geräten auf der Sendeseite.

3.8.1.2 Group of Pictures

Am Ausgang des Decoders wird, genau wie vor der Encodierung, eine Stufe zur Bildumsortierung durchlaufen. Diese ist erforderlich, weil bei MPEG i. d. R. nicht alle Bilder gleich codiert werden. Die effizienteste Form ist die, die auf bidirektionaler Prädiktion beruht, da die DCT ja auf Differenzbilder angewendet wird und die Differenzen hier am geringsten sind. Da die B-Codierung aber von den Nachbarbildern abhängen, kann sie nicht für alle Bilder durchgängig beibehalten werden, sondern im Datenstrom müssen auch unidirektional prädizierte P-Frames und auch I-Frames auftauchen, die nicht von den Nachbarbildern abhängen, da ein Decoder sonst keinen Einstiegspunkt finden kann. Bei MPEG werden die I-, B- und P-Frames periodisch wiederholt. Auf diese Weise entstehen zyklisch wiederkehrende Bildgruppen, die als Group of Pictures (GOP) bezeichnet werden. Die GOP-Struktur wird bei der Encodierung festgelegt, je mehr B- und P-Frames auftauchen, desto effektiver ist die Datenreduktion. In der Praxis wird oft mit einer Folge von 12 Bildern mit einem I-, drei P- und viermal zwei B-Frames gearbeitet (Abb. 3.76), sodass ein Einstieg in eine Sequenz bei 25 fps innerhalb einer halben Sekunde möglich ist. Die Bildreihenfolge wird dann für die Übertragung mit der erwähnten Bildumsortierungsstufe verändert, denn dem Decoder müssen zuerst die Nachbarbilder eines B-Frames bekannt sein, damit dieses decodiert werden kann [94]. Für die Verarbeitung mehrerer benachbarter B-Frames müssen viele Bilder gespeichert werden, das geht mit einer entsprechend großen Verzögerungszeit einher, die die Verwendung sehr langer GOP einschränkt.

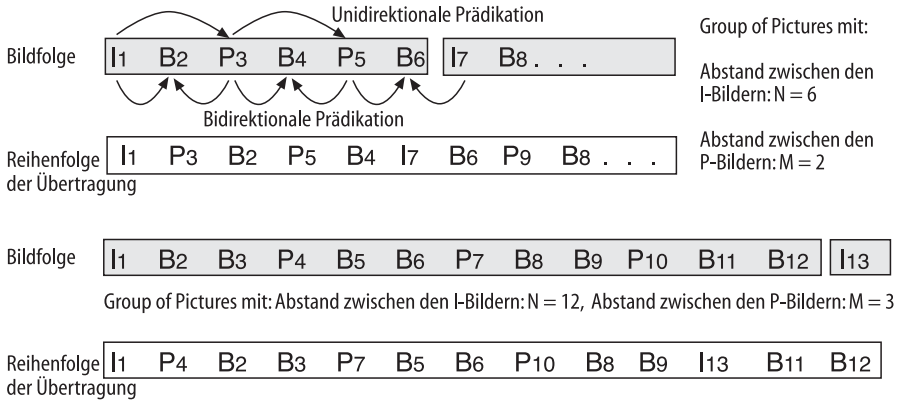


Abb. 3.76. Verschiedene Formen von Groups of Pictures

Um die Effizienz der Verwendung zeitlich abhängiger Bilder abzuschätzen, kann man z. B. den Vergleich mit einem DV-Datenstrom heranziehen. Hier liegen bei 25 Mbit/s nur I-Frames vor, die mit 5:1 reduziert sind, sodass jedes Bild 1 Mbit benötigt und dabei eine hohe Qualität aufweist. Wenn man nun annimmt, dass bei MPEG jedes P-Bild 1/3 des I-Frames und jedes B-Frame die Hälfte eines P-Bildes umfasst, so entsteht für die GOP 12 ein Datenumfang von $(1 + 3/3 + 8/6) \cdot 1$ Mbit = 3,3 Mbit bzw. eine Datenrate von nur 6,9 Mbit/s bei einer Bildqualität, die mit dem DV-Ausgangsmaterial vergleichbar ist.

Die GOP (Abb. 3.76) stellt nur eine von mehreren Organisationseinheiten dar. Der MPEG-codierte Videodatenstrom wird in weitere Schichten (Layer) aufgeteilt, damit die verschiedenen Elemente des Codierverfahrens, die in unterschiedlichster Weise miteinander kombinierbar sind, einfach getrennt und somit möglichst einfach aufgebaute Decoder realisiert werden können. Alle Schichten haben eigene Startcodes und Headerinformationen.

Die oberste Schicht ist der Sequenz-Layer (Abb. 3.77), in dem alle Bilder einer Videosequenz bzw. eines Programms enthalten sind. Darunter befindet sich der Layer, in dem die Group of Pictures definiert sind. Der Header die-

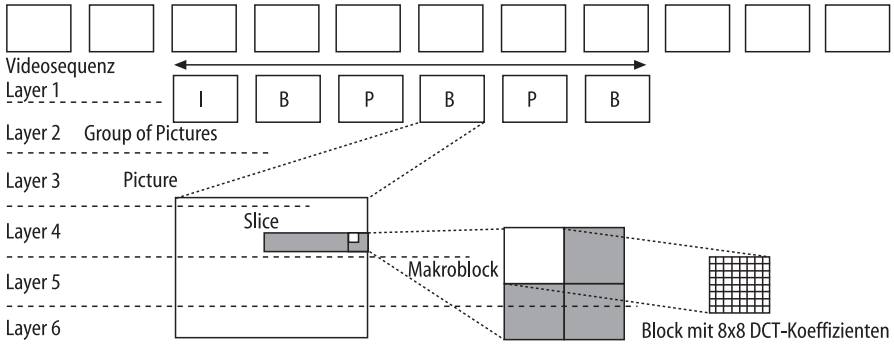


Abb. 3.77. Struktur des Datenstroms bei MPEG-2

Startcode: **00 00 01 B3**

PARAMETER	BINARY CODE	VALUE	MEANING
horizontal_size	000101100000	352	picture width
vertical_size	000011110000	240	picture height
aspect_ratio_information	1100	0.9132	pel width:height
frame_rate_code	0101	30	frames per second
bit_rate_value	00000010110011011	1150000	bits per second
marker_bit	1	1	always 1
vbv_buffer_size	0000010100	40960	buffer size (Bytes)
constrained_parameter_flag	1	1	flag
load_intra_quant_matrix	0	0	if 1, load new matrix
load_non_intra_quant_matrix	0	0	if 1, load new matrix

Abb. 3.78. Beispiel für einen Header im MPEG-GOP-Layer

ser Schicht enthält Informationen über die Videoparameter, d. h. Angaben zur Pixelanzahl für Bildbreite und -höhe, zum Pixel-Aspect-Ratio, also dem Seitenverhältnis des einzelnen Bildpunktes, sowie der (Halb-)Bildfrequenz und weiter Angaben zu Bitstream-Parametern wie Bitrate, Puffergröße und Angaben zur Quantisierungstabelle. Abbildung 3.78 zeigt ein Beispiel.

Die Einzelbilder befinden sich im Picture-Layer. Im zugehörigen Header gibt es Informationen über den Timecode, also zu der absoluten Bildnummerierung und GOP-Parameter. Das Einzelbild ist wiederum in Slices (Scheiben) unterteilt, die vertikal 16 Bildpunkte, also die Makroblockgröße umfassen. Horizontal beträgt die Slice-Größe wenigstens auch 16 Pixel, kann sich aber auch über die ganze Bildbreite erstrecken. Im Header befindet sich hier die Angabe über den Typ, d. h. ob der Makroblock I-, P- oder B-codiert ist. Auf der untersten Ebene liegt schließlich der Block-Layer, der neben den eigentlichen codierten Daten auch Angaben zur Quantisierung enthält.

Die strukturellen Eigenschaften einer MPEG-Bildsequenz seien hier anhand eines Beispiels noch einmal illustriert: Die dazu gewählte Bildfolge ist ein GOP-Teil mit vier Bildern, die I-, B-, B- und P-codiert sind. Sie entstammen der Videosequenz „Flowergarden“, die neben anderen, wie „Mobile and Calendar“ oder „Barcelona“, oft als Vorlagen zur Beurteilung der Qualität von Videocodierverfahren benutzt werden. Die hier verwendete Sequenz zeigt eine stetige Kamerafahrt nach rechts. Im Vordergrund befindet sich ein fein strukturiertes Blumenfeld und ein Baum, der Hintergrund zeigt einige Häuser (Abb. 3.79). Da die Bilder zeitlich direkt benachbart sind, ist die geringe örtliche Bewegung nicht ganz einfach, aber doch deutlich an der Relation zwischen dem Hintergrund und dem Baum im Vordergrund zu erkennen.

Die Einzelbilder haben eine Low-Level-Auflösung, verwendet wurde ein Bildausschnitt, der der linken oberen Bildecke entnommen ist. In dieser Darstellung werden die 16 x 16-Blöcke deutlich sichtbar. Neben der Blockdarstel-

Abb. 3.79. Beispielfolge aus der Bildsequenz „Flowergarden“

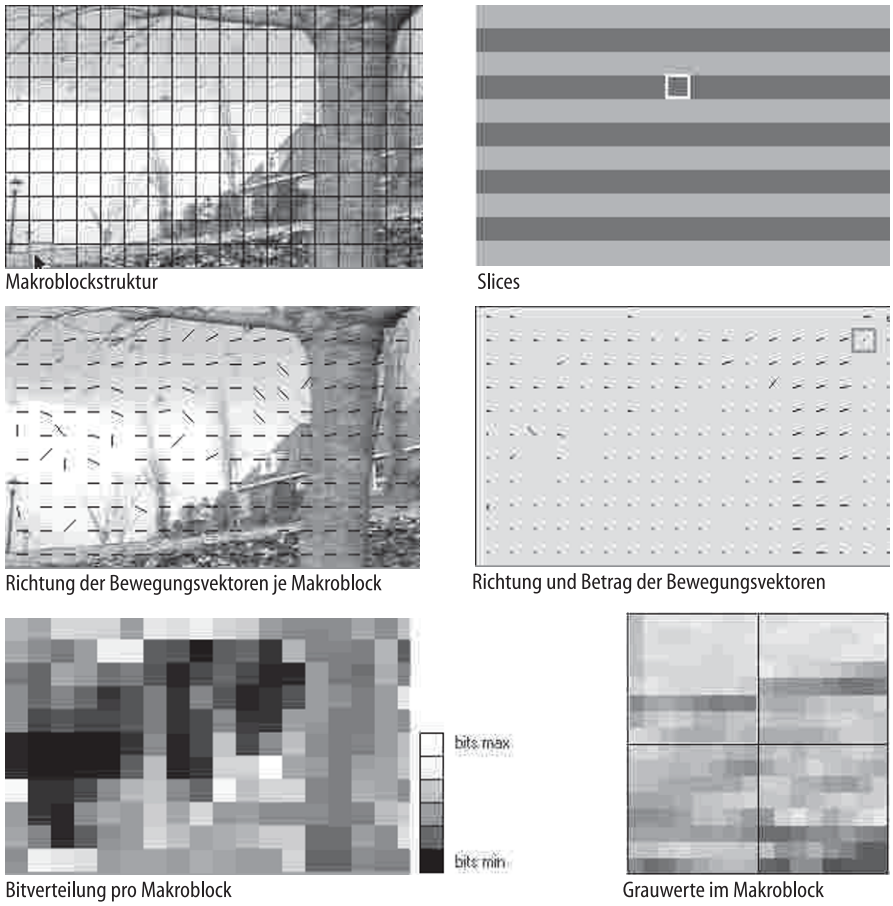


Abb. 3.80. Beispiel für MPEG-Stukturen in der Bildsequenz „Flowergarden“

lung befindet sich die Einteilung des Bildes in Slices (Abb. 3.80 oben). Darunter sind die Bewegungsvektoren für jeden Bildblock dargestellt, und zwar links nur die Vektorrichtung und rechts, ohne Hintergrundbild, Vektorrichtung und Betrag. Es wird deutlich, dass sich die Blöcke, die den Baum im Vordergrund zeigen, am stärksten verändern und mit dem größten Betrag verbunden sind (Abb. 3.80 mitte). Die Richtung ist aufgrund der Kamerafahrt für alle Blöcke gleich, doch kann sie in einigen Fällen aufgrund zu wenig strukturierter Blokkinhalte nicht eindeutig erkannt werden.

In Abbildung 3.80 wird unten schließlich die Verteilung der Bits, die pro Makroblock aufgewandt werden, als Grauwert mit steigender Helligkeit repräsentiert, und daneben ist der zweite Makroblock von links am unteren Bildrand (Pfeil im Bild oben links) mit seinen verschiedenen Grauwerten dargestellt. Dieser unterscheidet sich in allen vier Bildern nur wenig und benötigt im I-Bild 927 Bit, in den B-Bildern 151 Bit bzw. 141 Bit und im P-Bild 541 Bit.

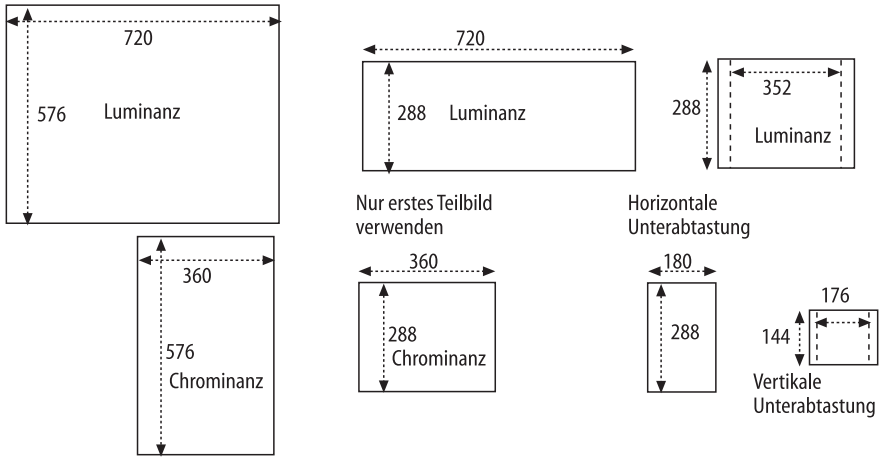


Abb. 3.81. Vorverarbeitung durch Unterabtastung bei MPEG-1

3.8.1.3 MPEG-1

Die MPEG-Codierung wurde zunächst als MPEG-1 im Hinblick auf kleine Datenraten von bis zu 1,5 Mbit/s ausgelegt. Eine Anwendung war z. B. das CD I-Format (Compact Disc Interactive). MPEG-1 arbeitet mit der hybrid-DCT mit Blöcken aus 8 x 8 Bildpunkten, die Datenrate erlaubte nur die Übertragung gering aufgelöster Bilder. In einer Vorverarbeitungsstufe werden die Abtastwerte daher horizontal und vertikal halbiert. MPEG-1 bezieht sich auf progressiv abgetastete Bilder mit 352 x 288 Bildpunkten bei 50 Hz Bildwechselfrequenz, bzw. 352 x 240 Punkten bei 60 Hz. Abbildung 3.81 zeigt, wie diese, als Source Input Format (SIF) bezeichnete, Struktur durch Unterabtastung aus dem digitalen Komponentensignal gewonnen wird. Die Abtrennung der Pixel am Bildrand ist erforderlich, um für die Makroblöcke eine durch 16 bzw. 8 teilbare Zahl zu erhalten. Auf der Wiedergabeseite werden die 720 x 576 Bildpunkte durch Interpolation aus den übertragenen Werten zurückgewonnen.

3.8.1.4 MPEG-2

Auch die Codierung nach MPEG-2 beruht auf der in diesem Abschnitt allgemein dargestellten MPEG-Videocodierung auf Basis der hybriden DCT mit Blöcken aus 8 x 8 Bildpunkten. Im Gegensatz zu MPEG-1 sind aber mehrere Bildauflösungen zwischen dem SIF von MPEG-1 bis hin zur HD-Auflösung zugelassen. Neben der am häufigsten angewendeten 4:2:0-Abtastung sind auch die Formate 4:2:2 oder 4:4:4 erlaubt.

Darüber hinaus wird die Verarbeitung von Halbbildern unterstützt. Wie beim DV-Algorithmus kann bei der Encodierung zwischen dem Voll- und dem Teilbildmodus gewählt werden. Der Vollbildmodus nutzt die Ähnlichkeiten der Nachbarzeilen aus zwei Halbbildern, die bei ruhender Bildvorlage auftreten und beruht daher auf 8 x 8-Bildblöcken, die aus beiden Halbbildern stammen. Bei starker Bewegung im Blockbereich sind die Halbbilder meist so stark unterschieden, dass es günstiger ist, auf die Teilbildverarbeitung überzugehen

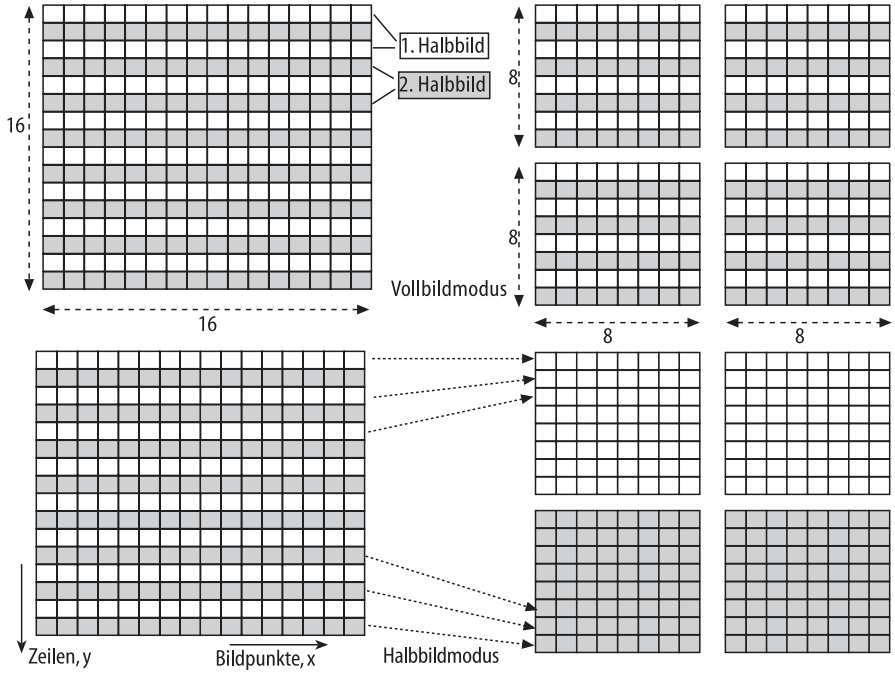


Abb. 3.82. Makroblockaufteilung in 8 x 8-DCT-Blöcke zur Halbbildverarbeitung

und die Zeilen entsprechend Abbildung 3.82 umzuordnen, sodass nur die Ähnlichkeiten innerhalb des Halbbildes ausgenutzt werden.

Als universeller Standard erlaubt MPEG-2 die Codierung von Videobildern mit Auflösungen auf dem Niveau MPEG-1, mit Standard-Auflösung (720 x 576 Bildpunkte) bis hin zu HD-Auflösung (1920 x 1152 Bildpunkte). Auch bezüglich weiterer Funktionen stellt MPEG-2 eine so reiche Auswahl an Möglichkeiten zur Codierung zur Verfügung, dass dafür gesorgt wurde, diese auf einige wenige, häufig verwendete Funktionalitäten (Profile) bei verschiedenen Auflösungsstufen (Levels) zu beschränken. Damit wird es möglich, den unterschiedlichen Aufgaben auf einfache Weise gerecht zu werden und klare Anforderungen an die Leistungsfähigkeit der Decoder zu formulieren. Fünf verschiedene Profile werden zusammen mit vier Stufen (Level) der Bildauflösung für Low bis High Definition-TV in einer Matrix angeordnet. Abbildung 3.83 zeigt, welche Kombinationen dabei zulässig sind. Aus der Abbildung ist auch die jeweils maximal erlaubte Datenrate (nicht die typische), sowie die Anzahl der verwendeten Bildpunkte ersichtlich. Weiterhin wird deutlich, wie die Farbdifferenzsignale relativ zum Luminanzanteil abgetastet werden (meist 4:2:0).

Der Low Level entspricht einer durch MPEG-1 festgelegten Bildauflösung. Es ist aber eine höhere Datenrate bis zu 4 Mbit/s erlaubt, jedoch wird keine biddirektionale Prädiktion unterstützt. Der Main Level zielt auf die Mehrzahl der Videoanwendungen ab, die Auflösung entspricht der heutigen Standardauflösung. Die beiden High Levels mit den hohen Auflösungen sind für HD-

Profile Level	Simple Profile 4:2:0 (no B-Frames)	Main Profile 4:2:0	Professional Profile 4:2:2	Scalable Profile 4:2:0	High Profile 4:2:0 or 4:2:2
High Level < 60 fps		1920 x 1152 < 80 Mbit/s			1920 x 1152 < 100(80) Mbit/s
High 1440 < 60 fps		1440 x 1152 < 60 Mbit/s		1440 x 1152 (Spat.) < 60 (40) Mbit/s	1440 x 1152 < 80 (60) Mbit/s
Main Level < 30 fps	720 x 576 < 15 Mbit/s	720 x 576 < 15 Mbit/s	720 x 576 < 50 Mbit/s	720 x 576 (SNR) < 15 (10) Mbit/s	720 x 576 < 20 (15) Mbit/s
Low Level < 30 fps		352 x 288 < 4 Mbit/s		352 x 288 (SNR) < 4 (3) Mbit/s	

Werte in Klammern geben die Typen (Spatially oder SNR) und Alternativen bezgl. der Skalierung an

Abb. 3.83. Profiles und Levels bei MPEG-2

Anwendungen konzipiert (MP@H14L und MP@HL). Das Simple Profile richtet sich auf einfachste Lösungen ohne die MPEG-2-spezifische bidirektionale Prädiktion. Das Main Profile at Main Level (MP@ML) mit einer Auflösung von 720 x 576 Pixeln hat die größte Bedeutung, es bietet sich bei mäßigem technischen Aufwand für eine Vielzahl von Anwendungen an. Ein mit Standard-PAL-Qualität vergleichbares Signal erfordert hier eine Datenrate von ca. 3 ... 5 Mbit/s, die maximal erlaubte Datenrate beträgt 15 Mbit/s. Das High Profile zielt auf hochqualitative Anwendungen, hier kann insbesondere mit erhöhter vertikaler Farbauflösung (4:2:2) gearbeitet werden.

Zwischen Main und High Profile sind die skalierbaren Profile angeordnet (Abb. 3.83). Skalierbare Profile werden unterschieden in eine Skalierung bezüglich des Signal-Rauschspannungsabstandes (SNR) und der Ortskoordinaten (Spacially). Die skalierbare SNR-Codierung ist für die terrestrische Ausstrahlung von Videobildern gedacht. Sie soll das typische »Alles oder Nichts« der Digitaltechnik verhindern, d. h., dass das System bei schwankender Übertragungsqualität entweder gut funktioniert oder gar nicht. Stattdessen soll sich ein Verhalten wie bei analogen Systemen einstellen, die Bildqualität soll also bei verschlechtertem Signal-Rausch-Verhältnis allmählich und nicht abrupt abnehmen (Graceful Degradation, Abb. 3.84). Die SNR-Skalierbarkeit wird erreicht, indem zwei Datenströme mit Bildern gleicher Auflösung aber un-

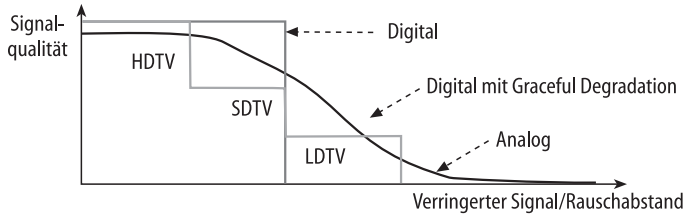


Abb. 3.84. Graceful Degradation

terschiedlicher Störsicherheit erzeugt werden. Zunächst wird eine DCT mit grober Quantisierung gewählt und damit ein Datenstrom für Bilder mit geringer Qualität erzeugt, für den eine sehr störsichere Übertragungsart gewählt wird. Daneben wird der Quantisierungsfehler berechnet und codiert in einem zweiten Datenstrom übertragen. Die Signalqualität am Empfänger ist davon abhängig, ob der Signal-Rauschabstand ausreicht, beide Datenströme auszuwerten oder nicht.

Die räumliche (Spatially) Skalierung bezieht sich auf unterschiedliche Bildauflösungen. Damit soll eine hierarchische Codierung ermöglicht werden, die aus nur einem Datenstrom in Abhängigkeit vom Empfänger mehrere Qualitätsstufen zwischen LDTV und HDTV ableiten kann. Bezüglich eines terrestrischen Übertragungskanals mit 8 MHz Bandbreite sind vier Qualitätsstufen vorgesehen: HDTV (High Definition), EDTV (Enhanced Definition), SDTV (Standard Definition) und LDTV (Low Definition).

Vor dem Hintergrund einer erfolgreichen Einführung von MPEG-2 und des darauf beruhenden DVB-Verfahrens (Digital Video Broadcasting) zur digitalen Signalverteilung an den Konsumenten entstand der Wunsch, den MPEG-Standard auch für Studioanwendungen nutzen zu können. Dafür ist eine Erweiterung vorgesehen, die mit 4:2:2- oder Professional Profile bezeichnet wird und in der üblichen Darstellung der Profiles und Levels oft nicht enthalten ist. Die wichtigsten Aspekte sind dabei die Unterstützung einer 4:2:2-Abtastung und die Nutzung der Zeilen in der vertikalen Austastlücke (720H x 608V) [80]. Um geringe Qualitätsverluste und den Zugriff auf jedes Einzelbild zu ermöglichen, werden die »I-Frame-only«-Codierung oder die »I-B-I-B«-Codierung mit Datenraten von bis zu 50 Mbit/s genutzt. (s. Kap. 10).

Die Videodatenübertragung mit MPEG-2 ermöglicht eine hohe Bildqualität bei geringer Datenrate. Im Laufe der Jahre wurde die Qualität der Coder stetig verbessert, sodass die erforderliche Datenrate auch mit der Aktualität der Implementierung zusammen hängen kann. Eine Qualität, die mit der Wiedergabe von im VHS-System aufgezeichneten Bildern vergleichbar ist, ist mit 2 ... 4 Mbit/s erreichbar. Für eine mit dem Standard-PAL-Signal vergleichbare Qualität rechnet man, wie gesagt, mit 3 ... 6 Mbit/s. MPEG-codierte HDTV-Signale erfordern eine Datenrate von ca. 12 ... 30 Mbit/s.

3.8.1.5 MPEG-4

Nach der Standardisierung von MPEG-1 im Jahre 1992 konnte bereits 1994 die Arbeit an MPEG-2 abgeschlossen werden, der Variante, die über die Anwendung bei DVB und DVD der erste weltweit erfolgreiche Codierstandard wurde. Da die für MPEG-3 anvisierten HD-Definitionen bereits in MPEG-2 eingegangen waren, arbeitete die Arbeitsgruppe an MPEG-4 weiter. Damit sollte zunächst vor allem ein Codierverfahren für kleine Datenraten entstehen, es wurde aber sehr schnell erheblich weiter ausgebaut. Niedrige Datenraten werden bei verbesserter Codiereffizienz erzielt, was das erste Ziel des neuen Standards war.

Als zweiter wichtiger Bereich kam die objekt- und inhaltsbasierte Codierung hinzu, die es den Anwendern erlaubt, interaktiv mit separaten Audio- und Videoobjekten umzugehen. Dabei sind örtliche und zeitliche Skalierbarkeit vor-

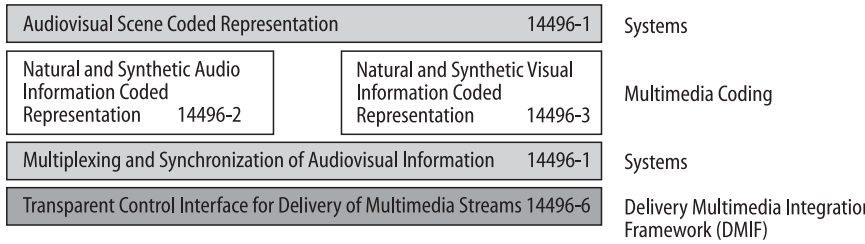


Abb. 3.85. Struktur bei MPEG-4

gesehen, sodass ein Videobild z. B. nicht auf eine rechteckige Form beschränkt sein muss. Neben natürlichen Bildern sind besondere Codierformen für Texturen und künstlich erzeugte, synthetische Video- und auch Audioobjekte definiert (Synthetic Natural Hybrid Codierung, SNHC).

Als drittes wurde besonderes Augenmerk auf die Übertragung der Daten gerichtet. Neben universellen Zugriffsmöglichkeiten soll eine Robustheit in stark gestörter Umgebung ebenso erreicht werden wie eine einfache Anpassung an verschiedene Netzwerke und Speichermedien. Hinzu kommt die Systemebene, die bei MPEG-4 eine große Bedeutung hat weil hier neben der Synchronisation auch die Zugriffsmöglichkeiten für den interaktiven Umgang abgedeckt werden müssen.

Abbildung 3.85 zeigt eine Übersicht über MPEG-4, in der diese Struktur sichtbar wird. Die verschiedenen Teile des Standards, der die offizielle Bezeichnung ISO/IEC 14496 trägt, sind darin nummeriert. Die Teile für die Audio- und Videocodierung (ISO/IEC 14496-2 und -3) umfassen die Definitionen für synthetische und natürliche Objekte. Die Systemebene (ISO/IEC 14496-1) ist zwei geteilt. Sie erfasst einerseits die Darstellung der audiovisuellen Szene, die aus mehreren Audio- oder Videoobjekten bestehen kann und bestimmt andererseits das Datenmultiplexing und die Synchronisation. Schließlich gibt es einen separaten Teil (ISO/IEC 14496-6) in dem über den DMIF-Layer (Delivery Multimedia Integration Framework) die Anpassung an verschiedene Netzwerke und Verteilsysteme geregelt wird.

Die Vielzahl und Komplexität Anforderungen bei MPEG-4 führte dazu, dass wie bei MPEG-2 versucht wird, eine Struktur der Funktionalitäten mit Hilfe von Profilen zu erzeugen. Damit muss nicht in alle Geräte die volle Funktionalität des Standards implementiert werden, sondern es kann auf eine Beschränkung auf die des Profils vorgenommen werden. Abbildung 3.86 zeigt eine Übersicht über die Profile für den Bereich MPEG-4-Visual, gültig für den Entwicklungsstand im Jahre 2004. Zukünftige Erweiterungen von MPEG-4 können sich über weitere Profile niederschlagen.

Das Simple Profile hat etwas Bedeutung bei der Verwendung von Handheld-Geräten etc. erhalten, denn es kann z. B. sehr gut für mobile Videodienste verwendet werden, da es für geringe Datenraten vorgesehen ist. Dabei werden nur einfache, rechteckige Bilder unterstützt, die nicht skalierbar sind. Coder und Decoder sind hier möglichst einfach gehalten, dafür wird für eine sehr robuste Übertragung Sorge getragen.

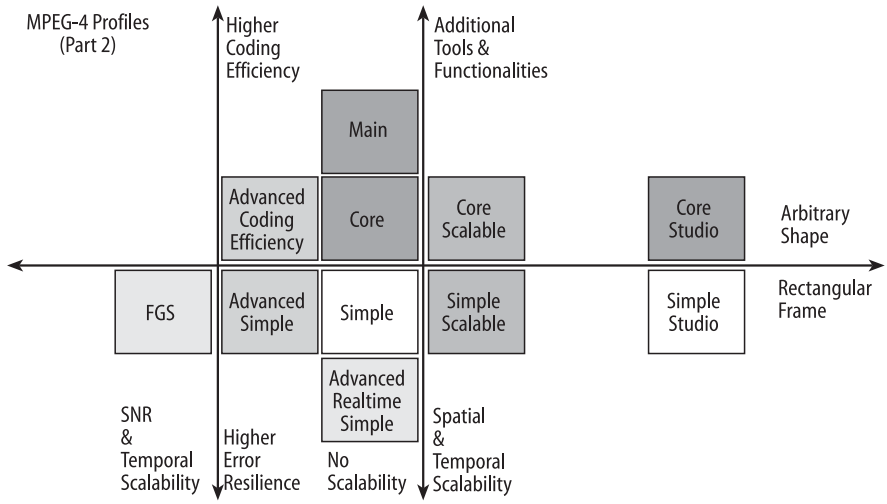


Abb. 3.86. Profiles bei MPEG-4 Visual

Die verschiedenen Core-Profile oberhalb der Mittelinie in Abbildung 3.86 unterstützen die erweiterten Möglichkeiten des Standards, wie die Abweichung von rechteckigen Bildformen und den interaktiven Zugriff. Das Advanced Simple Profile ist für Broadcastanwendungen gedacht, allerdings hat die Codierung nach H.264/AVC für diesen Zweck mehr Marktrelevanz erhalten, da sie eine höhere Codiereffizienz aufweist. H.264/AVC wurde später in MPEG-4 integriert. Als weiteres Profil, für das gegenwärtig Marktchancen gesehen werden, ist das Simple Studio Profile zu nennen, was für HD-Anwendungen mit Datenraten um ca. 400 Mbit/s angewandt wird.

Neben der Unterteilung in Profiles wird bei MPEG-4, wie bei MPEG-2, auch eine Unterteilung in Levels vorgenommen, die die verschiedenen Bildauflösungen und Datenraten betreffen. Tabelle 3.7 zeigt mögliche Levels für ausgewählte Profile, darin steht CIF (Common Interchange Format) für eine Bildauflösung von 352 x 288 Bildpunkten und QCIF für ein Viertel davon. Tabelle 3.8 zeigt eine Auswahl von Levels der Studioprofile, die für Auflösungen mit 720 x 576 (ITU-R BT.601) über 1920 x 1080 (ITU-R BT.709) bis zu 2k x 2k angegeben sind, wobei bis zu 12 Bits pro Pixel verwendet werden können. Die Zahlenfolge

Tabelle 3.7. Einige Profile mit zugehörigen Levels

Profile	Level	Typische Szenengröße	Datenrate (bps)	Maximale Objektanzahl
Simple	L1	QCIF	64 k	4
	L2	CIF	128 k	4
Core	L1	QCIF	384 k	4
	L2	CIF	2 M	16
Main	L2	CIF	2 M	16
	L3	ITU-R 601	15 M	32
	L4	1920 x 1080	38,4 M	32

Tabelle 3.8. Levels für Studio-Profile

Profile	Level	Typ. Formate	Max. Bits/pixel	Max. Objektzahl	Max. Bitrate (Mbps)
Simple	L1	ITU-R 601: 4224			
Studio		ITU-R 601: 444	10	1	180
Simple	L2	ITU-R 709. 60i: 422			
Studio		ITU-R 601: 444444	10	1	600
Simple	L3	ITU-R 709. 60i: 444			
Studio		ITU-R 709. 60i: 4224	12	1	900
Simple	L4	ITU-R 709. 60p: 444			
Studio		ITU-R 709. 60i: 444444			
		2K x 2K x 30p: 444	12	1	1800
Core	L1	ITU-R 601: 4224			
Studio		ITU-R 601: 444	10	4	90
Core	L2	ITU-R 709.60i: 422			
Studio		ITU-R 601: 444444	10	4	300
Core	L3	ITU-R 709. 60i: 444			
Studio		ITU-R 709. 60i: 4224	10	8	450
Core	L4	ITU-R 709. 60p: 444			
Studio		ITU-R 709. 60i: 444444			
		2K x 2K x 30p: 444	10	16	900

444444 beschreibt darin z. B. drei RGB-Kanäle mit drei Zusatzkanälen ohne Unterabtastung.

Für die eigentliche Videocodierung definiert MPEG-4 eine Hierarchie, an deren unterster Stelle einzelne Medien-Objekte (Media Objects) stehen. Media Objects sind alle erdenklichen audiovisuellen Datenformen wie unbewegte Bilder (z. B. Hintergründe), einzelne Video-Objekte (Personen, Möbel, Bäume etc.), Audio-Objekte (Musik, Sprache, synthetische Sounds), Texte und Grafiken. Avatare, also künstlich erstellte Repräsentanten tatsächlich agierender Personen, wurden u. a. für die Videotelefonie entwickelt. Synthetische Köpfe und Körper, die in Echtzeit mit Textvorgaben animiert werden, bilden damit eine weitere Objektvariante [96].

Zur codierten Form eines jeden Media Objects gehören immer beschreibende Elemente (Object Descriptors), die seine Zuordnung relativ zur Szene sowie sein Verhältnis zu etwaigen Datenströmen festlegen. Die Media Objects werden speziell im visuellen Bereich als Video Objects (VOs) bezeichnet, Audio und Video zusammen als Audio-Visual Objects (AVOs). Ihre Beziehung zueinander muss, wie bei allen interaktiven Anwendungen, im sogenannten Authoring-Prozess festgelegt werden. Auf der technischen Ebene ist die Beziehung durch die Komposition definiert. Eine oder mehrere fertige Kompositionen dieser Art werden bei MPEG-4 dann als Szene zusammengefasst. Dazu bedient sich MPEG-4 einer dynamischen Szenenbeschreibungssprache, die Binary Format for Scenes (BIFS) genannt wird und auf der ebenfalls von der ISO/IEC entwickelten Skriptsprache VRML (Virtual Reality Modeling Language) basiert, die für Interaktionsanwendungen im Internet genutzt wird. Der wesentliche Unterschied zur VRML besteht darin, dass das BIFS keinen Quellcode mit Texteingabe beinhaltet, sondern lediglich eine leicht handhabbare, binäre Zeichenfolge

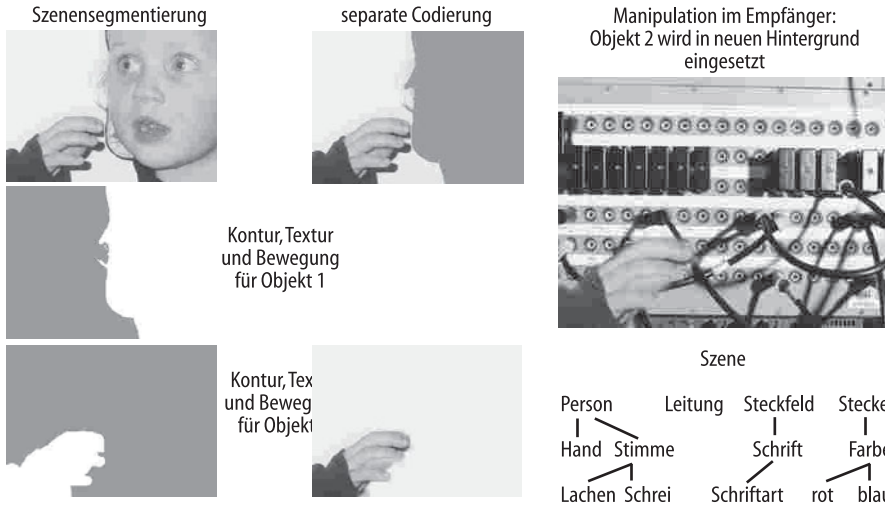


Abb. 3.87. Szenenaufbau und Video Objekt Planes bei MPEG-4

und die Fähigkeit zum Streaming bereithält. Der Begriff Streaming bezeichnet die Möglichkeit, audiovisuelle Daten betrachten zu können, ohne sie zuvor komplett übertragen zu müssen. Zusätzlich definiert MPEG-4 eine auf Java basierende, objektorientierte Sprache namens MPEG-J, die adaptiv eine Skalierung der Datenströme vornehmen kann, um sich aktuell vorhandenen Ressourcen wie Netzwerk-Bandbreite oder Arbeitsspeicher optimal anzupassen.

Eine Szene kann damit in inhaltsbezogene Bestandteile zerlegt und auf verschiedene VOP verteilt werden (Abb. 3.87). Dabei werden feste und variable Bildwiederholfrequenzen für die einzelnen Planes unterstützt und überlappende und nicht überlappende Bildbestandteile definiert. Die VOP werden getrennt codiert und übertragen. Erst beim Empfänger werden die Planes wieder zu einem Bild zusammengesetzt, wobei sich die Form und Position mit der Bewegung der Bildinhalte ändern kann. Auf diese Weise wird, ähnlich, wie es mit den Ebenen in Bildbearbeitungsprogrammen (Photoshop) oder beim Compositing (s. Kap. 9) geschieht, ein interaktiver Zugriff auf die Objekte möglich. Die Bildentstehung kann im Empfänger manipuliert werden.

In Abbildung 3.87 wird auch deutlich, dass es bei MPEG-4 nicht mehr nur möglich ist, herkömmliche rechteckige Bilder (Frames) zu codieren, sondern auch darin befindliche, willkürliche Formen (Shapes). Die freie Form kann je nach Verwendungszweck auf zwei verschiedene Arten bestimmt werden, entweder als binäre Form (Binary Shape) oder als Graustufenebene (Gray Scale, Alpha Shape). In beiden Fällen wird eine rechteckige Maske genau um die Grenzen der Video Object Plane (VOP) gelegt. Diese Maske muss dabei in horizontaler als auch in vertikaler Richtung ein Vielfaches von 16 Pixeln darstellen, sodass in einem nächsten Arbeitsschritt der Ausschnitt in 16 x 16 Pixel große Blöcke (Binary Alpha Blocks, BAB) aufgeteilt werden kann, von denen jeder einzeln codiert wird.

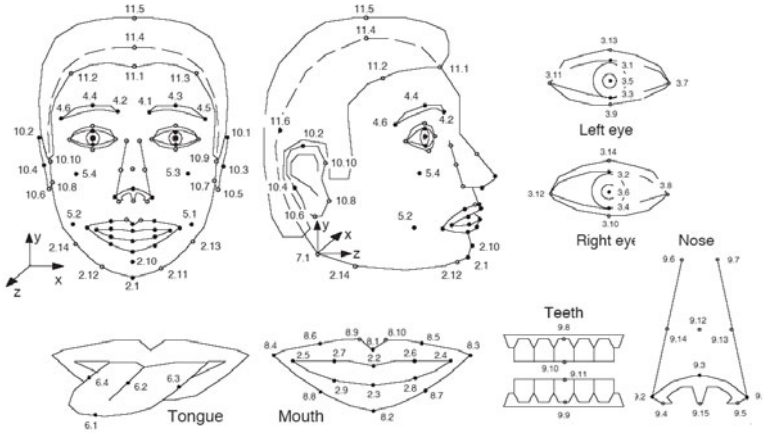


Abb. 3.88. Facial Animation Object bei MPEG-4

Beim Binary Shape wird einfach mit „on“ oder „off“ (1 oder 0) festgelegt, welche Blöcke als zum Objekt dazugehörig befunden werden und welche nicht. Für Anwendungen, bei denen höhere Qualität gefordert, wird das Binary-Shape-Verfahren auf eine 8 Bit große Graustufenbeschreibung der Blöcke erweitert, die als Alpha Shape oder Gray Scale bezeichnet wird. Dadurch wird ein Block mit einem Transparenzwert von 0 (100 % Opazität) bis 255 (100 % Transparenz) versehen, was einen weichen Übergang an den Kanten ermöglicht. Aus Gründen der Einfachheit bzw. der Abwärtskompatiblität bietet MPEG-4 auch die Möglichkeit, Video-Sequenzen wie von den früheren Standards gewohnt als rechteckige Frames zu codieren.

Eine Besonderheit stellen Objekte dar, die statisch sind, sich also über einen gewissen Zeitraum hinweg nicht verändern. Sie werden Sprites genannt und können aufgrund ihrer Eigenheit besonders Bit sparend codiert werden, da sie grundsätzlich nur einmal, zu Beginn, übertragen werden müssen und dann in einem Puffer gespeichert werden können.

Bei 2D- und vor allem 3D-Anwendungen kommen Texturen als austauschbare Grafik-Muster zum Einsatz. Diese werden auf statische und dynamische Polygon-Drahtgitter-Objekte (Meshes) aufgebracht, um ihnen ein individuelles Äußeres zu verleihen. Dabei kann es sich um Stein- oder Holz-Texturen für Wände und Böden, um Haut und Kleidung für virtuelle Charaktere handeln.

Ein besonderes Augenmerk wurde bei MPEG-4 auf die Erstellung von Avataren gelegt, künstliche Gesichts- und Körper-Repräsentationen, deren strukturelle Beschreibungseigenschaften sich für eine äußerst effiziente Form der Datenreduktion eignen. Ein Gesicht (Facial Animation Object) wird dabei z. B. als Drahtgitter mit neutralem Gesichtsausdruck vorbereitet (Abb. 3.88) und dann nach vorgegebenen Parametern mit Mimik, Gefühlsausdruck und Lippenbewegungen versehen, um Wut, Freude etc. auszudrücken. Dies bedeutet z. B. für die Anwendung in der Videotelefonie, dass ausschließlich die Merkmale des Sprechenden eingefangen, als Parameter mit dem Datenstrom übertragen und

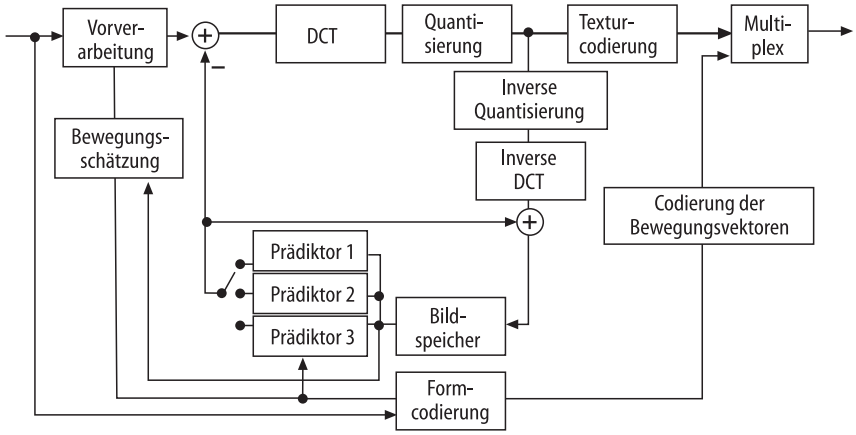


Abb. 3.89. MPEG-4-Videocoder

auf der Decoderseite in die entsprechende Animation umgesetzt werden. Gesichter wie auch Körper (Body Animation Objects) werden über das Binary Format for Scenes (BIFS, s. o.) gesteuert. Damit soll die umfassende Animation virtueller Charaktere von Internet-Spielen bis hin zu absolut realistischen Darstellungen und deren Verschmelzung mit herkömmlichem Video-Content ermöglicht werden.

MPEG-4 ist ein objektbasiertes Codierverfahren, bei dem Bewegungsschätzung und -kompensation zur effizienten Steigerung der Datenreduktion ebenso nötig und möglich ist wie bei MPEG-1 und -2. Sie beruht wieder auf der Bildung eines Gitters von Makroblöcken, nur dass die Beseitigung zeitlicher Redundanzen hier nun auf Basis von Objekten, genauer gesagt der Video Object Planes (VOP), funktioniert. Der Codieralgorithmus, der dabei zum Einsatz kommt, ist die sogenannte Shape-Adaptive DCT (SA-DCT), also die bekannte hybride DCT, die um die formangepasste Codierung erweitert wurde (Abb. 3.89). Dazu definiert MPEG-4 in Analogie zu den früheren Standards die sogenannten IntraVOP (I-VOP), die unabhängig von anderen VOPs codiert werden, die prädizierten VOP (P-VOP), deren Information aus der Differenzberechnung des zeitlich vorangegangenen VOPs beruht und die bidirektionalen VOP (B-VOP), deren Codierung aus der Vor- und Rückschau auf dasselbe, zeitlich benachbarte Objekt resultiert (Abb. 3.90). Jeder Makroblock enthält Luminanz- (Y) und Chrominanzblöcke (C_R, C_B), bei denen die Bewegungskompensation auf Blockgrößen von 16 x 16 Pixeln bzw. 8 x 8 Pixeln beruht. Mit Hilfe des Blockmatching-Verfahrens werden die in den zeitlichen Nachbarblöcken verschobenen Makroblöcke mit bis zu einem halben Pixel Genauigkeit gesucht und ein Bewegungsvektor für den gesamten Makroblock und je einer für jeden einzelnen Block darin bestimmt. Bei der sogenannten Advanced Coding Efficiency (ACE), welche seit MPEG-4 Version 2 bereit steht und bei der Bewegungskompensation hauptsächlich aufgrund ausgefeilterer Interpolationsfiltermethoden für 30 % bis 50 % mehr Effizienz gegenüber Version 1 sorgt, verläuft die Genauigkeit der Berechnung sogar bis auf 1/4 Pixel

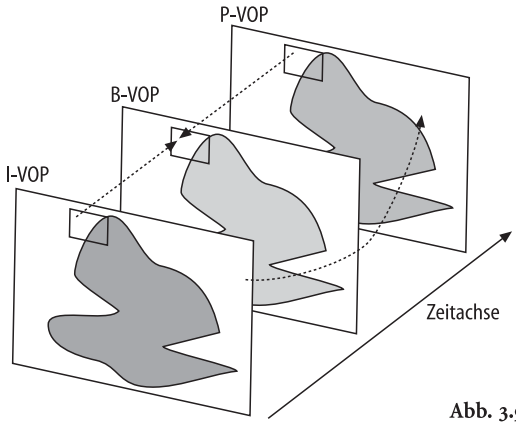


Abb. 3.90. VOP-Zeitbezüge

genau. Allgemein betrachtet unterscheidet sich der MPEG-4-Algorithmus der Bewegungsschätzung nicht wesentlich von der Anwendung bei MPEG-1 und -2. MPEG-4 ist abwärtskompatibel, denn bei der gezielten Festlegung eines Objekts als konventionelles, rechteckiges Frame wird die Formcodierung verworfen und für die Bewegungsschätzung auf die bekannte Standard-Block-DCT zurückgeschaltet. Obwohl die umfangreichsten Entwicklungen bei MPEG-4 in den vielen Defintionen bezüglich der Interaktion stecken, ist es bisher doch vor allem die durch ACE und mehr noch durch H.264 gesteigerte Codiereffizienz, die MPEG-4 bekannt gemacht hat [96].

3.8.1.6 MPEG-7 und MPEG-21

Mit MPEG-1, -2 und -4 werden interoperable AV-Inhalte (Content) zur Verfügung gestellt. Um eine Verwaltung (Management) der immens wachsenden Fülle dieser Inhalte und ihrer Quellen in der Unübersichtlichkeit des digitalen Zeitalters zu gewährleisten, arbeitet die ISO/IEC derzeit an weiteren Standards. Nach dem Wegfall von MPEG-3 gab es Spekulationen über die Fortführung der Namensgebung. Man hätte mit 5 (dem Nächstliegenden) fortsetzen oder mit 8 ein offensichtlich binäres Muster verfolgen können. MPEG hat sich jedoch entschieden, keiner logischen Erweiterung dieser Sequenz zu folgen und statt dessen die Nummer 7 gewählt. MPEG-5 und -6 sowie alles zwischen 7 und 21 ist daher nicht definiert.

MPEG-7 knüpft an die Errungenschaften von MPEG-4 an. Das Ziel ist dabei aber weder eine weiter erhöhte Kodiereffizienz noch die Repräsentation von Inhalten überhaupt. MPEG-7 verfolgt also nicht mehr das Ziel der (Video-)Datenreduktion, sondern das des Content Managements. MPEG-7 wird im Standard als „Multimedia Content Description Interface“ ausgezeichnet und bietet eine Beschreibung dessen, was die vorigen MPEG-Standards liefern. Basis dafür sind die sogenannten Metadaten (Daten über Daten), die entweder zusammen mit dem Programm oder unabhängig davon gespeichert werden können. Über sie wird in zahlreichen Anwendungen wie z. B. Archivierungsdiensten in Broadcast-Unternehmen per Schlagwortsuche eine leichte Identifikation

des gesuchten Materials ermöglicht. Im einfachsten Fall greifen Metadaten auf Beschreibungen zurück, die ohnehin in PCM, MPEG-1, -2 oder -4 vorhanden sind, erstellen gleichsam automatische Auszüge wie z. B. die Verwendung von Shape Descriptors aus MPEG-4, die Video-Objekte im Bild festlegen, die Verwendung von Bewegungsvektoren, die aus MPEG-2 bekannt sind oder die Formateigenschaften von Audio, die in PCM-Dateien als Beschreibung im Header der Datei integriert sind.

MPEG-7 hat aber den Anspruch, weit über die proprietäre Verwendung von Metadaten hinauszugehen. Wie bei den vorigen Standards werden auch hier, neben vielen weiteren, Parts wie »Video«, »Audio« und »Systems« definiert, deren Eigenheiten sich mit Hilfe sogenannter »Descriptors« auf unterschiedlichen Stufen beschreiben lassen. Low-level Descriptors geben Auskunft über grundlegende Eigenschaften wie z. B. Farbe, Formen, Größe und Bewegungen im Videobereich und Charakteristika wie spektrale, parametrische und zeitliche Eigenheiten im Audiobereich; high-level Descriptors gehen über zu Abstraktionen wie Timbre, Melodiestrukturen etc. Die Generierung der Descriptors, eventuell sogar aller gleichzeitig, ist keine Bedingung, je nach Wunsch des Nutzers soll aber jede für ihn relevante Information als solche bis zum Maximum darstellbar sein.

Ein weiterer Aspekt in diesem Kontext ist die umfassende Einbindung der Verwaltung digitaler Rechte (Digital Rights Management, DRM), die nicht ohne weiteres aus proprietären Lösungen standardisiert werden kann. Dies ist der Bereich von MPEG-21. Mit Hinblick auf Interoperabilität auf Anwenderseite, also die Kompatibilität in Bezug auf Formate, Codecs, Metadaten etc., zwischen kommerziellen Business-Systemen zielt MPEG-21 auf die Bereitstellung eines sogenannten »Multimedia Frameworks« ab. Hier wird versucht, die verschiedenen Elemente der Infrastruktur vom Content-Lieferanten bis zum Verbraucher und ihre Beziehungen zueinander zu beschreiben. Das Ziel ist die Schaffung einer heterogenen Landschaft von Geräten und Netzwerken, die eine globale Anwendbarkeit anstrebt, also in Bezug auf Multimedia einen weitaus größeren Rahmen anvisiert, als sich wie bisher auf den rein audio-visuellen Sektor zu beschränken [96].

3.8.2 MPEG-Audio

Neben Videosignalen erfasst der MPEG-Standard auch die Codierung von Audiosignalen und beschreibt im Teil Systems ihre Synchronisation. Hier geht es zunächst um die Audiocodierung. Da Video-Programme, z. B. auf DVD, alternativ zu MPEG-Audiosignalen oft eine Codierung nach Dolby AC3 enthalten, wird auch diese hier beschrieben.

Ebenso wie die Videocodierung beruht auch die Datenreduktion für Audiosignale auf der Ausnutzung der Eigenschaften bzw. Unzulänglichkeiten der menschlichen Wahrnehmung und wird im englischen Sprachraum entsprechend mit Perceptual Coding bezeichnet. Zum Verständnis der Reduktionsverfahren müssen daher zunächst einige Phänomene der menschlichen Hörwahrnehmung bekannt sein.

3.8.2.1 Grundlagen der Audiocodierung

Der wesentliche Vorgang beim Hören ist die Umsetzung von Schalldruckschwankungen in Nervenreize. Die Lautheitsempfindung steigt dabei nicht linear mit dem Schalldruck. Der Dynamikbereich, also der empfundene Unterschied zwischen laut und leise, ist sehr groß und umfasst 6 Größenordnungen, er liegt etwa bei Schalldrücken zwischen $2 \cdot 10^{-5}$ und 20 Pa. Die Grenzen werden als Hörschwelle und Schmerzgrenze bezeichnet. Der große Dynamikbereich wird vom Gehör nicht linear verarbeitet, daher ist es vorteilhaft, die Schallwechselgrößen in logarithmischer Form als Pegelwerte anzugeben, wobei der Bezugswert (Pegel 0 dB) etwa dem Schalldruck der Hörschwelle entspricht. Der Schalldruckpegel berechnet sich nach der Beziehung

$$L = 20 \log p/p_0.$$

Als Bezugswert gilt $p_0 = 2 \cdot 10^{-5}$ Pa. Ein mittlerer Schalldruck von $2 \cdot 10^{-2}$ Pa entspricht einem Schallpegel von 60 dB und wird als mittellaut empfunden.

Die Gehörempfindlichkeit hängt von der Frequenz ab. Nur bei mittleren Frequenzen (1 ... 5 kHz) liegt die Hörschwelle in der Größenordnung $2 \cdot 10^{-5}$ Pa, bei höheren und tieferen Frequenzen nimmt die Empfindlichkeit des Gehörs ab und erreicht die Grenzen der Unhörbarkeit bei 16 Hz bzw. 20 kHz, wobei die obere Grenze mit zunehmendem Alter der Hörer sinkt. Der Verlauf der Frequenzabhängigkeit der Hörempfindung bei verschiedenen Pegeln ist in Abbildung 3.91 als sogenannte Hörfläche dargestellt. Hier wird deutlich, welcher Schallpegel bei jeder Frequenz für eine bestimmte Lautheitsempfindung erforderlich ist. Dargestellt sind die Kurven gleicher Lautstärke mit der Einheit phon, die bei 1 kHz mit dem Schallpegel in dB übereinstimmen.

Zur Signalverarbeitung wird der Schalldruck mit Mikrofonen möglichst linear in elektrische Spannungen umgesetzt. Die geringen Ausgangsspannungen der Mikrofone (ca. 2 ... 20 mV/Pa Schalldruck) wird verstärkt und auf den Studiopegel angehoben, der in Deutschland bei +6 dBu (1,55 V) und international bei + 4 dBu (1,23 V) liegt. Für eine ggf. anschließende Digitalisierung wird die

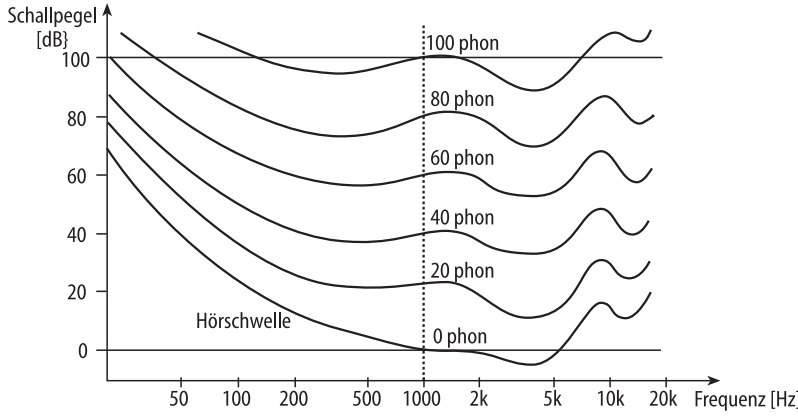


Abb. 3.91. Die Hörfläche

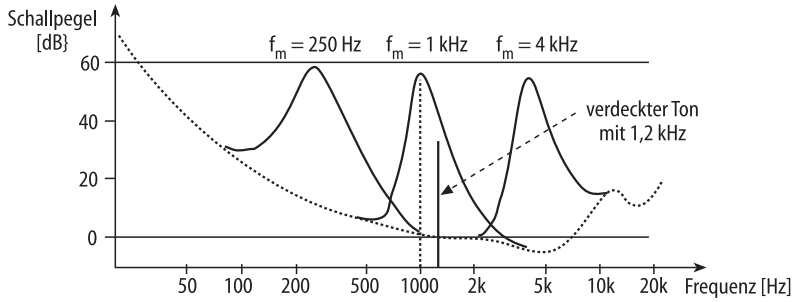


Abb. 3.92. Mithörschwellen resultierend aus Schmalbandrauschen dreier Mittenfrequenzen

Spannung für professionelle Anwendungen mit 48 kHz abgetastet (s. Abschn. 3.2). Um einen guten Signal-Rauschabstand zu erhalten, wird jeder Abtastwert oft mit 16 Bit dargestellt. Daraus folgt ein Dynamikbereich von 96 dB, der sich aber durch Head- und Feetroom auf ca. 70 dB verringert (s. Abschn. 3.2.1). Dieser Wert ist völlig ausreichend, da einerseits die Schallsignale selbst meist weniger Dynamik aufweisen und andererseits auch bei der Wiedergabe nicht mehr umgesetzt werden kann. So liegt z. B. schon bei der Abhörsituation keine absolute Stille vor, sodass Rauschsignale, die um 70 dB unter dem Maximalpegel liegen, nicht mehr wahrgenommen, sondern verdeckt werden [71].

Dieser Effekt, bei dem Hörereignisse durch andere überdeckt werden können (Verdeckungseffekt) ist es, der bei den Audio-Datenreduktionsverfahren ausgenutzt wird. Dabei interessiert besonders der Umstand, dass nicht nur Fremdsignale das Nutzsignal verdecken können, sondern auch Spektralanteile im Signal selbst. Durch Experimentieren mit definierten Tönen oder Rauschen, lässt sich feststellen, in welcher Form die Ruhehörschwelle bzw. die Kurven gleicher Lautstärke in der Hörfläche durch diese sogenannte Frequenz-Maskierung, verändert werden. Abbildung 3.92 zeigt als Beispiel das Verhalten des Gehörs bei Anregung durch Schmalbandrauschen mit drei verschiedenen Mittenfrequenzen bei einem Schallpegel von 60 dB [244]. Der sich ergebende gesamte Kurvenverlauf wird als Mithörschwelle bezeichnet. Signale, die unter dieser Schwelle liegen, sind unhörbar, sie werden von den Maskierern verdeckt. Dies gilt auch für das Quantisierungsrauschen. Daraus resultiert die Grundidee der Datenreduktion: jeder Abtastwert wird nur mit so wenig Bits quantisiert, dass das Quantisierungsrauschen gerade unter der Mithörschwelle liegt.

Neben der Frequenz-Maskierung kann auch die zeitliche Maskierung ausgenutzt werden. Der Begriff beschreibt ein Ereignis, bei dem laute Geräusche leisere, zeitlich kurz darauf folgende Geräusche in einem Bereich von wenigen Millisekunden verdecken können. Ein lauter Knacks ist z. B. in der Lage, ein nachfolgendes kurzes, leises Rauschen gänzlich zu unterdrücken. Interessant ist, dass unser Gehirn darüber hinaus sogar das kurze Rauschen unterdrückt, wenn es dicht genug vor dem Knacks einsetzt, sodass die zeitliche Verdeckung in beiden Richtungen funktioniert (Vor- und Nachverdeckung, Pre- und Post-Masking). Beide Varianten der Verdeckung sind immer von den jeweiligen Pegeln abhängig und ebenso von den aktuell beteiligten Frequenzen.

3.8.2.2 Audiocodierung nach MPEG 1/2

Als Abtastraten für das digitale Eingangssignal sind bei MPEG-1 folgende Standardwerte zugelassen: 32 kHz, 44,1 kHz und 48 kHz, bei MPEG-2 auch die jeweils halbierten Werte. Um die Signalverarbeitung flexibel durchführen zu können, wurde der MPEG-Audio-Standard in drei Layer eingeteilt. Layer 1 bietet die geringste Effektivität der Datenreduktion, erfordert dafür aber auch den geringsten Implementierungsaufwand, während für den Layer 3 das Umgekehrte gilt. Der Rechenaufwand schlägt sich auch in der Zeitverzögerung nieder, die für den Codiervorgang gebraucht wird, die Delay-Werte betragen 19 ms, 35 ms und 59 ms für die Layer 1, 2, 3. Bei Internetverbindungen, wo es auf minimale Datenrate ankommt, wird meist Layer 3 verwendet, der oft auch als MP3 bezeichnet wird. Für DVB wird Layer 2 eingesetzt.

Abbildung 3.93 zeigt das Blockschaltbild des Coders nach Layer 2. Das zu reduzierende digitale Audiosignal wird eingangs auf eine Filterbank und eine echtzeitfähige Fast Fourier Transformation (FFT) verzweigt. Die Analyse-Filterbank teilt eine Bandbreite von 24 kHz (bei einer Abtastfrequenz f_T = 48 kHz) mittels steilflankiger digitaler Bandpassfilter in 32 gleich große Filterbänder (Subbands) à 750 Hz auf. Jedes einzelne, 750 Hz umfassende Band wird einer Unterabtastung mit 1,5 kHz ($f_T/32$) unterzogen, welche die Signale an sich nicht verändert, sondern sie vielmehr mit genauer Zuordnung zum einstigen Originalbereich in die Basisbandlage hinuntermoduliert. Damit liegen 32 parallele Subband-Signale vor, deren gesamter Bandbreitenbedarf nicht den des Eingangssignals überschreitet. Dieser erste Arbeitsschritt ist noch vollkommen reversibel und verlustfrei. Layer-3 teilt als Besonderheit unter Zuhilfenahme einer direkt anschließenden Modifizierten Discreten Cosinus Transformation (MDCT) jedes der 32 Subbands zusätzlich noch einmal in 18 weitere auf, um mit dann insgesamt 576 Filterbändern die weiteren Arbeitsschritte noch feiner abstimmen zu können. Die MDCT wird auch als überlappende Blocktransformation bezeichnet (Overlapping Block Transform, OBT) und berücksichtigt die kritischen Bänder des menschlichen Hörens durch bis zu 50-prozentige Überlappung der Analysefenster zur besseren Beachtung der gegenseitigen Abhängigkeit benachbarter Bänder. Ebenfalls findet hierbei die differenzierte Betrachtung von Tonalitäten (Harmonien) und Atonalitäten (Rauschen) bessere Berücksichtigung.

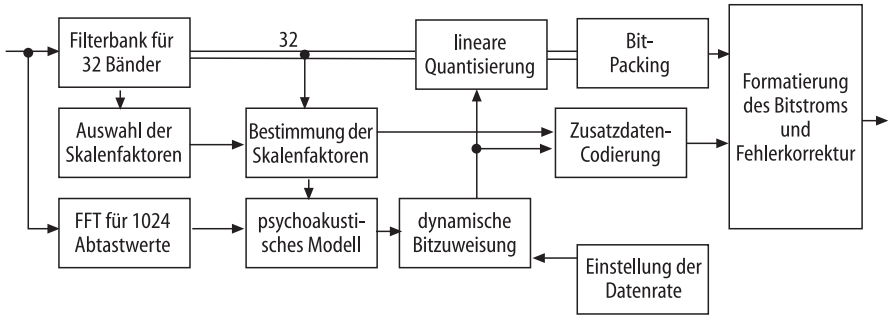


Abb. 3.93. Blockschaltbild der Audiocodierung nach MPEG-Layer 2

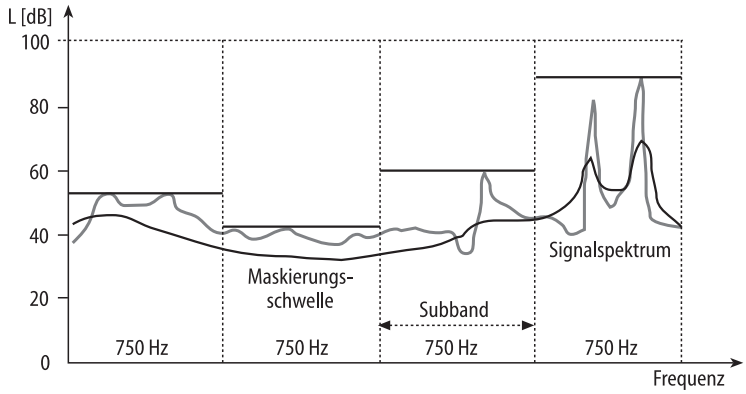


Abb. 3.94. Signalspektrum, Maskierungsschwelle und Maximalpegel in vier Subbändern

Zur effektiven Weiterverarbeitung können dann Gruppen von Samples gebildet werden, denn das Gehör hat auch nur eine beschränkte zeitliche Auflösung. Bei Layer-1 wird pro Arbeitsschritt eine Gruppe von 12 Samples mit einer zeitlichen Länge von zusammen 8 ms (bei 48 kHz) gebildet, bei Layer-2 und -3 sind es 36 Samples (24 ms).

Parallel zu diesem ersten Arbeitsschritt wird das Signal mittels FFT in den Frequenzbereich transformiert. Layer-1 arbeitet dabei mit einer Genauigkeit von 512 Punkten (512-FFT), Layer 2 und 3 verwenden eine 1024-FFT. Das psychoakustische Modell nutzt die FFT-Ergebnisse und errechnet anhand ausgefeilter Algorithmen, inwiefern die vorliegenden Frequenzanteile von der Maskierung betroffen sind. So kann im ersten Augenblick bereits bestimmt werden, in welchen Subbändern Frequenzanteile mit ihren Maximalamplituden gar nicht erst die untere Hörschwelle oder die Maskierungsschwelle erreichen und daher beseitigt werden können. Bei den Anteilen, die nicht komplett gelöscht werden, werden die für die weitere Bearbeitung bedeutenden Bereiche bestimmt (Abb. 3.94). Alles, was über den maximalen Signalpegel eines Subbands hinausragt und somit den Wert null aufweist, also die überflüssigen Most Significant Bits, kann ebenfalls gelöscht werden. Was übrig bleibt, ist der Bereich von der 0-dB-Linie bis zum aktuell maximalen Pegel im entsprechenden Frequenzbereich. Das Löschen der obersten Bits (MSB) mit ihren führenden Nullen entspricht einer Skalierung oder Normierung der Darstellung auf den relevanten Teil des Signals. Der für diese Division nötige Skalenfaktor kann aus einer Tabelle mit vordefinierten Möglichkeiten ausgewählt werden. Anschließend werden diese relevanten Bereiche dann von Block zu Block veränderbar mit 2 bis 15 Bit neu quantisiert.

Bei Layer-3 gibt es neben der Möglichkeit einer nichtlinearen Neu-Quantisierung (Kompandierung) eine zusätzliche Feinheit, indem der Block der Quantisierung aus zwei iterativ arbeitenden Schleifen aufgebaut ist, welche diesen Vorgang mehrfach auf ihre Qualität hin überprüfen. Weiterhin wird ab Layer-3 noch ein sogenanntes Bit-Reservoir angeboten. Dieser Baustein trägt dafür Sorge, dass unausgeschöpfte Bits bei grober Quantisierung nicht ver-

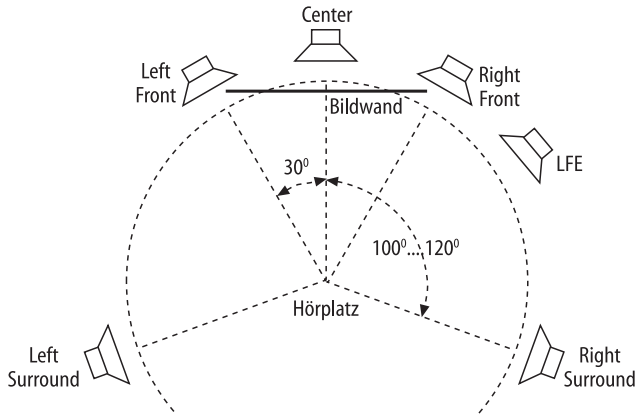


Abb. 3.95. Lautsprecheranordnung beim 3/2 Surround-Verfahren

schwendet werden, sondern für kritische Fälle zum Einsatz kommen, in denen der Coder mehr Bits benötigt als er aktuell zur Verfügung hat. Ein Signal muss somit nicht zwangsläufig zusammenhängend in einem Datenpaket untergebracht werden, sondern kann auf mehrere (maximal neun) verteilt werden.

Um die Daten beim späteren Decodieren wieder zu PCM-Signalen rekonstruieren zu können, bedarf es dann lediglich über den Skalenfaktor der Information, an welcher Position die Original-Bits einmal zu finden gewesen waren und der Angabe des Verhältnisses der ehemaligen Quantisierung zur Neu-Quantisierung. Der Decoder, der die Neu-Quantisierung umkehrt und die Subbands mit einer Synthese-Filterbank zusammensetzt, kann daher mit viel geringerem Aufwand betrieben werden. MPEG-Audio stellt also, wie MPEG-Video, eine asymmetrische Codierung dar.

Bei einem eingangseitigen Stereo-PCM-Datenstrom mit 1,536 Mbit/s, also einer zweikanaligen, ursprünglichen Abtastung mit 48 kHz bei 16 Bit Quantisierung, ergibt sich mit der geschilderten Methode eine typische Datenraten von 128 kbit/s bei Layer-3 (.mp3). Diese Datenreduktion um etwa den Faktor 10 ist die aktuell meistgenutzte im PC-Bereich und liefert recht gute Audioqualität. Eine Minute unkomprimierter Musik, die etwas mehr als 10 Mbyte auf einem Speichermedium einnimmt, kann also auf ca. 1 Mbyte reduziert werden. Bei Layer-1 sind insgesamt Datenraten von 32 bis 448 kbit/s möglich, bei einem typischen Wert von 384 kbit/s, bei dem im Broadcastbereich verwendeten Layer-2 beträgt der Bereich 32 bis 384 kbit/s (typisch 256 kbit/s) und bei Layer-3 schließlich 32 bis 320 kbit/s.

Bei MPEG-1 kann Audio-Information auf vier verschiedene Arten codiert werden: als eine Mono-Signal-Verarbeitung, als 2-Kanal-Codierung (zweimal Mono), Stereo und Joint Stereo. Bei Joint Stereo werden die großen Korrelationen (Ähnlichkeiten) zwischen den beiden Audiokanälen berücksichtigt und zur Einsparung der Datenmenge genutzt, indem z. B. der linke Kanal vollständig codiert und die Information des rechten Kanals nur als Differenz zum linken mit übertragen wird. MPEG-2 enthält als Erweiterung (Abb. 3.95) die

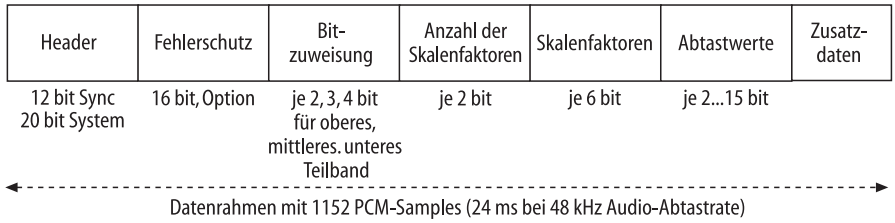


Abb. 3.96. Audio-Datenrahmen nach MPEG-Layer 2

Möglichkeit der stereokompatiblen Surround-Codierung. Der Begriff Surround steht für das fünfkanalige 3/2 Format, das mit drei Lautsprechern vor dem Hörer (Left, Center, Right) und zwei hinter dem Hörer platzierten Lautsprechern arbeitet (Left Surround, Right Surround) und gegenüber einfachen Stereosystemen eine deutliche Verbesserung der akustischen Raumabbildung gestattet (s. Kap. 5.5).

Die gewonnenen Daten werden in eine Frame-Struktur gebracht und mit einer Datenrate von z. B. von 128 kBit/s pro Kanal in den MPEG-Gesamtdatenstrom integriert. Abbildung 3.96 zeigt den Datenrahmen. Bei Layer-2 und -3 umfasst er 1152 Abtastwerte (384 bei Layer-1), die einer Dauer von 24 ms (8 ms) entsprechen, da ja bei einem 48 kHz-Eingangssignal mit 32 Subbands à 36 (12) Samples gearbeitet wird. Den codierten Daten werden ein Header und ein Prüfsummen-Check vorangestellt. Ein Frame umfasst immer 1152 Samples, unabhängig von der verwendeten Datenrate, die von Frame zu Frame wechseln darf. Der Header birgt insgesamt 32 Bit. 2 Bit stehen für Copyright-Informationen zur Verfügung, 12 Bit für die notwendige Synchronisation, 2 Bit geben Auskunft über den verwendeten Audio-Layer, und 4 Bit stehen für die aktuelle Datenrate, deren jeweiliger Wert aus einer fest definierten Tabelle entnommen werden kann (z. B. „1001“ für 128 kbit/s bei Layer-3). Zwei weitere Bit geben die eingangsseitige Abtastfrequenz an, und es folgt die Auskunft über den Kanal-Modus, Mono, Stereo etc. (z. B. „01“ für Joint Stereo).

Im Anschluss an die Multiplexierung erfolgt die Kanalcodierung (Fehlerschutz), die für Störresistenz auf dem Übertragungsweg Sorge trägt. Dazu stehen, falls erforderlich, weiterhin bis zu 16 Bit bereit. Nachdem dann die verwendete Anzahl und die Größe der Skalenfaktoren im Rahmen untergebracht ist, folgen die eigentlichen Nutzdaten, welche die mit je zwischen 2 und 15 Bit neu quantisierten Samples darstellen, die aus der 32-fach-Filterbank stammen. Abgeschlossen wird eine komplette Datei mit einem 128 Byte großen Datenblock, dem sogenannten ID3-Tag. Dies ist ein Etikett zur inhaltlichen Identifizierung, dessen Beginn ihn mit einem eindeutigen 3-Byte-Wort als solchen kennzeichnet. Das ID3-Tag stellt Metadaten bereit und liefert Computer-Benutzern oder Stand-Alone-Geräten wie MP3-Playern Informationen über das aktuelle Audio-File, in der Regel Interpret (30 Byte), Erscheinungsjahr (4 Byte) und Musiksparte (1 Byte). Es bleibt jedem Anwender selbst überlassen, dieses ID3-Tag teilweise oder überhaupt zu nutzen, die Informationen, auch mit falschen Angaben, zu ändern und zu überschreiben oder seinen Inhalt komplett zu löschen.

3.8.2.3 Advanced Audio Coding

Das 1997 eingeführte und nachträglich in den MPEG-2-Standard übernommene Advanced Audio Coding (AAC, ISO/IEC 13818-7) implementiert grundsätzlich die gleichen Funktionsprinzipien wie sein Vorgänger Layer-3, ist jedoch nicht mehr abwärtskompatibel. Zu den bereits dargestellten Bausteinen liefert AAC einige Zusätze. So wurde die Huffman-Codierung noch flexibler gestaltet, das Joint Stereo-Coding in seiner Effizienz gesteigert und die Aufteilung der Frequenzbänder für das Subband-Coding auf 1024 statt vorher 32 bzw. 576 erweitert. Die hybride Filterung von Layer-3, bestehend aus der 32-fachen Subband-Gewinnung mit anschließender MDCT, wurde auf eine reine MDCT-Transformation geändert, u. a. um störende Vorechos (Pre-Echos) zu vermeiden. Pre-Echos bezeichnen den Fall, dass Klangereignisse Geräusche erzeugen können, bevor sie eigentlich selbst zu hören sind. Diese Artefakte entstehen aufgrund der Tatsache, dass mit den einzelnen Samples stets kleinste Zeitbereiche (Fenster) analysiert werden. Da sich der Quantisierungsfehler aber stets über das gesamte Analysefenster verteilt, verursacht ein in dessen Mitte plötzlich einsetzendes Ereignis mit hohem Pegel beim Decodieren ein Quantisierungsgeräusch, welches sich deutlich von der relativen Stille vor dem eigentlichen Ereignis abhebt Dies kann durch den temporären Einsatz stark verkleinerter Fenstergrößen verhindert werden. Um störendes Rauschen, das insbesondere bei Sprachsignalen mit niedrigen Datenraten aufzufinden ist, zu vermeiden, liefert MPEG AAC zusätzlich eine zeitliche Rauschformung, das Temporal Noise Shaping (TNS).

Abbildung 3.97 zeigt das Arbeitsprinzip des Advanced Audio Coding. Eine wichtige Neuerung ist die Möglichkeit zur Rückführung der quantisierten Werte auf den nachrückenden Datenstrom. Insgesamt soll mit diesen neuen Implementierungen und einer noch weiter in der Leistung erhöhten Entropie-Codierung eine Steigerung der Datenreduktionseffizienz von 30 % bis 50 % gegenüber Layer-3 möglich sein, sodass AAC bei einer Datenrate von 64 kbit/s pro Kanal noch besser klingt als seine Vorgänger.

Zusätzlich sind nun auch die gegenüber Layer-1/2/3 ein Viertel so großen Abtastraten erlaubt, als wesentliche Neuerung also auch 8 kHz, sowie die Erweiterung auf bis zu 96 kHz. Maximal ist eine Datenrate von 348 kbit/s vorgesehen, in der von Mono und Stereo über 5.1 und 7.1 bis zu insgesamt 48 Kanäle untergebracht werden können [96].

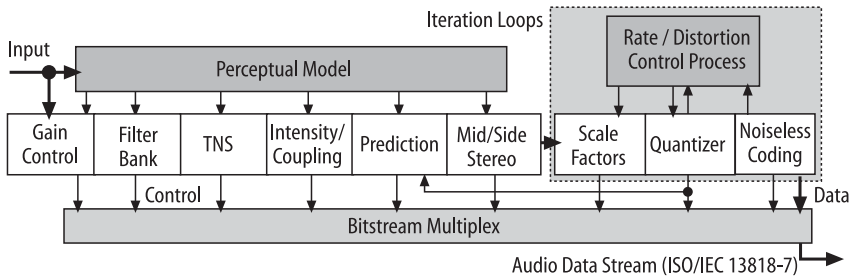


Abb. 3.97. Blockschaltbild für das Advanced Audio Coding

3.8.2.4 MPEG-4 Audio

Mit MPEG-4 AAC ist es möglich, gezielt unterschiedliche Typen von Audio-Daten (Audio Objects) in einem Bitstrom zu übertragen, wobei der Decoder in der Lage ist, diese unabhängig voneinander zu extrahieren und wiederzugeben. So darf der Anwender, insofern dies beim Authoring berücksichtigt wurde, z. B. Sprache leiser stellen, ausschalten oder zu einer anderen Fassung wechseln, während Musik und Hintergrundgeräusche davon nicht betroffen sind. Dabei wird erstmals in natürliche Audiosignale (Musik, Sprache) und synthetische unterschieden, also Audiodaten, die auf parametrisch beschriebenen Klängen basieren, wobei das Material abhängig von der benötigten Qualität jeweils getrennt nach der effizientesten Methode codiert werden kann.

Musik und Audiovorlagen, die keine separaten Objekte beinhalten, werden nach dem regulären MPEG-4 AAC oder dem Verfahren TwinVQ (Transform-Domain Weighted Interleaved Vector Quantisation) codiert. Darüber hinaus richtet MPEG-4 sein Augenmerk mit Verfahren wie CELP (Code Excited Linear Prediction), HVXC (Harmonic Vector Excitation Coding), HILN (Harmonic and Individual Line plus Noise) und TTS (Text to Speech) erstmals auch explizit auf die Codierung von Sprachsignalen mit Datenraten hinunter bis 2 kbit/s. CELP vergleicht, vereinfacht gesagt, das ankommende Signal mit vorgegebenen Codes aus einem Wörterbuch, die statt der eigentlichen Daten übertragen werden. Bei HVXC und HILN handelt es sich um Verfahren, die Audio-Signale, in der Hauptsache Sprache, nicht wie gewohnt übertragen, sondern in Parameter zerlegen. Gemäß dem Vocoder-Prinzip (Voice Coder) wird hier das Signal einer Analyse unterzogen, welche dieses aus tonalen (sinusähnlichen) Anteilen und Rauschanteilen nachkonstruiert und dann lediglich die entsprechenden Parameter übertragen. Zur differenzierteren Codierung sind auch Kombinationen der genannten Verfahren möglich.

3.8.2.5 Audiocodierung nach Dolby

Alternativ zur Audiocodierung nach MPEG wird, vor allem getrieben aus den USA, sehr oft die aus dem Filmbereich stammende Codierung von Dolby verwendet. Das zugehörige Datenreduktionsverfahren heißt AC3 und ist die Grundlage des in Abschn. 5.5.2 näher erklärten Dolby-Digital-Verfahrens, bei dem sechs separate, Audiokanäle für ein Surroundsignal verwendet werden. Nachdem Dolby Digital im Kinobereich eine sehr dominante Position erringen konnte, ist es auch in den Bereich der Heimanwendung vorgedrungen. Da dort unterschiedlichste Abhörkonfigurationen vorliegen, ist eine größere Flexibilität erforderlich. Dolby Digital erlaubt daher die Verwendung aller in Tabelle 3.9 dargestellten Kanalkonfigurationen. Die dort verwendete Bezeichnungsweise 3/2 anstelle von 5.1 verdeutlicht die Kanalzuordnung vorne/hinten. Dolby Digital unterstützt Datenraten zwischen 32 kbit/s bis 640 kbit/s, die den Kanälen

Tabelle 3.9. Kanalkonfigurationen bei Dolby Digital

Mode	1/0	2/0	3/0	2/1	3/1	2/2	3/2
Kanäle	Mono	Stereo	L - C - R	L - R - S	L - R C- S	L - R L_s - R_s	L - R - C L_s - R_s - SW

flexibel zugeordnet werden. Eine 2/0-Stereokonfiguration wird typischerweise mit 192 kbit/s betrieben. Bei 5.1-Anwendungen wird diese Rate meist auf 384 kbit/s verdoppelt, wobei eine Audiogrenzfrequenz von 18 kHz erreicht wird. Für den Einsatz bei der Digital Versatile Disk (DVD, s. Abschnitt 5.2.4) ist eine Maximaldatenrate von 448 kbit/s möglich. Damit können bei einer 5.1-Konfiguration auch 20 KHz erreicht werden.

Darüber hinaus enthält Dolby Digital eine Erweiterung des Audiodatenstroms durch sogenannte Metadaten, d. h. Hilfsinformationen über die Signale. Dieses Verfahren ist wertvoll für die Verbreitung von Mehrkanalprogrammen im Bereich des digitalen Fernsehens (DTV) und für die DVD. Für diese stehen auf der Wiedergabeseite Heimanwendergeräte unterschiedlichster Qualitätsstufen zur Verfügung. Mit den Metadaten können die Empfangsgeräte entsprechend ihren Möglichkeiten angesprochen werden. So braucht z. B. keine generelle Dynamikeinschränkung auf der Sendeseite vorgenommen zu werden, da sie im Endgerät durchgeführt werden kann. Oder eine aufwendige 5.1-Mischung kann mit Hilfe der Metadaten in optimaler Form in eine Stereomischung überführt werden und die mittlere Wiedergabelautstärke kann über das sogenannte Dialnorm auf einfache Weise an unterschiedliche Programmarten angepasst werden. Dialnorm bezeichnet eine Schaltung, die den verwendeten Pegel an der empfundenen Lautheit orientiert, die gemäß der menschlichen Wahrnehmung nicht vom Spitzenwert des Signals, sondern von seinem Mittelwert bestimmt ist.

Als weitere Bezeichnung in der Dolby-Codierungsfamilie (Tabelle 3.10) gibt es Dolby E. Dabei geht es nicht mehr um die Vermehrung oder Umordnung von Audiokanälen für die Wiedergabe, sondern nur um die Audiodatenübertragung für den Produktions- und Postproduktionsbereich. In diesem Bereich ist es üblich, Signale digital über die auch in der Analogtechnik benutzten symmetrischen Leitungen mit 3-poligen XLR-Steckern zu übertragen. Für eine zweikanalige, nicht datenreduzierte Übertragung der digitalen Audiodaten wird eine von der AES/EBU definierte Schnittstelle verwendet (s. Abschn. 3.3.4). Der Datenrahmen erlaubt die Verarbeitung von 24-Bit-Wortbreiten für jeden Abtastwert und jeden Kanal. Die Datenworte sind jeweils von 8 Bit für den Rahmen und Zusatzdaten umgeben, sodass bei einer Abtastrate von 48 kHz ein Datenstrom von 3,072 Mbit/s resultiert. Bei Dolby E wird nun diese Schnittstelle und diese Datenrate zur Übertragung von bis zu acht Audiokanälen mit Metadaten genutzt, mit einer Datenreduktion, die so milde sein kann, dass nach Ansicht von Dolby die Qualität auch nach mehrfacher Signalverarbeitung nicht leidet. Eine typische Anwendung der acht Kanäle wäre der gleichzeitige Transport einer 5.1- und 2-Kanal-Mischung für ein Stereosignal.

Tabelle 3.10. Übersicht über die Dolby-Varianten

Anwendung	professionell	heim
Rauschunterdrückung	Dolby A, SR	Dolby B, C
Mehrkanal analog	Dolby Stereo	Dolby Surround/Pro Logic
Mehrkanal Digital	Dolby Digital/EX, E	Dolby Digital/EX
Datenreduktion	Dolby AC3	Dolby AC3

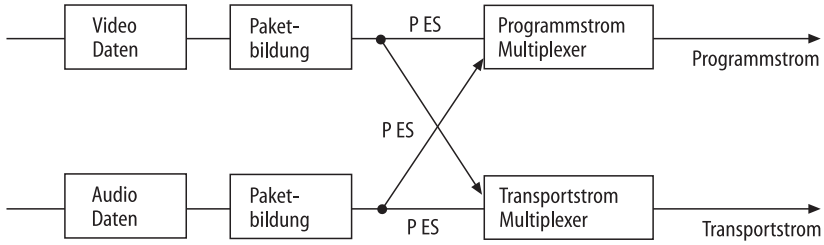


Abb. 3.98. Bildung des Packet Elementary Stream

3.8.3 MPEG-Systems

Die wesentliche Aufgabe des Teils »Systems« von MPEG, hier zunächst dargestellt für MPEG-2, ist die Definition der Multiplexstruktur für die verschiedenen Datenarten unter Berücksichtigung der Synchronisierung von Audio- und Videodaten mit ihren jeweils unterschiedlichen Datenraten.

Um die Synchronität zu erreichen, ist es erforderlich, die Datenströme in Portionen aufzuteilen, sie zwischenzuspeichern und im Multiplex zu übertragen. Das Multiplexverfahren bestimmt die Strukturierung der Daten, es ist dabei in der Lage, nicht nur mit festen, sondern auch mit flexiblen Datenraten zu arbeiten (constant and variable Bitrates, CBR und VBR). Weiterhin können die Übertragungskapazitäten mehreren Programmen zugeordnet werden.

Die Daten werden in zunächst in Pakete aufgeteilt, die jeweils nur eine Datenart (Bild, Ton, Daten) enthalten (Abb. 3.98). Dieser Grunddatenstrom wird als Packet Elementary Stream (PES) bezeichnet. Abbildung 3.99 zeigt den Aufbau der Datenpakete. Der Beginn wird mit einem Header von 6 Byte Umfang gekennzeichnet, hierin befindet sich die Information über die Art der Daten (Audio, Video, Daten) und über die PES-Länge, die variabel sein darf. Ein Videostrom mit Nummer »nnnn« wird z. B. durch die Bitfolge »1110 nnnn« gekennzeichnet. Darauf folgt ein Feld variabler Größe, das PES-spezifische Informationen enthält und schließlich die eigentlichen Nutzdaten (Payload) mit einem Umfang von maximal 65526 Byte, einer Zahl, die im Header gerade mit zwei Byte dargestellt werden kann. Die PES-spezifischen Informationen

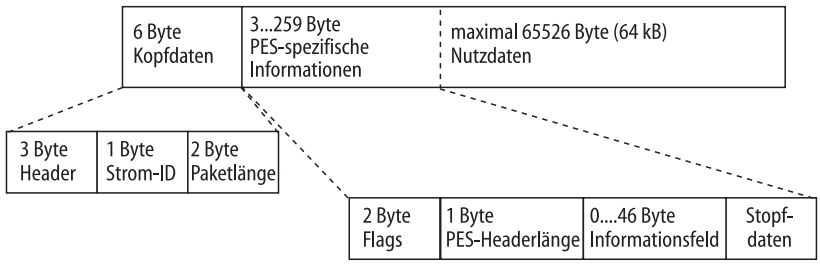


Abb. 3.99. Aufbau eines PES-Pakets

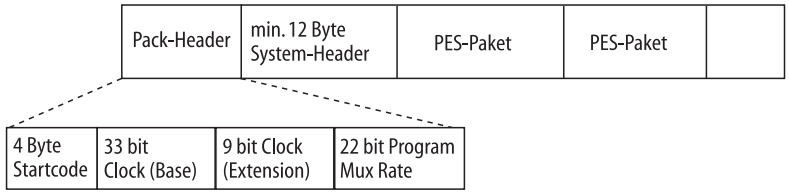


Abb. 3.100. Struktur des Programmstroms

enthalten sogenannte Flags, die die Übertragung von periodisch eingesetzten Zeitmarken zur Synchronisation (PTS und DTS, s. u.) anzeigen. Ein weiteres Flag wird z. B. gesetzt, wenn im PES die Daten eines Paketzählers übertragen werden, mit Hilfe dessen erkannt werden kann, ob bei der Übertragung ein Paket verloren ging.

Die Pakete des Elementarstroms dürfen nach der MPEG-Definition nicht direkt übertragen werden. Sie werden entweder in einen Programm-Strom oder einen Transport-Strom integriert. Der Program Stream ist für wenig fehleranfällige Anwendungen (z. B. CD-ROM, DVD) konzipiert, er erlaubt Pakete variabler Länge und arbeitet mit nur einer Zeitbasis. Der Programmstrom besteht aus einer Folge von PES, die alle zu einem Programm gehören. Die PES-Folge ist periodisch durch Pack- und System Header unterbrochen, die Synchroninformationen bereitstellen und einen Rahmen bilden, der den PES in Gruppen unterteilt (Abb. 3.100). Der Programmstrom erlaubt die Verwendung langer Pakete. Ein Datenverlust ist schwer auszugleichen, daher sollte er nur bei Übertragungsarten angewandt werden, bei denen sehr wenige Fehler auftreten.

Der Transport Stream ist dagegen für störanfällige Broadcast-Anwendungen gedacht. Es werden kurze Pakete verwendet, bei denen Störungen weniger Probleme bereiten. Im Hinblick auf einfach konstruierte Empfangsgeräte ist die Paketlänge fest vorgegeben, dabei sind mehrere Zeitbasen zugelassen. Abbildung 3.101 zeigt die Multiplexbildung. PSI steht darin für programmspezifische Informationsdaten, z. B. über Multiplexbestandteile. Es wird deutlich, dass mehrere Programme in einen Transportstrom integriert werden können.

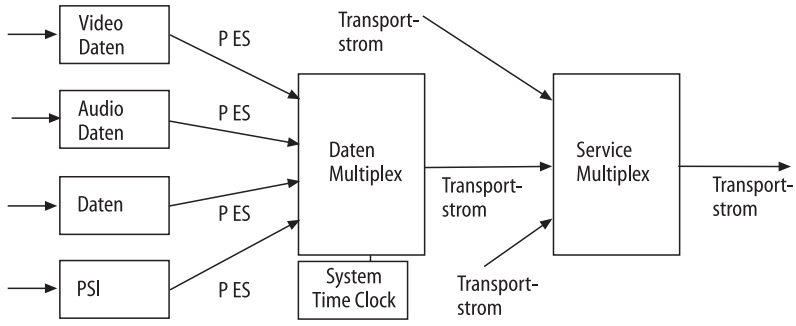


Abb. 3.101. Daten- und Transportstrom-Multiplex

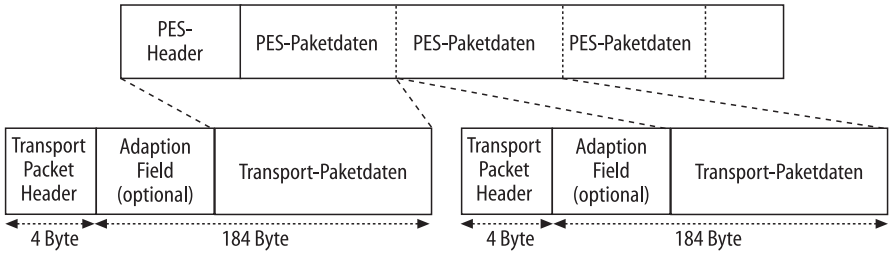


Abb. 3.102. Zuordnung von PES- und Transportstrom-Paketen

In dem für den Broadcast-Bereich wichtigen Transport Stream hat jedes Paket einen Umfang von 188 Byte, von denen 184 Byte für Nutz- oder Adaptionsdaten zur Verfügung stehen. Der im Vergleich lange Elementarstrom wird in kurze Stücke aufgeteilt, die in den Nutzdatenbereich passen (Abb. 3.102). Wenn das letzte TS-Paket nur noch einen Rest des PES aufnimmt, wird der freibleibende Teil durch sogenannte Stopf-Bytes gefüllt. Im Adaptionsfeld werden Steuerdaten untergebracht, die nicht in jedem Fall benötigt werden. Ein Paket wird immer von 4 Byte Kopfdaten eingeleitet. Abbildung 3.103 zeigt die Bedeutung der 32 Bits. Der Anfang des TS-Pakets wird vom Synchronwort mit dem Wert »0100 0111« bestimmt. Das TEI-Bit zeigt an, ob bei der Übertragung ein nicht korrigierbarer Fehler aufgetreten ist. Die 13 Bit der Packet Identification (PID) dienen der Identifikation der Nutzdaten und dem Auffinden eines einzelnen Elementarsignals im Multiplex. Der Contnuity Counter nummeriert schließlich alle Pakete mit derselben PID gemäß der gesendeten Reihenfolge.

Aufgrund der unterschiedlichen und variablen Laufzeiten für die verschiedenen Bestandteile des Transportstroms ist die Synchronisierung der Teile eine besondere Aufgabe, die alle Signalverarbeitungsstufen im Encoder und Decoder und den Übertragungskanal berücksichtigen muss. Zur Synchronisation wird eine System Time Clock (STC) als Zeitreferenz eingesetzt, sie arbeitet bei MPEG-1 mit 90 kHz und bei MPEG-2 mit 27 MHz. Ein Zeitraum von 24 Stunden wird mit 42 Bit festgelegt. Der Datenstrom wird regelmäßig mit einer aus der STC abgeleiteten Zeitmarke (Program Clock Reference, PCR) versehen, damit der Encoder-Takt auf der Empfangsseite regeneriert werden kann. Es ist festgelegt, dass die Referenz mindestens alle 100 ms im Datenstrom eingesetzt wird (bei DVB alle 40 ms). Wenn Audio- und Videodaten übertragen werden,

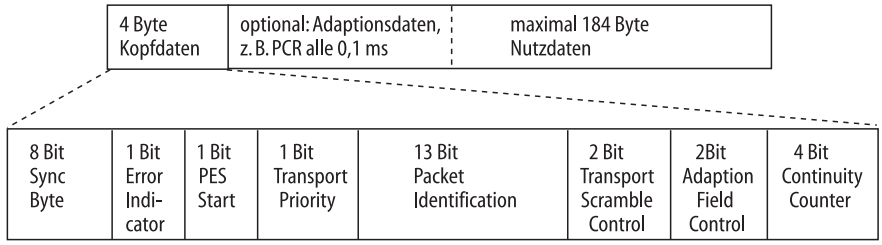


Abb. 3.103. Bedeutung der Kopfdaten im Transportstrom-Paket

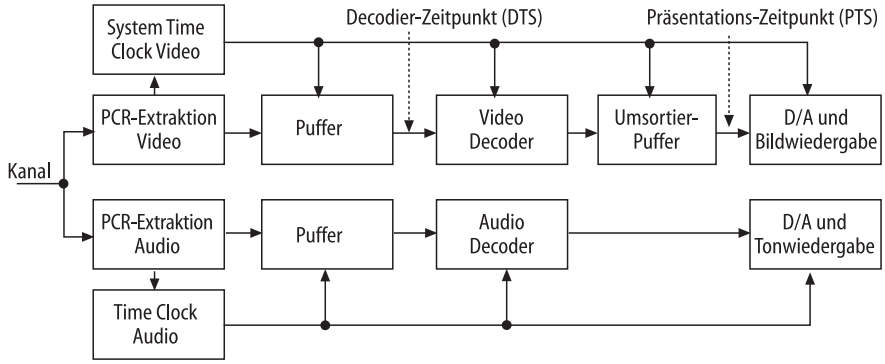


Abb. 3.104. Synchronisation im Decoder

die zu unterschiedlichen Programmen gehören, so ist es möglich, dass sie separate Zeitmarken aufweisen. Jede Änderung des Multiplexes ist also mit der Änderung der PCR-Markierung (Restamping) verbunden. Der Decoder erzeugt einen eigenen Takt, der fortwährend mit der PC-Referenz verglichen und ggf. korrigiert wird. Es wird also vorausgesetzt, dass auf der Sende- und Empfangsseite jeweils eine Systemuhr vorhanden ist, die die aktuelle Uhrzeit festlegt. Auf der Sendeseite wird den Datenpaketen beim Verlassen des Multiplexers die aktuelle (System-) Uhrzeit aufgeprägt. Durch Auswertung der Zeitmarken wird die Systemzeit auf der Empfangsseite rekonstruiert. Die Audio- und Videodaten sind bezüglich ihrer Abtastraten taktverkoppelt. Mit Hilfe der Zeitmarken ist so gewährleistet, dass die Ton- und Bildsignale synchron sind.

Neben der Systemzeitbasis sind relative Zeitmarken für die Decodierung und die eigentliche Wiedergabe definiert (Decoding und Presentation Time Stamp, DTS und PTS), die wie die PCR in regelmäßigen Abständen im Adaption Field nach den PES-Header-Daten übertragen werden. Diese Unterscheidung ist erforderlich, da sich, schon wegen der veränderten Bildreihenfolge für die bidirektionalen Prädiktion, der Decodierungszeitpunkt vom Präsentationszeitpunkt unterscheiden kann. Abbildung 3.104 zeigt die Lage der Decodier- und Präsentationszeitpunkte und weiterhin die Daten zur Systemkontrolle, die das Demultiplexing beeinflussen können.

Um auf die MPEG-Datencontainer zugreifen und die darin enthaltenen Bestandteile extrahieren zu können, sind im Systems-Teil programmspezifische Informationen (PSI) definiert. MPEG selbst legt vier Tabellen fest, für weitere Anwendungen, wie z. B. DVB, können eigene Tabellen definiert werden. Die erste der vier Tabellen enthält eine Liste der Programme, die im Transport-Multiplex übertragen werden (Program Association Table, PAT), die zweite einen Verweis auf die Paket-Identifikation für ein Programm (Program Map Table, PMT), damit die Video- und Audiodatenpakete dem richtigen Programm zugeordnet werden können. Zwei weitere Tabellen bieten Platz für Informationen zur Entschlüsselung von Programmen, die nur bedingten Zugriff erlauben (Conditional Access Table, CAT) und für Informationen zum Übertragungssystem (Network Information Table, NIT).

3.8.3.1 MPEG-4-Systems

MPEG-4-Systems ist erheblich komplexer als bei MPEG-2-System, da dieser Teil neben den dort beschriebenen Funktionen insbesondere die Interaktionsmöglichkeiten mit den Media Objects über BIFS unterstützen muss. Darüber hinaus tritt die Beachtung vorhandener Ressourcen sehr viel stärker in den Vordergrund als bisher. Die Skalierbarkeit der Datenströme berücksichtigt die Dienstgüte (Quality of Service, QoS), also die maximale Bitrate und die Bitfehlerrate des Übertragungsweges sowie die Leistung des jeweiligen Decoders.

Abbildung 3.105 zeigt, dass sich die MPEG-4-Systems-Architektur aus mehreren Layern (Ebenen) aufbaut. Links sind die Speicher- und Transportsysteme dargestellt, die einen MPEG-4-Datenstrom liefern. Die unterste Ebene der Architektur bildet der TransMux-Layer (Transport Multiplex). TransMux wird zusammen mit dem optional anwendbaren FlexMux als Einheit betrachtet, die dem Delivery Layer (der Auslieferungsebene) zugeordnet ist. MPEG-4-TransMux muss mit einer Vielzahl von Übertragungsprotokollen zurechtkommen, u. a. mit den bereits bekannten MPEG-2 Transport Streams (TS), dem UDP (User Daten Protokoll) über IP (Internet Protokoll), mit ATM AAL2 (Asynchroner Transfermodus Adaption Layer 2), MPEG-4 Files (MP4) und DAB-Streams (Digital Audio Broadcasting) etc.

Das sogenannte Delivery Multimedia Integration Framework (DMIF) stellt im Zusammenhang mit dem TransMux die Schnittstelle zwischen der Content-Anlieferung und dem SyncLayer dar, da bei MPEG-4 nicht mehr länger nur einfach auf die bei MPEG-2 Systems beschriebenen Programm- und Transport-Ströme reagiert werden muss. Im Regelfall erfolgt die Demultiplexierung daher bereits vor dem Eintritt in den eigentlichen MPEG-4-Decoder, da jede externe Anwendung ihre eigenen Paketierungs-, Multiplex- und Übertragungsmechanismen beinhaltet und somit auch dafür Sorge tragen muss, dass deren spezielle Anforderungen von ihnen aus erfüllt werden. Als abstrakte Kommunikationsschnittstelle für die verschiedenen Übertragungsprotokolle dient das DMIF Application Interface (DAI).

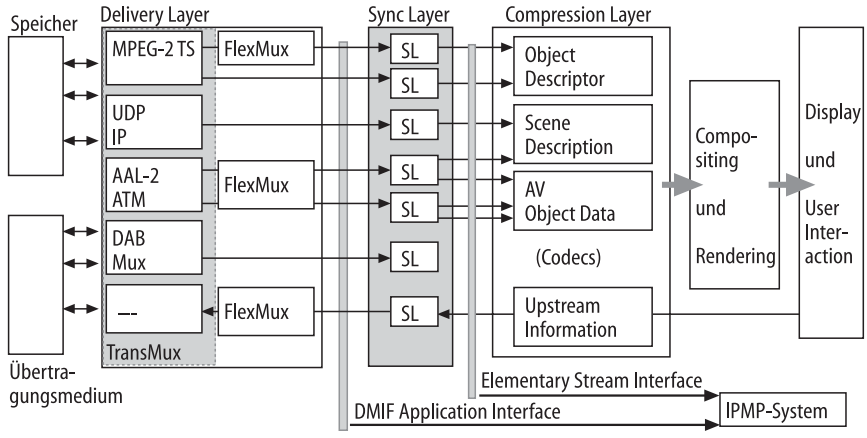


Abb. 3.105. Systems-Layer bei MPEG-4

Der FlexMux (Flexibler Multiplex) ist eine optionale Einrichtung und ist im Regelfall nicht notwendig, wenn die anliefernden Applikationen über eigene Mechanismen zur Demultiplexierung bzw. Multiplexierung verfügen.

Die demultiplexten elementaren Streams (ES) werden dem SyncLayer (Synchronisations-Ebene) zugeführt. Hier werden die Zeitreferenzen für die Objekte und die Szenenbeschreibungen ausgelesen und es wird für die Identifikation des jeweiligen Content-Typs gesorgt. Die zur Decodierung notwendigen Zeitstempel geben an, zu welchen Zeitpunkten die Datenströme vom Decoder-Puffer dem Darstellungspuffer übergeben werden sollen. Ein Zeitstempel wird bei MPEG-4 Object Time Base (OTB) oder Object Clock Reference (OCR) genannt. Nach ihm haben sich alle zu einem Objekt gehörigen Datenströme zu richten. Als Zeitbasis dienen entweder die traditionellen Clock-Referenzen und Zeitstempel oder die Verkopplung der Decodierung mit den Abtastraten der AV-Daten. Als dritte Möglichkeit kommt MPEG-4 Systems im Sonderfall auch gänzlich ohne Zeitinformation aus und nimmt eine strikte Abarbeitung in der angelieferten Reihenfolge vor. Anders als bei MPEG-2 liefert der neue Standard keinen Datenstrom mehr als Packet Elementary Streams (PES) sondern dient nur als paketbasierte Schnittstelle für den Delivery Layer [96].

Vom SyncLayer aus gelangen die Datenströme zu den jeweils zugehörigen Coder/Decoder-Paaren (Codecs). Dort werden zum einen die AV-Objekte an sich decodiert, zum anderen die entsprechenden Object Descriptors (OD) und die Scene Descriptors (SD), die alle in separaten Elementary Streams (ES) angeliefert und separat verarbeitet werden. Die Object Descriptors sind dabei das Bindeglied zwischen den Elementary Streams und den Szenenbeschreibungen. Sie verweisen auf die Daten-ES (Audio, Video) oder auf weitere OD-Streams. Wichtig ist in jedem Fall, dass mindestens ein Object Descriptor und i. d. R. eine Szenenbeschreibung vorliegen, da ohne diese keinerlei Decodierprozess in Gang gesetzt werden kann. Vor und nach dem SyncLayer können die Daten, über das DMIF Application Interface oder das Elementary Stream Interface abgegriffen und ggf. dem IPMP-System (Intellectual Property Management and Protection Descriptor Pointer) zugeführt werden, das Angaben zum Urheberrecht und zur Verschlüsselung enthält.

Nach der Decodierung werden die audiovisuellen Objekte gemäß der dazugehörigen Beschreibungen zur Gesamtszene zusammengesetzt, und nach Durchführung des abschließenden Renderings aus dem Kompositionsspeicher in die Präsentation übergeben. Das Rendering und die tatsächliche Wiedergabe müssen von Anwendungen außerhalb des Standards übernommen werden.

Zusammenfassend lässt sich für MPEG-4 feststellen, dass es ein überaus mächtiges Werkzeug für mulimediale Applikationen ist, das sowohl auf hohe Codiereffizienz als auch auf interaktiven Medienzugang abzielt. Bezüglich des letzten Punktes ist jedoch zu bedenken, dass die Nutzung der vielen Möglichkeiten mit einem hohen Arbeitsaufwand beim Authoring und entsprechenden Kosten verbunden sein kann. Dieser Umstand und die Komplexität von MPEG-4 bewirkt zur Zeit, dass in den meisten Fällen nur die verbesserte Codiereffizienz, insbesondere durch H.264 AVC (s. u.) von Interesse ist. Dabei wird zudem oft nicht einmal der aufwendige Systems-Bereich von MPEG-4 genutzt, sondern die Daten werden in einem MPEG-2-Transportstrom übertragen.

3.9 ITU-H.26x- und SMPTE-VC-Standards

Die Bezeichnung H.26x (das x steht für eine Zahl) kennzeichnet eine Familie von Videocodierstandards, die von der International Telecommunication Union (ITU) seit Beginn der 80er Jahre entwickelt wird. Der erste sehr verbreitete Standard trägt die Bezeichnung H.261. Er beruht auf der hybriden Codierung und wurde 1991 verabschiedet. Hier finden sich bereits viele Grundstrukturen, wie die Blockbildung und die Bewegungskompensation, die auch bei MPEG Verwendung finden. Die Weiterentwicklung führte zu H.262 (1994) mit bidirektionaler Prädiktion und Bewegungskompensation mit Halbpixelgenauigkeit und zu H.263 (1996), ein Standard, der für geringe Bitraten bis 30 kbit/s entwickelt wurde aber auch bei höherer Auflösung effizient arbeitet. H.262 entspricht weitgehend MPEG-2. H.263 wurde bei Videokonferenzsystemen eingesetzt. Dazu existieren zusätzlich die Standards H.320 bzw. H.323, für ISDN-Anwendung, in dem der Videoteilstandard mit einem Audiostandard wie z. B. G.722 und weiteren ITU-Übertragungsspezifikationen zusammengefasst ist. Das ISDN (Integrated Services Digital Network) beruht auf einer Datenrate von 64 kbit/s, was nur die Bildauflösungen CIF oder QCIF (176 x 144 Pixel) erlaubt.

Zusätzlich zu den Festlegungen der ITU sind hier die Standards dargestellt, die die SMPTE unter den Bezeichnungen VC 1 bis VC 3 herausgab. Sie waren zunächst nur herstellerspezifisch definiert und sind weniger universell [206].

3.9.1 H.264/AVC

Nachdem mit den ITU-Standards H.263+, H.263++ und H.263L an weiteren Verfeinerungen gearbeitet wurde, hat sich schließlich der Standard H.264/AVC etabliert. Das Kürzel AVC steht für den Part 10 des MPEG-4-Standards mit dem Namen „Advanced Video Coding". Dies bedeutet, dass sich die langjährige Zusammenarbeit zwischen der MPEG und der ITU-T Video Coding

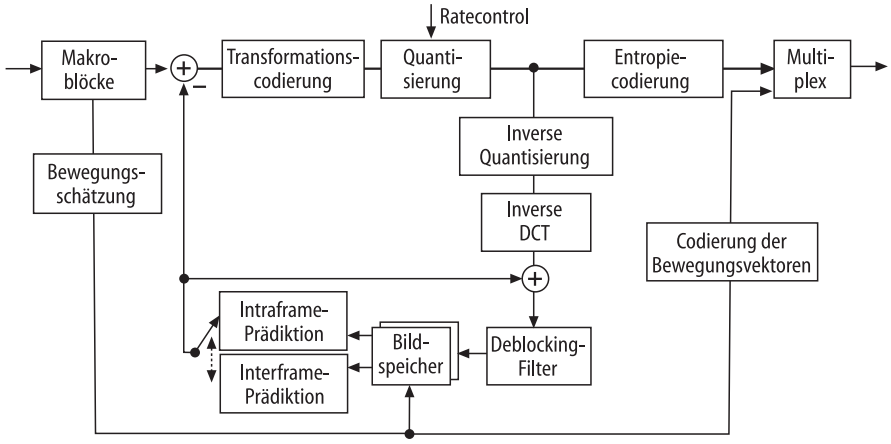


Abb. 3.106. Blockschaltbild des H.264/AVC-Encoders

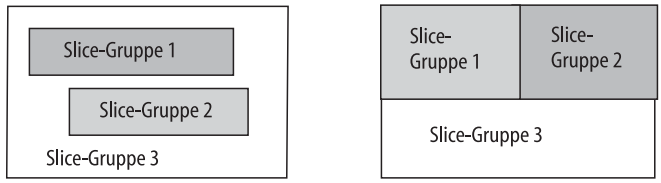


Abb. 3.107. Möglichkeiten zur Bildung von Slice-Gruppen bei H.264/AVC

Experts Group (VCEG) mit dem Joint Video Team (JVT) in einem gemeinsamen Standard niederschlägt. Mit einer Steigerung der Kompressionseffizienz von mehr als 2:1 gegenüber MPEG-2 kann digitales Fernsehen mit H.264/AVC noch bandbreitensparender gestaltet, die HDTV-Bitrate mit ihrem gegenüber SDTV mehr als fünffach größeren Bandbreitenbedarf weiter geschmälert werden. H.264/AVC wurde als Videocodec der Blu-ray Disc und zudem für die Verwendung in der zweiten Generation von DVB-S, -T und -C, LAN ausgelegt [56]. Es löst in vielen Bereichen MPEG-2 als dominanten Codec ab.

Im Gegensatz zu MPEG-4 konzentrierte man sich bei der Entwicklung von H.264/AVC auf die Steigerung der Codiereffizienz unter Vermeidung der Komplexität von MPEG-4. H.264 beruht wie MPEG-2 auf dem bewährten Konzept der hybriden Kodierung, d. h. dass bewegungskompensiert prädizierte Bildblöcke einer Transformation unterzogen werden, wobei mit einer geschlossenen Schleife dafür gesorgt wird, dass im Codierer die Rekonstruktion im Decodierer vorweggenommen wird (Abb. 3.106). Alle genannten Stufen sind optimiert, einige, wie die Intraframe-Prädiktion, sind neu hinzugekommen.

3.9.1.1 Interframe-Prädiktion

Ein Bild oder ein Halbbild wird zunächst in Makroblöcke der Größe 16 x 16 Pixel für Y und 8 x 8 Pixel für die Chroma-Komponenten unterteilt, ggf. können auch kleinere Blockgrößen verwendet werden, wobei eine flexible Makroblock Ordnung (FMO) unterstützt wird. Bezüglich der zeitabhängigen Aspekte ist die Organisation der Makroblöcke eine entscheidende Neuerung.

Die Makroblöcke werden aus unterschiedlichen Bereichen eines Bildes zu Slices (Scheiben) zusammengefasst (Abb. 3.107). Damit lässt sich die Bewegungskompensation optimieren. Slices können auffälligen Mustern im Bild entsprechen, aber auch willkürliche Formen darstellen. Mit I-Slices werden ausschließlich Prädiktionswerte innerhalb eines Frames gewonnen, dafür stehen mehrere Intra-Prädiktoren zur Verfügung. Mit P- und B-Slices gibt es des Weiteren die Vorhersage aus zeitlich benachbarten Bildern. Die Slice-Varianten SP (Switching P) und SI (Switching I) sind für Umschaltvorgänge zwischen Datenströmen unterschiedlicher Rate vorgesehen. SP ermöglicht eine effiziente Umschaltung zwischen verschieden codierten Bildern, SI eine exakte Anpassung eines Makroblocks in einen SP-Slice [137].

Die zu einem Slice gehörigen Makroblöcke müssen keine starren 16 x 16- bzw. 8 x 8-Blöcke sein, sondern können zur gezielteren Bewegungsschätzung bei der Inter-Frame-Codierung noch in weitere, rechteckige Untereinheiten aufgeteilt werden (Abb. 3.109). Mit der Unterteilung bis hinab zu 4 x 4 Bildpunkte kann eine wesentlich bessere Anpassung der Bewegungsvektoren an

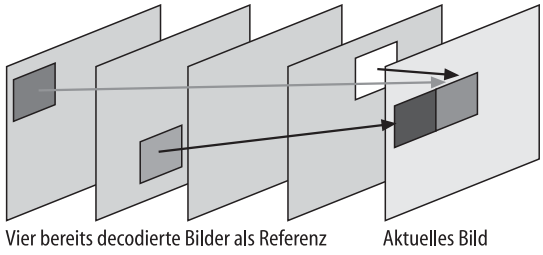


Abb. 3.108. Bewegungskompensation mit Bezug zu mehr als zwei Bildern

die Bildstrukturen erreicht werden als bei gröberen Blockstrukturen. Als Neuerung ist zudem die sogenannte bewegungskompensierte Langzeitprädiktion eingeführt worden, d. h. dass die Bildinhalte aus mehreren Referenzbildern, unterschiedlich gewichtet, gemischt werden können (Abb. 3.108).

Die zur Bewegungskompensation verwendeten Bewegungsvektoren können mit bis zu 1/4-Pixel-Genauigkeit bestimmt werden. MPEG-2 erlaubte dagegen nur eine Bestimmung mit 1/2-Pixel-Genauigkeit.

3.9.1.2 Intraframe-Prädiktion

Eine wesentliche Neuerung gegenüber MPEG-2 ist die Verwendung einer Intraframe-Prädiktion, also einer Prädiktion die sich nur auf ein Einzelbild bezieht. Ein Makroblock steht dabei in einem Bezug zu seinen Nachbarn innerhalb eines Bildes. Dies ist dadurch möglich, dass das Bild makroblockweise von links oben nach rechts unten decodiert wird und so in der Mitte entsprechend Nachbarblöcke zu Verfügung stehen, auf die Bezug genommen werden kann. Es werden Blockgrößen von 4 x 4-Pixels für detaillierte Bildinhalte und 16 x 16-Pixels für Bildbereiche mit gröberen Strukturen unterstützt.

Aufgrund der Tatsache, dass die Prädiktion bekannte und codierte Bildbereiche voraussetzt, ergeben sich Einschränkungen bei der Wahl der Intra-Prädiktionsmodi. Hier ist festgelegt, dass neben einer sogenannten DC-Prädiktion, bei der alle Pixel eines Blocks aus dem Mittelwert der Pixel des Nachbarblocks prädiziert werden, acht verschiedene Prädiktionsrichtungen unterstützt wer-

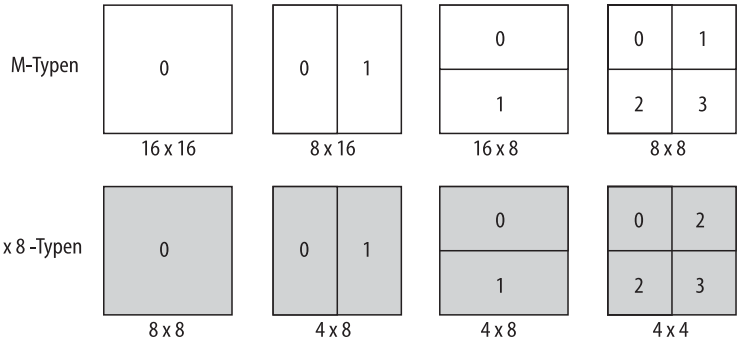


Abb. 3.109. Blocktypen bei H.264/AVC

den, d. h. dass mit gerichteter Prädiktion ein Pixel aus einem benachbarten Randpixel in einer von acht Richtungen prädiziert werden kann. Auf diese Weise gelingt es, verschiedenen Texturen im Bild gerecht zu werden, die die entsprechenden Richtungen haben.

3.9.1.3 Transformation

Wie bei den Vorgänger-Codecs wird auch bei H.264/AVC der Prädiktionsfehler (prediction residual) in den Ortsfrequenzbereich transformiert und anschließend werden die Koeffizienten gerundet und skaliert. Neu ist, dass i. d. R. auf der Basis von 4 x 4-Pixel-Blöcken gearbeitet wird. Kleinere Blöcke erfordern eine geringere Rechenleistung. Zudem werden statt Fließkomma- nur ganzzahlige Rechenoperationen (adds and shifts) verwendet (Integer-Transformation), die verhindern, dass bei der inversen Transformation Fehler entstehen.

Da das Bild aufgrund der inter- und intra-Prädiktion meist nur noch wenig räumliche Korrelation aufweist, ist im Allgemeinen ein 4 x 4-Block günstiger. Für Fälle, in denen die Korrelation vieler benachbarter Pixel zu höherer Effizienz führt, kann aber auch auf 8 x 8 Pixel umgeschaltet werden. Mit 4 x 4-Blöcken werden zudem Artefakte, die aus den Kanten der Blöcke resultieren (Moskito-Noise), verringert [137].

3.9.1.4 Entropiecodierung

Für H.264/AVC wurde auch die Entropiecodierung weiterentwickelt. Als erste Option steht das Variable Length Coding (VLC) in kontextadaptiver Form (CAVLC, context adaptive VLC) zur Verfügung. Das bedeutet, dass zwischen verschiedenen VLC-Tabellen, abhängig von den bereits übertragenen Syntax-Elementen, umgeschaltet wird. Damit kann beispielsweise der typische Umstand genutzt werden, dass nach dem Zick-Zack-Scanning beim Auslesen der Werte die DC-nahen Koeffizienten höhere Werte aufweisen als die hochfrequenten.

Als zweite Option kann das Context Adaptive Binary Arithmetic Coding (CABAC), eine leistungsfähigere arithmetische Codierung, verwendet werden, die jedoch eine hohe Rechenleistung erfordert. Bei der arithmetischen Codierung wird nicht jeder Wert einzeln, sondern eine Reihe von Symbolen verarbeitet, wobei für alle Syntaxelemente eine einzige Codeworttabelle benutzt wird, anstatt für jedes Element eine eigene zu entwerfen. Dabei kann jedem Symbol des Alphabets eine nicht-ganzzahlige Anzahl von Bits zugeordnet werden. Die adaptive arithmetische Codierung berücksichtigt, dass sich die Auftrittswahrscheinlichkeit der einzelnen Symbole von Szene zu Szene ändert und nutzt die Kenntnis der Nachbarschaftsbeziehungen, was einen Codiergewinn von bis zu 30 % bringen kann.

3.9.1.5 In loop Deblocking Filter

Eine zusätzliche Qualitätssteigerung wird bei H.264/AVC durch ein adaptives Schleifenfilter erreicht, das die Sichtbarkeit der Blockstrukturen ohne Schärfeverlust reduzieren kann. Dabei werden die Differenzen der Pixelwerte an den Blockgrenzen untersucht und das Filter so gesteuert, dass Blockingfehler von Sprüngen im tatsächlichen Bildgehalt unterschieden werden können. Das Filter wirkt nicht, wie sonst üblich, nach der Decodierung, sondern ist in die Rück-

Tabelle 3.11. Auswahl von Levels bei H.264/AVC

Level	Typ. Pixelzahl	Max. Bildrate	Bitrate			
	Profile		Main	High	High 10	High4:2:2/4:4:4
1	176 x 144	15	64 kbit/s	80 kbit/s	192 kbit/s	256 kbit/s
1.1	352 x 288	7,5	192 kbit/s	240 kbit/s	576 kbit/s	768 kbit/s
1.2	352 x 288	15	384 kbit/s	480 kbit/s	1,15 Mbit/s	1,54 Mbit/s
1.3	352 x 288	30	768 kbit/s	960 kbit/s	2,3 Mbit/s	3,07 Mbit/s
2	352 x 288	30	2 Mbit/s	2,5 Mbit/s	6 Mbit/s	8 Mbit/s
2.1	352 x 288	50	4 Mbit/s	5 Mbit/s	12 Mbit/s	16 Mbit/s
2.2	352 x 288	50	4 Mbit/s	5 Mbit/s	12 Mbit/s	16 Mbit/s
3	720 x 576	25	10 Mbit/s	12,5 Mbit/s	30 Mbit/s	40 Mbit/s
3.1	1280 x 720	30	14 Mbit/s	17,5 Mbit/s	42 Mbit/s	56 Mbit/s
3.2	1280 x 720	60	20 Mbit/s	25 Mbit/s	60 Mbit/s	80 Mbit/s
4	1920 x 1080	30	20 Mbit/s	25 Mbit/s	60 Mbit/s	80 Mbit/s
4.1	1280 x 720	60	50 Mbit/s	62 Mbit/s	150 Mbit/s	200 Mbit/s
4.2	2048 x 1080	60	50 Mbit/s	62 Mbit/s	150 Mbit/s	200 Mbit/s
5	2048 x 1080	60	135 Mbit/s	167 Mbit/s	405 Mbit/s	540 Mbit/s
5.1	4096 x 2048	30	240 Mbit/s	300 Mbit/s	720 Mbit/s	960 Mbit/s
5.2	4096 x 2160	60	240 Mbit/s	300 Mbit/s	720 Mbit/s	960 Mbit/s

führungsschleife integriert (Abb. 3.106), sodass das gefilterte Bild zur Bildung der Differenzen herangezogen werden kann. Tests haben ergeben, dass mit H.264/AVC insgesamt eine Effizienzsteigerung von ca. 60% gegenüber MPEG-2 erreicht wird (s. Tabelle 3.13).

3.9.1.6 Profile und Level

Um, wie bei MPEG, die Anforderungen an die Decoder variabel zu halten, definierte H.264/AVC zunächst elf Levels für Bildauflösungen von QCIF- bis zu Auflösungen für das digitale Kino-Format (Tabelle 3.11) bei einer Auswahl von drei Profilen: nämlich Baseline für die Videokommunikation ohne B-Frames, das Main Profile für Broadcast- und Unterhaltungsanwendungen und Extended für Streaming. Die Anpassung an verschiedene Netzwerke wird über einen Network Abstraction Layer (NAL) geregelt, der sich zwischen Coding- und Transportlayer befindet. Dieser ermöglicht z. B., die H.264-Daten in MPEG-2-Transportströme zu packen, sodass H.264-HD-Daten in der DVB-Infrastruktur oder via RTP im Internet übertragen werden können (s. Kap. 4.6).

Zu den H.264/AVC-Profilen, die zunächst für Distributionszwecke entwickelt worden waren, wurden weitere Profile hinzugefügt, die auf Studioanwendungen und hochqualitative Speicherung abzielen (Tabellen 3.11 und 3.12). Wie die vorher definierten Typen basieren die Profile High und High 10 auf einer 4:2:0-Abtastung. High 10 erlaubt die Verwendung von 10-Bit-Daten. Das gilt auch für High 4:2:2 und High 4:4:4, die das Abtastratenverhältnis im Namen tragen. High 4:4:4 erlaubt sogar bis zu 14 Bit Auflösung. Die obersten drei Profile, also

Tabelle 3.12. Varianten der Implementierung von H.264/AVC in Europa

Bezeichnung	Verwendung	HD-Parameter
AVCHD/NXCAM	semiprofessionell	720p/50; 1080i/25 4:2:0, 8 Bit Long GOP max. 24 Mbit/s
AVC-Intra50	professionell	720p/50; 1080i/p/25 4:2:0, 10 Bit 54 Mbit/s
AVC-Intra100	professionell	720p/50; 1080i/p/25 4:2:2, 10 Bit 112 Mbit/s
AVC-Ultra	professionell	720p/50; 1080p/24/25/50 4:4:4, 12 Bit bis 240 Mbit/s
XAVC	professionell	2k /4k /24/25/50 4:4:4, 12 Bit bis 600 Mbit/s

High 10, High 4:2:2 und High 4:4:4 haben auch Varianten für reine Intraframe-Codierung, die mit dem Zusatz „intra“ bezeichnet werden [137].

Das High-Profile der Level 3 und 4 des H.264/AVC-Codecs wurde für den Heimanwender- und den semiprofessionellen Bereich auch unter der Bezeichnung AVCHD adaptiert. Bei Geräten der Firma Sony wird dafür auch der Name NXCAM verwendet. AVCHD arbeitet mit 8-Bit-Daten bei 4:2:0-Abtastung bei Auflösungen von 1280 x 720 (24p, 50p oder 60p) sowie 1440 x 1080 und 1920 x 1080 Pixels (25 oder 30 frames interlaced). Die Datenrate beträgt bis zu 24 Mbit/s in einem Datencontainer der von MPEG-2 übernommen wurde. Die geringe Datenrate wird durch die Verwendung zeitlich abhängiger Bilder (Long GOP) erreicht. Auch Audiodaten sind berücksichtigt, wahlweise mehrkanalig, mit einer Kompression nach Dolby AC3, oder zweikanalig, unkomprimiert.

Die für Studioanwendungen konzipierte Intraframe-Variante von H.264/AVC wurde unter der Bezeichnung AVC-I zunächst von der Firma Panasonic favorisiert. Hier stehen die Varianten Class 50 für HD-Daten mit 4:2:0-Abtastung bei 54 Mbit/s und Class 100 für 4:2:2-abgetastete 10-Bit-HD-Daten bei 112 Mbit/s zur Verfügung. Als Datencontainer wird hier MXF verwendet. Das gilt auch für die mit AVC-Ultra bezeichnete Variante, die 12-Bit-HD-Daten mit 1080p/50 und einer 4:4:4-Abtastung bei bis zu 240 Mbit/s unterstützt (Tabelle 3.12). Die Firma Sony verwendet eine Intraframe-Variante von H.264/AVC für hohe Bildauflösungen unter der Bezeichnung XAVC [141]. Diese bezieht sich hauptsächlich auf den Level 5.2 und soll wiedergabekompatibel zu AVC-I sein. XAVC unterstützt Auflösungen bis zu 4k, mit 4:4:4-Abtastung bei bis zu 600 Mbit/s.

3.9.2 H.265/HEVC

In der Zusammenarbeit der Standardisierungsgremien ITU (VCEG) und ISO (MPEG) entstand mit der Doppel-Bezeichnung H.265/HEVC (H.265/High Efficiency Video Coding), bzw. MPEG-H Part 2, auch die Weiterentwicklung von H.264. Erneut steht hier vor allem die Steigerung der Codiereffizienz im Fokus wobei eine weitere Verdopplung, d. h. eine Effizienzsteigerung um den Faktor zwei gegenüber H.264/AVC, bzw. um den Faktor vier gegenüber MPEG-2 erreicht werden soll. Wiederum werden dabei die bewährten Grundkonzepte von MPEG-2 und H.264 übernommen und an diversen Stellen verfeinert (Tabelle 3.13).

Grundsätzlich soll der Codec wieder universell verwendbar sein. Er unterstützt daher alle gängigen Bildauflösungen bis hin zu 8k-UHD mit 7680 x 4320 Pixels und auch verschiedene Farbräume, von handheld- bis 8k-Kino-

Tabelle 3.13. Vergleich der Reduktionsmechanismen bei MPEG-2, H.264/AVC und H.265/HEVC

Bezeichnung	MPEG-2	H.264/AVC	H.265/HEVC
Transformation	DCT	Integer-DCT	Integer-DCT
Transfomationsblock	8 x 8	4 x 4 / 8 x 8	4 x 4 bis 32 x 32
Intraframeprädiktion	–	in bis zu 8 Richtungen	in bis zu 34 Richtungen
Interframeblockgröße	16 x 16	4 x 4 bis 16 x 16	4 x 4 bis 64 x 64
Loop-Filter	–	in loop deblocking	adaptive loop filter
Bewegungskompensation	1/2-Pixel	1/4-Pixel	1/4-Pixel advanced pred.
Entropiecodierung	VLC	CAVLC / CABAC	CABAC / LCEC
Relative Effizienz	1	2	4

anwendungen. Die Entwicklung basiert auf einem Test Modell in dem zwei Zielrichtungen mittels zweier Kategorien für (a) High Efficiency- und (b) Low Complexity-Anwendungen verfolgt werden. Die wichtigsten Aspekte verfeinern die Ansätze, die bei H.264/AVC implementiert sind. Dazu gehört vor allem die Unterteilung des Bildes in Blöcke, deren Größen noch variabler gestaltet sein können. Die übliche MPEG-Makroblockstruktur wird durch Coding Units (CU) ersetzt, die eine oder mehrere Transfomation Units (TU) oder Prädiktions-Units (PU) enthalten können, deren Größe variabel ist. Für die auf die Intraframe-Kompression bezogene Transformation sind als quadtree-strukturierte Blockgrößen 4 x 4 bis 32 x 32 Pixels erlaubt. Für die Interframe-Prädiktion werden Blockgrößen von 4 x 4 bis 64 x 64 in einer Baumstruktur unterstützt, wobei zusätzlich eine Advanced motion vector prediction zum Einsatz kommen soll. Auch die Intraframe-Prädiktion, die erstmals bei H.264 zu Einsatz kam, ist verfeinert und erlaubt die Prädiktion in bis zu 34 Richtungen. Das Filter, das in der Rückkopplungsschleife bei H.264 zur Verminderung der Blocking-Artefakte eingesetzt wird, soll in weiterentwickelter Form als Adaptive Loop Filter (ALF) verwendet werden (Tabelle 3.13).

Für die Entropiecodierung ist neben dem von H.264 bekannten CABAC zusätzlich ein Low Complexity Entropy Coding (LCEC) vorgesehen. Für dieses und auch für die anvisierte generelle Steigerung der Präzision der Berechnung aller Parameter von H.265/HEVC, wird gegenüber H.264 erheblich mehr Rechenleistung erforderlich. Erste Untersuchungen zeigen, dass die erwünschte Steigerung der Effizienz um den Faktor 2 gegenüber H.264 erreicht wird [190]. Insgesamt entsteht mit H.265/HEVC, das 2013 standardisiert wurde, nach MPEG-2 (1992) und H.264/AVC (2002) etwa alle 10 Jahre ein neuer Standard mit jeweils verdoppelter Codiereffizienz [28].

3.9.3 SMPTE VC-1

VC-1 ist die Bezeichnung der in SMPTE 421M beschriebenen Variante des Windows Media Encoders WM9. Mit Windows Media hat die Firma Microsoft im Jahre 2004 einen AV-Codec entwickelt der in das PC-Betriebssystem Windows integriert und für dieses optimiert ist. Wie bei H.264/AVC steht die effiziente Datenreduktion im Vordergrund. Auch die Codierungsmechanismen sind ähn-

lich, z. B. kann mit I, P und B-Frames sowie mit Blockgrößen von 8 x 8, 8 x 4 und 4 x 8 Pixels gearbeitet werden. Die Codierung unterstützt interlaced und progressive Formate mit fester und variabler Datenrate. Damit das Verfahren für Standard-PCs geeignet ist, wurde der Codec so entwickelt, dass er weniger komplex ist als H.264/AVC. Trotzdem ist er ähnlich effizient, die Bildqualität soll bei gleicher Bitrate mit der von H.264 vergleichbar sein. Im Unterschied zu H.264 hat VC1 nur zwei Modi für die Prädiktion, basierend auf Blöcken von 16 x 16 oder 8 x 8 Pixels. Als Referenzbilder sind wie bei MPEG-2 nur die direkten Nachbarbilder zugelassen. Mit Windows Media Audio (WMA) ist eine Audiodatenreduktion für bis zu acht Audiokanäle integriert, die mehr als doppelt so effizient arbeiten soll wie eine MP3-Codierung.

Bei VC-1 bzw. WM9 sind zunächst drei Profile definiert worden: Das Simple Profile ist bei geringem Implementierungsaufwand für Internetanwendungen und mobile Kommunikation konzipiert, meist bei Datenraten zwischen 96 und 384 kbit/s. Das Main Profile ist für die Wiedergabe am PC mit Bildraten bis 30 fps vorgesehen, wobei drei Level bestehen: Die Verteilung von Bildern mit CIF-Auflösung mit bis zu 2 Mbit/s, Bildern mit Standard-TV-Auflösung mit bis zu 10 Mbit/s und schließlich HD-Bildern mit bis zu 20 Mbit/s. Die gleichen Level stehen auch beim dritten, dem Advanced Profile zur Verfügung, das für Broadcastanwendungen mit Unterstützung des Interlaced-Verfahrens konzipiert ist. Für HD-Anwendungen sind vor allem die Advanced-Levels L2 bis L4 vorgesehen, die maximalen Datenraten betragen hier 20, 45 und 135 Mbit/s.

Die SMPTE definiert in den zugehörigen Recommended Practices RP 227 auch Verfahren für den Datentransport, die es z. B. ermöglichen, die Daten in die bei MPEG-2 definierten Transport- oder Programmströme einzubinden. VC-1 hat aber für professionelle Anwendungen bisher kaum Bedeutung.

3.9.4 SMPTE VC-2

VC-2 wurde ursprünglich von der BBC unter der Bezeichnung Dirac-Pro oder auch Mezzanine entwickelt. Das Ziel war hier zunächst, die Datenrate von mehr als 2 Gbit/s, die bei 1080p/50-HD-Signalen entstehen, über gewöhnliche HD-SDI-Strecken übertragen zu können. Dabei war eine milde Intraframe-Datenreduktion um den Faktor 2 bis 5 mit sehr geringen Qualitätsverlusten vorgesehen. Später wurden auch Datenraten bis hinab zu 50 Mbit/s unterstützt.

Das Verfahren basiert wie JPEG2000 auf der Waveletcodierung und nutzt ein 5/3-Filter. Um eine geringe Verzögerungszeit bei der Codierung zu erreichen, wird im Low Delay Profile die Möglichkeit bereit gestellt, nicht das ganze Bild, sondern Slices von 6 Zeilen Höhe zu codieren. Zudem wird eine einfache VLC für die Entropiecodierung eingesetzt.

VC-2 zielt auf eine hohe Bildqualität und unterstützt 4:2:2 und 4:4:4-Abtastmuster sowie Quantisierungen zwischen 8 und 12 Bit. Es beinhaltet fünf Levels mit verschiedenen Bildauflösungen. Der für HD-Anwendungen vorgesehene Level 3 erlaubt bis zu 840 Mbit/s. Level 4 zielt auf D-Cinema-Anwendungen mit Bildauflösungen bis zu 4k und Datenraten bis 3,2 Gbit/s, während Level 5 für Ultra-High Definition und 8k-Bilder bis zu 12,8 Gbit/s spezifiziert ist.

Tabelle 3.14. Varianten des DNxHD-Codec in Europa

Format	Avid	DNxHD36	DNxHD100	DNxHD145	DNxHD220	DNxHD444
Bittiefe		8 Bit	8 Bit	8 Bit	8/10 Bit	10 Bit
Sampling		4:2:2	4:2:2	4:2:2	4:2:2	4:4:4
Bitrate		36 Mbit/s	100 Mbit/s	145 Mbit/s	220 Mbit/s	440 Mbit/s

3.9.5 SMPTE VC-3, Avid DNxHD und Apple ProRes

Die Zielsetzung des von der Firma Avid unter der Bezeichnung DNxHD entwickelten Codecs VC-3 ist ähnlich wie bei VC-2, hier geht es aber darum, die hohen Datenraten von HD-Signalen auf SD-SDI-Niveau übertragen zu können.

Es ist somit zu diesem Zweck wieder nur eine recht geringe Datenreduktion von ca. Faktor 6 erforderlich. Sie basiert auf einer Intraframe-Codierung mittels einer DCT mit 8 x 8-Blöcken und einer anschließenden VLC. VC-3 war auf die Codierung von HD-Daten mit 4:2:2-Abtastung bei Datenraten von entweder 145 Mbit/s oder 220 Mbit/s eingeschränkt, der Avid-DNxHD-Codec erlaubt zusätzlich die Datenrate 36 Mbit/s für Vorschauzwecke und das Offline-Editing. Bei der höchsten Datenraten können neben 8-Bit- auch 10-Bit-Daten verwendet werden, wobei bei dem nachträglich hinzugefügten Format DNxHD444 auch RGB-Signale mit 4:4:4-Abtastung unterstützt werden (Tabelle 3.14) [10].

Mit einer ähnlichen Zielsetzung, nämlich einen hochwertigen Intermediate-Codec für hochqualitative Studio-HD-Anwendungen zur Verfügung zu stellen, hat neben Avid hat auch die Firma Apple, unter dem Namen ProRes, einen eigenen HD-Codec entwickelt. Dieser ist nicht unter VC-3 standardisiert, wird aber hier aufgeführt, da er DNxHD sehr ähnlich ist. Bei Apple ProRes werden SD-, HD- und die 2K-Auflösung mit 4:2:2-Abtastung mittels einer DCT so komprimiert, dass bei 8- oder 10-Bit-Auflösung meist Datenraten unter 270 Mbit/s entstehen. Im Profil ProRes 444 wurden später auch Signale bis zu 16 Bit mit Alphakanal bei Datenraten über 300 Mbit/s unterstützt (Tabelle 3.15).

DNxHD und ProRes sind Intermediate Codecs, die unabhängig von den sie entwickelnden Firmen eine breite Akzeptanz im Studiobereich gefunden haben. Sie sparen Speicherplatz und belasten Netzwerke beim Datenaustausch nur wenig. Bei sehr hoher Bildqualität arbeiten sie mit I-Frame-only sehr latenzarm.

Zum Abschluss dieses Kapitels zeigt Tabelle 3.16 die Vielzahl von Bezeichnungen, unter denen die verschiedenen Hersteller die wesentlichen Standards adaptieren. Die Abbildung 3.110 gibt schließlich die historische Entwicklung der in diesem Kapitel erläuterten Videocodierstandards wieder. Mit der Zeit

Tabelle 3.15. Varianten des Apple ProRes-Codec in Europa

Auflösung	Bildrate	ProRes (proxy)	ProResLT	ProRes422	ProResHQ	ProRes444
720 x 576	25i/p	12 Mbit/s	28 Mbit/s	41 Mbit/s	61 Mbit/s	92 Mbit/s
1280 x 720	25p	19 Mbit/s	42 Mbit/s	61 Mbit/s	92 Mbit/s	138 Mbit/s
1920 x 1080	25i/p	38 Mbit/s	73 Mbit/s	122 Mbit/s	184 Mbit/s	275 Mbit/s
2048 x 1152	25p	43 Mbit/s	97 Mbit/s	140 Mbit/s	210 Mbit/s	315 Mbit/s
2048 x 1024	50p	86 Mbit/s	194 Mbit/s	280 Mbit/s	419 Mbit/s	629 Mbit/s

Tabelle 3.16. Klassifizierung verschiedener Codecs und Herstellerbezeichnungen

DV-Codec	MPEG-2/4	H.264/AVC	Intermediate
DV SD, I-only 8 Bit 4:2:0 25Mbit/s	**IMX** SD, I-only 8 Bit 4:2:2 50 Mbit/s	**AVCHD/NXCAM** HD, Long GOP 8 Bit 4:2:0 24 Mbit/s	**ProRes** max. 2k, I-only 10 Bit 4:4:4 max. 629 Mbit/s
DVCAM SD, I-only 8 Bit 4:2:0 25Mbit/s	**HDV** HD, Long GOP 8 Bit 3:1,5:0 25Mbit/s	**AVC-I** HD, I-only 10 Bit 4:2:2 max. 110 Mbit/s	
DVCPRO SD, I-only 8 Bit 4:1:1 25Mbit/s	**XDCAMHD** HD, Long GOP 8 Bit 3:1,5:0 35 Mbit/s	**AVC-Ultra** HD, I-only max. 12 Bit 4:4:4 max. 240 Mbit/s	**DNXHD/VC3** HD, I-only 10 Bit 4:4:4 max. 440 Mbit/s
DVCPro50 SD, I-only 8 Bit 4:2:2 50 Mbit/s	**XDCAMHD422** HD, Long GOP 8 Bit 4:2:2 50 Mbit/s	**XAVC** max 4k , I-only 10 Bit 4:2:2 max. 600 Mbit/s	
DVCPROHD HD, I-only 8 Bit 3:1,5:1,5 100 Mbit/s	**MPEG4-SStP/HDCAM SR** HD, I-only 10/12 Bit 4:4:4 max. 880 Mbit/s		

wurde ein erheblicher Fortschritt bezüglich der Kompressionseffizienz erzielt. MPEG-2 ist der erste besonders erfolgreiche Standard. MPEG-4 setzt sich eher zögernd durch, da es aufgrund der Interaktionsmöglichkeiten sehr komplex ist. Dagegen übernahm das später in MPEG-4 integrierte H.264/AVC als doppelt so effizientes Codierverfahren, die Rolle von MPEG-2. Im Intra-Mode ist es studiotauglich. Als Nachfolgestandard ist H.265 vorgesehen. Die VC-Codecs werden nur im PC- oder Studiobereich eingesetzt.

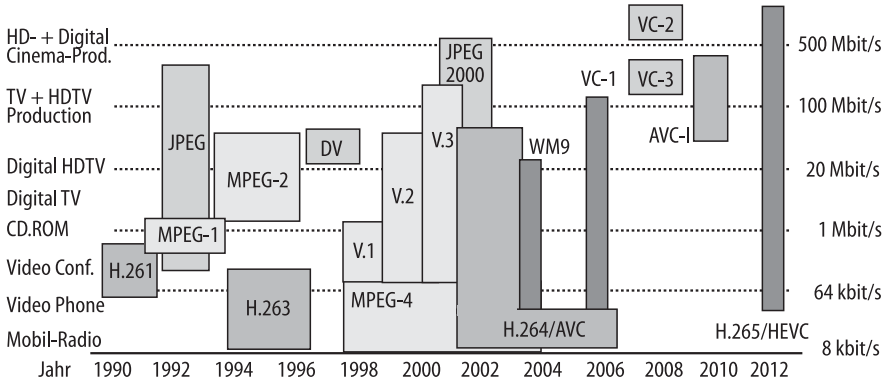


Abb. 3.110. Entwicklung der Videocodierstandards

4 Fernsehsignalübertragung

In diesem Kapitel werden die Signalformen und Verteilwege erörtert, die für die Übertragung des Fernsehprogramms zu den Zuschauern erforderlich sind. Für einen Zeitraum von ca. sechzig Jahren wurde das Fernsehprogramm analog übertragen. Basis des analogen Fernsehsignals ist das in Kap. 2 beschriebene FBAS-Signal. Neben der Vorstellung der zugehörigen Fernsehbegleitsignale geht es in diesem Kapitel zunächst um das analoge Fernsehübertragungssystem. Im zweiten Teil wird dann das aktuelle digitale Fernsehsystem vorgestellt, das die Bezeichnung Digital Video Broadcasting (DVB) trägt und das das analoge Verfahren ablöst. Die Übertragung analoger oder digitaler Programme zum Zuschauer geschieht über die gleichen Verteilwege, nämlich terrestrische Ausstrahlung, Kabel- und Satellitenübertragung, dabei ist in jedem Fall eine Modulation erforderlich, auf die hier zunächst eingegangen wird.

4.1 Analoge Modulationsverfahren

Bei der Basisbandübertragung ist pro übertragenem Signal eine eigene Leitung erforderlich. Für die Übertragung mehrerer Signale ist das nicht effektiv, da die auf der Leitung zur Verfügung stehende Bandbreite meist nicht ausgenutzt wird. Mit der Multiplextechnik wird die Übertragungseffektivität gesteigert. Um z. B. zwei Videosignale mit einer Bandbreite von je 5 MHz zu übertragen, kann das eine im Basisband belassen werden, während das zweite auf eine Trägerfrequenz moduliert wird, z. B so, dass das Frequenzband in den Bereich zwischen 6 und 11 MHz transformiert wird.

4.1.1 Multiplexverfahren

Die Versetzung von Frequenzbändern lässt sich ausweiten. Bei einer Übertragungsbandbreite von z. B. 30 MHz kann durch diese Frequenzmultiplextechnik die Übertragungskapazität gegenüber der Basisbandübertragung bei 6 MHz Kanalbandbreite verfünffacht werden.

Das Demultiplexing, d. h. die Rückgewinnung der Ausgangssignale, geschieht bei der Frequenzmultiplextechnik durch Filterung und Rücktransformation ins Basisband (Demodulation). Da die Filter nicht ideal arbeiten, müssen die Kanäle etwas mehr Bandbreite aufweisen als die übertragenen Signale.

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4_4, © Springer-Verlag Berlin Heidelberg 2013

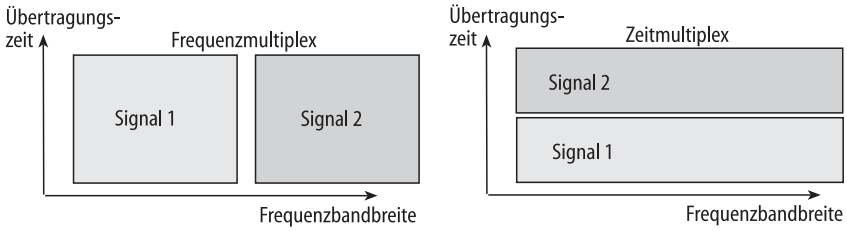


Abb. 4.1. Frequenz- und Zeitmultiplex

Neben dem angesprochenen Frequenzmultiplex- gibt es das Zeitmultiplexverfahren. Ein Signal belegt dabei den Kanal nicht während der gesamten Übertragungsdauer, sondern beispielsweise nur die Hälfte der Zeit, sodass die andere Hälfte für die Übertragung eines zweiten Signals zur Verfügung steht. Da man ein Signal ohne Qualitätseinbuße natürlich nicht einfach zeitweise abschalten kann, muss es bei der Zeitmultiplextechnik zunächst geeignet komprimiert und auf der Wiedergabeseite expandiert werden. Dieses Verfahren ist technisch aufwendiger als die Frequenzmultiplextechnik, allerdings auch qualitativ hochwertiger, da die Kanaltrennung in der Zeitebene so durchgeführt werden kann, dass die Basisbandsignale geringere gegenseitige Störungen aufweisen als bei einer Signaltrennung in der Frequenzebene.

Zwei Videosignale, die mit je 5 MHz Bandbreite im Zeitmultiplex auf einer Leitung übertragen werden, beanspruchen insgesamt aber nicht nur eine Bandbreite von 5 MHz. Die Übertragung eines um den Faktor zwei zeitkomprimierten Signals erfordert bei unveränderter Informationsmenge gerade die doppelte Bandbreite des Einzelsignals. Allgemein lässt sich eine gegebene Information bezüglich der Dimensionen Zeit und Bandbreite so darstellen, dass sich eine Fläche ergibt, deren Inhalt die Informationsmenge repräsentiert. Die konstante Fläche kann bei vergrößerter Breite (Bandbreite) in der Höhe (Zeit) verkleinert werden, d. h. die Zeitdauer eines Vorgangs und die Breite seines Spektrums stehen in einem reziproken Verhältnis. Dieser Umstand wird auch als Grundgesetz der Nachrichtentechnik bezeichnet (Abb. 4.1). Es besteht die Möglichkeit, die Fläche um eine dritte Dimension zu einem sogenannten Nachrichtenquader zu erweitern. Die dritte Dimension repräsentiert dann das Signal/Rauschverhältnis (Abb. 4.2). Bei großem Störabstand wird die im Quader liegende Informationsmenge groß, denn je kleiner das Rauschsignal ist, desto mehr Amplitudendifferenzen können aufgelöst werden. Wird die Nach-

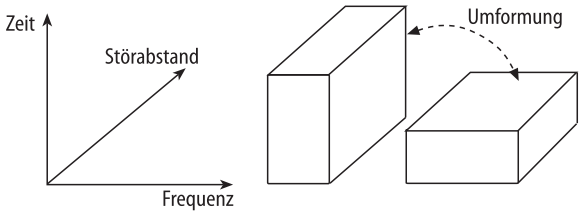


Abb. 4.2. Nachrichtenquader

richt durch Zahlen dargestellt, die nur die Werte 1 und 0 annehmen können, wie es z. B. in Digitalsystemen geschieht, so ist eine gegenüber der Analogübertragung vergrößerte Bandbreite erforderlich. Gleichzeitig werden aber geringere Ansprüche an den Störabstand gestellt, da die zwei Zustände auch bei Vorliegen eines relativ großen Störpegels noch leicht unterscheidbar sind.

4.1.2 Amplitudenmodulation (AM)

Nicht nur für das Frequenzmultiplexverfahren wird das Basisbandsignal in den Hochfrequenzbereich transformiert, sondern auch damit es effektiv über Antennen abgestrahlt werden kann. Dies geschieht durch Modulation eines hochfrequenten Trägersignals mit der Frequenz ω_T, das vom Nutzsignal der Frequenz ω_M verändert wird. Das modulierende Nutzsignal kann analoger oder digitaler Art sein. Als Trägersignals wird für die Fernsehübertragung ein Sinussignal benutzt:

$$a_T(t) = \hat{a}_T \sin(\omega_T t + \varphi_T).$$

Allgemein können sowohl die Amplitude $\hat{a}_T$, als auch die Frequenz ω_T und die Phase φ_T vom modulierenden Signal verändert werden. Für die analogen Fernsehübertragungsverfahren werden nur die Amplituden- oder Frequenzmodulation verwendet.

Bei den konventionellen Fernsehsystemen mit Antennen- oder Kabelübertragung wird die Amplitudenmodulation eingesetzt. Ein AM-Signal entsteht, indem die Amplitude der Trägerfrequenz gemäß dem Verlauf des zu übertragenden Signals verändert wird (Abb. 4.3). Das Trägersignal kann durch folgende Beziehung beschrieben werden [95]:

$$a_T(t) = \hat{a}_T \sin(\omega_T t + \varphi_T).$$

Die Gleichung für das modulierende Signal lautet:

$$a_M(t) = \hat{a}_M \cos(\omega_M t + \varphi_M).$$

Die Modulation entspricht mathematisch einer Multiplikation, damit folgt:

$$a_{AM}(t) = (\hat{a}_T + \hat{a}_M \cos(\omega_M t)) \sin(\omega_T t).$$

Mit dem Modulationsgrad $m = \hat{a}_M/\hat{a}_T$ ergibt sich :

$$a_{AM}(t) = \hat{a}_T (1 + m \cos(\omega_M t)) \sin(\omega_T t).$$

Mit der Beziehung $\sin\alpha \cos\beta = 1/2 (\sin(\alpha + \beta) + \sin(\alpha - \beta))$ folgt:

$$a_{AM}(t) = \hat{a}_T \sin(\omega_T t) + \hat{a}_M/2 \sin(\omega_T t + \omega_M t) + \hat{a}_M/2 \sin(\omega_T t - \omega_M t).$$

Die Frequenz ω_M repräsentiert das Basisfrequenzband mit den Grenzen ω_u und ω_o. Es tritt einmal als Summe und einmal als Differenz zur Trägerfrequenz auf, sodass zwei Seitenbänder, das untere (USB) und das obere Seitenband (OSB) entstehen (Abb. 4.4). Das AM-Signal benötigt damit insgesamt die zweifache Bandbreite (B_{NF}) des Basisbandsignals oder das Doppelte der höchsten Modulationsfrequenz: $B_{AM} = 2\, B_{NF}$.

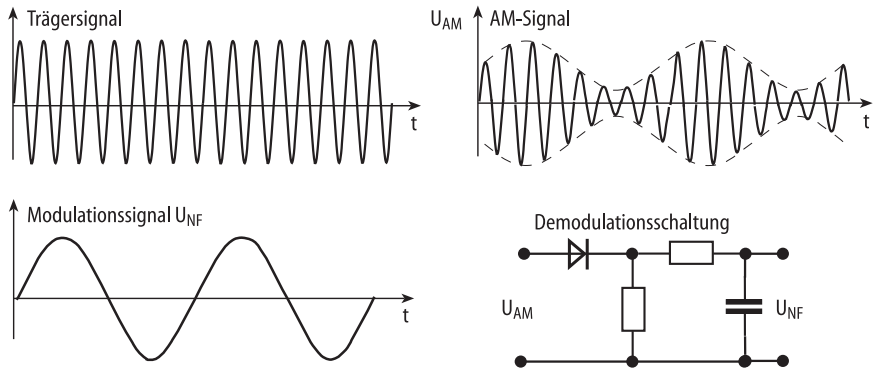


Abb. 4.3. Amplitudenmodulation

Für die Leistung einer harmonischen Schwingungen gilt allgemein, dass sie der Hälfte des Amplitudenquadrates proportional ist. Das gilt auch für die Sendeleistung P_T des Trägersignals:

$$P_T \sim 1/2\ \hat{a}_T^2.$$

Für die beiden Seitenbandsignale gilt:

$$P_S \sim 2 \cdot 1/2\ (\hat{a}_M/2)^2.$$

Daraus folgt für das Verhältnis von Seitenband- zu Trägerleistung:

$$P_S\ /P_T = m^2/2.$$

Hier wird ein wesentlicher Nachteil der AM deutlich: Selbst bei einem Modulationsgrad m = 100% (in der Praxis wird zur Vermeidung von Verzerrungen meist mit m < 70% gearbeitet) wird ein Großteil der Sendeleistung eines AM-Signals für den Träger benötigt, der keine Information enthält.

Der Vorteil der AM ist, dass die Demodulation auf der Empfängerseite sehr wenig Aufwand erfordert. Es reicht ein einfacher Hüllkurvendemodulator, bestehend aus einer Diode zur Unterdrückung einer Halbwelle und einem Tiefpassfilter zur Beseitigung der Trägerschwingung aus (Abb. 4.3). Die Demodulation kann ohne Rücksicht auf die Phasenlage vorgenommen werden (inkohärente Demodulation) und der Gleichanteil der Hüllkurve ermöglicht einen einfachen Ausgleich schwankender Empfangsfeldstärken.

Die Effektivität der AM kann durch Trägerunterdrückung gesteigert werden. Der Träger braucht nicht übertragen zu werden, da er sich aus den beiden Seitenbändern zurückgewinnen lässt:

$$(\omega_T - \omega_M) + (\omega_T + \omega_M) = 2\ \omega_T.$$

Da das zu übertragende Signal zweimal, nämlich jeweils im oberen und unteren Seitenband enthalten ist, kann bei der Übertragung auch ein Seitenband weggelassen werden. In diesem Fall liegt die Einseitenband-Modulation (EM), auch ESB oder SSB (Single Side Band) genannt, vor. EM erfordert einen relativ hohen Empfängeraufwand, da der Träger im Empfänger regeneriert werden

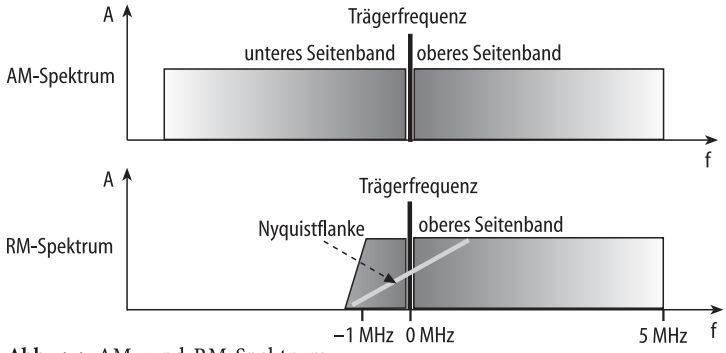


Abb. 4.4. AM- und RM-Spektrum

muss. Die Beseitigung eines Seitenbandes ist für die Übertragung von Videosignalen vor allem aus Gründen der Frequenzökonomie sehr erwünscht, denn EM erfordert im Vergleich zu AM nur die halbe Bandbreite. Diese Modulationsart ist für Videosignale jedoch praktisch nicht realisierbar. Die Trennung der Seitenbänder würde ein ideales Filter mit fast unendlicher Flankensteilheit erfordern, da das Videosignalspektrum bei nahezu 0 Hz beginnt.

Zur Fernsehsignalübertragung wird daher als Kompromiss die Restseitenband-AM (RSB-AM) verwendet, die auch als RM oder VSB (Vestigal Side Band) bezeichnet wird. Für ein 5-MHz-Standardvideosignal wird sendeseitig das untere Seitenband bis zu 0,75 MHz ungedämpft übertragen und dann steilflankig abgeschnitten, sodass bei 1,25 MHz das Signal quasi vollständig unterdrückt ist. Bei der Demodulation steht der Bereich bis 0,75 MHz doppelt zur Verfügung. Daher wird empfängerseitig mit einer sogenannten Nyquistflanke gefiltert, die eine Überlagerung der Teilspektren gerade so ermöglicht, dass die Restseitenbandlücke des oberen Seitenbandes von dem unteren Seitenbandrest aufgefüllt wird (Abb. 4.4).

4.1.2.1 Quadraturamplitudenmodulation (QAM)

Eine Steigerung der Effektivität der AM ist auch dadurch möglich, dass neben der Amplitude die Phase moduliert wird. Dies gelingt mit der sogenannten Quadraturamplitudenmodulation (QAM). Es werden hier zwei gleichfrequente, um 90° gegeneinander phasenverschobene Träger eingesetzt (Abb. 4.5). Die

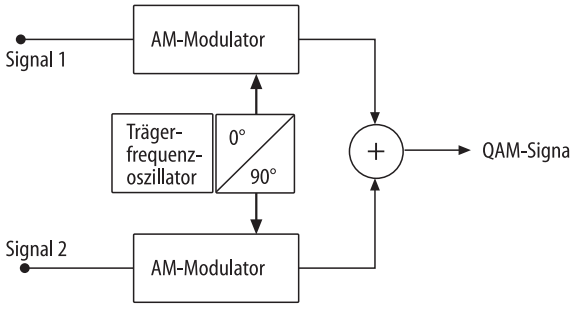


Abb. 4.5. QAM-Signalbildung

Träger werden von verschiedenen Signalen amplitudenmoduliert und dann additiv zusammengeführt. Zur Demodulation eines QAM-Signals ist der Zusatz eines Hilfsträgers erforderlich, der mit dem Originalträger in Frequenz und Phase synchronisiert ist (kohärente Demodulation). Bei dieser Art der Demodulation wird erreicht, dass das Signal für 0° Phasenverschiebung maximal wird, während es für 90° gerade verschwindet, d. h. die beiden Signale lassen sich trennen. QAM wird zur Übertragung der Farbdifferenzsignale im Luminanzspektrum verwandt, aber nicht als Modulationsart für die Sendetechnik.

4.1.3 Frequenzmodulation (FM)

Bei der Frequenzmodulation weicht der Augenblickswert f(t) der Trägerfrequenz von der Mittenfrequenz f_T um einen Betrag ab, der dem Momentanwert des modulierenden Signals proportional ist (Abb. 4.6). Die Frequenz des zu übertragenden Signals spiegelt sich in der Häufigkeit der Abweichung der Trägerfrequenz von ihrem Mittelwert. Der Maximalwert des modulierenden Signals führt zur größten Abweichung von der Trägerfrequenz, dem sogenannten Frequenzhub Δf. Bei sinusförmig frequenzmodulierten Signalen kann ein Modulationsindex η als Verhältnis des Frequenzhubes zur Modulationsfrequenz angegeben werden:

$$\eta = \Delta f / f_M.$$

Bei kleinem Hub liegt mit $\eta < 1$ eine Schmalband-FM vor und es gilt:

$$a_{FM}(t) \approx \hat{a}_T \, [\cos(\omega_T t) - \eta/2 \cos(\omega_T - \omega_M)\, t + \eta/2 \cos(\omega_T + \omega_M)\, t].$$

Es treten also auch hier wieder additiv und subtraktiv Seitenbänder auf, ohne die vorausgesetzte Näherung sind es unendlich viele Seitenschwingungen. Das FM-Spektrum kann aber unter Inkaufnahme kleiner, definierter Verzerrungen durch Filter beschränkt werden, da die Amplituden der Seitenschwingungen zu hohen Frequenzen hin schnell abnehmen. Je größer der Modulationsindex, desto feiner kann die modulierende Amplitude aufgelöst werden, desto größer ist der Störabstand. Bei großem Hub ($\eta > 1$) ergibt sich die Breitband-FM, für die hier erforderliche Bandbreite B_{FM} gilt:

$$B_{FM} = 2 \cdot (\Delta f + B_{NF}).$$

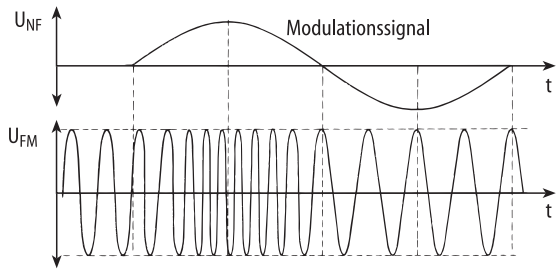


Abb. 4.6. Frequenzmodulation

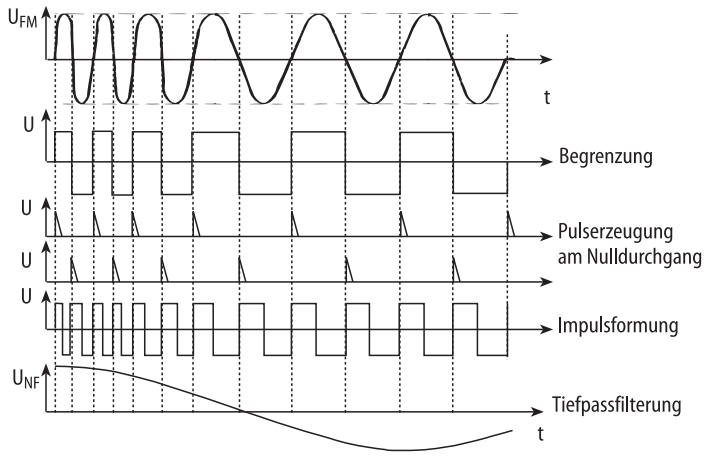


Abb. 4.7. Frequenzdemodulation

Die Frequenzmodulation kann mit einem Trägeroszillator vorgenommen werden, dessen Oszillatorfrequenz vom Modulationssignal gesteuert wird (Voltage Controlled Oscillator, VCO). Zur Demodulation eines FM-Signals kann ein Zähldiskriminator eingesetzt werden, der die bei FM veränderliche Anzahl von Nulldurchgängen auswertet. Zähldiskriminatoren erzeugen bei jedem Signal-Nulldurchgang einen Impuls und integrieren dann über die Impulsanzahl (Abb. 4.7). Eine weitere häufig verwandte Demodulationsart nutzt die PLL-Schaltung (Phase Locked Loop). Hier wird ein steuerbarer Oszillator (VCO) mit einem Regelkreis auf die veränderliche Phase des zu demodulierenden Signals abgeglichen. Bei FM muss der Oszillator ständig nachgestimmt werden, das Nachstimmsignal entspricht dem demodulierten Signal.

Breitband-FM erfordert eine größere Übertragungsbandbreite als AM, bietet dafür aber eine größere Störsicherheit, bzw. erfordert kleinere Störabstände. Der große Vorteil ist, dass die Information nicht in der Amplitude verschlüsselt ist. Die Amplitude ist der am stärksten von Störungen betroffene modulierbare Parameter. Bei FM können die Amplitudenfehler mit Amplitudenbegrenzern einfach beschnitten werden (Abb. 4.8). Damit wird, entsprechend dem Grundgesetz der Nachrichtentechnik, die Störsicherheit auf Kosten der Bandbreite erhöht. In der Videotechnik wird FM bei der analogen Magnetbandaufzeichnung verwendet und als Modulationsart für die Satellitenübertragung.

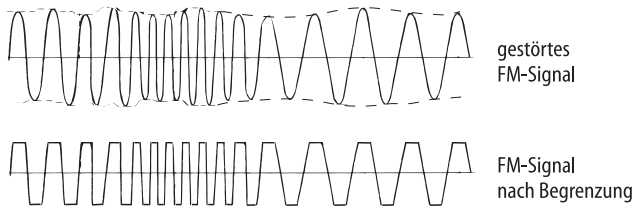


Abb. 4.8. Amplitudenbegrenzung bei FM

4.2 Fernsehbegleitsignale

Basis des analogen Fernsehsignals ist das in Kapitel 2 dargestellte FBAS-Signal. Das Fernsehprogramm besteht aber nicht nur aus dem Videosignal, sondern enthält den Begleitton und Zusatzdaten, die hier vorgestellt werden.

4.2.1 Fernsehbegleitton

Zur Verteilung des Videosignals an viele Empfänger (Broadcast) wird das FBAS-Signal auf einen Bildträger moduliert. Unabdingbar ist dabei die gleichzeitige Übertragung des Begleittons, die mittels der Frequenzmultiplextechnik mit einem eigenen Tonträger realisiert wird. Der Träger kann amplituden- oder frequenzmoduliert werden, in Deutschland wird mit FM gearbeitet.

Der Abstand zwischen Bild- und Tonträger liegt fest, jedoch gibt es hier auch innerhalb Europas einige Differenzen. Nach der in Deutschland gültigen CCIR-B/G-Norm beträgt der Bild-Tonträgerabstand 5,5 MHz, der Tonträger hat einen Abstand von 0,5 MHz zur oberen Videogrenzfrequenz. Dieser Abstand erlaubt den Einsatz der Frequenzmodulation, die gegenüber AM einen höheren Bandbreitenbedarf hat, aber auch eine höhere Qualität bietet. Um Störungen im Bild zu vermeiden, ist die Tonträgerleistung gegenüber dem Bildträger auf 1/20 reduziert (Abb. 4.9). Der Frequenzhub beträgt 50 kHz, die übertragene Audiobandbreite 15 kHz. Die Tonqualität entspricht damit der des UKW-Rundfunks.

Zur Übertragung eines zweiten Tonkanals wird ein weiterer Tonträger (TT_2) frequenzmoduliert. Die Leistung von TT_2 ist gegenüber dem ersten noch einmal auf 1/5 reduziert. Bei der Konzeption wurde großer Wert auf hohe Übersprechdämpfung zwischen den Tonkanälen gelegt, denn der zweite Tonkanal soll nicht nur für Stereosendungen sondern auch separat zu Verfügung stehen, um z. B. bei synchronisierten Beiträgen den Originalton im Zweikanalbetrieb senden zu können.

Der zweite Tonkanal arbeitet auch mit 15 kHz Bandbreite und gewährleistet die gleiche Qualität wie der erste. Um Störungen zu minimieren, ist die zweite Tonträgerfrequenz mit der Zeilenfrequenz derart verkoppelt, dass er etwa um 250 kHz höher liegt als TT_1. Es gilt:

$$f_{TT_2} = f_{TT_1} + 15{,}5 \cdot f_H = 5{,}742 \text{ MHz}.$$

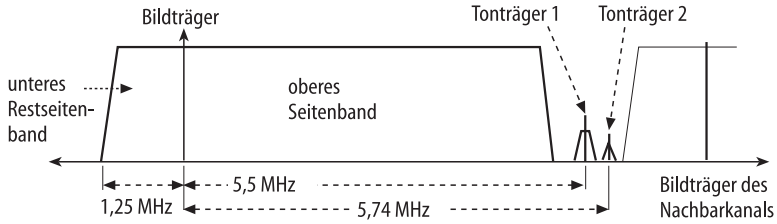


Abb. 4.9. Lage der Bild- und Tonträger

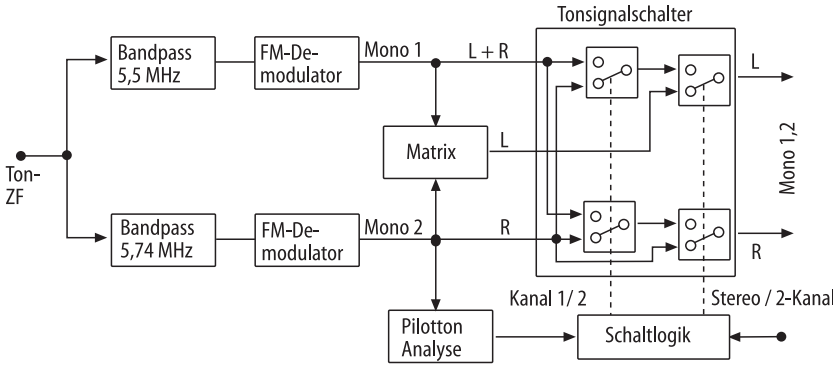


Abb. 4.10. Stereo- und Zweikanaltongewinnung

Empfänger, die beide Tonkanäle auswerten, müssen mit Verstärker- und Lautsprecherpaaren für die Stereowiedergabe ausgestattet sein und eine Umschalteinheit zur Anwahl des Stereo- oder Zweikanalmodus enthalten. Bei Zweikanalbetrieb wird wahlweise entweder der eine oder der andere Tonkanal auf beide Lautsprecher geschaltet, bei Stereobetrieb müssen die Lautsprecher den Rechts- Linkskanälen zugeordnet werden (Abb. 4.10).

Die Umschaltung zwischen Zweikanal- und Stereomodus erfolgt automatisch mit Hilfe eines Pilothilfsträgers der Frequenz $f_p = 3{,}5 \cdot f_H = 54{,}7$ kHz, der bei Stereosendungen mit 117,5 Hz und bei Zweikanalbetrieb mit 274,1 Hz amplitudenmoduliert wird. Bei Stereobetrieb wird wegen der Kompatibilität zu Monoempfängern über Tonträger 1 das Monosignal (L + R)/2 übertragen, Tonträger 2 trägt nur das R-Signal. Im Empfänger wird im Stereomodus der Linksanteil nach folgender Beziehung gebildet:

$$L = 2 \cdot (L + R)/2 - R.$$

Neben der Tonübertragung im Frequenzmultiplex gibt es weitere, weniger bedeutungsvolle Verfahren, von denen hier nur zwei kurz erwähnt werden sollen. Das COM-Verfahren nutzt für einen Tonkanal eine Zeile in der Vertikalaustastlücke eines jeden Halbbildes. Das kontinuierliche Audiosignal wird dabei in 20 ms-Sequenzen aufgeteilt, die, entsprechend der nutzbaren Zeilendauer, auf 50 µs komprimiert werden. Damit bei dieser Zeitkompression um Faktor 400 kein Qualitätsverlust stattfindet, wird das komprimierte Signal mit der 400-fachen Bandbreite übertragen, woraus sich bei einer Videobandbreite von 5 MHz eine maximale Audiofrequenz von 12,5 kHz ergibt. Die Audioqualität ist damit ausreichend und könnte ggf. durch Nutzung weiterer Zeilen gesteigert werden.

Beim SIS-Verfahren (Sound in Sync) wird ein digitales Audiosignal verwendet. Die A/D-Wandlung geschieht mit der Abtastrate $f_s = 2 \cdot f_H$, die Digitaldaten werden in der Horizontalaustastlücke, im Bereich des Synchronsignals, platziert. Das Verfahren kommt nur in speziellen Fällen zur Anwendung, da es in Heimempfängern zu Störungen führen kann.

4.2.2 Signale in der Vertikalaustastlücke

Die Vertikalaustastlücke enthält bei jedem Halbbild 25 für das Bild ungenutzte Zeilen, von denen nur wenige für die Auslösung des Bildwechsels benötigt werden. Schon früh wurde daher versucht, diese Zeilen zur Übertragung von Zusatzinformationen zu nutzen (Abb. 4.11). Zunächst wurde Teletext als textorientierter TV-Zusatzdienst eingeführt, später kamen die Prüfzeilensignale hinzu. Um eine internationale Kompatibilität zu gewährleisten, wurde festgelegt, dass die Zeilen 7...15, 20, 21 im ersten und 320...328, 333, 334 im zweiten Halbbild für den Teletext genutzt werden können. Die Zeilen 16/329 sind für die Datenübertragung und VPS reserviert, die Zeilen 17...19/330...332 sind Prüfzeilen.

4.2.2.1 Prüfzeilen

Zwei Zeilen pro Halbbild sind für messtechnische Zwecke definiert: Es sind die Zeilen CCIR 17, 18 und 330, 331, welche aber, abhängig von der Sendeanstalt, auch auf Zeile 19 ausgedehnt werden. Diese Prüfzeilen sind jeweils in 32 Schritte á 2 µs Dauer eingeteilt, die Schritte 6–31 werden für messtechnische Informationen genutzt. Mit den Prüfzeilen können Amplitudenfrequenzgang, Einschwingverhalten, Bandbreite etc. bestimmt werden. Sie können automatisch ausgewertet und Übertragungsstrecken entsprechend entzerrt werden.

Abbildung 4.12 zeigt eine Übersicht über die Prüfzeilen. Zeile CCIR 17 beginnt mit einem Weißimpuls von 10 µs Dauer, der als Amplitudenbezugswert für alle weiteren Signale dient. An der Dachschräge kann das NF-Verhalten des zu untersuchenden Übertragungssystems beurteilt werden. Es folgt dann ein 2T-Impuls. T steht für die Einschwingzeit eines Übertragungskanals der Bandbreite B. Bei einer Grenzfrequenz von 5 MHz beträgt T = 1/(2B) = 100 ns. Der 2T-Impuls hat sinusförmige Flanken und die Eigenschaft, dass keine Frequenzanteile oberhalb 5 MHz auftreten (Abb. 4.13). Aus dem Amplitudenabfall

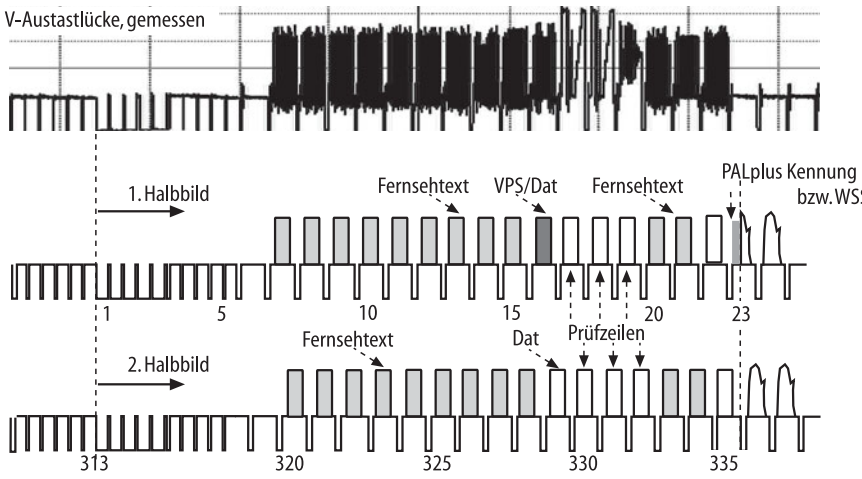


Abb. 4.11. Nutzung der V-Austastlücke

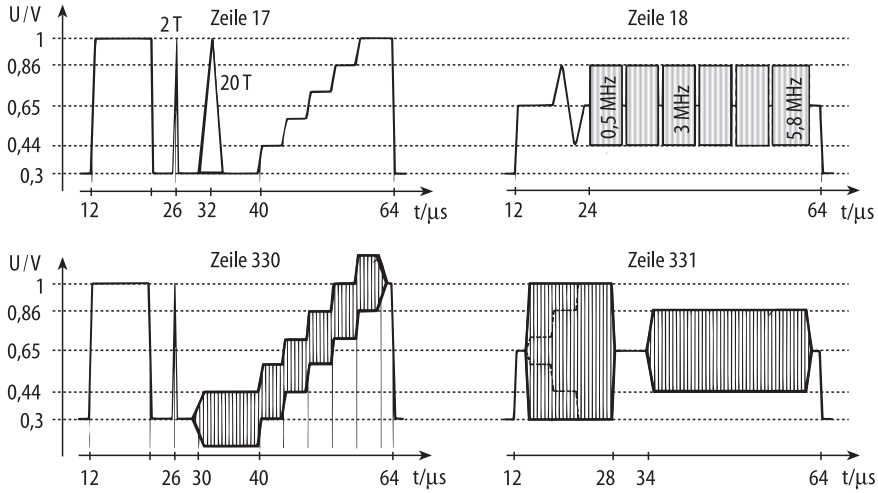
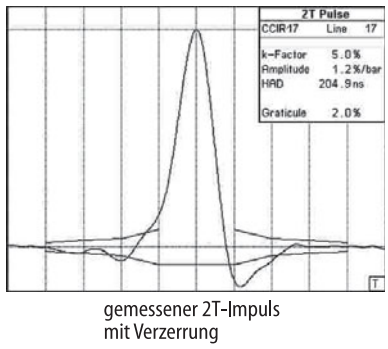


Abb. 4.12. Prüfzeilensignale

des 2T-Impluses können Schlüsse auf den Amplitudenfrequenzgang des Systems gezogen werden, dazu wird auch der k-Faktor benutzt (s. Kap. 2.4). Am folgenden 20T-Impuls können Dämpfungs- und Phasenverzerrungen beurteilt werden, denn diese führen zu einer Aushöhlung oder zu sinusförmiger Verzerrung am Boden des 20T-Impulses (Abb. 4.13). Mit der Grautreppe in Zeile 17 kann schließlich die Linearität des Systems überprüft werden.

In Zeile 18 wird ein Multiburstsignal, d. h. eine Folge von sechs verschiedenen Frequenzen (0,5 ... 5,8 MHz) übertragen, an der der Amplitudenfrequenzgang erkennbar ist. Die Amplitude beträgt 0,42 V_{ss}, der Mittelwert 0,65 V. Prüfzeile 330 enthält wie Zeile 17 den Weiß-, den 2T-Impuls und die Grautreppe, die hier mit einem Farbträgersignal mit 0,28 V_{ss} überlagert ist, um differen-



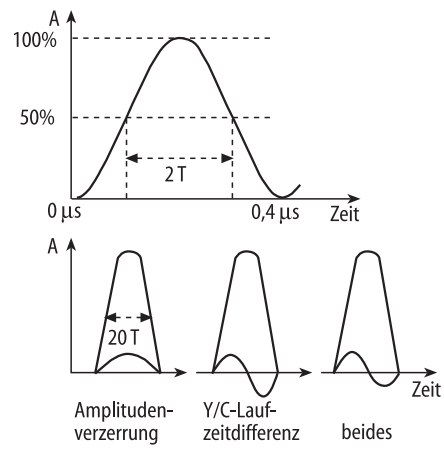


Abb. 4.13. 2T-Impuls und Verzerrungen am 20T-Impuls

Zeile	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	320	321	322	323	324	325	326	327	328	333	334
ARD	x	x	x		x	x	x	x	x	V	17	17	18	x	x		x	x	x		x	x	x	x	x	x	x
SAT1	x	x	x		x	x	x	x	x	V	17	17	18	x	x		x	x	x		x	x	x	x	x	x	x
N3	x	x	x		x	x	x	x	x	V	17	17	18	x	x		x	x	x		x	x	x	x	x	x	x

x = Fernsehtext, 17/18 = CCIR 17/18

Abb. 4.14. Nutzung der V-Austastlücke für Videotext und Prüfzeilen bei versch. Sendern

tielle Amplituden- und Phasenfehler feststellen zu können. Zeile 331 enthält schließlich zwei Farbträgerpakete, mit denen die Intermodulation zwischen Chrominanz- und Luminanzsignal bestimmt werden kann.

4.2.2.2 Teletext

Tele-, Fernseh- oder auch Videotext (bei ARD und ZDF) ist eine Übertragung von Zusatzinformationen als Text oder in grafischer Form in den Zeilen der Vertikalaustastlücke des Videosignals. Die Inhalte hängen von den jeweiligen Sendeanstalten ab. Beispiele sind: Programmhinweise, Untertitel, Nachrichten, Börsennachrichten und Computerdaten. Videotext wird üblicherweise anstelle des Fernsehbildes eingeblendet, wobei der Ton unbeeinflusst bleibt. Im Mix-Mode kann die Videotextwiedergabe auch gleichzeitig mit dem Fernsehbild erfolgen, um z. B. Untertitel darzustellen.

Das Videotextsignal wird in digitaler Form in die Austastlücke eingefügt. Die genutzten Zeilen sind von der Sendeanstalt abhängig (Abb. 4.14). Die Daten sind mit 8-Bit-Worten im ASCII-Code verschlüsselt und werden im NRZ-Code übertragen (Abb. 4.17). Die Taktfrequenz beträgt $f_{Bit} = 444 \cdot f_H = 6{,}9375$ MHz, für ein Bit stehen 144 ns zur Verfügung. Die Amplitude des Fernsehtextsignals beträgt 66% BA. In einer Zeile können insgesamt 360 Bit oder 45 Byte übertragen werden, die Übertragungskapazität beträgt bei zwei genutzten Zeilen also 36 kbit/s. Von den 45 Byte pro Zeile stehen effektiv nur 40 Byte für die Nachricht zur Verfügung, drei Byte werden für die Datensynchronisation verwandt, zwei weitere zur Magazinadressierung (Abb. 4.15).

Die Informationen werden in Videotextseiten aufgeteilt, die jeweils gerade den Bildschirm füllen. Eine Seite enthält 24 Textzeilen, die wiederum aus 40 Zeichen bestehen. Die Grafikdarstellung ist daher auf 40 x 24 Bildelemente pro Bildschirmseite beschränkt und demgemäß sehr grob (Abb. 4.16). Wenn

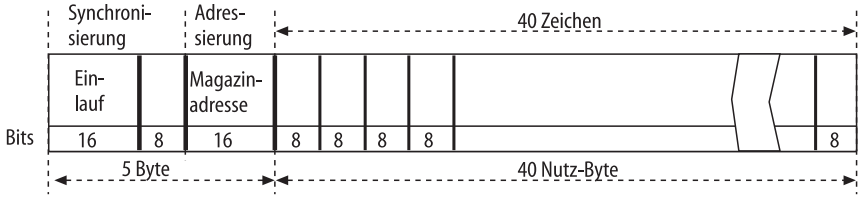


Abb. 4.15. Aufbau der Datenzeile bei Teletext

Abb. 4.16. Zuordnung von Videotextdaten und Bildzeile

100 Textseiten zu einem Magazin zusammengefasst werden, enthält dieses 96 kB Daten. Bei Nutzung von zehn Videotextzeilen pro Halbbild ergibt sich eine maximale Übertragungsdauer von $t_{max} = 100 \cdot 24\ Z \cdot 20\ ms/10Z = 4{,}8\ s$ für ein 100-Seiten-Magazin. Die Wartezeit ist in der Praxis geringer, da Textseiten im Empfänger gespeichert werden, und wichtige Seiten häufig wiederholt werden.

4.2.2.3 VPS

Die binär codierte Video Program System Information (VPS) wird in Zeile 16 der V-Austastlücke übertragen (Abb. 4.17). Durch Vergleich der VPS-Daten (Abb. 4.18) mit einer in einem Videorecorder gespeicherten Sollstartzeit wird ermöglicht, dass der Recorder genau ab dem tatsächlichen Programmbeginn aufzeichnet. Das VPS-Signal beinhaltet den Ist-Zeitwert, der bei zeitversetzt beginnenden Beiträgen simultan verschoben wird, sodass der Start des Recorders durch einen einfachen Soll/Ist-Vergleich ausgelöst werden kann. Sobald und solange Ist/Soll-Gleichheit vorliegt, wird aufgezeichnet. Die Datenzeile 16 enthält 15 Datenworte, VPS ist in den 8-Bit-Worten 11–14 verschlüsselt, die weiteren Worte werden von den Sendeanstalten innerbetrieblich genutzt [40].

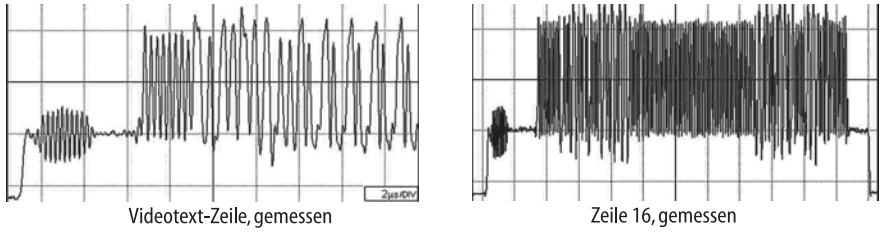


Abb. 4.17. Gemessene Zeilen in der V-Austastlücke

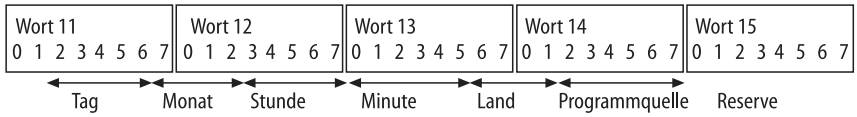


Abb. 4.18. VPS-Codeworte

4.2.2.4 PAL-plus-Kennung

Zur Übertragung einiger Zusatzinformationen zum PALplus-Signal (s. Abschn. 4.4.3) wird auch die vertikale Austastlücke benutzt. Trotz dichter Belegung durch die dargestellten Dienste fand sich ein noch nicht genutztes Gebiet, das zudem den Vorteil aufweist, dass es auch von Heim-Recordern aufgezeichnet wird. Es ist dies die Zeile 23, die nach CCIR-Norm noch zur Hälfte zur Austastlücke gehört. Hier werden 14 Bits im Bi-Phase-Mark-Code (s. Kap. 3.2) mit einer Taktfrequenz von 833 kHz übertragen, die Informationen über das Bildformat, Kamera- oder Filmmode, Colour Plus etc. enthalten. Unabhängig von PALplus ist dieses Wide Screen Signaling Signal (WSS) überall dort von Bedeutung, wo Formatänderungen auftreten können, wie z. B. auch bei der DVD.

4.2.2.5 VITC

Auch der Vertical Interval Timecode (VITC) ist ein Digitalsignal, das in der V-Austastlücke übertragen wird. Es wird aber nicht in das ausgestrahlte Fernsehsignal eingetastet, sondern nur bei der Nachbearbeitung innerhalb des Studios verwendet und dient der Nummerierung der Videobilder (s. Kap. 9.4).

4.3 Analoge Übertragungsverfahren

Die Verteilung (Distribution) von Fernsehprogrammen an die vielen Empfänger zuhause geschieht über die Verteilwege terrestrische Ausstrahlung, Kabel und Satellit. Terrestrische Ausstrahlung war bis in die 90er Jahre des 20.Jahrhunderts hinein die wichtigste Art, verlor dann aber gegenüber der Kabel- und Satellitenübertragung stark an Bedeutung. Dies lag u. a. an der begrenzten Kanalvielfalt, die in ländlichen Gebieten bis zu 5 und in Ballungsgebieten bis zu 12 terrestrisch verteilte analoge Programme bot. Abbildung 4.19 zeigt die Entwicklung der Nutzung der verschiedenen Verteilwege.

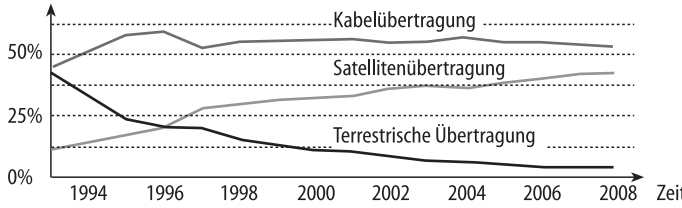


Abb. 4.19. Veränderung der Nutzung der drei Verteilwege für analoge Fernsehsignale

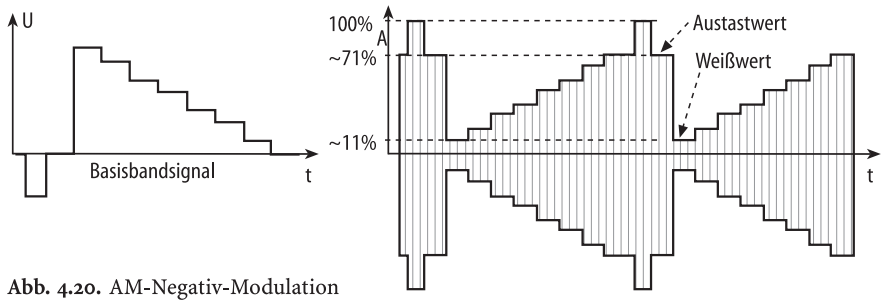


Abb. 4.20. AM-Negativ-Modulation

4.3.1 Terrestrische Ausstrahlung

Terrestrische Ausstrahlung bezeichnet die Signalübertragung mit Hilfe elektromagnetischer Wellen, die sich in der Nähe des Erdbodens ausbreiten. Das Videosignal wird auf eine Trägerfrequenz moduliert, zur Distribution wird als Modulationsart für das FBAS-Videosignal die RSB-AM (s. Abschn. 4.1.2) eingesetzt. Vor der Amplitudenmodulation wird das Videosignal invertiert (AM-Negativ). Das hat den Vorteil, dass der Mittelwert des Signals kleiner ist als beim nicht invertierten Signal, da die Signalspitzen nun durch die relativ kurzen Synchronpulse gebildet werden. Damit ist im Mittel eine geringere Sendeleistung als beim nicht invertierten Signal erforderlich. Positiv gerichtete Störimpulse erscheinen im Empfänger zudem als schwarze Flecke, die wenig störend in Erscheinung treten als weiße. Das invertierte Signal wird so eingestellt, dass, bezogen auf den Wert 100% für die Synchronimpulse, das Bildsignal zwischen 71% (Schwarz) und 11% (Weiß) liegt (Abb. 4.20).

Bei der terrestrischen Ausstrahlung werden Trägerfrequenzen zwischen 170 und 800 MHz im VHF- und UHF-Bereich verwendet (VHF: Very High Frequency, UHF: Ultra High Frequency). Die elektromagnetischen Wellen haben hier relativ kleine Wellenlängen (Größenordnung Dezimeter) und breiten sich wie Licht, also »quasioptisch« aus. Elektromagnetische Wellen im VHF- und UHF-Bereich werden nicht – wie Kurzwellen unter 30 MHz – in der Erdatmosphäre an der Ionosphäre reflektiert, sondern dringen durch diese hindurch. Das Signal muss also den Empfänger möglichst auf direktem Weg erreichen, Hindernisse führen zur Abschattung der Welle (Abb. 4.21).

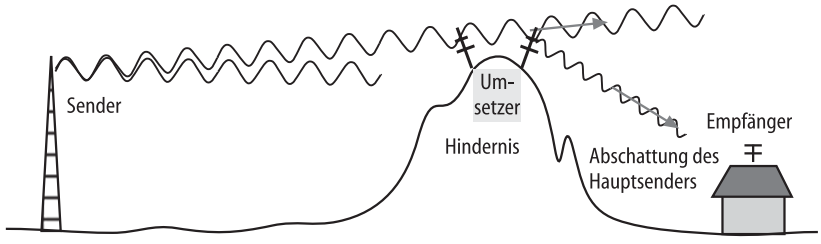


Abb. 4.21. Abschattung der terrestrischen Welle

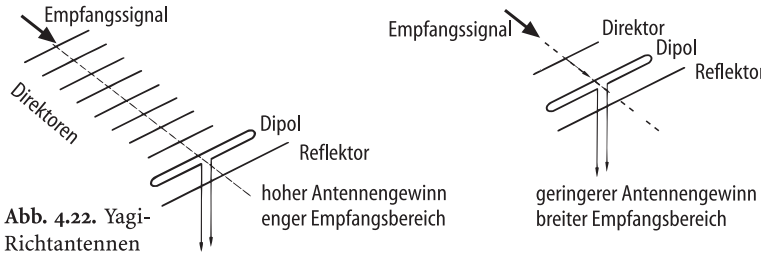


Abb. 4.22. Yagi-Richtantennen

Die Ausdehnungen von Sende- und Empfangsantennen liegen in der Größenordnung der Wellenlänge, also im Dezimeterbereich. Sendeseitig wird eine hohe Leistung an Rundstrahl- oder Richtantennen aufgewandt. Als Empfänger dienen vergleichsweise einfach ausgestattete Geräte mit Antennen (Yagi), die auf den Sender ausgerichtet werden (Abb. 4.22).

Die zur Verfügung stehenden Frequenzbänder im VHF- und UHF-Bereich werden mit der Frequenzmultiplextechnik zur Übertragung mehrerer Fernsehprogramme genutzt. Dazu wurden im VHF-Band nach der CCIR-B-Norm Kanäle mit 7 MHz Bandbreite und im UHF-Band, nach CCIR-G, Kanäle mit 8 MHz Bandbreite festgelegt. Das Videosignal beansprucht davon nur 5 MHz, wegen der Restseitenbandmodulation muss aber ca. 1 MHz Bandbreite zusätzlich für das untere Seitenband zur Verfügung stehen (s. Abschn. 4.1.2). Außerdem wird das Videosignal immer zusammen mit dem Fernsehbegleitton im Frequenzmultiplex übertragen. Die Tonkanäle liegen 5,5 MHz bzw. 5,74 MHz oberhalb des Bildträgers, sodass sich eine Mindestkanalbandbreite von 7 MHz ergibt. Abbildung 4.23 zeigt die Frequenzbelegung zweier Kanäle.

Bei der Erschließung der Frequenzbereiche wurde zunächst der VHF-Bereich belegt. Im Bereich 47–68 MHz liegen in Deutschland hier die Kanäle 2–4 und zwischen 174 MHz und 230 MHz die Kanäle 5–12. Später wurde auch der UHF-Bereich mit den Kanälen 21–60 genutzt. Eine Übersicht über die gesamte Kanalaufteilung bietet Tabelle 4.1.

Zur flächendeckenden terrestrischen Versorgung, z. B. bei der Verbreitung des Programms des Zweiten Deutschen Fernsehens, ist eine Vielzahl von Sendern erforderlich. Obwohl die Sender alle dasselbe Programm abstrahlen, können sie nicht alle die selbe Trägerfrequenz benutzen. Wenn ein Haushalt ein Programm aus mehreren Quellen empfängt, haben die Signale verschiedene Laufzeiten, und die Überlagerung der empfangenen Bilder führt zu Störungen

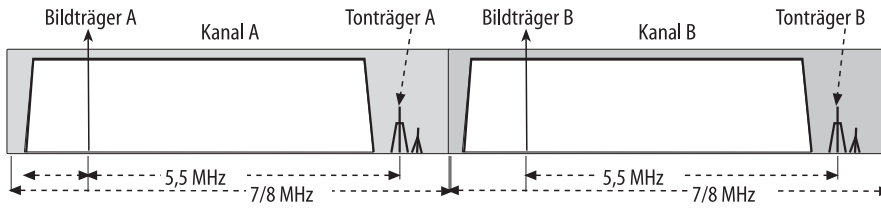


Abb. 4.23. Zwei Fernsehkanäle

Tabelle 4.1. Fernsehkanäle im VHF- und UHF-Band

Kanal	Frequenz (MHz) (untere Grenze)		Kanal	Frequenz (MHz)	Kanal	Frequenz (MHz)
K 2	47		K 21	470	.	.
K 3	54		K 22	478	.	.
K 4	61		K 23	486	K 58	766
UKW-Hörfunk		VHF UHF	K 24	494	K 59	774
			K 25	502	K 60	782
K 5	174		K 26	510	K 61	790
K 6	181		K 27	518	K 62	798
K 7	188		K 28	526	K 63	806
K 8	195		K 29	534	K 64	814
K 9	202		K 30	542	K 65	822
K 10	209		K 31	550	K 66	830
K 11	216		.	.	K 67	838
K 12	223		.	.	K 68	846

(Geisterbilder). Wegen des aus diesem Problem resultierenden hohen Bedarfs an verschiedenen Fernsehkanälen können in Deutschland nur 3–4 analoge Fernsehprogramme flächendeckend terrestrisch übertragen werden. Die Rundfunkanstalten ARD und ZDF brauchen dafür jeweils ca. 100 starke Grundnetzsender (100– 500 kW) und fast 3000 Füllsender (Tabelle 4.2) [19]. Füllsender dienen zur Versorgung abgeschatteter Gebiete. Es werden Transponder benutzt, die das Signal des Grundnetzsenders empfangen und in das betroffene Gebiet auf einer anderen Frequenz wieder abstrahlen (Abb. 4.21). Wegen der quasioptischen Wellenausbreitung werden die Transponder möglichst hoch angeordnet.

Terrestrische Ausstrahlung und Transpondertechnik wurden auch für Zuspielzwecke (Contribution) eingesetzt. Es existierte ein eigenes, von der Telekom betriebenes Richtfunknetz für den Überspielbetrieb der Rundfunkanstalten. Mobile Richtfunkanlagen überbrücken Entfernungen bis zu 30 km, sie arbeiten mit FM meist im 13 GHz-Bereich. Kleinrichtfunkgeräte (Window-Units) ermöglichen drahtlose Verbindungen zwischen Kamera und Ü-Wagen bzw. Kamera und Aufzeichnungsgerät. Bei Übertragungen aus ausgedehntem, hügeligen Gelände (z. B. bei Sportereignissen) werden auch Hubschrauber als Träger von Transpondern eingesetzt.

Tabelle 4.2. Terrestrische Versorgung

Sendeanstalt	Grundnetzsender	Umsetzer	Versorgungsgrad
ARD	94	2821	99,3%
ZDF	104	2869	98,9%
Dritte Progr.	119	3048	98,7%
RTL	88		
SAT 1	104		
VOX/Pro 7	44/37		
DSF/NTV	7/4		
AFN/SSVC/BRT	156		(Stand: 1999)

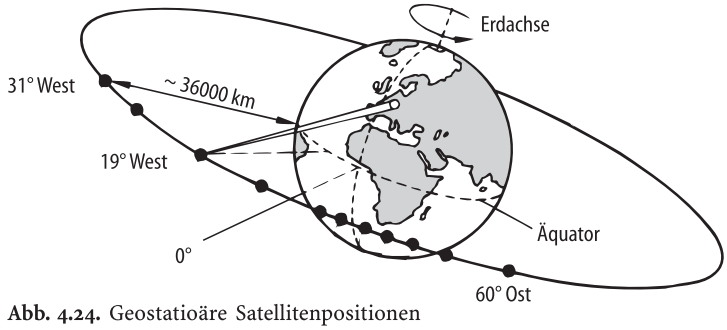


Abb. 4.24. Geostatioäre Satellitenpositionen

4.3.2 Satellitenübertragung

Optimal positionierte Transponder stehen in großer Höhe. Die höchsten befinden sich auf Satelliten im Weltraum, die die Erde umkreisen. Man unterscheidet die Orbitalpositionen nach dem Abstand zur Erde. Für Zwecke wie Erderkundung oder Überwachung werden sehr viele erdnahe Satelliten verwendet. Sie befinden sich im Low Earth Orbit (LEO) zwischen 200 und 1400 km über der Erde oder im Medium Earth Orbit (MEO) im Abstand von 1000 km bis 35 000 km (Abb. 4.25).

Die für die Fernsehübertragung wichtigste Umlaufbahn ist der geostationäre Orbit (GEO). Hier, bei einem Abstand von 35786 km zwischen Satellit und Erdoberfläche, ergibt sich aus der Gleichheit von Zentrifugal- und Gravitationskraft eine Umlaufzeit des Satelliten die der Eigenrotationszeit der Erde entspricht. Satelliten, die die Erde parallel zur Äquatorlinie pro Tag genau einmal umkreisen befinden sich damit auf einer geosynchronen Umlaufbahn und stehen quasistationär über der Erde. Kleine Abweichungen, die sich z. B. aus Gravitationskräften von Mond und Sonne ergeben, müssen im Laufe der Zeit immer wieder korrigiert werden. Dafür sind die Satelliten mit mehreren Düsenantrieben ausgestattet, die von der Erde aus gesteuert werden. Die dafür zur Verfügung stehende Treibstoffmenge begrenzt die Lebensdauer der Satelliten auf ca. 15 Jahre.

Die Satelliten werden quasi nebeneinander über dem Äquator aufgereiht, optimale Positionen sind nur beschränkt verfügbar (Abb. 4.24). Von der nördlichen Halbkugel aus erscheinen die Satelliten generell in südlicher Richtung.

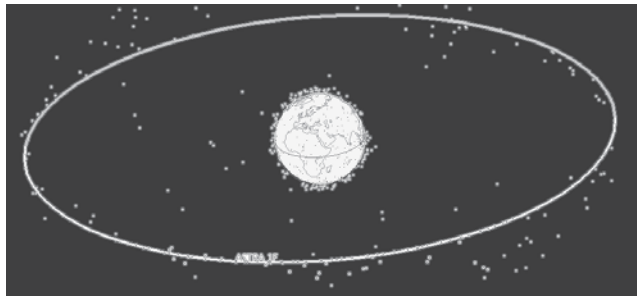


Abb. 4.25. Künstliche Satelliten im Bereich der Erde [180]

Abb. 4.26. Satellit mit Antennen und Solarpanels [180]

Ausgehend vom 0-Längengrad (Greenwich Medridian) wird die Position nach Osten oder Westen angegeben (Azimut-Winkel). Abhängig vom Breitengrad müssen verschiedene Elevationswinkel gewählt werden, um einen Satelliten anzupeilen. Das für den Direktempfang in Mitteleuropa wichtigste Satellitensystem der Firma SES-Astra hat z. B. die Position 19,2° Ost und muss von München aus unter einem Elevationswinkel von 34,2° und von Hamburg aus unter 28,3° anvisiert werden.

Frühere internationale Festlegungen der ITU sahen eine Unterscheidung in Fernmeldesatelliten (Fixed Satellite System, FSS) und Satelliten für den Direktempfang im Heim vor (Broadcasting Satellite System, BSS), die mit höherer Sendeleistung (100 W) arbeiten wie erstere. Diese Unterscheidung wurde zunehmend unwichtiger, nachdem sich SES-Astra 1988 entschied, ihr System für den Direktempfang mit FSS-Satelliten mittlerer Leistung (40 W) zu realisieren und andere Satellitenbetreiber folgten.

Die Satelliten selbst dienen als Realaisstation die mehrere Transponder tragen. Neben dem Direktempfang werden sie auch für die TV-Zuspielung während der Produktion, u. a. beim Satellite News Gathering (SNG) und zur Einspeisung der TV-Kanäle in Kabelnetze genutzt. Moderne Satelliten sind mit bis zu 32 oder mehr Transpondern ausgestattet, die meist mit Bandbreiten von 26,

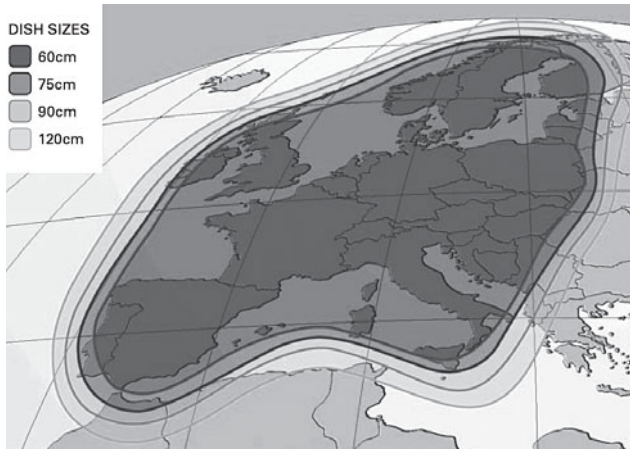


Abb. 4.27. Footprint des Satelliten Astra 1H [180]

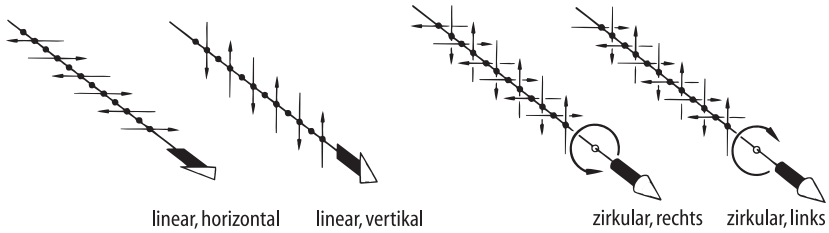


Abb. 4.28. Polarisationsrichtungen

33 und 36 MHz arbeiten. Die Trägerfrequenzen liegen im K_u-Band zwischen 11 GHz und 18 GHz. Die zu übertragenden Signale werden von der Erdefunkstelle mit relativ großer Leistung (500 W) und hoher Richtwirkung (Antennengewinn 63 dB) zum Satellitentransponder gestrahlt (Up-Link), wobei für das FSS der Frequenzbereich 12,75 ... 14,5 GHz und für das BSS der Bereich 17,3 ... 18,1 GHz genutzt wird. Satellitentransponder für den Direktempfang geben dann das Signal im K_u-Band mit dem Frequenzbereich 10,7 ... 11,7 und 11,7 ... 12,75 GHz wieder ab (Down-Link). Die Laufzeit des Signals beträgt für up- und downlink insgesamt ca. 250 ms.

Bei Analogübertragung wird das Signal frequenzmoduliert mit 8,5 MHz Hub abgestrahlt, daher sind Gesamtbandbreiten von mehr als 26 MHz pro Kanal erforderlich. Aufgrund der großen Störsicherheit von FM-Signalen kann die Sendeleistung der Transponder auf 1/10 der Sendeleistung gesenkt werden die für AM-Übertragung erforderlich wäre. Die Energieversorgung geschieht mit Hilfe von Solarzellen, die die Größe des Satelliten bestimmen. Bei modernen Satelliten mit 26 m Spannweite werden ca. 7 KW erzeugt (Abb. 4.26).

Über die Sendeantennen auf dem Satelliten wird festgelegt, welche Bereiche auf der Erde in welcher Form bestrahlt werden. Durch die Antennen wird der sogenannte Footprint des Satelliten bestimmt (Abb. 4.27). Anstelle der Feldstärke wird dabei oft der Durchmesser des Empfangsreflektors angegeben, der für einen Empfang bei gewöhnlichen Witterungsbedingungen ausreichend ist.

Um mit einem Transponder möglichst viele Fernsehkanäle übertragen zu können, werden die einzelnen Trägerfrequenzen doppelt genutzt, indem die Trägerwellen mit unterschiedlichen Polarisationsrichtungen abgestrahlt werden. Stehen die Polarisationsvektoren senkrecht aufeinander (horizontal/vertikal bzw. rechts/links zirkular), so lassen sich die Kanäle eindeutig trennen (Abb. 4.28). Beim Astra-Satellitensystem wird lineare Polarisation verwendet.

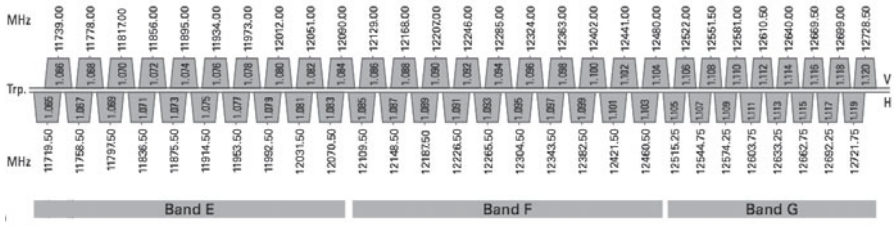


Abb. 4.29. Frequenzbelegung und Polarisation bei Astra 1H [180]

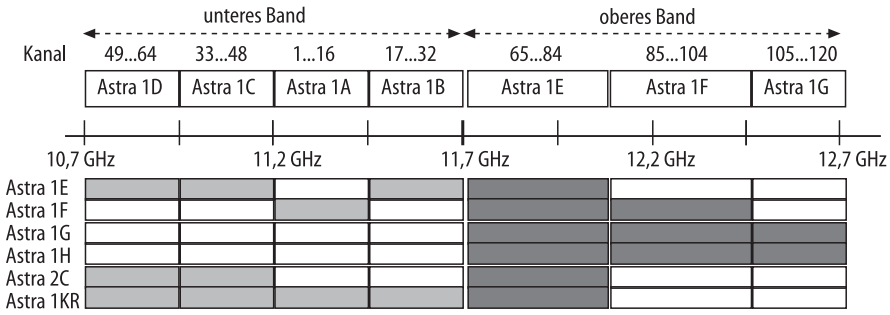


Abb. 4.30. Astra-Satellitensystem

Am Beispiel des Satelliten Astra 1H zeigt Abbildung 4.29, dass die Grenzen der Frequenzbänder für die eine Polarisationsrichtung in die Mitte der Bänder für die andere gelegt werden.

Abbildung 4.30 zeigt eine Übersicht über das Astra-Satellitensystem auf der Position 19,2° Ost. Im oberen Teil ist das ursprüngliche System zu sehen. Auf den Astra-Satelliten 1A – 1D standen früher bei einem 27 MHz-Kanalraster insgesamt 64 Transponder zur Verfügung. Jeder dieser Transponder übertrug ein analoges Fernseh- und mehrere Hörfunkprogramme im unteren Frequenzband bis 11,7 GHz. Astra 1A und 1B sind inzwischen außer Dienst gestellt, 1D ist verkauft. Auf den Satelliten Astra 1E – 1G stehen 56 Transponder zur Verfügung, die für die Übertragung digitaler Programme im oberen Frequenzband (11,7 ... 12,75 GHz) genutzt werden. Die Übertragungsbandbreiten betragen hier 33 MHz für Astra 1E und F und 27 MHz für Astra 1G [201].

Als Ersatz für Astra 1A–1D befinden sich auf dieser Position nun die Satelliten Astra 1H, 1KR und 1L zur Versorgung von Mitteleuropa. Im unteren Teil der Abbildung wird deutlich, wie die Frequenzen von den neuen Satelliten genutzt werden. Bei der Erweiterung der Flotte musste aus Platzgründen auf andere Positionen ausgewichen werden. Die Satelliten Astra 2A-D befinden sich bei 28,2° Ost und sind insbesondere für die Versorgung von Großbritannien vorgesehen.

Obwohl die Sendefrequenz von ca. 11 GHz für den Downlink so gewählt wurde, dass die Dämpfung der Welle durch die Erdatmosphäre ein Minimum hat, muss bei Satellitenverbindungen mit einer Übertragungsdämpfung von mehr als 110 dB gerechnet werden. Die Empfangsfeldstärke auf der Erde ist

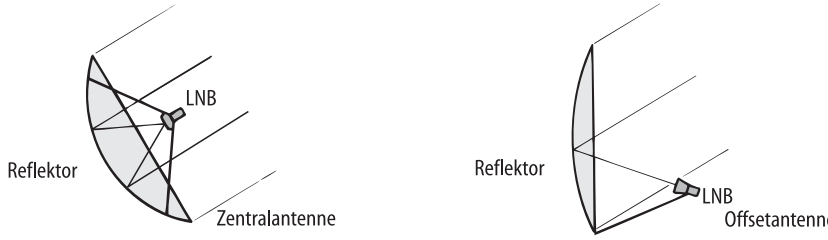


Abb. 4.31. Satellitenempfangsantennen

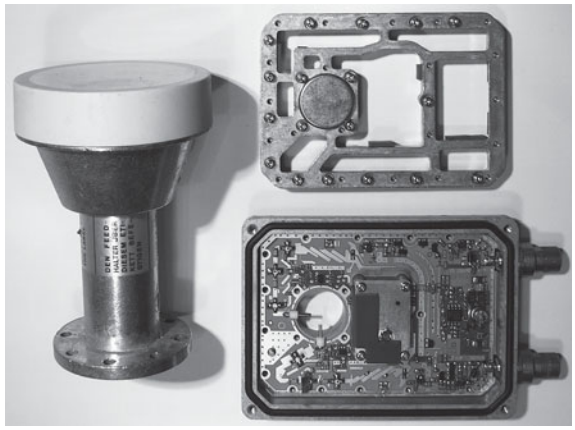


Abb. 4.32. Innenansicht eines LNB [180]

dementsprechend gering, zum Ausgleich müssen die Empfänger mit hochwertigen Antennen ausgestattet sein. Um eine gute Signalqualität zu erzielen, muss zudem quasi Sichtverbindung zum Satelliten bestehen. Eine Signaldämpfung durch Wände, Häuser etc. macht den Empfang unmöglich.

Für den Empfang werden Reflektorantennen verwendet (Schüsseln, engl. dishes), die nach dem Prinzip des Hohlspiegels arbeiten (Abb. 4.31). Im Brennpunkt des Hohlspiegels befindet sich die eigentliche Antenne (Feedhorn). Bei großen Spiegeldurchmessern ergibt sich eine hohe zum Horn reflektierte Intensität und damit eine gute Signalqualität. Moderne Direktempfangssatelliten sind so konzipiert, dass bei mittleren Qualitätsansprüchen ein Spiegeldurchmesser von 60 cm ausreicht. Kleinere Spiegel sind handlicher als große und haben den Vorteil, dass sie nicht so genau auf den Satelliten ausgerichtet zu werden brauchen.

Direkt an der Empfangsantenne ist in der Outdoor Unit ein Low Noise Converter (LNC, LNB) angebracht (Abb. 4.32). Mit Hilfe des LNB wird eine besonders rauscharme Umsetzung des Signals in den UHF-Bereich (Sat-ZF) vorgenommen, da eine leitungsgebundene Übertragung das 11-GHz-Signal zu stark dämpfen würde. In einer Indoor Unit (Abb. 4.33) wird dann das Signal auf

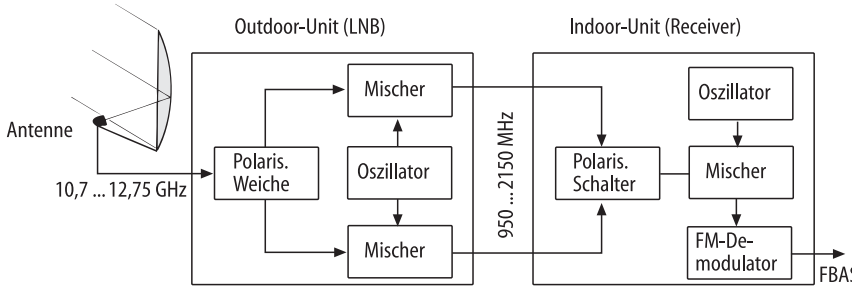


Abb. 4.33. Satelliten-Empfangssystem

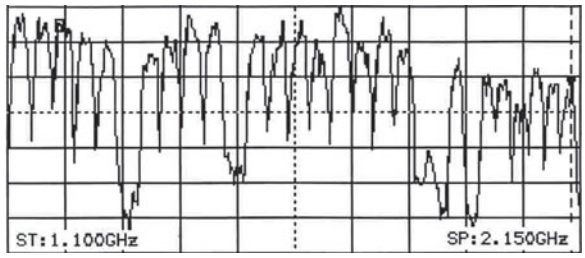


Abb. 4.34. Amplitudenspektrum am LNB-Ausgang [158]

eine zweite Zwischenfrequenz umgesetzt, frequenzdemoduliert und entweder direkt dem Videoeingang des Fernsehempfängers zugeführt oder per RSB-AM für den Antenneneingang des Empfängers aufbereitet.

Der gesamte Empfangsfrequenzbereich wird in das untere Band von 10,7 GHz bis 11,7 GHz und das obere Band zwischen 11,7 GHz und 12,75 GHz aufgeteilt. Das untere Band wird in den Sat-Zwischenfrequenzbereich 950 MHz bis 1950 MHz umgesetzt, das obere Band in den Bereich 1100 MHz bis 2150 MHz. Abbildung 4.34 zeigt ein Frequenzspektrum, das am Ausgang des LNB gemessen wurde. Die einzelnen Transponder erscheinen hier als Spitzen im Signalverlauf.

Es gibt verschiedene Arten von LNB. Ein Universal Single LNB ist für das obere und untere Band geeignet und wird von der Indoor-Unit mit einer Spannung versorgt, die zwischen 13 V und 18 V schaltbar ist, um damit die Umschaltung zwischen den Polarisationsebenen zu ermöglichen. Die Umschaltung zwischen dem oberen und unteren Band geschieht mit der Überlagerung eines 22-kHz-Signals, wenn das obere Band gewählt wird.

Ein Twin-LNB enthält zwei LNB in einem Gehäuse und wird zum gleichzeitigen Empfang mehrerer Programme, z. B. für Mehrteilnehmeranlagen genutzt. Single- und Twin-LNB sind nur für eine Satellitenposition ausgelegt. Um mit einer Empfangsanlage mehrere Satellitenpositionen zu erfassen, kann eine motorgetriebene Verstelleinrichtung des Reflektors vorgesehen werden. Wenn die Positionen der Satelliten nicht zu weit auseinander liegen, kann es ausreichen, nur den LNB seitlich zu verschieben (Motorfeed) oder zwei LNB in festem Abstand z. B. für Astra auf 19,2°-Ost und Eutelsat Hot Bird auf 13°-Ost zu montieren. Zum Ausgleich der nicht mehr optimalen LNB-Position wird ein größerer Reflektor gewählt. Die Umschaltung zwischen verschiedenen Satellitenpositionen erfordert ein weiteres Schaltsignal. Für diese Information und zukünftige Erweiterungen wurde Digital Satellite Equipment Control (DiSEqC) entwickelt, das darauf beruht, das 22-kHz-Signal zu tasten (ASK) und die gewünschten Informationen in digital codierter Form über diese Art der Modulation zu übertragen [178].

Die Satellitenübertragung hat heute eine große Bedeutung für die Fernsehsignalübertragung, das gilt nicht nur für den Direktempfang, sondern auch für die Einspeisung in Kabelnetze und die Zuspielung. Große Entfernungen werden überbrückt, indem mehrere Satelliten nacheinander genutzt werden (doppel-Hop). Die Transponder können für analoge Signale ebenso genutzt werden wie für digitale.

4.3.3 Kabelfernsehen

Seit Anfang der 80er Jahre hat die deutsche Bundespost/Telekom begonnen in fast allen Städten die einzelnen Häuser über Koaxialkabel oder Glasfaser zu verbinden, um darüber Fernsehsignale zu verteilen. Dieses Breitbandkommunikationsnetz (BK-Netz) wurde zunächst in zwei Ausbaustufen für Frequenzen bis 300 MHz (BK300) und 446 MHz (BK450) ausgelegt. Ab 2001 begann dann die Erweiterung bis zu 846 MHz (BK860). Das Netz erlaubt prinzipiell die Kommunikation in beide Richtungen, wobei die Rückkanäle unterhalb 60 MHz liegen. Die Übertragungsqualität im Kabelnetz ist sehr gut, da keine atmosphärischen Störungen auftreten können. Außerdem kann der gesamte Frequenzbereich ohne Rücksicht auf andere Dienste genutzt werden, sodass im Kabel 28 Kanäle mit 7 MHz Bandbreite und 18 Kanäle mit 8 MHz Bandbreite zur Verfügung gestellt werden können.

Aus Kompatibilitätsgründen, um mit vorhandenen Geräten empfangbar zu sein, werden die Programme im Frequenzmultiplex wie bei der terrestrischen Ausstrahlung übertragen. Bei der Kabelkanalbelegung wurde das Kanalraster des terrestrischen Fernsehens und auch des Hörfunks übernommen und durch Sonderkanalbereiche erweitert. Diese Erweiterung ist möglich, da im Kabel keine Rücksicht auf rundfunkfremde Dienste genommen werden muss. Mit den üblichen Frequenzen sowie einem unteren und einem oberen Sonderkanalbereich kann ein Angebot von mindestens 30 Programmen bereitgestellt werden.

Der untere Sonderkanalbereich mit den Kanälen S 2 bis S 10 umfasst die Frequenzen 111 ... 174 MHz und der obere 230 ... 300 MHz (S 11 bis S 20). Die Bandbreiten betragen 7 MHz, Tabelle 4.3 bietet eine Übersicht über die Kabelkanäle. In der zweiten Ausbaustufe wurde für das BK450 der erweiterte Sonderkanalbereich eingeführt, der auch als Hyperband bezeichnet wird. Hier

Tabelle 4.3. Fernsehkanäle im Breitbandkabel

Kanal	Frequenz (in MHz, untere, obere Grenze)	Kanal	Frequenz	Kanal	Frequenz	Kanal	Frequenz
K 2	47 54						
K 3	54 61	S2	111 118	S 11	230 237	S 21	302 310
K 4	61 68	S3	118 125	S 12	237	S 22	310
		S 4	125 132	S 13	244	S 23	318
UKW Hörfunk	87,5 108	S 5	132 139	S 14	251	S 24	326
		S 6	139 146	S 15	258	S 25	334
		S 7	146 153	S 16	265	.	
K 5	174 181	S 8	153 160	S 17	272	.	
K 6	181	S 9	160 167	S 18	279	.	
K 7	188	S 10	167 174	S 19	286	S 35	414
K 8	195	unterer Sonderkanalbereich		S 20	293 300	S 36	422
K 9	202			oberer Sonderkanalbereich		S 37	430
K 10	209					S 38	438 446
K 11	216					erweiterter Sonderkanalbereich (Hyperband)	
K 12	223 230						

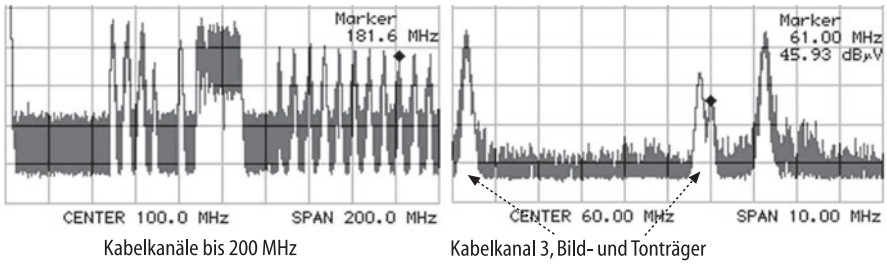


Abb. 4.35. Amplitudenspektrum der Signale im Breitbandkabel

betragen die Bandbreiten, wie auch bei BK860, 8 MHz. Die hochfrequenten Bereiche werden vorzugsweise zur Digitalübertragung DVB-C genutzt (s. Abschn. 4.6.3). Abbildung 4.35 zeigt das gemessene Amplitudenpektrum einer Kabelübertragung.

Für den analogen Empfang aus Kabelanlagen können Standardfernsehempfänger unmodifiziert verwendet werden. Um die Sonderkanäle empfangen zu können, müssen die Geräte aber einen gegenüber dem Standardsystem erweiterten Frequenzbereich verarbeiten können, also »kabeltauglich« sein.

In Deutschland wird das Kabelnetz von verschiedenen Unternehmen betrieben, die für die Nutzung eines Kabelanschlusses Gebühren verlangen. Fernseh- und Radioprogramme werden direkt, über Richtfunkstrecken oder aus Satellitenempfangsstationen in das Breitbandverteilnetz eingespeist und über Übergabepunkte an die Privathaushalte verteilt [243].

Bisher kommen im Kabelnetz meist Koaxialkabel zum Einsatz, in Zukunft wird das Netz zum großen Teil mit Glasfasern (s. u.) aufgebaut werden. Bei Koaxialleitungen muss mit einer großen, frequenzabhängigen Dämpfung gerechnet werden. Diese wird mit Verstärkern ausgeglichen, die in geringen Abständen installiert werden. Glasfasern weisen dagegen eine wesentlich kleinere Dämpfung auf, sodass sich entsprechend größere Verstärkerabstände ergeben.

Die gesamte kabelgebundene Übertragungsstrecke wird in Deutschland in vier Abschnitte aufgeteilt, die als Netzebenen bezeichnet werden (Abb. 4.36). Die Ebenen 1 und 2 dienen zur Verbindung zwischen Studios und den Schalt-

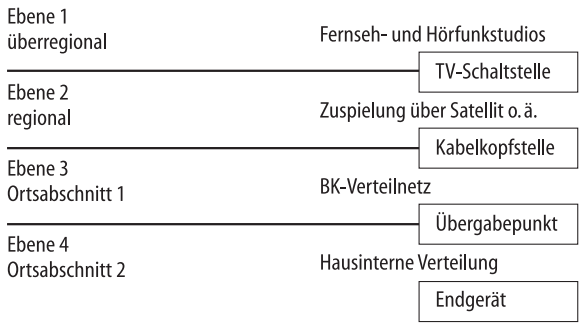


Abb. 4.36. Netzebenen im Breitbandkabel

stellen des öffentlichen Netzes. Ebene 3, bzw. der Ortsabschnitt 1, entspricht dem eigentlichen Breitbandkabelnetz, während Ebene 4 die Signalverteilung zum Endgerät betrifft. Das eigentliche BK-Verteilnetz in Ebene 3 beginnt bei der Kabelkopfstation (BK-Verstärkerstelle), wo z. B. ein Satellitensignal eingespeist wird, und endet am Hausübergabepunkte.

Das Netz hat einen sternförmigen Aufbau und ist wiederum in vier Teile aufgeteilt, die mit A ... D bezeichnet werden. Im Bereich A können höchstens 20 Verstärker hintereinander geschaltet werden, die eine maximale Distanz von je 412 m aufweisen dürfen. An jedem Verstärkerpunkt kann das Signal zur Ebene B ausgekoppelt werden. Über die ähnlich aufgebaute B-Ebene wird das Signal zur C-Ebene verteilt, wo D-Abzweiger den für die Hausverteilung erforderlichen Signalpegel bereitstellen. In diesen Bereichen sind noch geringere Verstärkerabstände erforderlich.

4.3.3.1 Lichtwellenleiter

Im vorigen Abschnitt wurde deutlich, dass für eine flächendeckende Verteilung der Signale mit Koaxialkabeln eine sehr große Zahl von Verstärkern erforderlich ist und trotzdem die Gesamtbandbreite meist auf ca. 500 MHz beschränkt bleibt. Diesbezüglich bietet die Signalführung über Lichtwellenleiter oder Glasfasern viele Vorteile.

Das Signal wird einer Lichtwelle aufgeprägt, die in der Faser geleitet wird. Zur eigentlichen Übertragung ist kein elektrischer Strom erforderlich. Lichtwellenleiter werden aus bis zu 2000 Quarzglas- oder Kunststoff-Fasern (Fiber) gefertigt. Typische Fasern haben einen Durchmesser von ca. 125 µm. Die Mitte der Faser (Kern) weist einen größeren Brechungsindex als der Mantel auf, deshalb wird das Licht total reflektiert, wenn der Einfallswinkel zur Faserachse nicht zu groß wird. Die Lichtenergie bleibt dann im Kern (Abb. 4.37) und das Signal wird nur sehr wenig gedämpft (z. B. 0,5 dB/km bei 1,3 µm Wellenlänge). Bezüglich des Verlaufs des Brechungsindex über den Faserradius wird dabei zwischen einer Stufenfaser und einer Gradientenfaser unterschieden, die wiederum verschiedene Leitungseigenschaften aufweisen.

Darüber hinaus wird zwischen Multi- und Monomodefasern unterschieden. Der Begriff mode beschreibt dabei verschiedene Schwingungs- bzw. In-

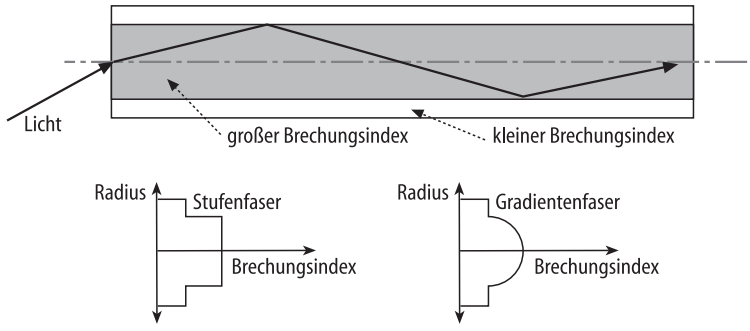


Abb. 4.37. Totalreflexion im Lichtwellenleiter

terferenzzustände. Diese können bei der Multimodefaser mehrfach auftreten, da sie einen relativ großen Kerndurchmesser von ca. 50 µm aufweisen. Bei Monomodefasern (Single mode) wird der Kern auf 3 µm bis 9 µm verkleinert. In diesem Falle kann sich nur ein Schwingungszustand ausbreiten, sodass die problematische Mehrwegausbreitung mit der damit verbundenen Modendispersion nicht bzw. nur reduziert auftreten kann. Mit Monomodefasern können ohne Verstärkung Distanzen bis zu mehr als 100 km überbrückt werden, während es bei Multimodefasern nur einige km sind. Der Nachteil der Monomodefaser ist der Umstand, dass die Einkopplung der Lichtwelle aufgrund des kleinen Kerns in der Praxis problematisch sein kann.

Als Sender werden Leuchtdioden oder Laserdioden eingesetzt und als Empfänger eine in Sperrrichtung gepolte, spezielle Empfangsdiode (PIN). Die verwendeten Wellenlängen liegen meist im Infrarotbereich. Hier werden bestimmte Wellenlängenbereiche (Fenster) zwischen 1,3 µm und 1,6 µm verwendet, in denen eine geringe Dämpfung vorliegt. Die Signalübertragung geschieht im einfachsten Fall durch die Steuerung der Lichtintensität in Abhängigkeit vom zu übertragenden Signal.

Lichtwellenleiter haben gegenüber elektrischen Leitungen folgende Vorteile:

- Eine große Datenübertragungskapazität (Bandbreite im GHz-Bereich, gegenüber ca. 500 MHz beim Koaxialkabel). Das System bleibt schmalbandig, denn gemessen an der hohen Lichtfrequenz von ca. 10^{15} Hz ist eine Bandbreite von 10 GHz gering. Schmalbandsysteme haben den Vorteil, dass keine aufwendigen Entzerrungsmaßnahmen erforderlich sind.
- Das System weist geringe Dämpfung auf, es sind relativ wenige Verstärker erforderlich, dazwischen sind große Abstände möglich.
- Das System unterliegt keinen Störungen durch elektromagnetische Felder.
- Die Sende- und Empfangseinheiten können auf unterschiedlichem elektrischen Potenzial liegen, es entstehen keine Erdungsprobleme.

In der Videotechnik werden Lichtwellenleiter neben der Kabelfernsehübertragung auch bei Kamerasystemen eingesetzt. Allgemein haben sie eine große Bedeutung für digitale Kommunikationsnetze, über die vermehrt auch Videosignale übertragen werden. Die Weiterentwicklung ermöglicht eine Steigerung der Datenrate durch Nutzung verschiedener Wellenlängen in einer Faser (Wavelength Division Multiplex, WDM).

4.3.4 Verbesserte analoge Fernsehsysteme

Bevor hier die Übertragung digitaler Fernsehprogramme erörtert wird, soll ein Aus- bzw. Rückblick auf analoge Fernsehsysteme (Sende- nicht Produktionsstandards) geworfen werden, die mit erheblichem Aufwand zu qualitativen Verbesserungen führen sollten, die aber vor dem Hintergrund der digitalen Fernsehübertragung keine Bedeutung mehr gewinnen konnten.

Der Wunsch nach verbesserten Fernsehsystemen resultiert aus den Mängeln der konventionellen Systeme. Hier wären zu nennen: mäßige Schärfe und Detailauflösung, Großflächenflimmern, 25-Hz-Flackern an horizontalen Kanten aufgrund des Zeilensprungverfahrens, Cross Colour, Cross Luminance, etc.

Praxisrelevante Vorschläge zur Verbesserung analoger Fernsehsysteme sind kompatibel zum bestehenden System und ermöglichen ein verbessertes Bild mit neuen Empfangsgeräten, jedoch auch eine unverminderte Standardqualität mit Geräten, die die Zuschauer bereits besitzen. Diese Ansätze zielen meist auf eine verbesserte Trennung der Luminanz- und Chrominanzkanäle ab, weiterhin auf Vor- und Nachfilterung zur besseren Reduktion der Aliasanteile, auf verbesserte sogenannte digitale Empfänger (s. Abschn. 6.1.5), die z. B. mit Bildspeichern ausgerüstet sind, und auf Techniken für die kompatible Umstellung auf ein Breitbildformat.

Eine wesentliche Verbesserung des analogen TV-Systems ist natürlich auch eine Großbilddarstellung, ähnlich wie im Kino; Telepräsenz statt Television, also der Übergang zu hochaufgelöstem Fernsehen (HDTV). Ein wichtiger Punkt ist dabei die Umstellung des Bildseitenformates von 4:3 auf 16:9. Die Einführung des neuen Bildseitenverhältnisses war bereits für konventionelle PAL-Verfahren vorgesehen und zwar mit dem PAL-kompatiblen PALplus-System.

4.3.4.1 PALplus

Das PALplus-Verfahren bietet einen zu Standard-PAL kompatiblen Übergang vom Bildformat 4:3 zum Breitbildformat 16:9. Neben dieser offensichtlichen Funktion enthält PALplus Verfahren zur optimierten Chrominanzsignalbearbeitung, die als Colour-Plus oder Motion Adaptive Colour Plus (MACP) bezeichnet werden. Mit der Ausstrahlung von PALplus sollte der Übergang zu neuen Übertragungsverfahren und HDTV geebnet werden, indem zunächst der Kauf von Empfängern mit dem neuen Bildseitenverhältnis stimuliert würde. Im Jahre 1994 wurde von einigen Sendeanstalten mit Fördergeldern der regelmäßige Sendebetrieb mit PALplus aufgenommen, später jedoch wieder eingestellt.

Das PALplus-Verfahren bietet bei Sendungen im neuen Format auf 16:9-Empfängern ein verbessertes Bild, demgegenüber ist bei solchen Sendungen die Auflösung bei Empfang mit Standardgeräten verschlechtert. Die Kompatibilität erfordert, dass Bilder im 16:9-Format auf Standardempfängern wiedergegeben werden können. Hierzu gibt es zwei Möglichkeiten, die mit Side Panel (Abb. 4.38) und Letterbox bezeichnet werden.

Beim Side Panel-Verfahren wird aus dem 16:9-Bild ein 4:3-Ausschnitt gebildet und dieser wird konventionell übertragen. Bei diesem Verfahren gehen seitlich Bildteile verloren, evtl. muss der Bildausschnitt dem Bildinhalt angepasst und dementsprechend geschwenkt werden (Pan Scan). Dieser Umstand

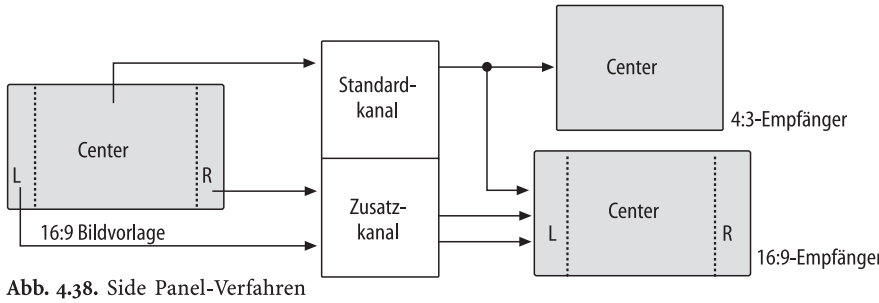


Abb. 4.38. Side Panel-Verfahren

erfordert den ständigen Einsatz eines Mitarbeiters in der Bildregie. Die seitlich abgeschnittenen Bildteile können in einem Zusatzkanal übertragen und im 16:9-Empfänger dem 4:3-Bild wieder zugefügt werden. Side Panel bietet den Vorteil, dass beide Bildformate vollständig gefüllt sind, aber den wesentlichen Nachteil, dass beim 4:3-Empfang Information verloren geht. Ein weiterer Nachteil ist, dass die Zusammensetzung der Bildteile im 16:9-Empfänger störanfällig ist, da die Side Panels separat hinzugefügt werden.

PALplus arbeitet wegen der genannten Nachteile mit dem Letterbox-Verfahren [50]. Hier gehen keine Bildteile verloren. Allerdings ist die Vertikalauflösung des 16:9-Bildes bei 4:3-Empfängern verringert, denn es erscheinen auf einem 4:3-Empfänger am oberen und unteren Bildrand schwarze Streifen, ähnlich denen, die bei der Wiedergabe von Spielfilmen auf 4:3-Empfängern in Erscheinung treten (Abb. 4.39). Die Streifen ergeben sich aus der Reduktion der aktiven Zeilenzahl des 16:9-Bildes um den Faktor 1,33. Würde die gesamte Breite des 16:9-Bildes unverändert auf einem 4:3 Empfänger dargestellt, so erschiene die Horizontale gestaucht (Abb. 4.39). Durch die Zeilenzahlreduktion wird erreicht, dass 16:9-Bilder im richtigen Format auf 4:3-Empfängern erscheinen. Bei PALplus werden dafür die 576 aktiven Bildzeilen um den Faktor 3/4 auf 432 aktive Zeilen reduziert. Durch eine Dezimations-Tiefpassfilterung werden 144 Zeilen entfernt. Das 432-Zeilenbild kann im 16:9-Format auf 4:3-Empfängern wiedergegeben werden. Die dabei entstehenden schwarzen Streifen an den Bildrändern haben eine Breite von jeweils 72 Zeilen (Abb. 4.40).

Die vertikale Tiefpassfilterung arbeitet nicht einfach mit Unterdrückung jeder vierten Zeile, sondern beinhaltet eine Interpolation, die darauf beruht, dass zunächst eine Verdreifachung der 576 aktiven Zeilen vorgenommen wird. Aus einem Speicher wird anschließend nur jede vierte Zeile ausgelesen.

Die Signale der entfernten Zeilen werden nicht unterdrückt, denn sie müssen für die hochwertige Wiedergabe mit PALplus-Empfängern zur Verfügung stehen. Dazu werden sie auf die Farbträgerfrequenz moduliert und als sogenannte Vertical Helper mit Phasenlage 0° und einem Pegel von max. 0,3 V_{ss} symmetrisch zum Schwarzbezugswert übertragen (Abb. 4.41). Der Mittelwert dieses Signals entspricht dem Schwarzwert, sodass auf konventionellen Empfängern nur schwarze Streifen zu sehen sind (falls die Helligkeit nicht zu hoch eingestellt ist). In einem hochqualitativen 16:9-Empfänger kann das Helpersignal ausgewertet und die senderseitig entfernten 144 Zeilen können wieder

Abb. 4.39. Gestauchte und richtige Darstellung des 16:9-Bildes im 4:3-Format

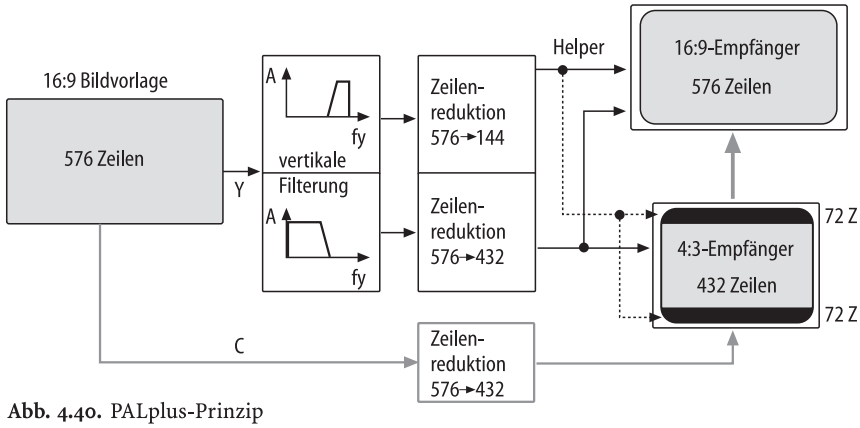


Abb. 4.40. PALplus-Prinzip

eingefügt werden, sodass das ursprüngliche 16:9-Bild mit voller Auflösung wiedergegeben wird [51]. Diese Verarbeitung des Helpersignals betrifft nur den Luminanzanteil. Das Weglassen hochfrequenter vertikaler Signalanteile im Chrominanz-Anteil ist für das menschliche Auge irrelevant.

Das 16:9-Bild erfordert, verglichen mit dem Standardsignal, bei gleicher Auflösung eine um den Faktor 1,33 größere Bandbreite. Da aus Kompatibilitätsgründen die Standardbandbreite von 5 MHz bei PALplus beibehalten wird, hat das Breitbild eine verringerte Horizontalauflösung. Es kann allerdings dafür gesorgt werden, dass diese für die meisten Konsumenten nicht sichtbar wird, denn die Konsumenten sind an Standardempfänger gewöhnt, die praktisch, insbesondere wegen der schlechten Y/C-Trennung, nur einen Frequenzbereich bis ca. 3,7 MHz nutzen. PALplus-Empfänger sind mit einer verbesserten Signalverarbeitung mit MACP (Motion Adaptive Colour Plus) ausgestattet, daher ergibt sich trotz breiterem Bild eine mit heutigen Standardsystemen vergleichbare Horizontalauflösung und eine weitgehende Befreiung von PAL-Artefakten. Nicht alle Empfangsgeräte mit 16:9-Bildschirmen waren PAL-plus-Empfänger, oft wurde aus Kostengründen auf die MACP-Verarbeitung verzichtet.

Die Verarbeitung eines 4:3-Bildes erfolgt beim 4:3-Empfänger wie üblich. Auf dem 16:9-Empfänger ergeben sich bei Vorliegen eines 4:3-Bildes am linken und rechten Bildrand schwarze Streifen, alternativ wird der Bildschirm horizontal gefüllt und vertikal überschrieben oder das Bild wird bewusst verzerrt.

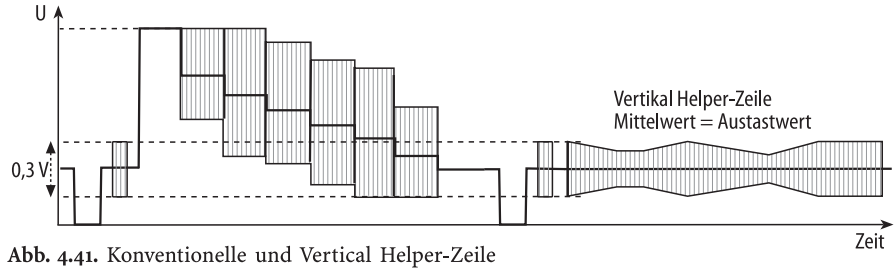


Abb. 4.41. Konventionelle und Vertical Helper-Zeile

4.3.4.2 MAC

Die Abkürzung MAC steht für Multiplexed Analogue Components. Genauer müsste es heißen Time Multiplexed, denn die Nutzung des Zeitmultiplexverfahrens ist der wesentliche Unterschied zu den Standardfarbcodierarten NTSC, PAL und SECAM. Bei MAC werden das Luminanzsignal Y und die Farbdifferenzkomponenten C_R und C_B zeitlich nacheinander übertragen. Die Übertragungsqualität ist gegenüber dem Frequenzmultiplexverfahren verbessert, da kein Übersprechen (Cross-Effekte) zwischen den Signalen auftritt. Auch der Begleitton ist im Zeitmultiplex erfasst, damit werden Intermodulationsstörungen vermieden.

Im Gegensatz zur Frequenzmultiplextechnik steht beim Zeitmultiplexverfahren allen Komponenten die gesamte Frequenzbandbreite zur Verfügung, dafür wird die Übertragungszeit aufgeteilt. Ein kontinuierliches Signal muss bei MAC also zeitlich komprimiert und im Empfänger wieder expandiert werden. Entsprechend der Bandbreitenreduktion bei herkömmlichen Farbcodierungsarten werden dabei die Farbdifferenzkomponenten in stärkerem Maße als die Luminanzkomponente komprimiert. Wie bei SECAM werden bei MAC die Farbdifferenzkomponenten dann zeilenweise abwechselnd übertragen, die vorhergehende ersetzt die jeweils nicht übertragene.

MAC basiert auf dem konventionellen Bildabtastsystem mit 625 Zeilen und mit 64 µs Zeilendauer, von der 52 µs für den Bildinhalt zur Verfügung stehen. Entsprechend der Empfehlung ITU-R BT.601 werden die Komponenten Y, C_R und C_B im gegenseitigen Verhältnis 4:2:2 (s. Kap. 3.3.2) komprimiert, da die Kompression auf digitaler Ebene erfolgt und das System kompatibel zu bestehenden Digitaltechniken sein sollte. Es existieren, historisch begründet, die MAC-Varianten A bis D, die sich im Wesentlichen bezüglich der Toncodierung unterscheiden. C-MAC bezeichnet eine breitbandige, für Satellitensysteme konzipierte Signalform, D-MAC ist für eine Bandbreite von 12 MHz konzipiert. Bei C- und D-MAC werden alle Bildkomponenten während 52 µs übertragen. Die 12 µs, die konventionell der Austastlücke entsprechen, sind für die Tonübertragung, das Synchronsignal und digitale Daten reserviert. Da somit der Synchronpegelbereich von 0,3 V wegfällt, steht für das Bildsignal anstatt 0,7 V_{ss} der gesamte Pegelbereich von 1 V_{ss} zur Verfügung (Abb. 4.42).

Ton- und Synchrondaten werden digital innerhalb der ersten 10,3 µs der Zeile übertragen. Es folgt eine kurze Klemmphase auf 0,5 V, die zur Farbstabilisierung dient, dann stehen 17,2 µs für die Differenzkomponenten C_R bzw. C_B und weitere 34,4 µs für die Luminanzkomponente zur Verfügung. Das Y-Signal ist damit um den Faktor 1,5 zeitkomprimiert, die Signale C_R und C_B um den Faktor 3. D- und D2-MAC verwenden eine Duobinärcodierung der Digitaldaten, d. h. es werden drei Pegelzustände benutzt. Die Variante D2-MAC arbeitet gegenüber D-MAC mit halbierter Digitaldatenrate und ist an ein 8-MHz-Kanalraster angepasst, sodass eine terrestrische Übertragung möglich ist. Der Datenbereich bei D2-MAC umfasst in jeder Zeile 105 Bit. Der digitale Datenstrom wird in Pakete von 751 Bit Umfang eingeteilt, die mit eigenen Adressen versehen sind. Der Datenbereich lässt sich flexibel für Ton- oder weitere Daten nutzen. Es ist z. B. möglich, zwei Tonkanäle sehr hoher Qualität oder bis zu acht Kanäle minderer Qualität zu verwenden.

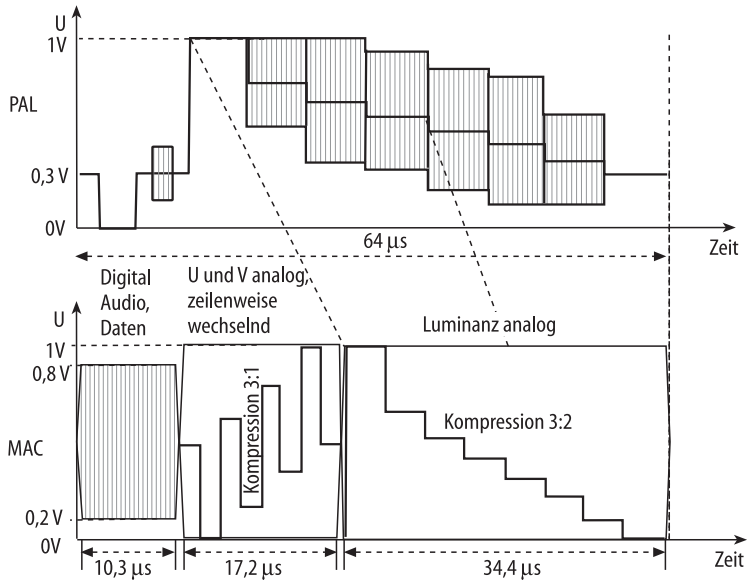


Abb. 4.42. Vergleich von je einer Bildzeile bei PAL und D2-MAC

In den 80er Jahren sollte MAC als kompatibler Übergang zu einem HDTV-System (HD-MAC) eingeführt werden. Dabei wurde die Übertragung im Hyperband oberhalb 302 MHz vorgesehen. D2-MAC wurde allerdings nie allgemeiner Standard für die Konsumenten.

4.3.4.3 HDTV analog

Mit High Definition Television (HDTV) wird eine Bildwiedergabe angestrebt, die sich qualitativ der Kinobildwiedergabe nähert. Die Verbesserung gegenüber herkömmlichen Systemen umfasst die Einführung eines Breitbildformates wie im Kino mit dem Seitenverhältnis 16:9 und die Vergrößerung des Bildfeldwinkels durch einen kleineren Betrachtungsabstand. Voraussetzung ist hier eine Auflösungserhöhung, ein TV-System, das mit ca. 1200 Zeilen vertikal und ca. 2000 Bildpunkten horizontal arbeitet. Die Bildqualität sollte verbessert sein, vor allem hinsichtlich des Chrominanz-Luminanzübersprechens.

Bereits in den 70er Jahren wurde in Japan mit der Entwicklung eines hochzeiligen TV-Systems begonnen, dementsprechend ist Japan bei Entwicklung analoger HDTV-Technik führend. Seit Anfang der 90er Jahre wird in Japan mit MUSE ein regelmäßiger HDTV-Dienst betrieben. In Europa wurde mit HD-MAC ein eigenes System entwickelt aber nie zum Zuschauer gebracht.

Multiple Subsampling Encoding (MUSE) ist das Übertragungsformat für die japanische HDTV-Variante, die auf 60 Hz-Bildwechselfrequenz beruht und für die Übertragung eines 1125-Zeilen-Signals in Schmalbandkanälen konzipiert ist. Auf der Sendeseite wird ein hochqualitatives Signal mit ca. 20 MHz Bandbreite erzeugt, denn ausgehend von einer NTSC-Bandbreite von 4,2 MHz, dem Übergang auf 16:9 und verdoppelter Horizontal- und Vertikalauflösung ergeben sich

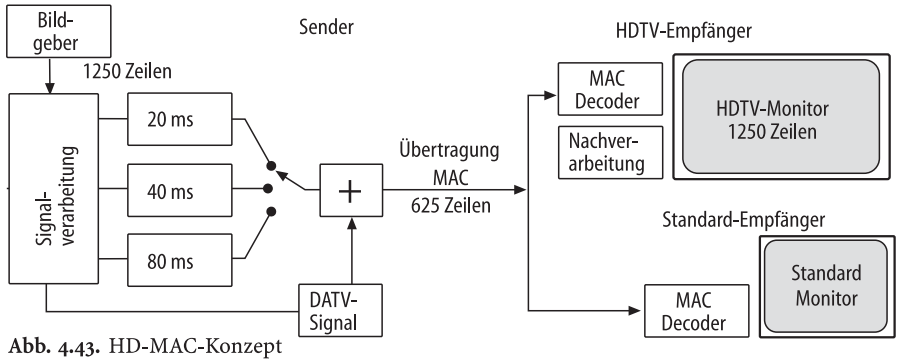


Abb. 4.43. HD-MAC-Konzept

$4 \cdot 1{,}33 \cdot 4{,}2$ MHz = 22,4 MHz. Dann werden die Luminanz- und Farbdifferenzsignale gewonnen, die hier im Zeitmultiplex übertragen werden.

Vor der Übertragung wird das Bildsignal bewegungsadaptiv bearbeitet, damit die Übertragungsbandbreite auf 8 MHz reduziert werden kann. Bei langsam veränderlichen Signalen werden dazu wenige Bilder mit hoher Auflösung gebildet, und bei schnellen Bewegungen wird die Bildrate auf Kosten der Ortsauflösung erhöht. Der Begleitton wird digital in der Vertikalaustastlücke übertragen, es werden vier Tonkanäle hoher Qualität unterstützt.

Für die europäische HDTV-Variante High Definition-MAC war das Ziel, eine möglichst kompatible Umstellung des Fernsehsystems, basierend auf MAC zu erreichen. Für High-Definition TV wurden die 625 Zeilen des Standardsystems auf 1250 Zeilen verdoppelt. Es wurde eine bewegungsadaptive Signalverarbeitung zur Datenreduktion eingesetzt. Die Tonübertragung war für vier digitale Kanäle hoher Qualität ausgelegt. Die Kompatibilität zu MAC erforderte eine Anpassung des hochaufgelösten Signals an ein 625-Zeilensystem. Das HDTV-Signal wurde dazu von 1250 Zeilen auf 625 Zeilen abwärts konvertiert, indem mittels »Line Shuffling« die Abtastwerte zweier Zeilen in einer Zeile zusammengefasst wurden [233]. Nach der MAC-Codierung konnte das Signal einem MAC-Empfänger zugeführt oder über eine Nachverarbeitungsstufe für einen HDTV-Empfänger wieder aufwärts konvertiert werden.

Die Konvertierung erfolgte wie bei MUSE bewegungsadaptiv, die dazu nötige Signalverarbeitung beruhte auf der Mehrteilbildcodierung. Je nach Bewegungsgeschwindigkeit wurden bei HD-MAC vier oder zwei Bilder zusammengefasst, oder jedes Bild wurde einzeln verarbeitet (4/2/1). Bei sehr langsam bewegten Bildinhalten wurden die Informationen auf vier Bilder verteilt, die Ortsauflösung erhöhte sich um den Faktor 4 (HDTV), während die Bewegungsauflösung auf 25% sank. Zwischen den Mehrteilbildcodierungsarten wurde hart umgeschaltet und der Umschaltvorgang wurde bewegungsadaptiv gesteuert (Abb. 4.43). Die Umschaltung zwischen den Teilbildcodierungen musste im Sender und im Empfänger synchron ablaufen. Dazu konnte entweder im Sender und im Empfänger je ein Bewegungsdetektor eingebaut werden, oder die Information über die Teilbildcodierungsart wurde dem Empfänger als Zusatzsignal übermittelt. Da dieses Signal digitaler Art ist, sprach man von Digitally Assisted TV (DATV). Wie D2-MAC wurde auch HD-MAC nie ein Sendestandard.

4.4 Digitale Modulationsverfahren

Nachdem die weiterentwickelten analogen TV-Systeme bereits viele digitale Elemente enthielten, war das nahe liegende Ziel, im nächsten Schritt ein volldigitales Fernsehsystem zu etablieren. Dabei sollten die Strukturen, die für die Analogübertragung geschaffen wurden, möglichst beibehalten werden. Die Versorgung mit Fernsehprogrammen geschieht also mit Digitalsignalen über die gleichen Wege wie bei Analogsignalen, über Kabel, Satellit und terrestrisch.

Die dafür erforderlichen Bandbreiten hängen nicht nur von der Datenrate, sondern, wie auch bei Analogsystemen, von der Modulationsarten ab, die hier zunächst erörtert werden. Wie bei Analogsystemen werden auch die Digitaldaten auf Sinusträger moduliert und im Frequenzmultiplex übertragen.

Die maximale Übertragungskapazität ist außer von der Bandbreite B auch vom Signal-Rauschabstand S/N abhängig, der auch als Rauschabstand C/N bezüglich des Trägers (Carrier/Noise) angegeben wird. Die Nutzung eines großen Störabstands ist durch mehrwertige Modulation eines Trägersignals möglich (Abb. 4.44). Das bedeutet, dass mehr als zwei Zustände zu einem Zeitpunkt zugelassen sind. Die Gruppe von Bits, die zu einem Modulationszustand gehört, wird als Symbol bezeichnet. Damit ergibt sich die maximale Datenrate aus der Symbolrate (mit der Einheit baud) multipliziert mit dem Informationsgehalt eines Symbols.

Jedes Bit entspricht der Verdopplung der Zustände und erfordert eine entsprechende Erhöhung des Signal-Rauschabstands um S/N = 6 dB = 20 log 2. Bei einem großen Signal/Rauschabstand im Kanal können viele Amplitudenzustände zugelassen werden und es ergibt sich eine hohe Bandbreiteneffizienz ε. Liegt z. B. eine Kanalbandbreite von 8 MHz vor, über die eine maximale Datenrate von 40 Mbit/s übertragen werden kann, so folgt daraus eine Bandbreiteneffizienz von ε = 40 Mbit/s/8 MHz = 5 bit/s/Hz. Die verschiedenen Übertragungswege weisen unterschiedliche Eigenschaften und Störabstände auf, sodass jeweils die Modulationsart eingesetzt wird, die dem Übertragungsweg am besten angepasst ist. Insgesamt sollten die Kanalcodierung und die Modulationsart so gewählt werden, dass ohne Fehlerschutz eine Bit Error Rate von maximal $2 \cdot 10^{-4}$ und mit Fehlerschutz der Wert 10^{-11} erzielt wird.

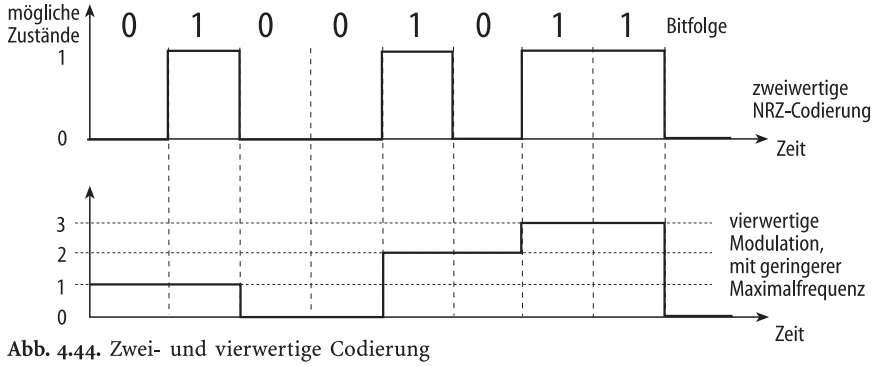


Abb. 4.44. Zwei- und vierwertige Codierung

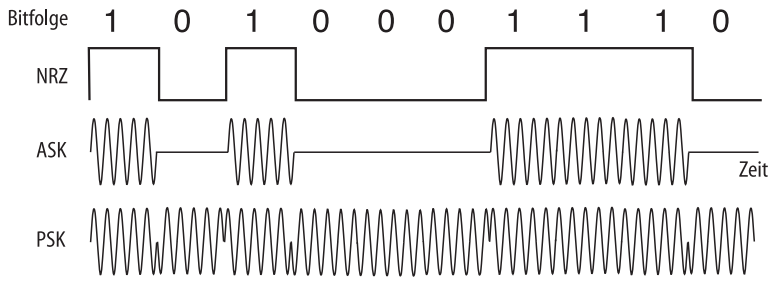


Abb. 4.45. ASK und PSK

4.4.1 Phase Shift Keying PSK

Wenn das modulierende Signal – wie in der Digitaltechnik – nur wenige diskrete Werte annimmt und als Träger ein Sinussignal benutzt wird, so wird der Zustand des Trägersignals den diskreten Werten entsprechend umgeschaltet. Man spricht von Umtastung, englisch: Shift Keying (SK). Wie bei der analogen Modulation kann auch hier die Amplitude (ASK), die Frequenz (FSK) oder die Phase (PSK) beeinflusst, d. h. umgetastet werden (Abb. 4.45). PSK hat aus Störabstandsgründen die größte Bedeutung.

Die PSK-Varianten unterscheiden sich in der Anzahl der erlaubten Zustände. Die am häufigsten angewandte Art ist die 4-PSK oder Quad-PSK, bei der vier Zustände, entsprechend zwei Bit (2^2), unterschieden werden. Für die Modulation werden zunächst je zwei Bit zusammengefasst (Dibits) und diese werden dann auf zwei um 90° verschobene Inphasen- und Quadraturträgerkomponenten (I, Q) moduliert (Abb. 4.46). Abbildung 4.47 zeigt die Zuordnung der Bits, die daraus resultiert, dass das MSB auf den I-Zweig gegeben wird.

Jede Bitkombination ist einer festen Phasenlage zugeordnet, die beim Empfang per Synchrondemodulation durch Vergleich mit der Phase eines Referenzträgers bestimmt wird. Der sich auf der Übertragungsstrecke möglicherweise

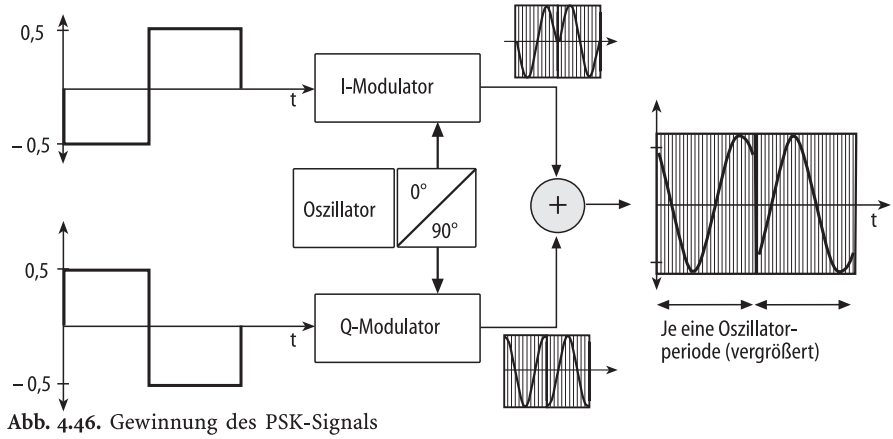


Abb. 4.46. Gewinnung des PSK-Signals

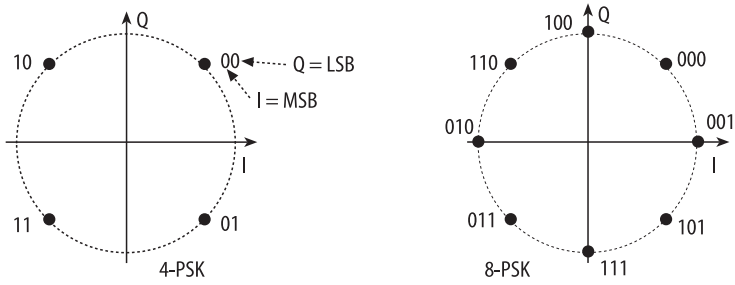


Abb. 4.47. Zustände des PSK-Signals

einstellende Phasenfehler muss bei der 4-PSK kleiner als ± 45° sein, bei der 16-PSK sind höchstens ± 11° erlaubt.

Bei der 4-PSK beträgt der theoretische Wert der Übertragungseffizienz ε = 2 bit/s/Hz, bei 16 PSK beträgt er 4 bit/s/Hz. Die für die PSK erforderliche Übertragungsbandbreite folgt aus dieser Übertragungseffizienz und der Datenrate. Für DVB wird QPSK bei der Satellitenübertragung eingesetzt, da für eine Bitfehlerrate von $2 \cdot 10^{-4}$ (ohne weiteren Fehlerschutz) nur ein C/N von ca. 14 dB erforderlich ist [59].

4.4.2 Quadraturamplitudenmodulation QAM

Zur Quadraturamplitudenmodulation lässt sich prinzipiell der gleiche Modulator einsetzen wie bei PSK. Die digitale QAM nutzt aber neben der Umtastung der Phasenzustände auch verschiedene Amplitudenzustände. Ein Teil der Informationen wird zu einem Signal mit mehreren Amplitudenstufen zusammengefasst und auf ein Sinussignal moduliert, der ebenso zusammengefasste Rest moduliert eine cos-Schwingung und anschließend werden beide addiert. Mögliche Zustände lassen sich wieder in Phasenzustandsdiagrammen mit den Koordinaten I und Q darstellen (Abb. 4.48). Die Modulation der Träger mit zwei 4-wertigen Signalen ergibt die 16-QAM, aus zwei 8-wertigen Signalen folgt die 64-QAM mit 64 Amplituden- und Phasenkombinationen.

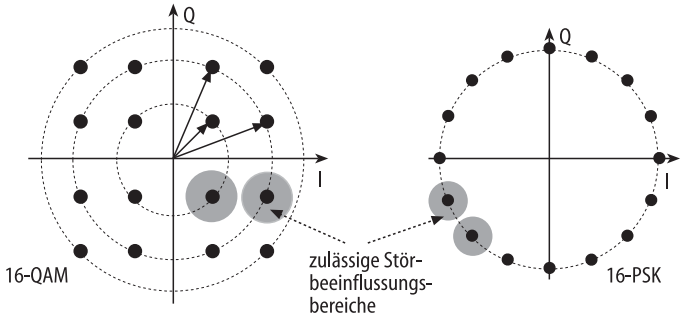


Abb. 4.48. Zeigerdiagramme von 16-QAM und 16-PSK

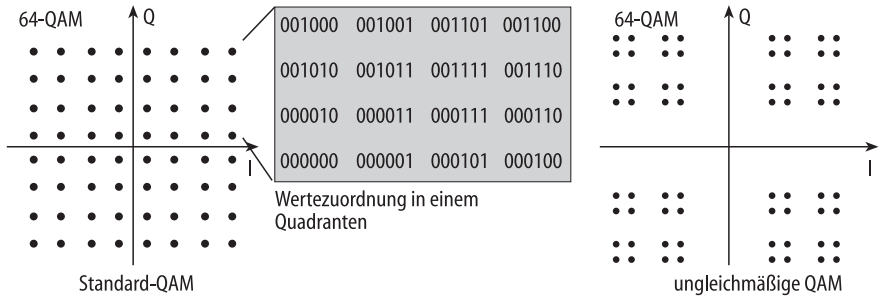


Abb. 4.49. Konstellationsdiagramm der 64-QAM und Vergleich mit Multiresolution-QAM

Die QAM ist ein effektives mehrwertiges Modulationsverfahren und bei Verwendung einer hohen Zahl von Zuständen ist QAM günstiger als PSK (Abb. 4.48), da der zulässige Störeinfluss größer sein kann. Die 16-QAM bietet eine Bandbreiteneffizienz von 4 bit/s/Hz, die 64-QAM 6 bit/s/Hz. Auch hier gilt: Je mehr Zustände erlaubt werden, desto größer muss der Signal-Rauschabstand sein. Die 16 QAM erfordert für eine BER von $2 \cdot 10^{-4}$ ein C/N = 20 dB, die 64 QAM ca. C/N = 27 dB.

Wenn die QAM-Zustände nicht in eine gleichabständige Konstellation gebracht werden, entsteht die Multiresolution QAM (Abb. 4.49). Um alle 16, in Abb. 4.49 rechts, dargestellten Zustände decodieren zu können, ist gegenüber der gleichförmigen Zustandsverteilung ein höherer Störabstand erforderlich. Dafür ergibt sich aber der Vorteil, dass bei schlechtem Störabstand die vier Zustände in einem (Teil-) Quadranten gemeinsam als ein einzelner Zustand interpretiert werden können. Ein Decoder könnte dann mit MR-QAM bei verschlechtertem Störabstand z. B. von 64-QAM auf 16-QAM und auf 4-QAM umschalten und damit ein qualitativ schlechteres statt gar kein Signal ausgeben. Diese Verkopplung wird als hierarchische Codierung bezeichnet und kann bei schlechten Übertragungsbedingungen genutzt werden, die z. B. bei DVB-T auftreten können.

4.4.3 Amplitude Phase Shift Keying APSK

Eine zweite kombinierte Umtastung von Amplituden- und Phasenzuständen ist, als Alternative zu QAM, das Amplitude Phase Shift Keying (APSK). Abbildung 4.50 zeigt, dass hier bei gleicher Zahl von Zuständen weniger Amplitudenstufen als bei der QAM genutzt werden. Dieser Umstand ist besonders dann wichtig, wenn die Amplitude ein besonders störempfindlicher Parameter ist. Dies ist z. B. bei der Satellitenübertragung der Fall, da hier die Senderöhren im Übersteuerungsbereich betrieben werden. Die 16-APSK bietet eine Bandbreiteneffizienz von 4 bit/s/Hz, die 32-APSK 5 bit/s/Hz. Die Modulationsformen 8-PSK (Abb. 4.47), 16- und 32-APSK sind für die erweiterten Spezifikationen der digitalen Satellitenübertragung, DVB-S2, spezifiziert. Sie sollen allerdings vorwiegend für stör-robuste Zuspielstrecken verwendet werden.

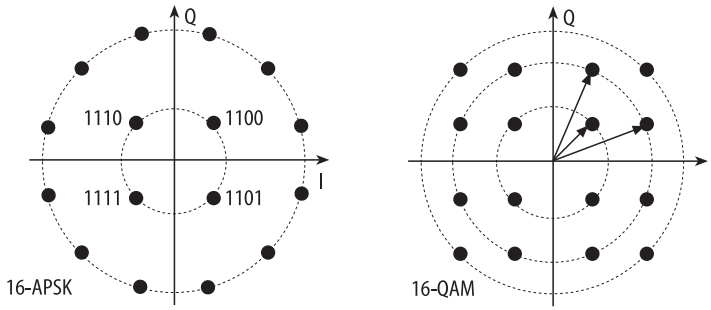


Abb. 4.50. Konstellationsdiagramme von 16-APSK und 16-QAM

4.4.4 Orthogonal Frequency Division Multiplex OFDM

OFDM (mit Fehlerschutz auch COFDM, coded OFDM) ist kein Modulations-, sondern ein Multiplexverfahren, das speziell für die Übertragung in terrestrischen Funkkanälen entwickelt wurde. Die Ausbreitungsbedingungen in Erdnähe sind im Vergleich zu Kabel- oder Satellitenstrecken besonders ungünstig. Insbesondere unterliegt das Signal starken zeitlichen Schwankungen aufgrund veränderlicher Dämpfung, aber vor allem durch Interferenzen mit Signalanteilen, die an Gebäuden oder Bergen reflektiert werden und sich je nach Phasenlage bzw. Laufzeitdifferenz am Empfänger mit dem direkten Signal konstruktiv oder destruktiv überlagern. Das Problem tritt auch bei der Analogübertragung auf, im Extremfall zeigen sich die Echostörungen hier als „Geisterbilder".

Die Auswirkungen sind bei der Digitalübertragung noch stärker, da Digitalsysteme ein »Alles oder Nichts«-Verhalten zeigen, d. h. dass das Signal solange fast einwandfreie Qualität erzeugt, bis eine Schwelle unterschritten wird, ab der es dann ganz ausfällt. Signalschwankungen um diese Schwelle bewirken damit einen ständig wiederkehrenden Bildausfall, der viel störender wirkt als ein sich langsam verschlechterndes Analogsignal. Aus diesem Grund wurde für die terrestrische Übertragung die oben erwähnte hierarchische Codierung entwickelt, und darüber hinaus das OFDM-Verfahren, mit dem die Übertragung weitgehend immun gegen Überlagerungen mit verzögerten reflektierten Anteilen gemacht werden kann. Damit wird zusätzlich ein weiterer Vorteil erzielt: die terrestrischen digitalen Sender können als Gleichwellennetz (Single Frequency Network, SFN) betrieben werden, da sich die Signale verschiedenen Sender mit ihren unterschiedlichen Laufzeiten zwar weiterhin am Empfänger überlagern, aber im Gegensatz zu analoger Übertragung keine Störungen verursachen.

Die Eigenschaften der drahtlosen Verbindungen werden in einfachen Kanalmodellen klassifiziert: AWGN (Additive White Gaussian Noise) beschreibt einen relativ unproblematischen Kanal, in dem eine direkte Verbindung zwischen Sender und Empfänger besteht und als Störung nur ein Rauschen hinzukommt, dessen Spektrum als gaußverteilt angenommen wird.

Falls zu dem Direktsignal reflektierte Anteile hinzukommen, spricht man vom Rice-Kanal. Eine Verbindung, die nur die reflektierten Anteile und nicht

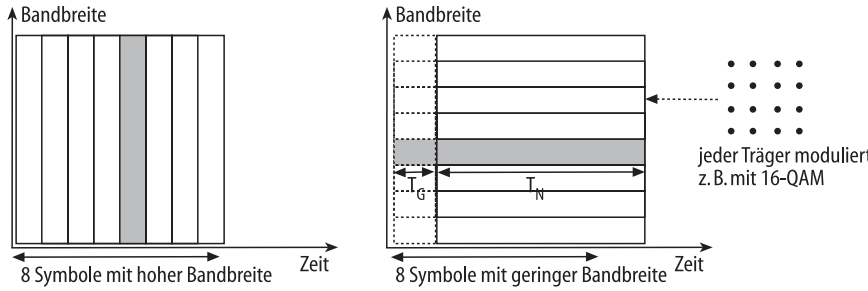


Abb. 4.51. OFDM-Grundprinzip

mehr das direkte Signal nutzen kann, wird als Rayleigh-Kanal bezeichnet. Mit OFDM wird die Nutzung von Rice- und auch Rayleigh-Kanälen möglich.

Das Problem bei den Echostörungen und Gleichwellennetzen ist die hohe Laufzeit des Echosignals. Eine Wegstreckendifferenz von s = 3 km führt zu einer Zeitdifferenz von ca. $\Delta t = s/c = 10\ \mu s$ (c = Lichtgeschwindigkeit). Diese Zeit ist länger als die Originalsymboldauer, denn diese beträgt z. B. 0,25 μs, wenn 4 MSymbole/s (z. B. bei einer Datenrate v = 20 Mbit/s und $\varepsilon = 5$ bit/s/Hz) übertragen werden. Wenn dies nicht der Fall wäre, so gäbe es die Möglichkeit, die Überlagerung von Originalsymbol und Echosymbol einschwingen zu lassen und nach einer gewissen Dauer (Schutzintervall) das Signal trotz der Verzerrung richtig zu rekonstruieren, da nicht die Signalform als Ganzes eine Rolle spielt, sondern die Werte zu den Rekonstruktions-Abtastzeiten.

Die Grundidee von OFDM ist daher, die Symboldauer zu verlängern und ein Schutzintervall hinzuzufügen in der die Störung abklingt, wobei sich natürlich die Symbolfolge pro Zeiteinheit entsprechend verringert.

Nach dem Zeitgesetz der Nachrichtentechnik ist die zeitliche Streckung des Nachrichtenquaders bei gleichem Informationsinhalt mit einer entsprechenden Verringerung der erforderlichen Bandbreite verbunden, sodass die aufgrund der Verlängerung der Symbole auftretende Verringerung der Datenrate durch die Nutzung der frei werdenden Bandbreite weitgehend ausgeglichen werden kann (Abb. 4.51). Dazu wird eine Parallelverarbeitung von N Symbolfolgen vorgenommen, die alle im Vergleich mit dem Original die 1/N-fache Symbolrate bei N-facher Dauer haben. Diese Folgen werden dann parallel im Frequenzmultiplex jeweils digital auf eine von N Trägerfrequenzen moduliert (Frequen-

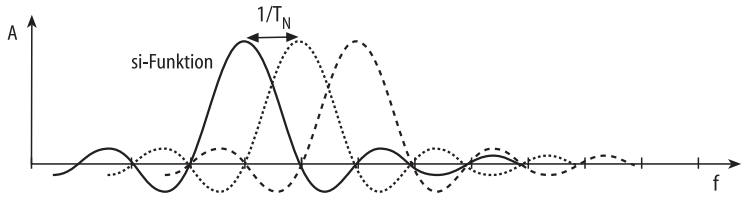


Abb. 4.52. Zur Orthogonalität von OFDM-Symbolen

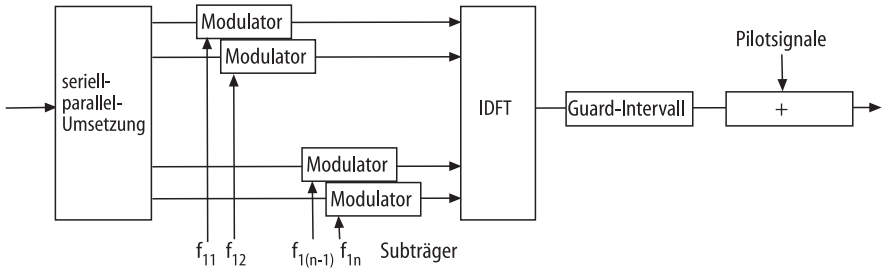


Abb. 4.53. Signalverarbeitung bei OFDM

cy Division Multiplex), sodass die Signale auf den sogenannten Subträgern schließlich zu einem Multisymbolsignal addiert werden können.

Die zeitliche Beschränkung des Nutzintervalls bewirkt ein Frequenzspektrum entsprechend einer si-Funktion (Abb. 4.52), die mit einem Übersprechen zwischen Nachbarsymbolen einhergeht (ISI, s. Abschn. 3.2.3). Im Demodulator sind die Frequenzbänder am einfachsten zu trennen, wenn das Übersprechen minimiert ist und das ist dann der Fall, wenn der Nullpunkt des Spektrums mit dem Maximum des Nachbarträgers zusammenfällt. Deshalb werden die Sub-Träger als ganzzahlige Vielfache von $f_o = 1/T_N$ gewählt. Die so gewählten Subträgerfrequenzen werden als orthogonal zueinander bezeichnet.

Die Addition ganzzahliger Vielfacher der Grundfrequenz f_o entspricht der Fouriersynthese, die Signalverarbeitung beruht daher auf Algorithmen der diskreten Fouriertransformation. Im Modulator wird zunächst dafür gesorgt, dass die Symbole parallel vorliegen, dann wird die Modulation auf die Subträger mit inverser diskreter Fouriertransformation (IDFT) vorgenommen (Abb. 4.53). Die Symbole weisen anschließend die verlängerte Nutzsignaldauer T_N auf. An dieser Stelle wird die Dauer verlängert und dem Nutz- das Guard-Intervall T_G hinzugefügt, das es dem Empfänger ermöglicht, die Echostörungen abklingen zu lassen, bevor das eigentliche Nutzintervall ausgewertet wird (Abb. 4.54).

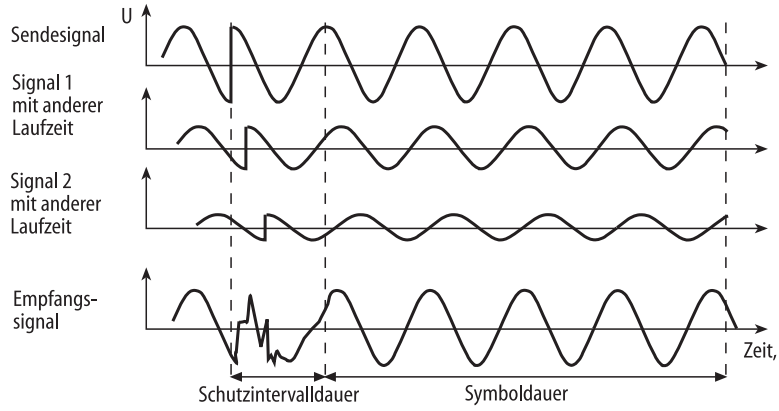


Abb. 4.54. Einschwingvorgang von OFDM-Symbolen

Je größer das Schutzintervall T_G gewählt wird, desto größere Zeitdifferenzen zwischen Nutz- und Echosignalen können ausgeglichen werden. Wegdifferenzen von 60 km erfordern ein T_G = 0,2 ms. Um die Effektivität der Übertragung nicht zu stark einzuschränken, sollte aber auch der Anteil des Schutzintervalls am Gesamtintervall erheblich kleiner sein, als der des Nutzintervalls T_N. Als Anteil von T_G an der Gesamtintervalldauer wird ca. 20%, als Nutzanteil ca. 80% gewählt. Damit erfordert ein großes Schutzintervall auch ein großes Nutzintervall und damit eine große Anzahl von Unterträgern, was wiederum mit komplexer Signalverarbeitung einhergeht.

Wenn zum Beispiel ein terrestrischer Kanal mit 8 MHz Bandbreite zur Verfügung steht und mit OFDM Wegdifferenzen bis zu 15 km ausgeglichen werden sollen, beträgt die Signallaufzeit $\Delta t = s/c$ auf 15 km ca. 50 µs. Damit ist die erforderliche Größe des Schutzintervalls festgelegt. Bei 20% Schutzintervallanteil am Gesamtintervall $T_N + T_G$, ergibt sich eine nutzbare Symboldauer T_N = 200 µs, entsprechend einem Frequenzabstand Δf = 5 kHz zwischen den Einzelträgern. Im 8-MHz-Kanal könnten damit ca. 1500 Unterträger mit 5 kHz Bandbreite verwendet werden. Zur Verarbeitung wird eine Unterträgerzahl als Potenz von 2 gefordert, hier also eine IDFT mit 2048 Werten (2k). Bei Modulation jedes Trägers mit einer 64-QAM ergibt sich für dieses Beispiel eine maximale Brutto-Datenrate von $6\ \text{bit/s/Hz} \cdot 1500 \cdot 5\ \text{kHz} \cdot T_N / (T_N + T_G) = 36$ Mbit/s. Wenn auch Echostörungen aus Wegdifferenz von bis zu 60 km ausgeglichen werden sollen, so sind die Schutz- und Nutzintervalldauern zu vervierfachen. Der Frequenzabstand der Unterträger und ihre Bandbreite wird geviertelt, d. h. es wird mit ca. 8000 Subträgern gearbeitet was eine 8k-IDFT erfordert. Bei 64-QAM-Modulation ergibt sich wieder eine Brutto-Datenrate von 36 Mbit/s.

Der Vorteil des OFDM liegt in der Unempfindlichkeit gegen Mehrwegempfang, sodass Echostörungen unwirksam gemacht und Gleichwellennetze aufgebaut werden können. Im Vergleich zur Analogübertragung, bei der Gleichwellennetze wegen des Mehrwegempfangs nicht realisiert werden können, stellt OFDM also eine Übertragungsform dar, bei der Frequenzen sehr ökonomisch genutzt werden können.

Während OFDM in Europa bei DVB-T (s. Abschn. 4.5.4) und in Japan eingesetzt wird, hat man sich in den USA früh auf die Verwendung der durch das ATSC (Advanced TV System Committee) bestimmten VSB-Modulation festgelegt. Hier wird für die terrestrische und die Kabelübertragung ein Restseitenbandverfahren mit 8 bzw. 16 Amplitudenstufen verwendet, um die sich im 6-MHz-Kanal ergebenden Datenraten von 19,3 Mbit/s bzw. 38,6 Mbit/s für die Übertragung MPEG-2 codierter Programme mit AC3-Begleitton zu nutzen.

Aber nicht nur für die Verteilung von Fernsehprogrammen, sondern auch für Zuspielzwecke ist OFDM sehr gut geeignet. Insbesondere bei kritischen Situationen mit bewegten Objekten, wie z. B. bei einer Videoübertragung des Kamerabildes aus einem Rennwagen, ist mit starken Signalstörungen zu rechnen [29]. COFDM kann für diese Zwecke mit großen Schutzintervallen, robusten Modulationsverfahren und robustem Fehlerschutz konfiguriert werden. In Kombination mit einer vernetzten Zusammenschaltung mehrerer Empfänger kann dann dafür gesorgt werden, dass trotz der bei hohen Geschwindigkeiten auftretenden Dopplereffekte das Signal noch decodierbar bleibt [30].

4.5 DVB

Mit dem Begriff Digital Video Broadcasting (DVB) oder digitales Fernsehen (digital Television, DTV) wird die standardisierte Übertragung digitaler Programme zu den Zuschauern zuhause beschrieben, die das analoge PAL-Fernsehsystem ablöst. Bei der Entwicklung des digitalen Fernsehens müssen vor allem die Aspekte Quellencodierung, Kanalcodierung und Verbreitungswege berücksichtigt werden. Die Quellencodierung betrifft die Festlegung der zu übertragenden Signalform, die Kanalcodierung die Verfahren zum Fehlerschutz und für die Verbreitungswege müssen verschiedene Formen der Modulation auf die Trägerfrequenzen bestimmt werden.

Das digitale Fernsehen nutzt die gleichen Übertragungswege wie das analoge: nämlich Kabel, Satellit und terrestrische Funkstrecken mit bodennaher Ausbreitung. Die Bezeichnungen dafür lauten DVB-C, DVB-S und DVB-T. Jeder dieser Verteilwege hat besondere Eigenschaften, auf die bei der Wahl des Fehlerschutzes und der Modulation Rücksicht genommen wird. Die drahtlosen Verbindungen DVB-S und -T unterliegen z. B. mehr Störungen als DVB-C und werden mit einem aufwendigeren Fehlerschutz versehen.

Die Quellencodierung ist unabhängig von den Verbreitungswegen. Sie beruhte für die erste Generation von DVB, die in den 1990er-Jahren entwickelt wurde, auf den Definitionen von MPEG-2, wo bereits komplette datenreduzierte Programme mit Bild und Ton definiert sind. DVB schränkt nun einerseits die durch MPEG gegebenen Möglichkeiten ein. Andererseits geht es mit zusätzlichen Definitionen für die Kanalcodierung mit Fehlerschutz und die Verbreitungsformen wieder über die Festlegungen bei MPEG hinaus. Dieses Wechselspiel von Einschränkungen und Erweiterungen findet bei der Einführung aller Standards statt und lässt sich über MPEG und JPEG bis zu den Grundlagen der Datenreduktion zurückverfolgen (Abb. 4.55).

DVB ist eine europäische Entwicklung (unter Federführung von Prof. Reimers von der TU Braunschweig) und ist in Europa eingeführt. Es wurde aber

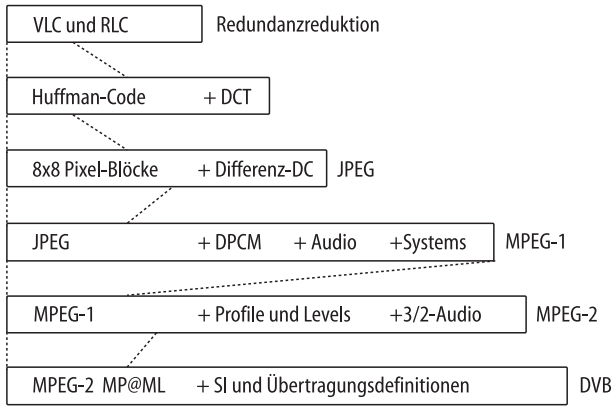


Abb. 4.55. Einschränkungen und Erweiterungen für die Datenreduktion bei DVB

auch international ein großer Erfolg. Neben Europa haben auch die großen Staaten Russland, China, Indien und Australien dafür entschieden. Die hier im Folgenden beschriebenen technischen Grundlagenaspekte für DVB gelten auch für die anderen Systeme. Größere Unterschiede gibt es nur in Teilbereichen wie der terrestrischen Variante, die in Nordamerika unter der Bezeichnung ATSC (Advanced TV-System Committee) und in Japan ISDB (Integrated Services Digital Broadcasting) eingeführt wurde.

Seit Beginn des neuen Jahrhunderts wird an der zweiten Generation von DVB gearbeitet. Für diese werden alle oben genannten Aspekte optimiert. Es gibt sowohl Verbesserungen bei den Fehlerschutzverfahren als auch die Optimierung der Modulationsparameter. Die Veränderungen betreffen aber vor allem die Quellencodierung. Während die erste Generation von DVB nur MPEG-2-codierte Signale, meist mit SD-Auflösung, unterstütze, wobei typische Datenraten zwischen 3 und 5 Mbit/s verwendet werden, ist für die Weiterentwicklung auch der Einsatz des effizienteren Datenreduktionsverfahrens H.264/AVC vorgesehen.

Darüber hinaus wird die Verteilung von HD-Signalen ermöglicht. Die deutschen Rundfunkanstalten haben sich hier für die HD-Variante 720p/50 entschieden. Die Gründe dafür liegen in dem Umstand, dass moderne Displays progressiv und nicht interlaced arbeiten und zudem verschiedene Auflösungen unterstützen. Das für die Displays vergebene HD-Ready-Zeichen erfordert, dass 720p- und 1080i-Signale akzeptiert werden und das Gerät mindestens das 720p-Signal unskaliert wiedergeben können muss. Wenn nun das 720p-Signal auf einem Display mit 1080 Zeilen dargestellt werden soll, muss es skaliert werden, während ein 1080i-Signal immer deinterlaced und im Falle einer 720-Zeilendarstellung noch zusätzlich skaliert werden muss. Deinterlacing bringt jedoch immer größere Qualitätsverluste mit sich als eine Skalierung. Die Auflösung mit 720 Zeilen wird zudem für Displaygrößen mit Bilddiagonalen bis zu ca. 1 m (40 Zoll) als ausreichend betrachtet. Hier wird davon ausgegangen, dass in Europa relativ große Wohnzimmer zur Verfügung stehen, die entsprechende Betrachtungsabstände über 2,5 m ermöglichen.

Ein zweites Argument für 720p gegenüber 1080i ist die geringere erforderliche Datenrate. In einem von der EBU durchgeführten Vergleich verschiedener Bildsequenzen und unterschiedlicher H.264-Codecimplementierungen mit den gleichen nach MPEG-2-codierten HD-Sequenzen, für die 24 Mbit/s aufgewandt wurden, zeigte sich, dass bei gleicher Bildqualität im Mittel für die 720p-Codierung ca. 10,7 Mbit/s und für 1080i-Signale 13,4 Mbit/s (bei 1920 Pixel/Zeile) bzw. 12,5 Mbit/s (bei 1080 Zeilen aber nur 1440 Pixel/Zeile) erforderlich sind. Entsprechend empfiehlt die EBU für das digitale HD-Fernsehen 720p/50-Signale bei H.264-Codierung für DVB der zweiten Generation mit mindestens 10 Mbit/s zu übertragen und für 1080i/25-Signale mindestens 14 Mbit/s aufzuwenden.

Bereits seit der ersten Generation unterstützt DVB die Verwendung von Zusatzinformationen, die vor allem einen komfortablen Programmzugriff ermöglichen. Dafür werden Tabellen in den Datenstrom integriert. Zunächst ist es wichtig, aus dem Transportstrom, der mehrere Programme enthalten kann, eines zu extrahieren und dann zu dem Bild einen von ggf. mehreren Tonkanälen auszuwählen zu können. Dazu dienen die Program Map Table (PMT), d.

Tabelle 4.4. Service Information

Tabelle	Name	Funktion
PAT	Program Association Table	Liste der Programme im Transport-Multiplex
PMT	Program Map Table	Verweis auf die Packet- ID eines Programms
CAT	Conditional Access Table	Verweis auf Entschlüsselungsdaten
NIT	Network Information Table	Daten einzelner Netzbetreiber, z. B. Satellitenposition
BAT	Bouquet Association Table	Informationen über das Angebot einzelner Anbieter
SDT	Service Description Table	Beschreibung der angebotenen Programme
EIT	Event Information Table	Programmtafeln und -kennungen, z. B. für Jugendschutz
TDT	Time and Data Table	Enthält die augenblickliche Uhrzeit
RST	Running Status Table	Angabe der laufenden Sendung zur Steuerung von VCR

h. die Liste aller Programme im Multiplex und die Program Association Table (PAT), die die Identifizierung einzelner Pakete im Transportstrom erlauben. Diese Tabellen sind, unabhängig von DVB, neben den Tabellen CAT und NIT bereits bei MPEG-2 festgelegt.

DVB legt nun die Inhalte für CAT und NIT fest und fügt fünf weitere Tabellen, als Service Information (SI) hinzu, die im Einzelnen in Tabelle 4.4 dargestellt sind. Die hier definierten Zusatzinformationen gehen wesentlich über das hinaus was beim analogen System möglich ist. Bei DVB sind die Informationen vielfältiger aber auch notwendiger, denn es muss den Zuschauern eine Hilfe an die Hand gegeben werden, mit der sie sich in der durch die Digitalübertragung ermöglichten Flut verschiedener Programme zurecht finden können. Aus diesem Grund werden Programme einzelner Anbieter oft zu sogenannten Bouquets zusammengefasst, sodass man zunächst eines der mit der Bouquet Association Table (BAT) definierten Bündel wählen kann, um sich anschließend mit Hilfe der Event Information Table (EIT) und des dort gegebenen Programmführers (Electronic Program Guide, EPG) auf ein gewünschtes Programm mit Zusatzinformationen, Untertitelung etc. zu verzweigen. Abbildung 4.56 zeigt ein Beispiel.

Abb. 4.56. Programmführung beim Digitalprogramm ZDF.vision [158]

4.5.1 Fehlerschutz

DVB ist für die Übertragung vieler Programme konzipiert, dabei muss darauf geachtet werden, dass jedes eine möglichst kleine Bandbreite belegt. Die Programme werden deshalb datenreduziert übertragen, d. h. dem Signal wird Redundanz entzogen, was wiederum bedeutet, dass es fehleranfälliger wird. Um Bandbreite zu sparen, wird darüber hinaus nach Möglichkeit ein mehrwertiges Modulationsverfahren gewählt. Durch den verringerten Abstand zwischen den Stufen steigt die Störanfälligkeit weiter. Aus diesem Grund wird bei der Kanalcodierung für DVB ein Fehlerschutz in Form der FEC (Forward Error Correction) eingesetzt. Dabei wird den Daten wieder gezielt Redundanz hinzugefügt, mit deren Hilfe auch ein fehlerhaftes Signal im Empfänger decodiert werden kann (Abb. 4.57). Die Digitaltechnik bietet hier die Möglichkeit, ein als fehlerhaft erkanntes Bit einfach zu invertieren, gleichzeitig stehen für die erforderliche Codierung effektive Rechenvorschriften zur Verfügung.

Vor dem Einsatz des eigentlichen Fehlerschutzes wird bei DVB mit Hilfe einer Energieverwischung (Energy Dispersal Scrambling) dafür gesorgt, dass der aus den 188 Bytes großen MPEG-Paketen bestehende Transportdatenstrom ein möglichst gleichmäßiges Leistungsdichtespektrum aufweist. Ähnlich wie beim Verwürfeln der NRZI-Daten bei der Basisbandübertragung werden hier die Daten mit Hilfe eines rückgekoppelten, 15-stufigen Schieberegisters, das Abgriffe bei den Stufen 14 und 15 hat, in eine Pseudozufallsfolge gebracht. Die Synchronisationsinformationen im ersten Byte (Sync-Byte) des Headers der TS-Pakete werden dabei ausgelassen. Zur Kennzeichnung des Verwürfelungsmusters wird das Sync-Byte bei jedem achten TS-Paket invertiert.

Die Grundfunktion der Fehlerkorrektur kann an einem einfachen Beispiel verdeutlicht werden: Zunächst muss ein Fehler vor einer möglichen Korrektur erkannt werden. Das kann bereits durch die Bildung einer einfachen Quersumme (Checksum) geschehen, die mit im Datenstrom übertragen wird. Soll z. B. ein digital codierter Zahlenblock mit den Werten 3, 1, 0, 4 gesendet werden, so ergibt sich nach Hinzufügung der Prüfsumme (Quersumme) die Zahlenfolge: 3, 1, 0, 4, 8. Nun könnte es sein, dass ein Fehler auftritt, der bewirkt, dass beim Empfang anstatt der 0 eine 3 erscheint. Die Quersumme weicht dann um 3 von der übertragenen Prüfsumme ab und damit ist erkannt, dass ein Fehler vorliegt.

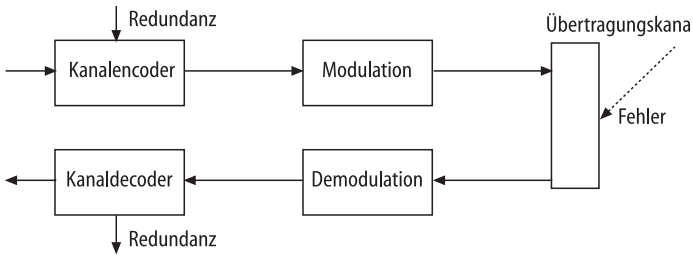


Abb. 4.57. Prinzip der Forward Error Correction

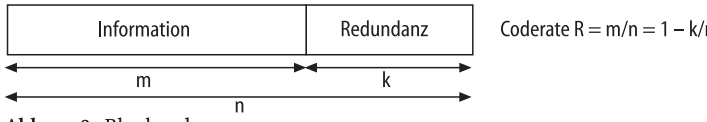


Abb. 4.58. Blockcode

Wenn der Fehler auch korrigiert werden soll, muss darüberhinaus die Position der fehlerhaften Zahl ermittelt werden. Dies ist durch Bildung einer gewichteten Quersumme möglich, deren Gewichtungsfaktoren Primzahlen sind. Bei der Gewichtung der oben genannten Werte mit den einfachen Primzahlen 2, 3, 5, 7 ergibt sich mit $2 \cdot 3 + 3 \cdot 1 + 5 \cdot 0 + 7 \cdot 4$ als Quersumme 37, die sich bei Vorliegen des genannten Fehlers auf 52 erhöht. Auf der Wiedergabeseite braucht nur die Differenz zwischen der übertragenen und der fehlerhaften gewichteten Quersumme durch den bekannten Fehlerbetrag geteilt zu werden, und die Fehlerposition ist bestimmt. In dem Beispiel ergibt sich mit $(52 - 37)/3$ die Zahl 5, also die Primzahl, die zu der entsprechenden Fehlerposition gehört. Die Erkennung und die Korrigierbarkeit der Fehler wird erhöht, wenn die Daten in zweidimensionalen Feldern angeordnet werden und die Codierung auf Zeilen und Spalten angewandt wird.

In der Praxis wird das Fehlerschutzprinzip in aufwendigen Codes umgesetzt. Bei DVB kommen Blockcodes (Block Coding), Faltungscodes (Convolution Coding) und Low Density Parity Check Codes (LDPC) zum Einsatz. Die Bezeichnung Block Code rührt daher, dass der Datenstrom in Blöcke mit fester Anzahl von Symbolen aufgeteilt wird, wobei ein Symbol mehrere Bits oder auch nur eines umfassen kann. Den m Nutzdatenbits wird eine Anzahl von k Redundanzbits angehängt, als Summe aus Nutz- und Redundanzbits ergibt sich dann die Bruttobitzahl n = m + k. Das Netto-Brutto-Verhältnis m/n wird als Coderate R bezeichnet und das Verhältnis fehlerhafter Bits zur Gesamtzahl als Bit Error Rate (BER) (Abb. 4.58).

In kritischen Situationen wird oft die Faltungscodierung gemeinsam mit einem Blockcode verwendet, denn die Effektivität des Korrekturmechanismus kann durch die Verkettung zweier Codes gesteigert werden. Dabei ist das Coder/Decoderpaar, das näher an der Signalquelle bzw. -senke liegt, für den sogenannten äußeren Fehlerschutz zuständig und das näher am Übertragungskanal liegende für den inneren (Abb. 4.59).

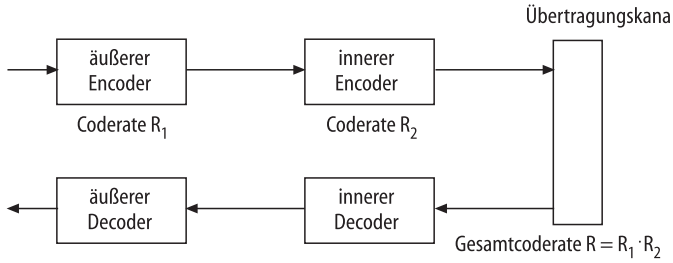


Abb. 4.59. Äußerer und innerer Fehlerschutz

4.5.1.1 Blockcodes

Es gibt verschiedene Block-Codes, bei DVB kommt der Reed-Solomon-Code zum Einsatz, der besonders dafür ausgelegt ist, Symbolfehler bei der Wiedergabe zu erkennen und zu korrigieren. Der Code gehört zu den zyklischen Codes, mit der Eigenschaft, dass sich durch zyklische Verschiebung der Elemente eines Codewortes ein neues ergibt. Das Grundprinzip entspricht dem oben dargestellten einfachen Beispiel. Statt mit einzelnen Primzahlen wird jedoch mit Polynomen für Nutzdaten und Prüfstellen gearbeitet. Bei DVB wird der Bockcode auf die jeweils 188 Bytes der Transportpakete angewandt. Durch Hinzufügung von 16 Redundanzbytes für den Fehlerschutz können bis zu acht Fehler erkannt und korrigiert werden. Die Coderate beträgt damit R = m/n = 188/204 = 0,92.

Die Fehlerkorrektur mit Blockcodes funktioniert nur bis zu einer Maximalzahl gleichzeitig auftretender Fehler innerhalb des Datenblockes. Nun können jedoch sogenannte Burstfehler leicht auftreten, also der Ausfall einer großen zusammenhängenden Anzahl von Bytes, z. B. wenn ein Signal bei der Übertragung kurzzeitig völlig ausfällt. Das Problem ist einfach dadurch zu lösen, dass die logisch zusammengehörigen Daten physikalisch voneinander getrennt werden. Dazu wird vor Anwendung des Blockcodes eine systematische Datenumverteilung vorgenommen, die als Interleaving bezeichnet wird. Burst- oder Bündelfehler werden auf diese Weise aufgebrochen und erscheinen als Einzelfehler (Abb. 4.60). Bei der Wiedergabe wird die Umordnung in entsprechender Weise wieder rückgängig gemacht.

4.5.1.2 Faltungscodes

Im Gegensatz zu den Blockcodes ist der Faltungscode bitorientiert und erfordert keine Aufteilung in feste Segmente. Hier werden m Eingangsdatenbits in einem Schieberegister gespeichert, welches verschiedene Abgriffe besitzt (Abb.

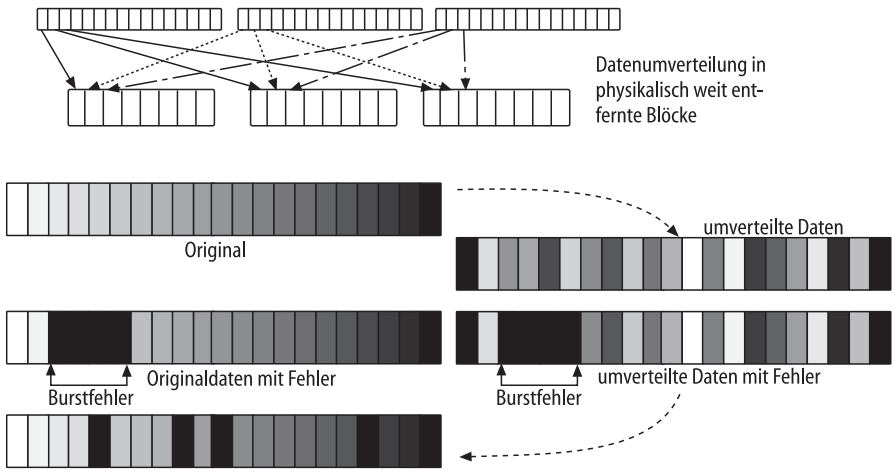


Abb. 4.60. Interleaving

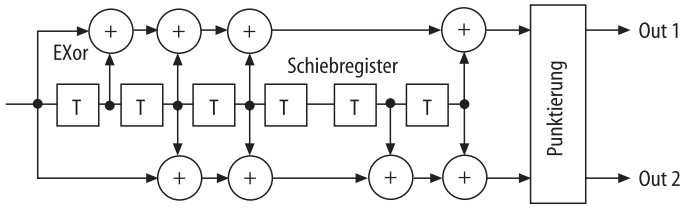


Abb. 4.61. Prinzip der Faltungscodierung

4.61). Die Eingangsinformation wird dabei auf mehrere Ausgangszweige verteilt. Im einfachsten Falle entstehen zwei Ausgangsdatenströme, wodurch die Gesamtinformation verdoppelt wird, d. h. die Hälfte der Daten sind redundant. Der Algorithmus der Faltungscodierung führt dazu, dass sich über mehrere Zeitpunkte ausgewählte Zustände ergeben. Diese können in einem Trellis-Diagramm dargestellt werden (Abb. 4.62). Ein Bitfehler kann daran erkannt werden, dass ein nicht erlaubter Zustandswechsel erfolgen soll, vergleichbar mit einem Fehler in dem regelmäßigen Muster eines Spaliers (Trellis). Bei der Decodierung wird das ursprüngliche Muster erkannt und der Fehler kann korrigiert werden und zwar umso wahrscheinlicher, je länger das Muster (Spalier) ist. Für die Decodierung wird der sogenannte Viterbi-Algorithmus eingesetzt, der einen großen Rechenaufwand erfordert.

DVB nutzt eine Faltungscodierung mit zwei Ausgängen an einem sechsstufigen Schieberegister mit jeweils fünf Auskoppelpunkten. Die große dabei zunächst entstehende Redundanz (Verdopplung der Daten bzw. R = 1/2) kann durch Punktierung gemindert werden. Das bedeutet, dass einzelne Bits nach einem regelmäßigen Muster weggelassen werden. Beim Empfang erscheint dies als Fehler, der vom Algorithmus repariert wird. Die Fehlerschutzqualität sinkt entsprechend. Die Punktierung führt zu ganzzahligen Verhältnissen, bei DVB sind Coderaten von 1/2, 2/3, 3/4, 5/6 und 7/8 erlaubt. Bei 7/8 ist der Fehlerschutz also erheblich geringer als bei 1/2 (Abb. 4.63).

Bei DVB wird die Faltungscodierung für den für den inneren und ein Reed-Solomon-Code mit der Symbolgröße 8 Bit für den äußeren Fehlerschutz verwendet. Dazwischen wird ein Interleaving durchgeführt, um zu den Bit- und kurzen Burstfehlern auch lange Burstfehler korrigieren zu können. Dabei ist eine Interleavingtiefe von 12 festgelegt, d. h. dass je 12 Transportpaketen immer zuerst die ersten, dann die zweiten Bytes etc. herausgenommen und nach-

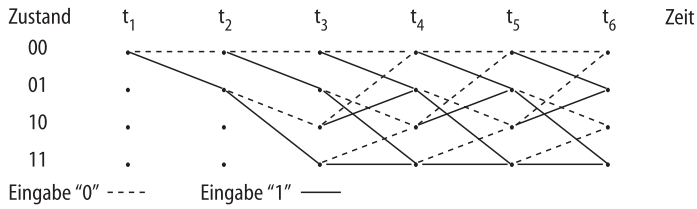


Abb. 4.62. Trellis-Diagramm

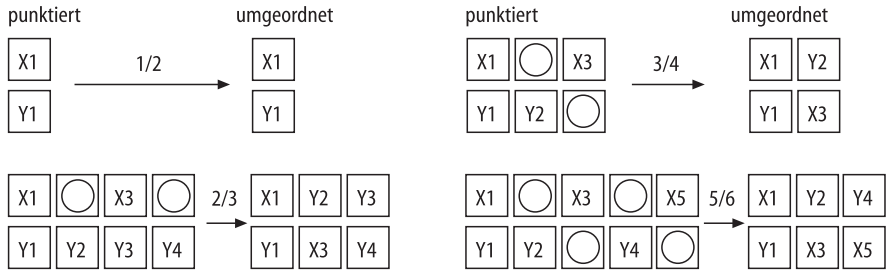


Abb. 4.63. Punktieung

einander übertragen werden, sodass ursprünglich benachbarte Bytes nach der Verschachtelung einen Mindestabstand von 205 Bytes aufweisen.

Um einen guten Schutz zu erreichen, kann die Coderate des Faltungscodes 1/2 betragen, d. h., dass keine Punktierung eingesetzt wird. Mit R = 1/2 kann auf diese Weise insgesamt eine Bit Error Rate von 10^{-2} im ersten Schritt auf 10^{-4} und dann auf einen Werte von ca. 10^{-11} gesenkt werden. Damit wird das Signal quasi fehlerfrei, was zur Decodierbarkeit des MPEG-Elementarsignals auch erforderlich ist.

4.5.1.3 LDPC-Code

Der Low Density Parity Check Code (LDPC) gehört zu den Blockcodes. Er arbeitet mit einer dünn besetzten Kontrollmatrix zur Paritätsprüfung, d. h. mit vielen Nullen und wenigen Einsen. Die Decodierung erfordert einen aufwendigen Rechenprozess auf Basis sogenannter Tanner-Graphen. Aufgrund dieses Aufwands wird dieser Code erst in neuerer Zeit verwendet, z. B. bei der Weiterentwicklung von DVB-S zu DVB-S2. Der Vorteil eines LDPC-Codes ist seine Effizienz. Er arbeitet näher an der Shannon-Grenze, d. h. näher an dem von Shannon ermittelten theoretisch möglichen Übertragungslimit des Übertragungskanals als die gewöhnlichen Block- oder Faltungscodes [57].

4.5.2 Digitales Satellitenfernsehen DVB-S

Bei Satellitenempfang ist der Umstieg von analoger zu digitaler Empfangstechnik einfach: wenn bereits ein rauscharmer LNB vorhanden ist, braucht nur der Satellitenreceiver ausgetauscht zu werden. Das Prinzip der Satellitenübertragung ist in Abschn. 4.3.2 dargestellt und für analoge und digitale Übertragung gleich. Die Übertragung ist nicht an spezielle Satelliten gebunden, freie analoge Satellitenkanäle können für Digitalsignale genutzt werden.

Die Transponder eines Satelliten sollen für mehrere unabhängige Dienste genutzt werden können (Transpondersharing), dazu müssen die Daten im Multiplex zusammengefasst werden. Es ist diesbezüglich sowohl das Frequenzmultiplex- (Frequency Division Multiple Access, FDMA) als auch das Zeitmultiplexverfahren (Time Division Multiple Access, TDMA) einsetzbar (Abb. 4.64). TDMA erfordert eine Zusammenfassung aller Programme in der Up-Link-Sta-

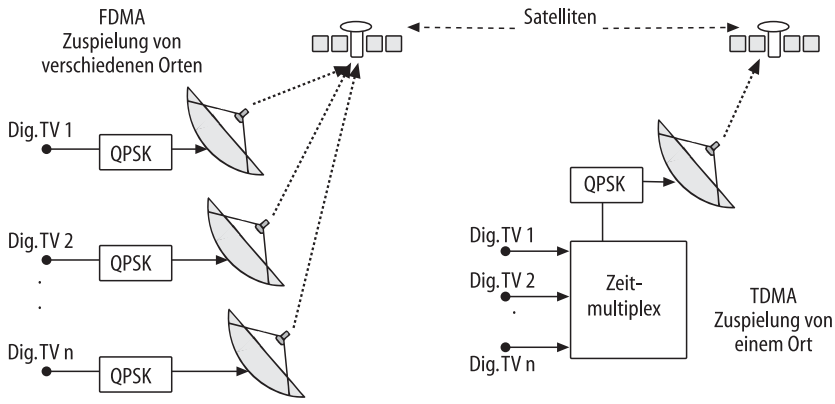


Abb. 4.64. Digitale Satellitenübertragung im Frequenz- und Zeitmultiplex

tion, bietet aber den Vorteil, dass die Transponderkapazität besser ausgenutzt werden kann, da die Signaltrennung im Zeitbereich effektiver ist als im Frequenzbereich.

DVB-S arbeitet für den einzelnen Transponder mit dem Zeitmultiplex-Verfahren. Die weiteren Parameter von DVB-S wurden so gewählt, dass bei einem Empfang mit Reflektoren von 0,5 m Durchmesser eine Bitfehlerrate von 10^{-12} bei einer mittleren zeitlichen Verfügbarkeit von 99,9% erreicht wird. Die Quellencodierung erfolgt nach MPEG-2, die damit definierten Transportströme werden in Datencontainer gepackt und mit einem verketteten Fehlerschutz versehen, wie es in Abschn. 4.6.1 dargestellt ist. Dabei ist die Wahl verschiedener Coderaten erlaubt, sodass z. B. eine höhere Datenrate durch Senkung der zugesetzten Redundanz gewählt werden kann, wenn ein Transponder mit hoher Sendeleistung benutzt wird.

Durch die Nutzung der Network Information Table (NIT), in der alle für die Einstellung auf einen Satellitentransponder erforderlichen Daten wie Transponderfrequenz, Orbitalposition, Polarisation gespeichert sind, ist der Empfänger eines zu einem Netzwerk gehörigen Transponders (z. B. Astra) sofort über die anderen Transponder im Netz informiert.

Abbildung 4.65 zeigt das Blockschaltbild der Signalcodierungsstufen [174]. Da die Satellitenstrecke einen stark gestörten Übertragungskanal darstellt, wurde als Modulationsart für DVB-S die robuste QPSK vorgesehen. Die Bitfehlerrate ist von den Störungen bestimmt, die den PSK-Zuständen als Phasenrauschen überlagert sind. Die vor dem inneren Fehlerschutz tolerierbare Fehlerrate von $2 \cdot 10^{-4}$ erlaubt eine durch das Phasenrauschen hervorgerufene Verschmierung der Zustände mit einer sogenannten I, Q-Closure von ca. 30%.

Die theoretische Bandbreiteneffizienz der QPSK von $\varepsilon = 2$ bit/s/Hz reduziert sich in der Praxis durch Filterung. Grundsätzlich ist für DVB ein roll-off-Faktor von $\alpha = 0{,}35$ festgelegt. In praktischen Fällen wird er ggf. etwas verkleinert. Bei der Übertragung über den Satelliten Astra 1E wird z. B. mit einer durch die Filterung beschränkten Bandbreiteneffizienz von $\varepsilon_{Filt} = 1{,}53$ bit/s/Hz gearbeitet.

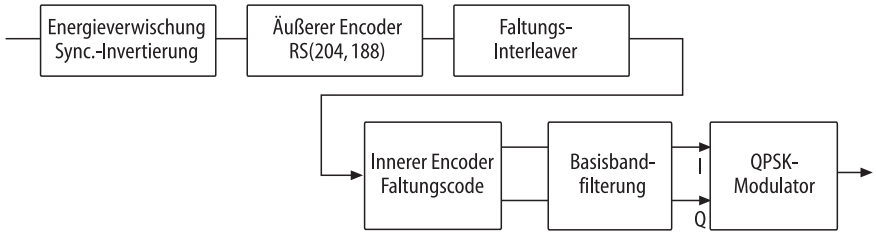


Abb. 4.65. Blockschaltbild für die digitale Satellitenübertragung

Zur Berechnung der übertragbaren Nutz-Datenrate müssen nun die Bandbreiteneffizienz, die Übertragungsbandbreite $B_{ü}$ selbst und die Coderaten R_1 und R_2 des inneren und äußeren Fehlerschutzes berücksichtigt werden (Abb. 4.65). Insgesamt gilt:

$$v_{Bit} = \varepsilon_{Filt} \cdot R_1 \cdot R_2 \cdot B_{ü}.$$

Mit R_1 = 188/204 = 0,922 für den Reed-Solomon-Code und z. B. R_2 = 3/4 für den Faltungscode ergibt sich in einem Satellitenkanal mit $B_{ü}$ = 36 MHz Bandbreite und ε_{Filt} =1,53 bit/s/Hz eine Nutz-Bitrate von 38 Mbit/s bei einer Brutto-Bitrate von 55 Mbit/s. Ausgehend von 4,5 Mbit/s für ein nach MPEG-2 MP@ML codiertes SDTV-Programm folgt daraus, dass 8 Programme in jedem Satellitenkanal untergebracht werden können. Wenn auf einem Satelliten 18 Transponder mit je 36 MHz zur Verfügung stehen, bietet ein einzelner Satellit somit eine Übertragungskapazität von 144 TV-Programmen in Standardqualität.

Die Datenrate reicht aber auch aus, um ein oder zwei HDTV-Signale zu übertragen. Zu Beginn des Jahres 2004 wurde unter dem Namen Euro 1080 in Belgien der erste echte HDTV-Dienst Europas aufgenommen. Über den Satelliten Astra 1H wurden zwei Programme mit je 19 Mbit/s übertragen. Dabei wurde jedes HDTV-Signal mit 1920 x 1080 Bildpunkten interlaced nach MPEG-2 MP@HL codiert.

Wie im analogen Falle ist die digitale Satellitenübertragung aber nicht nur für Broadcastzwecke von Bedeutung, sondern auch für Zuspielungen und ggf. Verteilung in Kabelnetze (Abb. 4.66) und für die aktuelle Berichterstattung (Digital Satellite News Gathering, DSNG). Hier wird oft mit robusterem Fehlerschutz und höheren Datenraten, z. B. von 8 Mbit/s gearbeitet. Auch unter diesen Umständen können immer noch mehr Programme über einen Transponder übertragen werden als im analogen Fall, sodass die Digitalübertragung erheblich preiswerter angeboten werden kann als die analoge.

4.5.2.1 DVB-S2

Die Weiterentwicklung von DVB-S zu DVB-S2 ermöglicht eine bis zu 30%-ige Effizienzsteigerung der Übertragungskapazität. Dies wird vor allem durch Veränderungen beim Fehlerschutz erreicht. Für DVB-S2 kommt der LDPC-Code mit sehr großen Blocklängen von 64800 Bits oder 16200 Bits zum Einsatz. Er wird als sogenannter Turbocode implementiert, d. h. mit einer Rückkopplung zum Eingang. Das erlaubt beim Einsatz der QPSK mit einem Roll-off-Faktor von α = 0,2 zu arbeiten. Ein 36 MHz-Transponder ermöglicht auf diese Weise

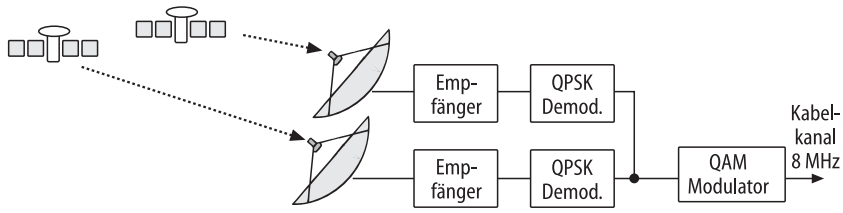


Abb. 4.66. Einspeisung digitaler Signale in ein Kabelnetz

mit R = 3/4 eine Nutzdatenrate von 46 Mbit/s statt der oben berechneten 38 Mbit/s bei DVB-S. In Kombination mit der nun zulässigen Quellencodierung nach H.264/AVC können bei DVB-S2 über den Transponder in diesem Falle 21 SD-Programme mit je 2,2 Mbit/s oder 5 HD-Programme mit je 9 Mbit/s übertragen werden.

Bei der Modulation sind neben QPSK auch 8-PSK, sowie 16- und 32-APSK zugelassen. Die 8-PSK kann bei hoher Sendeleistung und entsprechend vergrößertem Störabstand eingesetzt werden, 16- und 32-APSK sind jedoch nicht für Broadcastzwecke vorgesehen, sondern eher für DSNG etc. Die Schwierigkeit ist hier, dass unterschiedliche Amplitudenstufen erkennbar sein müssen, sodass die in den Satellitentranspondern verwendeten Wanderfeldröhren nicht mehr soweit in die Nichtlinearität getrieben werden können wie üblich [49].

In professionellem Zusammenhang kann die Datenrate noch weiter gesteigert werden, indem auf das übliche CCM (Constant Coding and Modulation) verzichtet und statt dessen das bei DVB-S2 mögliche ACM (Adaptive Coding and Modulation) eingesetzt wird. Die Steigerung wird dadurch möglich, dass die Qualität der Signalstrecke ständig überprüft und die Datenrate entsprechend angepasst wird. Insbesondere Punkt-zu-Punkt-Verbindungen, wie DSNG, könnten zukünftig erheblich von dem Verfahren profitieren.

4.5.3 Digitales Kabelfernsehen DVB-C

Auch in den in Abschn. 4.3.3 beschriebenen Kabelkanälen soll die Verbreitung von digitalen Fernsignalen in Datencontainern mit bis zu 38 Mbit/s möglich sein [175]. Die Kabelübertragung hat den Vorteil, dass hier ein relativ ungestörter Übertragungskanal vorliegt. Unter Berücksichtigung aller Einflüsse kann mit einem Träger/Rauschabstand von C/N = 28 dB gerechnet werden. Daher können für DVB-C die höherwertigen Modulationsarten 16-QAM, 64-QAM bis zu 256-QAM eingesetzt werden. Um in den für das Kabelsystem definierten Kanal-Bandbreiten im 8 MHz-Raster die Datenrate 38 MBit/s zu erreichen, muss die 64-QAM gewählt werden. Die 32-QAM mit der Bandbreiteneffizienz von 5 bit/s/Hz reicht hier nicht aus, da wieder der Fehlerschutz und die Wirkung des Matched Filters im Decoder berücksichtigt werden muss.

Da häufig Satellitenprogramme in Kabelnetze eingespeist werden (Abb. 4.66), sollten möglichst viele Gemeinsamkeiten zwischen den technischen Parametern der beiden Übertragungsstrecken bestehen. Auch im Hinblick auf die Kosten

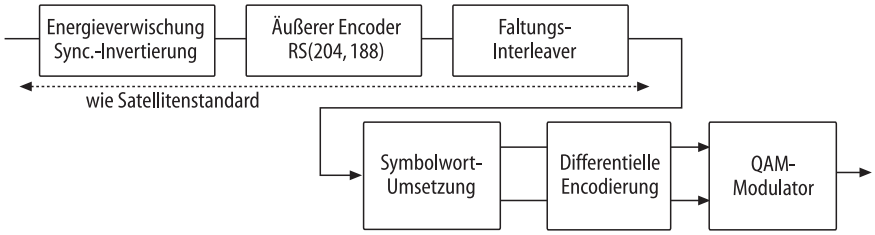


Abb. 4.67. Blockschaltbild für die digitale Kabelübertragung

für Empfänger, die für die Decodierung von Signalen nach DVB-C und DVB-S ausgelegt sind, wurde die Signalverarbeitung von DVB-S für DVB-C weitgehend übernommen (Abb. 4.67). Im Unterschied zum stark gestörten Satellitenkanal kann bei der Kabelübertragung jedoch auf den inneren Fehlerschutz verzichtet werden. Damit ergibt sich eine hohe Nutzdatenrate, die Bruttodatenrate ist nur durch den Faktor 188/204 des RS-Fehlerschutzes und die Filterung mit Roll-off-Faktor r = 0,15 beschränkt. Es gilt:

$$v_{Bit} = \varepsilon_{Filt} \cdot R_1 \cdot B_{\ddot{u}} = \varepsilon/1{,}15 \cdot 188/204 \cdot 8 \text{ MHz}.$$

Bei Einsatz der 64-QAM mit ε = 6 ergibt sich eine Nutz-Datenrate von 38,47 Mbit/s, die es erlaubt, die 38 Mbit/s eines Satellitenkanals vollständig zu übernehmen. Dabei steht noch ein Spielraum für das Einfügen von sogenannten Stopfbits zur Verfügung, die für die Synchronisierung bei der veränderten Zusammenfassung von Transportströmen erforderlich sind, wenn das Programmangebot aus verschiedenen Transpondern übernommen wird.

Bei der Einführung von digitalem Kabelfernsehen wird zunächst der Hyperband-Bereich zwischen ca. 300 MHz und 450 MHz benutzt, der mit 18 Kanälen

Tabelle 4.5. Beispiel für digitale TV-Programme im Kabelnetz Quelle: Deutsche TV-Plattform

S26-Fremdsprachen	S27-Premiere	S28-Premiere	S29-Premiere	S30-Premiere	S31-Premiere
ATV (türkisch)	Blue Movie 1	Cinedom 1.1	Cine Action	Comedy&Co	Sport 2
Kanal D (türkisch)	Blue Movie 2	Cinedom 1.2	Cine Comedy	Discovery	Sport 3
TV Polonia (polnisch)	Blue Movie 3	Cinedom 1.3	Premiere	Junior	Sport 4
RTP (portugiesisch)	Blue Channel	Cinedom 1.4	Romantic	K-Toon	Sport 5
ERT SAT (griechisch)	Cinedom 2.1	Cinedom 1.5	Sci-Fantasy	Krimi&Co	Cinedom 1
ZEE-TV (indisch)	Cinedom 2.2		Star Kino	Planet	Cinedom 2
CNE (chinesisch)	Cinedom 4.1		Studio Univ.	Sports World	Cinedom 3
	Heimatkanal		13th Street	Sunset	Cinedom 8

S32-ZDF.Vision	S33-Vorschlag	S34-ARD 1		S36-ARD 2		S37-Vorschl.
3 Sat	Seasons	Das Erste	Hörfunk:	B1	Hörf.:	Fashion TV
arte/Kinderkanal	Planet	BR-Alpha	Bayern 4, 5	Eins Extra	Fritz!	Bet on Jazz
ORF	Studio Universal	BR	Bremen 2	Eins Festiv.	MDR	Bloomberg
Phoenix	Disney Channel	HR	hr 1, 2, 2+	Eins Muxx	SFB	Landscape
ZDF	Simulcast 1 – 3	N3	NDR 4	MDR	SWR	Extr. Sports
ZDF-Info Box	Hörf.: Pop/Rock	SWR	Radio 3	ORB	WDR	Einstein
	Hörf.: Klassik	WDR				Single TV

von je 8 MHz Breite belegt ist, sodass das Hyperband, wie auch ein Satellitentransponder, eine Übertragungskapazität von bis zu 144 Programmen in Standardqualität bietet. Ein Beispiel für eine digitale Programmzuordnung ist in Tabelle 4.5 dargestellt.

4.5.3.1 DVB-C2

Auch für DVB-C wird an einer zweiten Generation, also DVB-C2 gearbeitet. Das Ziel ist, aktuell zu bleiben und mit den modernen technischen Parametern mitzuhalten. Das gilt insbesondere für die Maßstäbe, die durch DVB-S2 gesetzt wurden. Wie dort soll die Übertragung um ca. 30 % effektiver werden, sodass mit modernen Codecs wie H.264/AVC eine erhebliche Steigerung der übertragbaren Programme erreicht wird. Neben CCM soll auch hier ACM bzw. VCM (Variable Coding and Modulation) unterstützt werden, als Fehlerschutz ist wieder LDPC vorgesehen. Neben der üblichen QAM sollen für die Modulation auch Mehrträgerverfahren wie OFDM verwendet werden können, mit bis zu 4096-QAM [49].

4.5.3.2 MMDS

Der Kabelstandard von DVB soll auch für die Digitaldatenübertragung mittels Mikrowellen im Bereich unter 10 GHz benutzt werden. Dieses Multichannel/Multipoint Distribution System (MMDS) ist für eine lokale terrestrische Verteilung mit Reichweiten unter 50 km vorgesehen (wireless cable). Die vorgesehenen Frequenzbereiche liegen zwischen 2 und 42 GHz. Während für Frequenzen unterhalb 10 GHz der bandbreiteneffiziente Kabelstandard DVB-MC (Microwave Cable Based) favorisiert wird, ist oberhalb 10 GHz die Verwendung des Satellitenstandards DVB-MS (Microwave Satellite Based) vorgesehen, da dort die Dämpfung ansteigt und der Übertragungskanal als stark gestört betrachtet werden muss.

4.5.4 Digitales terrestrisches Fernsehen DVB-T

Die Übertragungsparameter zur digitalen terrestrischen Übertragung wurden erst nach denen für die Satelliten- und Kabelübertragung definiert. Wie bei DVB-C soll auch dieser Standard weitgehende Ähnlichkeit mit DVB-S aufweisen, dabei sind die Kanäle auch an einer Bandbreite von 8 MHz orientiert, die wieder möglichst effizient genutzt werden soll.

Ein wesentlicher Unterschied zwischen Kabel- und terrestrischer Fernsehübertragung besteht bei Programmen, für die eine flächendeckende Versorgung erzielt werden soll, z. B. beim Programm des ZDF. Im Kabel ist dafür nur ein Kanal erforderlich, denn in jedem BK-Netz kann der gleiche Kanal mit dem Programm des ZDF belegt werden.

Dagegen ist es mit terrestrischer Ausstrahlung bei Anwendung konventioneller analoger Modulationsarten nicht möglich, ein Programm flächendeckend mit nur einer Trägerfrequenz zu verteilen, denn bei diesem Gleichwellenbetrieb stören sich die vielen Sender, die zur Versorgung eingesetzt werden müssen, gegenseitig.

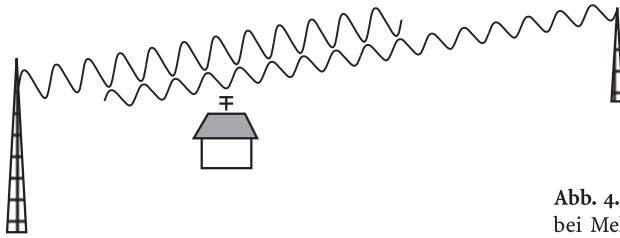


Abb. 4.68. Destruktive Interferenz bei Mehrwegempfang

Die Störung resultiert aus der Überlagerung der Wellen (Abb. 4.68) von mehreren Sendern oder aus Reflexionen. In den meisten Fällen laufen die Wellen über unterschiedlich lange Strecken, sodass Phasenverschiebungen auftreten, die zu destruktiver Interferenz, also Dämpfung oder Auslöschung der Welle, oder zu sogenannten Geisterbildern führen können.

Bei Digitalsignalen ist es möglich, mit Hilfe der OFDM (s. Abschn. 4.4.4) die genannten Probleme beim Mehrwegempfang zu umgehen, die Störsicherheit zu erhöhen und Gleichwellennetze zu errichten. Bei mehrwertiger Modulation besteht das Digitalsignal aus einer Folge von Symbolen, die eine bestimmte Anzahl von Bits repräsentieren. Der wesentliche Gedanke ist, jeder Symboldauer ein Schutzzeitintervall hinzuzufügen, das so bemessen ist, dass die störenden Überlagerungen aus dem Mehrwegempfang während der Schutzintervalldauer abklingen können (Abb. 4.54).

DVB-T macht sich die genannten Vorteile der OFDM zunutze, die hier wegen der Fehlerschutzcodierung auch als Coded OFDM (COFDM) bezeichnet wird. Die Festlegung der Parameter geschieht in Hinblick auf den Einsatz portabler Empfänger, wobei auch die hierarchische Codierung verwendet werden kann. Das in Abbildung 4.69 dargestellte Blockschaltbild des Encoders zeigt wieder eine weitgehende Übereinstimmung mit dem von DVB-S. Dem inneren Fehlerschutz folgt eine, bei DVB-S nicht vorhandene, innere Interleaving-Stufe, die dazu dient, Störungen aus Interferenzen dadurch zu umgehen, dass zeitlich aufeinanderfolgende Symbole auf verschiedene Träger verteilt werden. Als Modulationsarten für die Einzelträger können QPSK, 16-QAM und 64-QAM verwendet werden. Abbildung 4.70 zeigt im Vergleich zu dem Spektrum eines Analogsignals wie im Digitalkanal dafür gesorgt wird, dass sich die Energie spektral gleichmäßig verteilt.

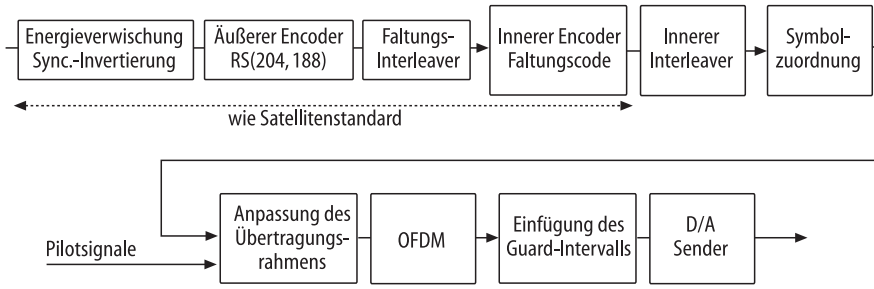


Abb. 4.69. Blockschaltbild für die digitale terrestrische Übertragung

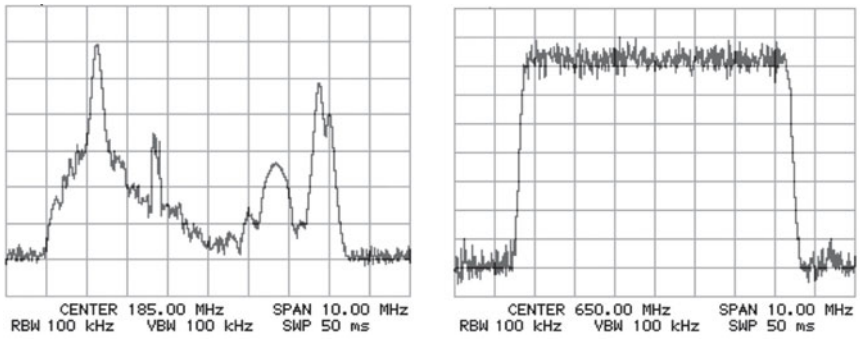


Abb. 4.70. Vergleich der Signalspektren eines analogen und eines digitalen TV-Kanals

Die Wahl der OFDM-Parameter für DVB-T hängt zunächst mit den größten Wegdifferenzen der sich überlagernden Wellen zusammen, woraus wiederum die erforderliche Schutzintervalldauer resultiert. Für die Errichtung nationaler Gleichwellennetze in Deutschland müssen Laufzeitdifferenzen von s = 60 km ausgeglichen werden können. Dafür sind Schutzintervalle der Dauer $\Delta t = s/c = 200\ \mu s$ erforderlich. Um die Nutzdatenrate nicht zu sehr zu beschränken, wird meist ein Verhältnis der Intervalldauern $T_G/T_N = 1/4$ gewählt, woraus eine Nutzintervalldauer von ca. 0,8 ms und ein Einzelträgerabstand von ca. 1,25 kHz folgt. Diese geringe Frequenz erfordert eine aufwendige Signalverarbeitung der ca. 6000 Subträger, die in einem 8-MHz-Kanal Platz finden, mit der 8k-Variante der IDFT (s. Abschn. 4.5.3). Um eine weniger aufwendige Signalverarbeitung realisieren zu können, wurde mit der Common 2k/8k-Specification bei DVB-T festgelegt, dass auch mit einer 2k-Variante gearbeitet werden kann. Die wesentlichen Parameter unterscheiden sich von der 8k-Variante um den Faktor vier. Mit der 2k-IDFT können daher nur Laufzeitdifferenzen aus einem Wegunterschied von 15 km ausgeglichen werden.

Für eine hierarchische Übertragung kann eine Multiresolution-QAM (MR-QAM) verwendet werden. Unter Einsatz einer 64-QAM liefert jedes Modulationssymbol 6 Bit. Bei Empfang mit Dachantennen sollte der Störabstand ausreichen, um alle Zustände auswerten zu können. Beim Empfang mit portablen Geräten ist dagegen der Störabstand schlechter. Im ungünstigen Fall können nur vier Zustände als »Wolken« unterschieden werden (Abb. 4.49). Als Zwischenstufe können die 16 kleineren »Wolken« detektiert werden (4 Bit/Symbol).

Die genaue Festlegung bei DVB-T bestimmt für das 8k-System eine Abtastfrequenz von 64/7 MHz, damit eine einfache Umschaltung der Parameter für Kanäle mit 7 bzw. 8 MHz Bandbreite erfolgen kann [57]. Die Angabe 8k bezieht sich auf die Dualzahl 2^{13}, die genaue Anzahl von Trägern beträgt daher 8192. Da im Übertragungskanal mit 8 MHz Bandbreite etwa 7,6 MHz praktisch genutzt werden können, ist die Verwendung von 6817 Subträgern festgelegt, die genau 7,609 MHz Bandbreite belegen. Um diesen Faktor 6817 wird die Symboldauer verlängert, bzw. die Einzelträgerbandbreite verringert, sodass der Einzelträgerabstand bei der 8k-Variante 1,116 kHz und bei der 2k-Variante 4,464 kHz beträgt.

Von den 6817 Trägern gibt es 6048, die mit gleicher Leistung übertragen werden. Für die Übertragung des eigentlichen, fehlergeschützten MPEG-Transportstroms können davon wiederum nur 5980 genutzt werden, da 68 Träger als sogenannte TPS-Träger zur Signalisierung von Übertragungsparametern verwendet werden. Die restlichen 769 Träger dienen als Referenzsignale, die mit einer gegenüber den Datenträgern um 4/3 erhöhten Leistung übertragen werden. 177 davon werden als sogenannte ständige Piloten verwendet, die ohne Modulation in jedem OFDM-Symbol an derselben Trägerposition erscheinen und zur Grobeinstellung des Empfängeroszillators dienen. Die restlichen Referenzsignale werden als verteilte Piloten (Scattered Pilotes) bezeichnet, da ihre Position bezüglich jedes Symbols so verändert wird, dass sie einer Sequenz folgen, die sich alle 4 Symbole wiederholt. Die hohe Zahl dieser Piloten ermöglicht die Feinabstimmung des Empfängers sowie eine ständige Abschätzung der Eigenschaften des Übertragungskanals [174].

Die mit DVB-T übertragbare Datenrate hängt wesentlich von der gewählten Modulationsart ab. Als höchstwertige kann die 64-QAM mit der Effizienz ε = 6 bit/s/Hz verwendet werden. Im 8k-System beträgt die Symboldauer 1/1,116 kHz = 896 µs. Die Schutzintervalldauer T_G kann in den Verhältnissen 1/4, 1/8, 1/16 und 1/32 zur Symboldauer T_N gewählt werden. Mit T_G/T_N = 1/4 ergibt sich die größte zulässige Wegdifferenz zwischen den sich überlagernden Signalen von 67 km bei einer Gesamtintervalldauer von 1,12 ms. Mit 16-QAM folgt unter diesen Umständen für die Bruttobitrate

$$v_{Bit\ (Brutto)} = 6817 \cdot 4\ \text{bit}/1{,}12\ \text{ms} = 24{,}35\ \text{Mbit/s}.$$

Für die Nettobitrate gilt:

$$v_{Bit\ (Netto)} = R_1 \cdot R_2 \cdot 6048/6817 \cdot v_{Bit\ (Brutto)} = 13{,}27\ \text{Mbit/s},$$

und zwar nach Abzug der Pilotsignale (Restanteil 6048/6817) und wieder unter Berücksichtigung des Fehlerschutzes mit den Coderaten R_1 und R_2, wobei eine

Tabelle 4.6. Erreichbare Datenraten bei DVB-T

Modulation	Coderate	1/4	1/8	1/16	1/32	Schutzintervallanteil T_G/T_N
QPSK	1/4	4,98	5,53	5,85	6,03	Netto-Datenrate [Mbit/s]
	2/3	6,64	7,37	7,81	8,04	
	3/4	7,46	8,29	8,78	9,05	
	5/6	8,29	9,22	9,76	10,05	
	7/8	8,71	9,68	10,25	10,56	
16-QAM	1/2	9,95	11,06	11,71	12,06	
	2/3	**13,27**	14,75	15,61	16,09	
	3/4	14,93	16,59	17,56	18,10	
	5/6	16,59	18,43	19,52	20,11	
	7/8	17,42	19,35	20,49	21,11	
64-QAM	1/2	14,93	16,59	17,56	18,10	
	2/3	19,91	22,12	23,42	24,13	
	3/4	22,39	24,88	26,35	27,14	
	5/6	24,88	27,65	29,27	30,16	
	7/8	26,13	29,03	30,74	31,67	

Coderate R_2 = 2/3 für die Faltungscodierung angenommen wurde. Die Nettobitrate von 13,27 Mbit/s ist für DVB-T ein typischer Wert, der durch 64 QAM, geringeren Fehlerschutz und kürzere Schutzintervalle noch auf die in Tabelle 4.6 angegebenen Datenraten gesteigert werden kann. Bei Verwendung einer 64 QAM ist ein Störabstand von mindestens 20 dB und Empfang mit gerichteter Antenne erforderlich. Andererseits kann durch eine robuste Modulation mit QPSK und R_2 = 1/2 der erforderliche Störabstand auf unter 6 dB reduziert werden.

DVB-T wird in Deutschland im Regelbetrieb seit 2003 eingesetzt. Der Umstieg von analoger zu digitaler Verbreitung erfolgte dabei in sogenannten digitalen Inseln im Simulcastbetrieb, d. h. die analoge Ausstrahlung, zumindest der öffentlich-rechtlichen Programme, blieb zunächst bestehen. Die erste Insel war der Raum Berlin/Brandenburg, es folgten die Regionen Hannover/Braunschweig und Bremen/Unterweser (Abb. 4.71).

In den Inseln wurden zunächst vier oder fünf 8 MHz-Kanäle für die Digitalübertragung freigemacht, die mit COFDM im 8k-Modus arbeitet. Der Schutzintervallanteil beträgt 1/4, sodass sich mit einer Coderate 2/3 eine Nettobitrate von 13,27 Mbit/s ergibt. Diese wird für vier Programme genutzt. Mit Hilfe moderner MPEG-2 Coder lässt sich mit der hieraus folgenden Beschränkung auf 3,1 Mbit/s für das Video- und 192 kbit/s für ein Stereosignal eine recht gute Bildqualität erzielen. Das gilt auch mit portablen oder mobilen Geräten.

Die Modulationsparameter sind so gewählt, dass ab einem Signalpegel von ca. 30 dB/µV am Empfängereingang mit stabilem DVB-T-Empfang gerechnet werden kann. In der Nähe des Senders ist ein sogenannter portable Indoor-Empfang mit kurzen Stabantennen möglich, bei größeren Abständen ist eine Außenantenne erforderlich, wie sie bereits für analoge Geräte verwendet wird. Mobiler Empfang mit einfachen Antennen gelingt im Stadtgebiet bis zu Geschwindigkeiten von 60 km/h, bei größeren Geschwindigkeiten müssen mehrere

Technische Parameter DVB-T:
Bandbreite: 8 MHz
Multiplex: COFDM, Modus: 8k
(6817 Einzelträger, davon 6048 für Nutzdaten)
Modulation: 16-QAM (4 bit pro Einzelträger)
Coderate: 2/3 (1/3 der Datenrate für Fehlerschutz)
Schutzintervall: 1/4
(erlaubt Laufzeitunterschiede bis zu 67 km)
Datenrate in Mbit/s: 13,27
(genutzt für 4 Fernsehprogramme)

DVB-T-Programme in der Region Unterweser
K 22: Das Erste, N3, Radio Bremen
K 32: ZDF, KiKa, ZDF-Doku/Info, 3Sat
K 42: RTL, RTL 2, VOX, Super RTL
K 49. Sat 1, Pro7, Kabel 1, N24
K 55: NDR, WDR, MDR, HR
K 29: Das Erste, 1+ Phönix, Arte

Abb. 4.71. DVB-T-Inseln, technische Parameter und Beispiel für ein Programmangebot

Antennen im Diversity-Betrieb verwendet werden, d. h. dass automatisch immer die Antenne mit dem höchsten Pegel an den Empfänger geschaltet wird [64].

Mit vier bis sechs freigegebenen Frequenzen können 16 bis 24 Programme gesendet werden, wobei die Programmmultiplexe der öffentlich-rechtlichen Anstalten und die wichtigsten privaten Sender verbreitet werden (Abb. 4.71). Dieses Angebot ist für viele Haushalte ausreichend, sodass durch DVB-T mit einer Verschiebung der Marktanteile vor allem zu Lasten des Kabelempfangs, für den ja zusätzliche Gebühren anfallen, gerechnet wird. Insgesamt wird prognostiziert, dass durch die Digitalisierung der Satellitenempfang von 35 % auf 40 % und der terrestrische von 7 % auf 20 % steigen werden, während der Kabelempfang von 58 % auf 40 % zurückgehen wird.

4.5.4.1 DVB-T2

Die zweite Generation von DVB-T soll wieder weitgehende Ähnlichkeiten zu DVB-S2 aufweisen und durch die verbesserte Codierung einen Effizienzgewinn der Frequenzausnutzung von ca. 30% gegenüber DVB-T erzielen.

Für den Fehlerschutz kommt ein LDPC-Code zum Einsatz. Die Modulation bedient sich weiter der OFDM, mit Trägerzahlen bis 32k und flexiblerer Skalierbarkeit. Für die Träger können QAM-Verfahren bis zu 256 Stufen verwendet werden. Als Neuerung wird das MISO-Verfahren (multiple input/single output) implementiert. Während konventionelle Systeme als SISO (single input/single output), also mit einer Sende- und einer Empfangsantenne arbeiten, kann MISO die Signalqualität durch die Verwendung mehrerer Sendeantennen verbessern [13]. Bei zwei Antennen könnten diese am gleichen Standort (co-located) die Welle mit zwei verschiedenen Polarisationsrichtungen oder von zwei verschiedenen Orten (distributed) abstrahlen. Eine weitere Verbesserung kann diesbezüglich das MIMO-Verfahren bieten, das zusätzlich mit mehreren Empfangsantennen arbeitet, diese Variante, also der Diversity-Betrieb, ist in DVB-T2 aber nicht integriert.

Neben den genannten Verbesserungen ist auch wieder die Nutzung des effizienten Videocodecs H.264 vorgesehen, sodass die übertragbare Programmanzahl auch hier erheblich gesteigert werden kann. Ein 8-MHZ-Kanal bietet bei einem Guard-Intervall von 1/8 und einem 5/6-LDPC-Code auch ohne MISO bereits eine Datenrate von 19,93 Mbit/s, was eine Steigerung von ca. 50% gegenüber den bei DVB-T üblichen 13,27 Mbit/s entspricht. Mit H.264 kann die Datenrate z. B. für zwei HD- oder für 8 SD-Programme genutzt werden [75].

4.5.4.2 DVB-H

Die technischen Parameter von DVB-T sind, vor allem wegen der Nutzung von OFDM, so gewählt, dass ein mobiler Empfang mit portablen Geräten möglich ist. Neben Geräten für Fahrzeuge sind bereits kleine portable Geräte am Markt verfügbar, die aber das Problem haben, den gesamten DVB-Datenstrom entschlüsseln zu müssen, was mit einigem Energieaufwand und entsprechend kurzen Akkulaufzeiten verbunden ist. Dieses Problem soll, auf der Basis von DVB-T mit DVB-H (für Handheld-Geräte) gelöst werden. Die Geräte haben ein kleines Display, sodass eine CIF-Auflösung mit 352 x 288 Bildpunkten ausreicht. Bei geringem Energiebedarf soll der DVB-Transportstrom genutzt wer-

den, um mit einer Videocodierung nach H.264 und via Internet Protokoll (IP) in einem DVB-T-Kanal bei 13,3 Mbit/s ca. zwanzig Programme unterzubringen.

Ein ähnlicher Ansatz existiert für das sogenannte DXB, nämlich der Verschmelzung von DAB (Digital Audio Broadcasting) und DVB-T, da wegen der sehr effizienten Codierung mit H.264 die DAB-Datenrate auch für Videosignale mit SDTV-Auflösung ausreicht.

Hinter dem Kürzel DVB-SH verbirgt sich schließlich ein Hybridsystem aus DVB-H und DVB-S das den portablen Empfang entweder über den terrestrichen oder den Satellitenkanal ermöglicht und damit auch in Gebieten in denen die Senderdichte von DVB-H nicht ausreicht.

4.5.5 Data Broadcasting

Neben der klassischen Nutzung der TV-Verteilwege für die Fernsehprogrammausstrahlung wird der Einsatz dieser Wege zur Übertragung fernsehfremder Dienste immer interessanter. Ein großes Interesse besteht besonders daran, mit Hilfe der Fernsehverteilwege über Satellit und Kabel (Cable TV, CATV) höhere Datenübertragungsraten für das Internet zu erzielen. Das Problem ist, dass dafür eine bidirektionale Kommunikation erforderlich ist. Die Nutzung der Satellitenverbindung ist sehr aufwendig, geeignet erscheint dagegen das Kabelnetz der Telekom AG und anderer Netzbetreiber. Als Breitbandnetz bietet es hohe Datenübertragungsraten. Da es in Deutschland weitgehend flächendeckend verlegt ist, stellt es eine hochwertige Infrastruktur dar. Das BK-Netz ist direkt als bidirektionales Netz konzipiert, die Rückkanäle sind bisher allerdings praktisch nicht überall nutzbar. In ersten Konzepten zur Nutzung für fernsehfremde Zwecke ist als Upstream-Kanal der Frequenzbereich zwischen 5 und 65 MHz vorgesehen, der Downstream, d. h. die Datenübertragung von der Zentrale zum Endkunden, erfolgt im Bereich von 70 MHz bis 862 MHz (Abb. 4.72), mit verschiedenen Modulationsarten.

Mit der Nutzung des BK-Netzes lassen sich eindrucksvolle Datenraten erzielen, allerdings ist zu bedenken, dass das Kabel ein sogenanntes shared Medium ist. Alle angeschlossenen Teilnehmer müssen sich die Datenrate teilen, was für den Einzelnen bedeutet, dass er vom Nutzungsverhalten seiner Nachbarn abhängig ist. Dieser Umstand stellt gegenüber xDSL einen Nachteil dar, da dort für alle Teilnehmer eine eigene Telefonleitung zur Verfügung steht.

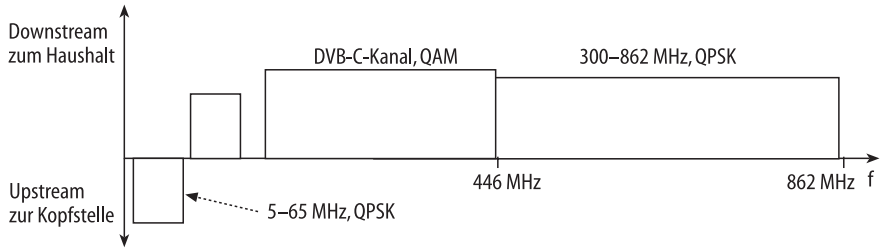


Abb. 4.72. Mögliche Frequenzbelegung im Kabel für bidirektionale Kommunikation

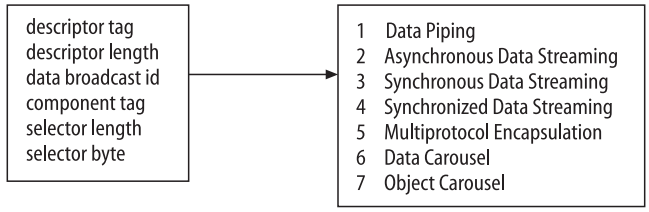


Abb. 4.73. Signalisierung von Datendiensten

Die neuen Verteilwege und die Nutzung der Wege für fernsehfremde Zwekke wurden auch bei den Definitionen für DVB berücksichtigt und werden als DVB-Datenrundfunk (Data Broadcasting Service, DBS) bezeichnet. Grundlage ist auch hier der MPEG-2-Transportstrom, dessen Datencontainer mit der Kapazität 184 Byte aber nicht mit AV-Daten gefüllt werden. Dabei wird die in DVB unter Service Information (SI) definierte Service Description Table (SDT) genutzt, in die für Zwecke des Datenrundfunks ein sogenannter Data Broadcast Descriptor eingefügt wird, mit Hilfe dessen der jeweilige Datendienst signalisiert wird. Der Descriptor beschreibt die Art der Anwendung und verweist über die PMT auf die zugehörigen Elementarströme und die spezifischen Anwendungsfelder.

Die Anwendungsgebiete beziehen sich auf typische Nutzungsprofile, Abbildung 4.73 zeigt einen Überblick. Der einfachste Fall ist das Data Piping, das sich für jede Art der asynchronen Datenübertragung nutzen lässt. Diese Offenheit erfordert aber die Übermittlung der Spezifikations- und Fehlerschutzparameter zwischen den Teilnehmern am Datentransfer. Das zweite Anwendungsfeld nennt sich Data Streaming und wird für eine Datenübertragung genutzt, die zu anderen Elementarströmen synchronisiert ist und an den Systemtakt gekoppelt wird. Das dritte Profil ermöglicht die Datenübertragung unter Verwendung verschiedener Kommunikationsprotokolle, insbesondere auch des Internet-Protokolls (IP) und nennt sich Multiprotocol Encapsulation. Das vierte Anwendungsfeld nutzt das auch bei Videotext verwendete Verfahren der zyklischen Wiederholung von Daten und wird als Data Carousel bezeichnet. Eine Erweiterung stellt hier das Objekt Carousel dar, bei dem ein Nutzer nicht nur auf Daten, sondern auch auf Verzeichnisse zugreifen kann, ähnlich wie es bei einer interaktiven Verbindung zu einem Server der Fall ist.

4.5.6 DVB Endgeräte

Die Nutzung neuer Übertragungsverfahren erfolgt in vielen Fällen zunächst mit Zusatzgeräten, die die neuen Signalformen in die etablierten überführen, sodass das eigentliche TV-Gerät nicht ausgetauscht werden muss. Das Zusatzgerät für DVB heißt Set Top Box (STB), ihre Funktionen sind in moderne TV-Geräte integriert. Die Set Top Box enthält ein Empfangsmodul, mit einem Tuner, dem Demodulator und dem Kanaldecoder, der auch die Fehlererkennung und -korrektur durchführt. Dieses Modul ist das einzige, das speziell auf den

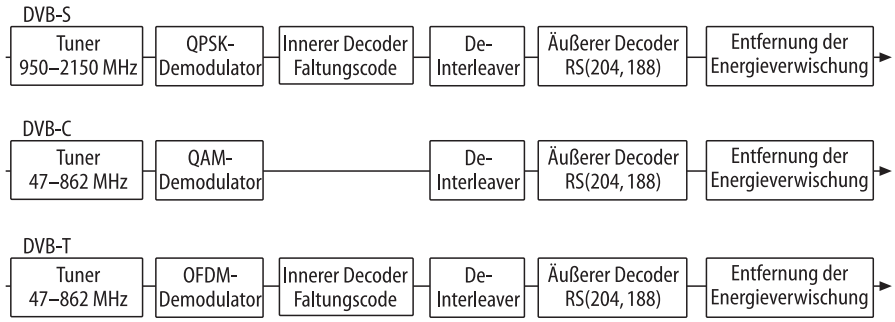


Abb. 4.74. Blockschema eines Empfangsmoduls in einer Set Top Box

Übertragungsweg (DVB-S, -C oder -T) und die dafür verwendete Modulationsart (QPSK, 64-QAM oder COFDM) abgestimmt ist (Abb. 4.74).

Am Ausgang des Decoders steht unabhängig von der Modulationsart der MPEG-Transportstrom zur Verfügung, der zum MPEG-Modul geleitet wird, das die Aufgabe als Demultiplexer und Decoder für Video und Audiodaten übernimmt. Der Transportstrom wird zwischengespeichert und anhand der Packet Identifieres (PID) aufgetrennt, sodass die Teildatenströme zu den richtigen Decodern gelangen. Die Teildatenströme werden mit Hilfe der DTS und PTS-Signale synchronisiert und anschließend wird die Quellencodierung rückgängig gemacht.

Danach liegen die Audio- und Videosignale in der ursprünglichen digitalen Form vor und werden mit D/A-Wandlern in die analoge Form überführt (Abb. 4.75) oder an Digitalschnittstellen wie z. B. HDMI ausgegeben. Das analoge Audiosignal steht meist im Stereo-Format an zwei separaten RCA-Buchsen zur Verfügung oder zusammen mit dem Videosignal an einer SCART-Buchse. Teilweise wird es auch in digitaler Form im Consumer-Audioformat S/PDIF ausgegeben. Das Videosignal liegt in analoger Form meist als FBAS-Signal vor.

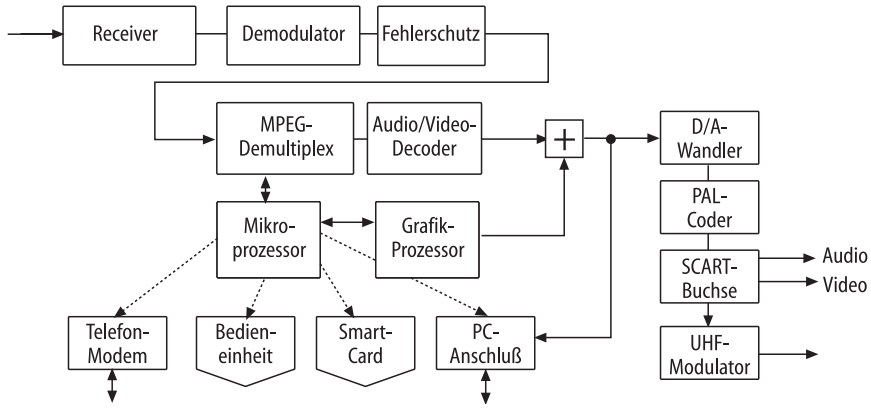


Abb. 4.75. Blockschema einer Set Top Box

Abb. 4.76. Artefakte bei zu schwachem Signal am digitalen Empfänger und HD TV-Logo

Eine Set Top Box enthält einen Mikroprozessor, der alle Funktionen steuert. Er ist wie ein PC mit Speicherbausteinen und Verbindungen zur Außenwelt über serielle oder SCSI-Schnittstellen ausgestattet. Die Betriebs-Software befindet sich oft auf einem Flash-RAM. Aufgrund der vorliegenden Ablage in elektrisch löschbaren Speichern kann die Software entweder über die serielle Schnittstelle oder direkt über das Sendesignal von außen aktualisiert werden. Die SI-Daten, wie die Programm-Übersicht, Zusatzdaten etc. werden während des Betriebs in einem DRAM-Speicher abgelegt, sodass auf sie ohne Zeitverzögerung zugegriffen werden kann.

Bei Signalqualitätsminderung zeigen digitale Empfangsgeräte das typische digitale »Alles oder Nichts«-Verhalten. Wenn sich die Signalqualität gerade an der Entscheidungsschwelle für die MPEG-2-Decodierung befindet, kann es zu Standbild- und Blocking-Artefakten kommen (Abb. 4.76).

Mit dem steigenden Angebot von HD-Programmen kommen auch immer mehr Set Top Boxes für HD-Signale auf den Markt. Da dieser Bereich für die Konsumenten neu und verunsichernd ist, u. a. weil die HD-Videosignale mit 1080 oder 720 Zeilen ausgestrahlt werden können, hat die EICTA (European Information, Communications and Consumer Electronics Industry Technology Association) für HD-Set Top Boxes, wie beim HD-ready-Logo für Displays, ein HD TV-Logo entwickelt, dessen Verwendung bestimmte Mindeststandards garantiert (Abb. 4.76). Geräte mit dem HD TV-Zeichen müssen DVB-Signale der ersten Generation, bei STB für den Satellitenempfang auch DVB-S2 unterstützen. Auf der System-Ebene müssen sie kompatibel zu MPEG-2-Systems sein und bezüglich der Videocodecs müssen sie mit MPEG-4 AVC HP@L4- sowie mit MPEG-2 codierten Daten arbeiten. Auf der Audioseite wird die Unterstützung von Dolby-AC3-, AAC- und MPEG-1-Layer II-codierten Daten gefordert.

Bezüglich der Interfaces müssen Digitalreceiver mit dem HD TV-Logo einen analogen Komponentenausgang, eine DVI- oder eine HDMI-Schnittstelle aufweisen. An den HD-Videoausgängen muss entweder ein 720p- oder ein 1080i-Signal ausgegeben werden, wobei sichergestellt werden muss, dass eines der beiden unskaliert auf einem HD-ready-Display dargestellt werden kann.

4.5.6.1 Conditional Access

Zur Realisierung besonderer Formen der Fernsehversorgung, die weder über Gebühren, noch über Werbung finanziert werden, ist es erforderlich, dass das Programm nur Kunden zugänglich gemacht wird, die für den Empfang bezahlt haben (Pay TV). Bei DVB wurde dafür ein bedingter Programmzugang (Conditional Access) vorgesehen. Dabei werden die Nutzdaten im MPEG-Transportstrom verschlüsselt, während die Header, Tabellen und Synchroninformationen unverschlüsselt bleiben. Ein verschlüsseltes Programm wird anhand zweier Transport Scrambling Control-Bits erkannt.

Die Verschlüsselung beruht auf zwei kaskadierten Verfahren, bei denen zuerst 8-Bit-Datenblöcke systematisch verwürfelt werden und anschließend die verschlüsselten Daten noch einmal. Zur Entschlüsselung werden zwei Control Words verwendet, die Entitlement Control Messages (ECM), die wiederum selbst verschlüsselt sind. Um das System möglichst sicher zu machen, wird der Schlüssel von Zeit zu Zeit gewechselt, der neue muss dazu im Datenstrom übertragen werden, daher sind immer zwei Schlüssel erforderlich. Neben den ECM gibt es EMM (Entitlement Management Messages), mit denen die Freischaltung der Programmbereiche geregelt wird, für die der Kunde auch bezahlt hat. Mit der EMM lässt sich eine sogenannte Smartcard adressieren, die der Kunde erwerben und in das Gerät einführen muss, damit die Berechtigung geprüft werden kann.

Es gibt kein einheitliches System für den Conditional Access. Damit zur Anwahl verschiedener Pay-TV-Anbieter nun nicht verschiedene Set Top Boxen verwendet werden müssen, wurde mit dem Common Interface eine Möglichkeit geschaffen, die Zugangsdaten zu verschiedenen Anbietern gänzlich auf die Smartcard zu verlegen, sodass nur noch diese ausgetauscht werden muss.

4.5.6.2 Betriebssysteme

Für den in der Set Top Box befindlichen Computer werden keine Standardbetriebssysteme verwendet, sondern proprietäre Verfahren, mit denen nur speziell abgestimmte Software lauffähig ist. Die wichtigste Applikationssoftware für das Betriebssystem ist der Electronic Program Guide (EPG), hinzu kommen Teletext-Programme, Steuerungen der Heimrecorder etc. Um nun mit einem einzelnen Empfangsgerät die Software verschiedener Anbieter, insbesondere den EPG, wiedergeben zu können, müssen die Hersteller eine standardisierte Software-Schnittstelle, das Application Programming Interface (API) offenlegen [178]. Wenn dies nicht geschieht, sind zwar die TV-Programme fremder Anbieter mit einer Set Top Box decodierbar, doch der EPG wird nicht wiedergegeben. Gerade der Programmführer hat aber besonders große Bedeutung, da sich beim Einsatz digitaler Übertragung die Programmzahl vervielfacht und die Beiträge zu einem Bouquet zusammengefasst werden, für die eine Navigationshilfe sehr erwünscht ist. Das Bouquet eines Anbieters wird erst durch den EPG erkennbar. Um also einen diskriminierungsfreien Zugang zu den Programmen und EPG aller Anbieter zu gewährleisten, ist ein Betriebssystem mit offenem API erforderlich. Das diesbezüglich erfolgreichste Betriebssystem ist OpenTV, das Applikationen von allen Digital-TV-Anbietern als Free TV oder Pay TV verarbeitet.

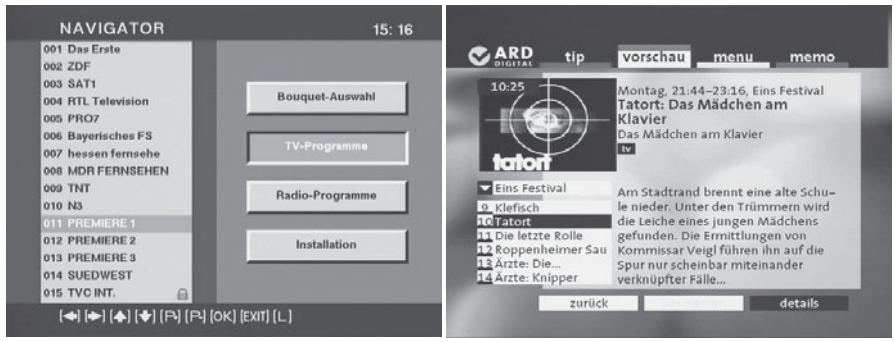


Abb. 4.77. Beispiele für einen Navigator einer Set Top Box und eines EPG [158]

Viele Pay TV-Anbieter setzten jedoch auf geschützte Schnittstellen, sodass OpenTV heute große Bedeutung für den Bereich unverschlüsselter Ausstrahlung, wie z. B. bei den Digitalprogrammen der öffentlich-rechtlichen Anstalten, hat, während die Bereiche im Pay-TV z. T. von proprietären Systemen beherrscht werden, die nicht in der Lage sind, den EPG der öffentlich-rechtlichen Programme darzustellen. Angesichts der beschriebenen Probleme hat sich, nachdem die technische Infrastruktur weitgehend geschaffen ist, das DVB-Entwicklungskonsortium auch den Anwendungen zugewandt, um mit der Multimedia Home Platform (MHP) eine übergreifende, hardware-unabhängige Struktur zu fördern.

Die Initiative ist dem DVB-Commercial Module zugeordnet und hat die Aufgabe, den Schritt zu einem horizontalen Markt zu gehen, in dem sich verschiedenste Dienste Konkurrenz machen können, ohne dass für jeden von ihnen eine separate Hardware erforderlich ist. Danach umfasst die Navigation im Programmangebot zwei Stufen. Nach dem Einschalten des Empfangsgerätes erscheint zunächst eine Übersicht über alle Angebote, die im angeschlossenen Netzwerk verfügbar sind. Dieser Navigationsteil ist herstellerspezifisch, Abbildung 4.77 zeigt links ein Beispiel. Anschließend können die Nutzer ein Programm direkt wählen oder ein Bouquet zusammengehöriger Programme, deren Auswahl vom EPG unterstützt wird. Dieser Navigationsteil ist anbieterspezifisch, d. h. er sieht auf allen Empfangsgeräten gleich aus. Abbildung 4.77 zeigt rechts zum Vergleich den EPG eines Bouquets.

Die Spezifikationen von MHP beruhen auf der Programmiersprache JAVA mit der eine Entkopplung von Hardware und Systemsoftware erreicht wird. Verschiedene Anbieter können sich über die besondere Gestaltung der Benutzeroberflächen profilieren, ohne sich um die Hardware kümmern zu müssen. Bisher etablierte Systeme werden über Plug-in-Schnittstellen unterstützt. Ein zweiter wichtiger Punkt ist die Einbindung der für das Internet verwendete Hypertext Markup Language (HTML). Damit wird ein Internetzugang via DVB unterstützt. Übergreifend kann MHP als Mittler-Software betrachtet werden, die digitale Empfänger in die Lage versetzt, herstellerunabhängig interaktive Dienste zu ermöglichen.

4.5.7 Neue digitale Videodienste

Aufgrund der umfassenden Digitalisierung aller Aspekte der Fernsehübertragung wird eine tiefgreifende Veränderung des Fernsehens bzw. der gesamten audio-visuellen Dienste erwartet. Durch die Digitalisierung werden Audio- und Videosignale prinzipiell ununterscheidbar von Signalen zur Textübertragung. Mit dem durch die nichtlineare Speicherung ermöglichten schnellen Zugriff werden die verschiedenen Medienarten gemeinsam zu dem, was oft mit dem Stichwort »Multimedia« bezeichnet wird.

Eine wichtige qualitative Umstellung ist die Entwicklung der Datenübertragung unter Einbeziehung eines Rückkanals, eine Entwicklung weg von Broadcast, hin zu interaktivem Fernsehen. Durch die Datenreduktionstechniken brauchen statt etwa 160 Mbit/s für ein SD-Videosignal nur 1 bis 5 Mbit/s übertragen zu werden, womit sich die Kanalvielfalt drastisch erhöhen lässt. So wird es kleinen Anbietern möglich sein, spezielle Videodienste, sogenannte Spartenprogramme, anzubieten. Welche Dienst-Typen sich etablieren werden, ist nicht genau zu bestimmen, bisher sind vor allem Konzepte wie das für die firmeninterne Kommunikation konzipierte Business-TV, Pay TV, Video on Demand und Interaktives TV entwickelt worden.

Pay TV-Programme werden den Kunden in Rechnung gestellt und daher verschlüsselt ausgestrahlt, d. h. der Videodatenstrom wird gezielt verwürfelt, und das Programm ist nur für die Zuschauer empfangbar, die ein Entschlüsselungssystem gekauft haben. Man unterscheidet Pay per Channel, bei dem monatlich ein Vollprogramm bezahlt, und Pay per View, bei dem nur für die gewählte Sendung bezahlt wird.

Auch Video on Demand ist eine Form von Pay per view, die Besonderheit besteht darin, dass die Zuschauer aus einem Programmangebot die sie interessierende Sendung zu jeder beliebigen Zeit abrufen können. Eine Vorform ist Video near Demand. Hier wird das gleiche Programm zeitversetzt, z. B. alle 10 Minuten erneut über mehrere Kanäle parallel ausgestrahlt, sodass fast jederzeit ein Einstieg in laufende Beiträge möglich ist.

Bei der Nutzung neuer Fernsehdienste sollen die Zuschauer Einfluss auf das Geschehen auf dem Bildschirm nehmen können. Die einfachste Form ist die Auswahl aus einem Programmangebot. Diese Form der Interaktion erfordert keinen Rückkanal zum Sender. Echte Interaktionsformen erfordern einen Rückkanal, dies kann im einfachsten Fall eine Telefonleitung sein. Die aufwendigste Variante ist der vollwertige Rückkanal, bei dem das Datenmaterial über dieselbe Schnittstelle in beiden Richtungen transportiert wird. Beim interaktiven Fernsehen im engeren Sinne ist vorgesehen, dass die Zuschauer über den Rückkanal z. B. Einfluss auf den Fortgang einer Spielfilmgeschichte nehmen können oder dass interaktive Regie möglich wird, indem ein Zuschauer beispielsweise bei einem Fußballspiel über die Kameraauswahl selbst die Kameraperspektive bestimmt.

Den hier angesprochenen Erweiterungen der Fernsehdienste erwächst zunehmend Konkurrenz durch die Videoverbreitung im Internet. Da hier der Rückkanal bereits vorhanden ist, ist zu erwarten, dass sich interaktive Videoanwendungen und die hier angesprochenen Dienste eher dort entwickeln.

4.6 Internet und IPTV

So wie die Verteilwege für die Fernsehdienste immer interessanter für den Austausch von Computerdaten werden, so werden umgekehrt noch mehr auch die Übertragungswege für Computerdaten für die Videodatenübertragung immer bedeutungsvoller. Dies gilt vor allem für die Datenübertragung auf kurzen Strecken im Produktionsbereich (s. Kap. 9.3), zunehmend aber auch für das Broadcasting, vor allem seit sich mit der xDSL-Technik (s. u.) die Verfügbarkeit von höheren Datenraten bei der Übertragung zum Endkunden gewährleistet wird. Da die Datenverbindungen, die hohe Bandbreiten ermöglichen, inzwischen in sehr vielen Haushalten verfügbar sind, hat sich die Videoübertragung im Internet erheblich verbreitet.

Wie das Internet selbst ist die Videosignalverteilung unter Nutzung des Internetprotokolls wenig standardisiert. Abhängig vom Grad der Struktur und damit vom Grad der Garantierbarkeit eines bestimmten QoS (Quality of Service) können vier Bereiche unterschieden werden:

Im einfachsten Fall werden einfach Videodateien per File Transfer aus dem Netz geladen. Dieser Fall bietet die schlechteste QoS, da der Inhalt erst nach abgeschlossener Datenübertragung betrachtet werden kann und diese also nicht echtzeitfähig ist. Auf der zweiten Stufe steht das Streaming, d. h., dass nach kurzer Wartezeit für die Pufferung der Daten ein Videoprogramm meist ohne Unterbrechung betrachtet werden kann. Die QoS ist hier dadurch beschränkt, dass bei großer Netzbelastung die Darstellung stoppt, weil Daten nachgeladen werden müssen. Diese beiden Verfahren sind meist mit beschränkter Bildauflösung sowie begrenzter Länge der Filme verbunden und damit recht deutlich vom gewöhnlichen Fernsehen unterschieden. Das gilt nicht zuletzt auch deshalb, weil davon ausgegangen wird, dass diese Programme eher am PC (lean forward) als auf einem Großdisplay in weiter Entfernung (lean backward) betrachtet werden. Zudem ist Übertragung mit einer Vielzahl verschiedener Codecs verbunden und nicht, wie bei TV-Anwendungen auf MPEG-2 oder H.264-AVC beschränkt. Trotz der geringen technischen Qualität hat das Streaming der kurzen Filme viel Interesse vor allem bei jungen Menschen gefunden, insbesondere seit Portale wie Youtube es erlauben, eigene Videofilme auf die Server hoch zu laden und im Netz zur Verfügung zu stellen. Dieser User Generated Content (UGC) ist eine wesentliche Neuerung und bietet zudem den Vorteil, dass keine Produktions- oder Lizenzkosten anfallen, sodass die Filme auch kostenfrei betrachtet werden können [237].

Die dritte und vierte Stufe kommen der klassischen TV-Übertragung näher. Sie verwenden typischerweise MPEG-Transportströme und MPEG-2- bzw. H.264-Codecs und erlauben die Übertragung langer Programme. Um hier dem Bedarf nach Video-Streaming Rechnung zu tragen, werden besondere Protokolle entwickelt, die auf dem Internet-Protokoll (IP) aufsetzen. Die dritte Stufe kann als Internet-TV und die vierte als IPTV bezeichnet werden. In der vierten Stufe kann die beste Service-Qualität garantiert werden, denn der Begriff IPTV sollte Programmen vorbehalten sein, die von Netzbetreibern angeboten werden, die die letzte Datenübertragungsstrecke bis zum Endkunden (last mile) kontrollieren und damit die technische Qualität sicherstellen können [205].

4.6.1 Internet-Grundlagen

Das Internet ist das weltweit größte Computernetz. Es zeichnet sich dadurch aus, dass es ein Netzwerk ohne Zentralcomputer ist. Die Kommunikation geschieht über verteilte, gleichberechtigte Teilnetze, in die einzelne Server eingebunden sind. Die Teilnehmer müssen bei dem Betreiber eines Servers (Provider) angemeldet sein, der in das Netz eingebunden ist und darüber den Zugang bietet. Ein großer Provider ist z. B. die Deutsche Telekom AG.

Ein sehr bedeutender Teil in diesem Rechnerverbund ist das World Wide Web (WWW), das sich dadurch auszeichnet, dass die WWW-Dokumente Verweise zu anderen Dokumente (Pages) im Netz enthalten, die durch einen einfachen Mausklick erreicht werden können. Die angemeldeten Nutzer stellen über die Provider ihre Inhalte als sogenannte Websites zur Verfügung und erlauben damit allen anderen Nutzern den Zugriff. Dabei wird die Sprache HTML (Hypertext Markup Language) und das Hypertext Transfer Protokoll eingesetzt, die den Verweis zu anderen Seiten ermöglichen. Der Zugriff geschieht über sogenannte Webbrowser mittels grafischer Benutzeroberflächen auf sehr einfache Weise. Aus diesem Grund hat die Internet-Nutzung, über die große Verbreitung in allen Bevölkerungsschichten, seine enorme Bedeutung erhalten.

Jede Website hat ihre eigene Adresse, die als Uniform Ressource Locator (URL) bezeichnet wird. Eine typische Adresse lautet z. B.: http://www.Meinname.de. Darin steht http für das Hypertext Transfer Protokoll, das für die Übertragung gewöhnlicher Web-Dokumente benutzt wird. Alternativen sind an dieser Stelle das File Transfer Protokoll (ftp) oder das Simple Mail Transfer Protokoll (smtp). Nach den zwei Schrägstrichen folgt die Angabe der eigentlichen Adresse. In dem angegebenen Beispiel kennzeichnet www einen Webserver. Darauf folgen, durch Punkte getrennt, der originäre Name und eine Länderkennung oder ein US-Organisationscode, wie (.de) für Deutschland oder (.com) für den kommerziellen Bereich. Weitere, durch Schrägstriche getrennte Bezeichnungen verweisen auf den Pfad zum Dokument auf diesem Server.

4.6.2 Netzwerke

Die Internet-Daten werden meist über Zugangsnetze wie TV-Kabel oder digital genutzte Telefonleitungen und z. T. auch drahtlos zum Endkunden übertragen, wobei eine bidirektionale Verbindung erforderlich ist. Diese ist aber nicht zwangsläufig symmetrisch, in vielen Fällen reicht es aus, für die Übertragung Downstream, also von der Zentrale zum Kunden, eine höhere Datenübertragungskapazität bereitzustellen als Upstream, vom Nutzer zur Zentrale.

Mit ADSL (Asymmetrical Digital Subscriber Line) wurde ein Verfahren entwickelt, dass die Bezeichnung asymmetrical im oben genannten Sinne enthält und das so leistungsfähig ist, dass es auch für die Übertragung von bitratenintensiven Videodatenströmen geeignet ist. Dabei ist die Grundidee, die Zweidrahttelefonleitung nicht nur für Telefongespräche mit einer Frequenzbandbreite bis 3,4 kHz zu nutzen, sondern mit Grenzfrequenzen, die weit darüber liegen [176]. Da die Leitung als relativ ungestörter Kanal betrachtet

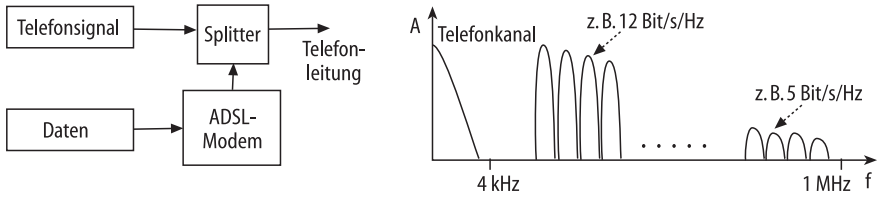


Abb. 4.78. ADSL-Prinzip

werden kann, sind höherwertige Modulationsverfahren einsetzbar, sodass erheblich höhere Datenraten möglich werden als bei ISDN, das nur für extrem datenreduzierte Videoströme einsetzbar ist.

Bei ADSL werden Frequenzen unterhalb 4 kHz für den üblichen Telefonverkehr genutzt. Mit Hilfe eines Splitters wird dieser Bereich vom darüberliegenden Frequenzband getrennt, das mit Hilfe eines Modems mit Daten belegt wird. Zur Datenübertragung wird dabei das Discrete Multitone-Verfahren (DMT) eingesetzt. Die Nutzbandbreite hängt von der Entfernung zur nächsten Vermittlungsstelle (Last Mile) ab. In vielen Fällen kann mindestens eine Bandbreite von 1 MHz genutzt werden. Das DMT-Verfahren belegt nun diesen Bereich mit z. B. 256 einzelnen Unterträgern, die alle unterschiedlich moduliert werden können, sodass z. B. im unteren Teil des Bandes eine Modulationsart mit einer Bandbreiteneffizienz von 12 bit/s/Hz und im oberen Teil eine mit nur 5 bit/s/Hz verwendet werden kann (Abb. 4.78). Mit ADSL werden Downstream-Datenraten bis zu 8 Mbit/s und Upstream-Raten bis 1 Mbit/s bei einer Leitungslänge bis 2 km erreicht. Auf kürzeren Leitungen können diese Werte bei VDSL und SDSL gesteigert werden. Die verschiedenen Varianten werden auch unter der Bezeichnung xDSL zusammengefasst. xDSL ist insgesamt ein Verfahren, das hauptsächlich für Internet-Anwendungen entwickelt wurde. Wenn die hohen Datenraten auch hinter der Vermittlungsstelle, im sogenannten Backbone aufrecht erhalten werden können, so wird eine weiterentwickelte xDSL-Technologie auch per Glasfaserverbindung realisiert werden.

4.6.3 Internet-Protokolle

Grundlage für den gesamten Datenaustausch im Internet ist die Adressierung der Teilnehmer mit dem Internet-Protokoll (IP). Der IP-Kopf umfasst bei Protokollversion 4 (IPv4) mindestens 5 Datenworte von je 32 Bit Umfang (Abb. 4.79). Im ersten Wort befindet sich die Versionsnummer, eine Prioritätenangabe (Type of Service), und die Angabe der Paketlänge. Falls die Länge den Adressraum von 16 Bit überschreitet, werden die Daten in Fragmente aufgeteilt, wobei die Informationen im nächsten Datenwort erforderlich sind. Im dritten Wort gibt TTL (Time to Live) an, wie lange ein Datenpaket im Netz verbleiben darf, dann folgt die Anzeige des verwendeten Transportprotokolls über einen Zahlencode (z. B. No. 6 für das Transmission Control Protokol, TCP oder No. 17 für das User Datagram Protocol, UDP). In den Datenworten 4 und 5 befinden sich schließlich die Quell- und Zieladressen mit jeweils 32 Bit Umfang.

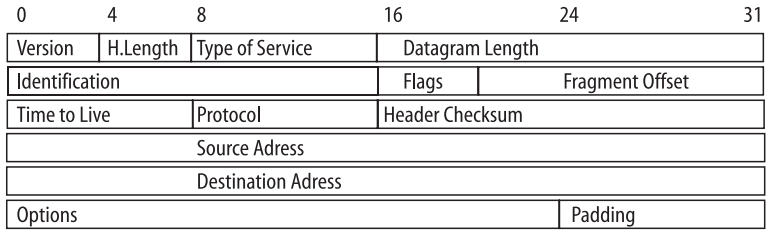


Abb. 4.79. IP-Protokollkopf

Als Übertragungsprotokoll kommt für gewöhnliche Aufgaben das erwähnte TCP zum Einsatz, mit dem eine duplexfähige Ende-zu-Ende-Verbindung aufgebaut wird. TCP sichert die Übertragung durch erneute Anforderung von ggf. verloren gegangenen Datenpaketen. Diese Funktion ist hilfreich für den üblichen Transfer abgeschlossener Datenpakete, aber hinderlich für die broadcast-ähnliche Streaming-Technik. Abbildung 4.80 zeigt, wie das TC-Protokoll auf IP aufgesetzt ist und es selbst wieder Träger der gängigen File Transfer- und Hypertext Transfer-Protokolle ist.

Für Video-Streaming via IP wird statt TCP meist das einfachere User Datagram Protocol (UDP) verwendet, das die Datenpakete ohne größere Kontrolle einfach weiterleitet [193]. Darauf aufbauend finden sich die für Echtzeitverbindungen entwickelten Real Time Protokolle (RTP), die eine Synchronisation verschiedener Teildatenströme (z. B. Audio und Video) mittels Zeitmarken unterstützen [197]. Die beteiligten Datenströme werden separat übertragen und am Empfänger mit Hilfe des Real Time Control Protokolls (RTCP) synchron zusammengeführt. Dabei wird die Bestimmung der aktuellen Datenübertragungsqualität und eine Reaktion darauf, d. h. eine dynamische Änderung der Datenrate unterstützt. Für übergeordnete Funktionen, wie den sprunghaften Übergang zu bestimmten Positionen im Datenstrom, steht zusätzlich das Real Time Streaming Protocol (RTSP) zur Verfügung, das auf RTCP aufsetzt.

Die klassische Internet-Verbindung ist eine Punkt zu Punkt Verbindung. Um broadcast-ähnliche Strukturen zu entwickeln, müssen geeignete Mechanismen zur Vervielfältigung eines einzelnen Datenstroms zur Verfügung stehen, also ein Übergang von der Unicast- zur Multicast-Technik geschaffen werden. Zu diesem Zweck wurde IP-Multicast entwickelt. Dabei werden einzelne Empfänger zu Gruppen zusammengefasst und ein Router sorgt dafür, dass alle Gruppenmitglieder eine Kopie des Datenstroms erhalten. Diese Verfahren verwaltet die Gruppenbildung inklusive der An- und Abmeldung der Teilnehmer über das Internet Group Management Protocol (IGMP).

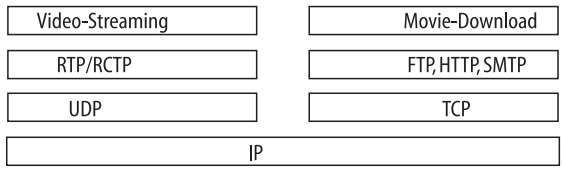


Abb. 4.80. Protokolle für Streaming und Datentransfer via IP

5 Filmtechnik

5.1 Film als Speichermedium

Ebenso wie das Fernsehen stellt Film stellt ein Medium dar, mit dem Bewegungsvorgänge wiedergegeben werden können (Motion Picture Film). Die Bewegung wird zeitlich diskretisiert und in einzelne Phasen zerlegt, die jeweils in einem Einzelbild festgehalten werden. Bei einer Präsentation von mehr als 20 Bildern pro Sekunde kann der Mensch die Einzelbilder nicht mehr trennen, und es erscheint ihm ein Bewegtbild. Der wesentliche Gegensatz zum Fernsehen ist der Umstand, dass der Film auch ein Speichermedium darstellt, d. h. zu jedem Zeitpunkt steht das ganze Bild bzw. die Bildfolge zur Verfügung. Aufgrund dieses Umstands hatte der Film vor der Einführung elektronischer Speichermedien auch eine große Bedeutung im Fernsehbereich.

Das Einzelbild entsteht über den fotografischen Prozess. Die Fotografie nutzt einen Effekt, bei dem sich Silberverbindungen unter Lichteinwirkung so verändern, dass in Abhängigkeit von der örtlich veränderlichen Intensität unterschiedliche Schwärzungen auftreten. Das lichtempfindliche Material wird auf ein transparentes Trägermaterial aufgebracht, das für Filmanwendung sowohl geschmeidig als auch sehr reißfest sein muss und zudem über lange Zeit formbeständig bleibt. Diese Forderungen werden sehr gut von Zellstoffmaterialien erfüllt, die mit einem Weichmacher behandelt werden. Bis in die 50er Jahre des 20. Jahrhunderts hinein wurde Zellulose-Nitrat, der Nitrofilm, als Trägermaterial verwendet. Dieser erfüllte die Anforderungen, hat aber die Eigenschaft, leicht entflammbar zu sein, was häufig zu sehr schweren Unfällen führte. Anschließend wurde der sogenannte Sicherheitsfilm aus Zellulose-Triazetat verwendet, oder der Träger besteht aus Polyesterkunststoff. Polyester ist formstabiler und reißfester als das Zellulose-Material, lässt sich aber nicht mit gewöhnlichen Mitteln kleben, sodass Polyester (bei Kodak Estar genannt) gut bei Endprodukten verwendet werden kann, die nicht mehr bearbeitet werden.

Auf das Trägermaterial von etwa 0,15 mm Stärke wird die lichtempfindliche Schicht aufgetragen. Die Schicht hat eine Dicke von ca. 7 µm, die bei neueren Filmen eine Toleranz von maximal 5% aufweist. Darüber wird eine dünne Schutzschicht aufgebracht, die Beschädigungen der Oberfläche verhindern soll. Auf der Filmrückseite befindet sich ebenfalls eine Schutzschicht. Sie ist bei Aufnahmefilmen grau eingefärbt, um zu verhindern, dass Licht von der Filmrückseite wieder zur lichtempfindlichen Schicht reflektiert wird und dort sogenannte Lichthöfe bildet. Als Hersteller von Filmmaterialien hat seit langer Zeit die Firma Eastman/Kodak eine sehr große Bedeutung.

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4_5, © Springer-Verlag Berlin Heidelberg 2013

5.1.1 Die Filmschwärzung

Zunächst sei hier der Schwarz/Weiß-Film betrachtet. Die lichtempfindliche Schicht besteht aus einer Emulsion aus Gelatine in die als Lichtrezeptoren Silbersalze, meist Silberbromid, eingemischt sind, also eine molekulare Verbindung von Ag^+ und Br^-. Das Silberbromid liegt in kristalliner Form vor und weist eine eigene Gitterstruktur auf. Unter Einwirkung von Licht kann sich ein Elektron vom Bromion lösen und ein Silberion neutralisieren. Damit entsteht undurchsichtiges metallisches Silber, das das Kristallgefüge an dieser Stelle stört. Bei sehr langer Belichtung geht schließlich das gesamte Silberbromid in seine Bestandteile über und macht das Material undurchsichtig. Bei kurzer Belichtung entsteht das Silber in so geringen Mengen, dass kein sichtbares, sondern nur ein latentes Bild entsteht. Das Material kann aber mit Hilfe von Substanzen auf Benzolbasis anschließend einer chemischen Behandlung, der Entwicklung, unterzogen werden, wobei die geringe Kristallstörung des latenten Bildes so verstärkt wird, dass der gesamte Kristall zu Silber und Brom zerfällt und eine neue Gitterstruktur aufweist. Durch den Entwicklungsprozess wird die Wirkung der Belichtung um einen Faktor zwischen 10^6 und 10^9 verstärkt, was die heute verwendbaren geringen Belichtungszeiten ermöglicht. Dabei bildet sich in hellen Bildpartien schneller Silber als in dunklen, d. h. diese Bereiche werden weniger transparent, und es entsteht ein negativer Bildeindruck. Der Grad der Silberbildung bzw. Schwärzung ist vom Grad der Beleuchtungsstärke abhängig und weiterhin durch die Art der Entwicklung beeinflussbar. Unbelichtete Stellen bleiben nicht völlig transparent, auch hier bildet sich ein wenig Silber. Dieses mindert den Kontrast und wird als Schleier bezeichnet.

Durch die Entwicklung allein entsteht noch kein dauerhaftes Bild, denn das Silberbromid, das noch nicht zerfallen ist, ist weiter lichtempfindlich, sodass Lichteinfall zu weiterer Schwärzung führt. Vor dem und während des Entwicklungsprozesses darf das Filmmaterial also nicht dem Licht ausgesetzt werden, da sonst der gesamte Film geschwärzt wird. Kritisch ist dabei vor allem energiereiche kurzwellige elektromagnetische Strahlung, die dem Auge blau erscheint, rotes Licht ist im Arbeitsprozess akzeptabel.

Um die Filme lichtecht zu machen wird durch einen sogenannten Fixiervorgang in einer Thiosulfatlösung das überschüssige Silberbromid abgelöst und durch die folgende Wässerung herausgewaschen. Anschließend wird der Film getrocknet, was einen großen Teil der Gesamtbearbeitungsdauer in Anspruch nimmt.

Zur Erzeugung eines Positivs gibt es zwei Möglichkeiten: das Negativ/Positiv- und das Umkehrverfahren. Bei Ersterem wird zum zweiten Mal negiert, indem das erste Negativ mit Hilfe gleichmäßiger Beleuchtung auf einen zweiten Film kopiert wird. Nachdem auch dieser der beschriebenen Entwicklung unterzogen wurde, entsteht schließlich das Positivbild, das im Idealfall die gleiche Leuchtdichteverteilung wie die Originalszene hervorruft.

Beim Umkehrverfahren wird kein zweiter Film benötigt. Hier wird zunächst auch das Negativ entwickelt, statt aber anschließend das unbelichtete Silberbromid zu beseitigen, wird in einem chemischen Bleichprozess das metallische Silber entfernt und das Silberbromid bleibt zurück. Anschließend wird der

Film diffusem Licht ausgesetzt, sodass nach einer zweiten Entwicklung und anschließender Fixierung die ursprünglich dunklen Bildpartien geschwärzt erscheinen [229].

Aufgrund der Trennung in zwei Schritte erlaubt der Positiv/Negativ-Prozess mehr Spielraum bei der Belichtung als das Umkehrverfahren, außerdem ist er gut geeignet, wenn von einem Negativ mehrere Positive kopiert werden sollen. Der Vorteil des Umkehrfilms ist die Zeitersparnis, da der aufwendige Kopierprozess entfällt. Dieser Vorteil kommt z. B. zum Tragen, wenn Filmmaterial für aktuelle Fernsehberichterstattung verwendet wird, was im Zeitalter der elektronischen Berichterstattung jedoch nicht mehr der Fall ist. Heute wird im Fernsehbereich Negativfilm nur noch sehr selten für szenische Produktionen verwendet. Bei der Umsetzung in ein elektronisches Signal mittels Filmabtastung kann das Positiv einfach durch Signalinvertierung gewonnen werden.

Die Empfindlichkeit des Filmmaterials wird neben der Anzahl der in der Emulsion befindlichen Silberbromid-Kristalle wesentlich von deren Größe bestimmt. Große Kristalle führen zu hoher Empfindlichkeit, denn sie fangen zum einen mehr Licht auf als kleine und bilden zum anderen anschließend auch mehr Silber. Die Silberbildung geht mit großen Kristallen schneller, im Englischen sagt man, der Film habe mehr speed. Die Kristallgröße kann bei der Herstellung der Emulsion beeinflusst werden, und damit können Filme verschiedener Empfindlichkeit produziert werden. Je größer die Kristalle werden, desto stärker werden sie als sogenanntes Filmkorn wahrnehmbar. Auch wenn das Filmkorn (Grain) oft im Einzelnen nicht sichtbar ist, sind doch die Körner statistisch unregelmäßig verteilt und führen so zu einer örtlich veränderlichen Dichte, die besonders bei Grautönen als unregelmäßige, rauschartige Überlagerung des eigentlichen Bildes sichtbar wird (Abb. 5.2) und ein wesentliches Charakteristikum des sogenannten Filmlook darstellt. Das Filmnegativ hat hier den größten Einfluss, da es meist relativ empfindlich ist. Das Kopiermaterial darf dagegen feinkörnig sein, da zum Ausgleich der geringeren Empfindlichkeit mit intensivem Kopierlicht gearbeitet werden kann.

Die Filmempfindlichkeit wird außerdem von der Energie der Lichtwelle bestimmt. Da kurzwellige Strahlung energiereicher ist, liegt vor allem Blauempfindlichkeit vor. Wie bereits angedeutet, kann die Emulsion aber mit Farbstoffen verändert werden, sodass sie auch für andere Wellenlängenbereiche sensibel wird. Dieser Umstand ist für die Entwicklung des Farbfilms von großer Bedeutung (Abb. 5.1).

Abb. 5.1. Querschnitt durchNegativ-Filmmaterial mit drei farbsensitiven Schichten [113]

5.1.2 Farbfilm

Farbfilme erfordern lichtempfindliche Schichten, die nur auf bestimmte Wellenlängenbereiche ansprechen. Aus der in Kapitel 2 dargestellten Theorie der Farbmischung ist bekannt, dass sich Farben aus nur drei Anteilen ermischen lassen, die in ihrer Intensität variiert werden. Ein großer Bereich natürlicher Farben wird erfasst, wenn die Grundfarben Rot, Grün und Blau additiv gemischt werden. Subtraktive Mischung liegt dagegen vor, wenn Farbanteile aus weißem Licht herausgefiltert werden. Die dazu gehörigen Grundfarben sind dann die Komplementärfarben zu Rot, Grün, Blau, also Blaugrün (Cyan), Purpur (Magenta) und Gelb (Yellow).

Farbfilme sind so aufgebaut, dass drei voneinander getrennte Emulsionen übereinander liegen (Abb. 5.1). Die Emulsionen werden so sensibilisiert, dass sie jeweils für einen der drei genannten Anteile des sichtbaren Lichtspektrums, also Rot, Grün und Blau, empfindlich werden. Im Negativmaterial ist die oberste Schicht blauempfindlich, darunter folgen die grün- und rotempfindlichen Schichten, die von Ersterer durch eine Gelbfilterschicht getrennt sind, die blaues Licht von ihnen fernhält.

Das Problem beim Farbfilm ist, dass die Silberbildung zur Schwärzung und nicht zur Färbung führt. Um Farbstoffe bilden zu können, werden Farbkuppler in die Emulsionsschicht eingebaut. Damit entstehen bei der Belichtung wie beim Schwarzweißfilm latente Bilder in den Schichten, die für die jeweilige Farbe empfindlich sind. Die Farbstoffe bilden sich erst im anschließenden Farbentwicklungsprozess. Bei der Umwandlung des Silbers aus dem Silberbromid wird der Farbentwickler oxidiert. Dieser kann eine Verbindung mit den Farbkupplern eingehen und es bilden sich Farbstoffe, die als Farbstoffwolken die Silberkörner einhüllen. Die Farbstoffe sind komplementär zu den ursprünglichen Lichtfarben, für die der Film empfindlich war, und ihre Intensität hängt von der Belichtung ab. Je mehr rotes Licht beispielsweise vorhanden ist, umso mehr blaugrüner Farbstoff entsteht. Bei weißem Licht bilden sich Farbstoffe in allen drei Schichten und mindern die Transparenz über den gesamten Spektralbereich. Das Farbnegativ beinhaltet somit ein helligkeitsinvertiertes Bild in Komplementärfarben. Die Silbersalze, die von den Farbstoffen umhüllt werden, sind jetzt störend und werden in einem Bleichbad entfernt und nachdem im Fixiervorgang auch das unbelichtete Silberbromid entfernt ist, liegt das lichtechte Farbnegativ vor. Abbildung 5.2 zeigt links den Farbfilm nach der Farb–

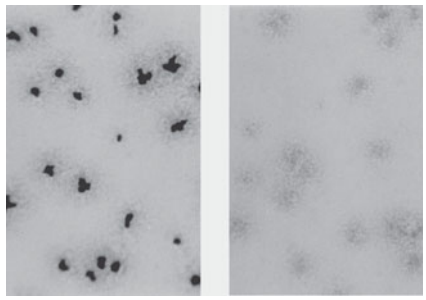


Abb. 5.2. Farbstoffwolken in der Filmemulsion vor (links) und nach (rechts) dem Bleichvorgang [113]

entwicklung allein und rechts mit zusätzlicher Behandlung im Bleich- und Fixierbad. Links sind die von den Farbstoffwolken umhüllten Silberkörner zu sehen, rechts die Farbstoffwolken allein [128].

Der Positivfilm verhält sich ähnlich wie der Negativfilm, auch er speichert wiederum die jeweiligen Komplementärfarben. Bei einem Kopiervorgang mit weißem Licht entsteht bei diesem Farb-Negativ-Positiv-Prozess im Positivfilm also ein Farbbild, das nach der zweiten Invertierung wieder weitgehend der Originalabbildung entspricht. Als Beispiel für einen Negativ-Positiv-Prozess sei eine Szene betrachtet, die ein rotes Objekt enthält: Das vom Objekt reflektierte Licht erzeugt im Film einen cyanfarbenen Farbstoff. Beim Kopierprozess durchdringt das Licht das entwickelte Negativ, wobei die Rotanteile herausgefiltert und im Positivfilm nur die Farbschichten angeregt werden, die für Blau und Grün empfindlich sind. Bei der Farbentwicklung werden hier nun wiederum die Farbstoffe Gelb und Magenta erzeugt. Diese filtern schließlich das Projektionslicht so, dass auf der Leinwand Rot zu sehen ist, da das für Rot wirkende Cyanfilter als einziges im Positiv nicht vorhanden ist.

Auch beim Farbfilm kann mit dem Umkehrverfahren gearbeitet werden. Abgesehen vom Bleichvorgang zur Entfernung des Silbers, entspricht das Verfahren dem Schwarzweiß-Umkehrprozess.

5.2 Filmformate

Filmformate beinhalten Angaben über die Filmbreiten, die Bildfeldgrößen, die Perforationen und die Orientierung. Die Bezeichnungen werden über die äußeren Filmbreiten festgelegt. Heute werden vornehmlich 16-mm- und 35-mm-Filme verwendet, seltener 8-mm-, und 65-mm- bzw. 70-mm-Filme. Die Breiten und Bildfeldgrößen haben sich historisch früh etabliert und sind anschließend nicht wesentlich verändert worden, sodass Film heute den großen Vorteil hat, ein international austauschbares Medium zu sein, eine Eigenschaft, die in Zeiten von Multimedia und ständig wechselnden Daten- und Fileformaten eine wichtige Besonderheit darstellt. Aufgrund der Tatsache, dass die Bildgrößen konstant blieben, die Auflösungs- und Farbqualität im Laufe der Zeit aber immer weiter gesteigert werden konnten, stellt Film hinsichtlich dieser Parameter ein hervorragendes Medium dar, dessen Eigenschaften auf elektronischer Basis erst mit sehr hoch entwickelter Digitaltechnik erreicht werden.

Bezüglich der verwendeten Filmbreiten gibt es eine Entwicklung zu immer schmaleren Filmen. Das größte heute verfügbare Format ist der 70-mm-Film. Dieser wird aber nicht für den Massengebrauch, sondern meist für Spezialfälle mit sehr intensivem Erlebniswert (Imax, Futoroskop, Expo) und in Museen verwendet. 70-mm-Film ist ein reines Wiedergabeformat. Um 70-mm-Kopien zu erzeugen, wird aufnahmeseitig mit 65-mm-Film gearbeitet (Abb. 5.3). Die Differenz wird auf dem Wiedergabematerial als Raum für Tonspuren genutzt.

Die Halbierung jedes 70-mm-Streifens führt zum 35-mm-Filmformat, das gegenwärtig für hochwertige Produktionen die größte Bedeutung hat und das Standardformat für die Kinoprojektion ist. Im Verlauf der Entwicklung sollte

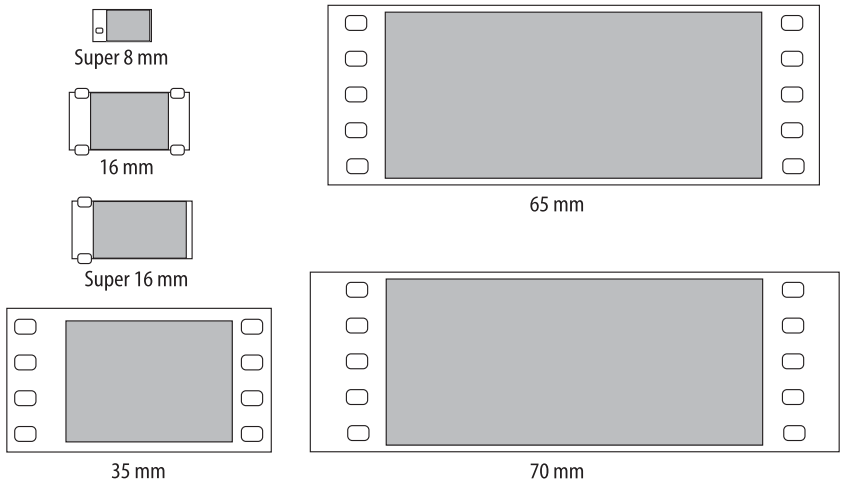


Abb. 5.3. Filmformate im Maßstab 1:1

dann ein Format für den Amateurbereich folgen. Für diesen Zweck wurde die Breite des Filmstreifens auf 5/11 des 35 mm-Formats verringert und es entstand mit 5/8“ der 16 mm-Film. Die Verbesserung der lichtempfindlichen Schichten erlaubte später die nochmalige Halbierung zum 8 mm-Film, der sich in den 70er und 80er Jahren tatsächlich als Amateurformat etablierte, während der 16 mm-Film so hohe Qualität bietet, dass er sich als ein relativ kostengünstiges Medium für Schulungs- und Industriefilme und für Fernsehproduktionen anbot. Aus Kostengründen wird 16-mm-Negativfilm auch verwendet, um durch Vergrößerung eine sogenannte Blow-up-Kopie für die Wiedergabe auf 35 mm zu erzeugen.

Filmmaterial jeglicher Breite ist perforiert. Der Film wird schrittweise transportiert und muss bei der Belichtung oder Projektion sicher in einer Position verharren. Deshalb wird der mechanische Transport mit Schrittschaltwerken durchgeführt, bei denen Greifer in Perforationen im Film einfallen und ihn weiterziehen. Um einen stabilen Bildstand gewährleisten zu können, war bereits bei den Anfängen der Kinotechnik eine Perforation mit vier Löchern pro Bild vorgesehen, die 1891 zum Patent angemeldet wurde und noch heute beim 35 mm-Film verwendet wird. Zuerst waren die Perforationslöcher rund, später rechteckig mit 2,8 mm Lochbreite und 2 mm Höhe für den 35-mm-Film [128]. Bei 16-mm-Film betragen die Breiten/Höhen-Maße 1,8 mm und 1,3 mm.

Die Perforation von 70-, 65- und 35 mm-Film erfolgt beidseitig, der Abstand der Löcher beträgt 4,74 mm. Er ist mit 4,75 mm aber beim Positivmaterial etwas größer als beim Negativ, um eine Kontaktkopierung an einer Transportrolle zu ermöglichen (s. Kap. 5.4.2). Im Gegensatz zum 35-mm-Film hat das 70-mm- (65-mm)-Format 5 Perforationslöcher pro Bild (Abb. 5.3). Bei 16-mm- und 8-mm-Film wird nur ein Perforationsloch pro Bild verwendet, das ein- oder zweiseitig eingestanzt sein kann. Der Lochabstand beträgt beim 16-mm-Format 7,62 mm und bei 8-mm-Film 3,6 mm bzw. 4,23 mm (Super 8).

5.2.1 Bildfeldgrößen

Die Kenntnis der Bildfeldgröße ist von großer Bedeutung, denn sie hat erheblichen Einfluss auf die verwendete Brennweite des Objektivs und damit auch auf wichtige bildgestalterische Parameter wie die Schärfentiefe. Die Größe des Bildfeldes folgt aus der Filmbreite, der Orientierung, der Breite der Perforationen und dem Bildseitenverhältnis sowie ggf. der Berücksichtigung von Platz für Tonaufzeichnung. Beim 35 mm-Film steht zwischen den Perforationslöchern eine Breite von 25,4 mm zur Verfügung. Zu Zeiten des Stummfilms wurde ein Bildfeld von 24 mm Breite und 18 mm Höhe genutzt. Der Bildfeldabstand beträgt bis heute 19 mm, der alten amerikanischen Längeneinheit 1 foot sind damit beim 35 mm-Film genau 16 Bilder zugeordnet. Bei Einführung des Tonfilms wurde die Bildbreite eingeschränkt, um Platz für die Tonspuren zu schaffen. Die Bildfeldgröße des 35 mm-Normalfilmformats beträgt seit dieser Zeit 22 mm x 16 mm. Es besteht damit ein Bildseitenverhältnis 1,37:1, das als Academy-Format bezeichnet wird. Obwohl aufnahmeseitig kein Platz für Tonspuren zur Verfügung stehen muss, wurde die Bildfeldeinschränkung auch hier vorgenommen, damit der Kopierprozess in einfacher Form ohne Größenveränderung ablaufen kann. Später wurde für die Aufnahme auch das Format Super 35 definiert, das mit 24,9 mm x 18,7 mm und dem Seitenverhältnis 1,33:1 die maximale Fläche ausnutzt. In Abbildung 5.4 wird deutlich, dass bei Super 35 nicht nur die Fläche vergrößert, sondern auch das Bildfeld seitlich verschoben ist. Dieser Umstand muss bei Kameras berücksichtigt werden, die sowohl für 35 als auch für Super 35 nutzbar sind. Die Darstellung in Abbildung 5.4 bezieht sich auf das 35 mm-Negativ, die über das Positiv projizierte Fläche ist in beiden Dimensionen immer um ca. 5% kleiner und beträgt beim Normalformat z. B. 20,9 mm x 15,2 mm.

Im Laufe der Zeit wurden die gebräuchlichen Bildseitenverhältnisse mehrfach verändert. Dies geschah sehr radikal, als dem Kinofilm in den 50er Jahren durch die stärkere Verbreitung des Fernsehens in den USA eine ernste

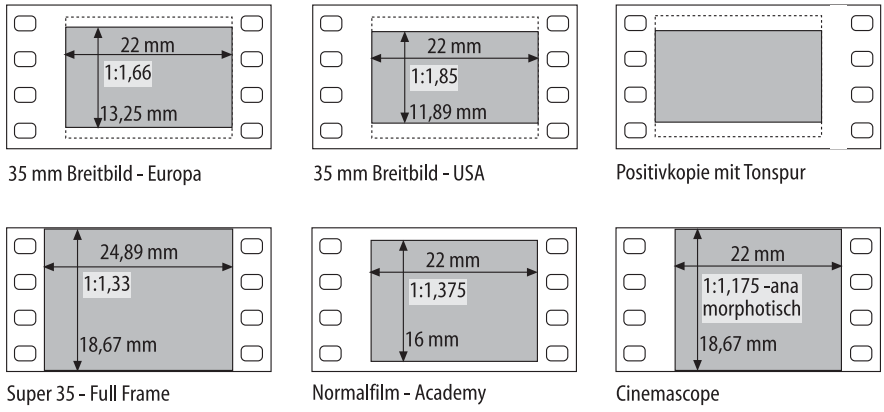


Abb. 5.4. Bildfeldgrößen beim 35 mm-Film

Gefahr erwuchs. Man begegnete ihr durch die Einführung eines sehr breiten Projektionsbildes, dessen Wirkung durch einen hochwertigen Mehrkanalton unterstützt wurde. Das bekannteste ist das Cinemascope-Verfahren, das 1953 eingeführt wurde.

Beim Cinemascope-Format beträgt die Bildhöhe 18,67 mm bei einer Breite von 22 mm (Abb. 5.4). Die Besonderheit ist, dass bei der Wiedergabe ein Breitbildformat vorliegt, das mit einem Verhältnis von 2,39:1 projiziert wird [109] (anfänglich 2,35:1), obwohl für die Aufnahme nur 1,195:1 zur Verfügung steht. Das gelingt bei diesem und ähnlichen Verfahren durch die Verwendung einer anamorphotischen Kompression, bei der das Bild mit Hilfe einer besonderen optischen Abbildung nur in der Horizontalen um den Faktor 2 gestaucht wird, während die Vertikale unbeeinflusst bleibt (Abb. 5.5). Die Verzerrung wird bei der Wiedergabe entsprechend ausgeglichen. Das Breitbild wurde international ein Erfolg, und in kurzer Zeit wurden viele Filmtheater mit entsprechenden Objektiven und Breitbildwänden ausgerüstet. In den 50er und 60er Jahren kamen viele weitere Formate mit Bezeichnungen wie Superscope, Vistavision, Technirama, Techniscope und Todd-AO heraus, die alle auf dem gleichen Prinzip beruhen und zusammenfassend als Scope-Verfahren bezeichnet werden.

Die anamorphotische Kompression kann nicht nur direkt bei der Aufnahme eingesetzt werden, sondern auch bei der Erzeugung einer Zwischenkopie, bei der das Bild in der Horizontalen entsprechend verzerrt und vertikal beschnitten wird. Für diesen Prozess ist die Aufnahme auf Super 35-Negativ besonders geeignet. Der Weg über die Zwischenkopie hat den Vorteil, dass keine Spezialobjektive für die Kamera erforderlich sind und dass das unverzerrte Originalnegativ auch für eine Fernsehauswertung genutzt werden kann, wenn darauf geachtet wird, dass der Bereich außerhalb des Breitbildausschnitts frei von störenden Elementen wie Lampen, Mikrofonen und Stativen gehalten wird.

Das Breitbild beeinflusste die weitere Entwicklung nachhaltig und wird heute bei den meisten Produktionen verwendet. Allerdings wird gewöhnlich auf den Einsatz der Anamorphoten verzichtet und statt dessen die genutzte Bildfläche vertikal eingeschränkt. Das auf diese Weise gewonnene Bildseitenverhältnis stellt einen Kompromiss aus Academy und Cinemascope dar, es beträgt in Europa 1,66:1, in den USA 1,85:1 (Abb. 5.4) [12].

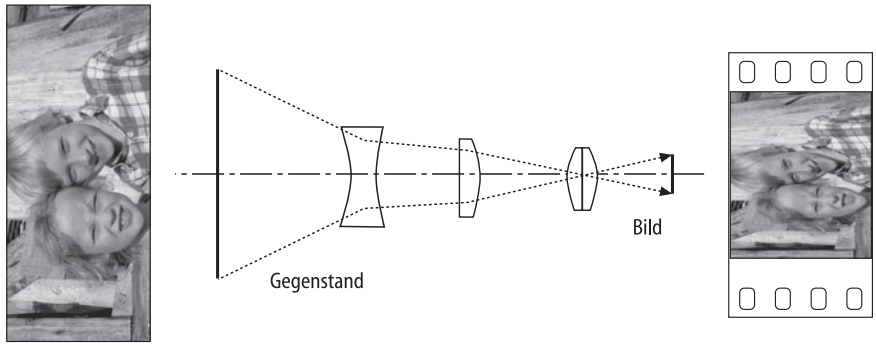


Abb. 5.5. Das Cinemascope-Prinzip

Bei konsequenter Nutzung des vertikal beschränkten Bildbereichs kann Filmmaterial gespart werden, wenn jedem Bild drei statt vier Perforationslöcher zugeordnet werden. Diese Möglichkeit wird als 3perf bezeichnet, in der Praxis aber wenig verwendet. Standard war ein ähnliches Verfahren dagegen bei Techniscope. Hier waren dem 2,35:1-Breitbild nur zwei Löcher zugeordnet.

Die TV-Auswertung der Breitbildformate ist nur dann unkritisch, wenn das HD-Bildseitenverhältnis B:H = 16:9 (1,78:1) verwendet wird. Für die TV-Auswertung mit dem gewöhnlichen Seitenverhältnis von 1,33:1 ist dagegen eine Anpassung erforderlich. So könnten die Bildseitenbereiche beschnitten und der Bildausschnitt gegebenenfalls dem bildwichtigen Teil angepasst werden (Pan and Scan), doch wird, um die Bildkomposition zu erhalten, die Alternative, nämlich die Sichtbarkeit schwarzer Streifen am unteren und oberen Bildrand (Letterbox), eher akzeptiert als der Verlust der Seiteninformationen (Abb. 5.6). Als dritte Option kann auch das aufnahmeseitige Format 1,37:1 übernommen werden, wenn das Bildfeld nicht bereits bei der Aufnahme maskiert war. In diesem Fall wird bei der TV-Wiedergabe mehr Bildinhalt zu sehen sein als im Kino. Dieser Umstand muss dann aber bei der Aufnahme berücksichtigt werden. D. h., dass auch hier darauf zu achten ist, dass keine Mikrofone oder andere störende Elemente in den größeren Bildbereich hineinragen.

Beim 16-mm-Film sind der Länge von einem Fuß (foot) 40 Bilder zugeordnet, der Bildfeldabstand beträgt damit 7,62 mm. Die Standard-Bildfeldgröße hat die Maße 10,3 mm x 7,5 mm mit einem Seitenverhältnis von 1,37:1. Auch hier entstand für die Aufnahme mit Super 16 ein Breitbildformat das ebenfalls den Tonspurbereich verwendet. Mit einseitiger Perforation kann dann ein Bildfeld der Größe 12,3 mm x 7,4 mm genutzt werden, das gegenüber Normal 16, ähnlich wie bei Super 35, etwas seitlich versetzt ist. Das Bildseitenverhältnis beträgt hier 1,66:1. Die Aufnahmefläche ist bei Super 16 mit 91 mm² ca. 20% größer als bei Normal 16, beträgt aber weniger als ein Viertel der Fläche von Super 35 bzw. ein Drittel der Fläche von 35 mm (Breitbild).

Noch kleinere Flächen weist das 8-mm-Format auf. Bei Super 8 beträgt die Bildfeldgröße 5,7 mm x 4,1 mm. Dagegen hat der 65 mm-Aufnahmefilm eine Bildfeldgröße von 23 mm x 52,5 mm bei 5 Perforationslöchern pro Bild. Die größten Bilder werden beim IMAX-System gespeichert. IMAX verwendet 70-mm-Film mit den Maßen 71 mm x 51 mm für die Projektion extrem großer Bilder. Dabei werden 15 Perforationslöcher pro Bild benutzt und der Film läuft mit bis zu 60 Bildern pro Sekunde im Gegensatz zu 16-mm- und 35-mm-Filmen horizontal durch den Projektor.

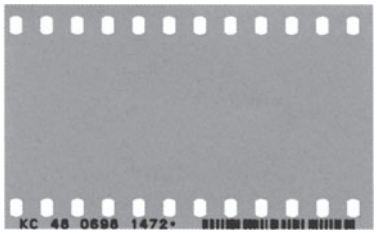
Abb. 5.6. Möglichkeiten zur Umsetzung des Filmbreitbildes in das 4:3-TV-Format

5.2.2 Filmkennzeichnung und Konfektionierung

Der endgültige Film entsteht erst, nachdem die gewünschten Szenen aus dem Negativmaterial ausgeschnitten und aneinander geklebt werden. Für den Schnitt ist eine eindeutige Kennzeichnung der Bilder erforderlich. Die Identifizierung geschieht im einfachsten Fall durch Zählung der Perforationslöcher oder der Bilder selbst. Zur Vereinfachung der Orientierung belichten die Filmhersteller bereits bei der Herstellung eine Filmkennzeichnung bestehend aus einer Folge von Buchstaben und Zahlen auf den Film, die bei 16-mm-Film zwischen den Perforationslöchern und bei 35-mm- und 65-mm-Film außerhalb am Filmrand liegt. Die Kennzeichnung ist nach der Entwicklung direkt lesbar und wird Fuß- oder Randnummer genannt, englisch Edge Code. Zusätzlich wird seit 1990 bei Filmen der Firma Kodak der Keycode verwendet, der einen maschinenlesbaren Strichcode enthält, aus dem auch Herstellerangaben über die Filmemulsion hervorgehen (Abb. 5.7). Damit ist es möglich, für größere Produktionen Material mit weitgehend ähnlichen Eigenschaften zu verwenden. Fußnummer und Keycode wiederholen sich beim 65- und 35-mm-Film im Abstand von 16 und beim 16-mm-Film im Abstand von 20 Bildern. Bei Fuji-Filmen gibt es ein ähnliches System mit der Bezeichnung MR-Code.

Eine weitere Vereinfachung, insbesondere für die Synchronisation von Bild und Ton, ergibt sich durch die Verwendung von aufbelichtetem Timecode. Nach einem System des Kameraherstellers Arri kann z. B. der sogenannte Arricode dicht neben den Perforationslöchern bei der Filmaufnahme als ein Strichmuster aufbelichtet werden, in dem digital ein 112-Bit-Zeitcode (Typ B) verschlüsselt ist, der in den wesentlichen 80 Bit mit dem Standard-SMPTE-Timecode vom Typ C identisch ist (s. Kap. 8.9.3). Ähnliche Systeme gibt es auch bei anderen Kameraherstellern wie z. B. der Firma Aaton. Der relative Keycode wird auf diese Weise mit einem absoluten Zeitcode ergänzt, was die Nachbearbeitungsschritte insgesamt erheblich vereinfacht.

Das perforierte Filmaufnahmematerial wird in Standardlängen geschnitten und auf Wickelkerne (Bobby) oder auf Spulen aufgewickelt und anschließend in lichtdichte Metalldosen gepackt. Die verwendeten Filmlängen basieren auf geraden Zahlen nach alter amerikanischer Metrik mit der Einheit Foot (Fuß). Bei 16-mm-Film entspricht ein Fuß genau 40 Bildern, bei 35 mm sind es 16 Bilder. Standardlängen sind bei 35-mm-Negativfilm 200, 400 und 1000 feet, entsprechend 61 m, 122 m und 305 m. Als Maximallänge wird 615 m angeboten. Für 16-mm-Material werden meist 30 m und 122 m als Standardlängen



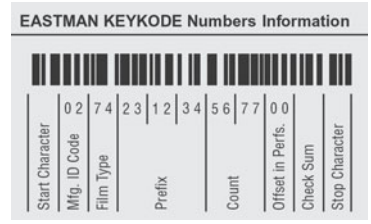


Abb. 5.7. Randkennzeichnung und Bedeutung des Keycodes [113]

Tabelle 5.1. Filmlaufzeiten und Filmlängen

	24 fps:	35 mm	16mm	25 fps:	35mm	16 mm
Min.		Meter	Meter		Meter	Meter
1		27,36	10,97		28,5	11,43
2		54,72	21,94		57,0	22,86
3		82,08	32,92		85,5	34,29
5		136,80	54,86		142,5	57,15
10		273,60	109,73		285,0	114,30
20		547,20	219,46		570,0	228,60
30		820,80	329,19		855,0	342,90

verwendet. Bei einer Filmgeschwindigkeit von 24 frames per second folgt für den 35-mm-Film bei einer Länge von 305 m eine Aufnahmedauer von ca. 11 min und für 16-mm-Film eine Dauer von ca. 28 Minuten. Tabelle 5.1 zeigt eine Übersicht, in der die Laufzeiten in Bezug zur Filmlänge angegeben sind.

Die Wickelkerne für Negativfilm haben in den meisten Fällen einen Außendurchmesser von 2“ (50,8 mm) mit einem 1“-Loch. Im Außenrand ist eine Kerbe angebracht, in die der Film eingefädelt wird. Dieser universellste Bobbytyp trägt bei 35-mm-Kodakmaterial die Bezeichnung U-Core, für 16-mm-Film T-Core. Filmaufnahmematerial, das auf Bobbies gewickelt ist, darf in keinem Falle dem Licht ausgesetzt werden. Für das Einlegen der Rollen in die Kassette einer Filmkamera muss ein lichtdicht abgeschlossener Dunkelsack verwendet werden. Für kurze Längen stehen sogenannte Tageslichtspulen zur Verfügung bei denen der Film eng zwischen die Seitenteile der Spule gewickelt ist.

Die Angaben der Parameter des verwendeten Films, also Angaben zur Filmbreite, Perforation, Filmlänge etc. finden sich auf dem Film Label. Abbildung 5.8 zeigt ein typisches Beispiel für einen 35-mm-Film von 122 m Länge. Kodak verwendet bei 35-mm-Film die Ziffern 52 und bei 16-mm-Film die Ziffern 72 als Beginn der Filmartkennzeichnung [128]. Die Filmempfindlichkeit (Exposure Index, EI) beträgt bei diesem Beispiel 500 ASA bei Kunstlicht (Tungsten Rating, T). Dabei ist auch angegeben, dass mit Hilfe eines Konversionsfilters Wratten Type 85 eine Abstimmung auf die Farbtemperatur des Tageslichts (Daylight, D) vorgenommen werden kann, wobei sich dann ein EI von 320 ASA ergibt.

Abb. 5.8. Filmlabel [113]

5.3 Filmeigenschaften

Für die verschiedenen Anwendungsgebiete, in denen Film verwendet wird, stehen unterschiedliche Filmtypen zur Verfügung. Hinsichtlich der Kamera und des Projektors muss das Filmformat festgelegt sein. Hier gibt es heute im Wesentlichen die Entscheidung zwischen 16-mm- und 35-mm-Film. Des Weiteren muss entschieden werden, ob in Farbe oder Schwarzweiß gedreht wird und ob auf dem Film Raum für Tonspuren vorgesehen wird. Dann kann der Film hinsichtlich der Frage ausgewählt werden, ob eher die Lichtempfindlichkeit oder die detailgetreue Abbildung im Vordergrund steht und ob unter Tages- oder Kunstlichtbedingungen gearbeitet wird.

Schließlich ist die Frage der Vervielfältigung sehr wichtig. Für die schnelle Verfügbarkeit von Einzelstücken stehen Umkehrverfahren zur Verfügung, für die Erstellung von Massenkopien Negativ- und Positivfilme. Die Aufnahme-Negativfilme werden englisch auch als Camerafilms und die Positive als Printfilms bezeichnet. Bei großen Stückzahlen stammen nicht alle Positive von demselben Negativ ab, sondern von Zwischenkopien, sogenannte Intermediate-Materialien, d. h. die Massenkopien kommen von kleineren Mengen von Internegativen, die von einer geringen Zahl von Interpositiven gewonnen werden, die wiederum von dem Einzelnegativ herrühren. Intermediate-Material, auch Duplikat- oder Laboratory Films genannt, werden auch für die Erstellung von Titeln, Effekten etc. gebraucht.

5.3.1 Belichtung und Schwärzung

Die Schwärzung des Filmmaterials unter Lichteinfall hängt von der Beleuchtungsstärke E und auch von der Beleuchtungsdauer t ab. Beide Ursachen werden unter dem Begriff Belichtung H (englisch: Exposure) mit der Einheit Luxsekunden (lxs) zusammengefasst. Es gilt:

$$H = E \cdot t \,.$$

Die Schwärzung eines Films mindert seine Transparenz T, d. h. das Verhältnis zwischen der Lichtintensität, die vom Film durchgelassen wird, und der auftreffenden Gesamtintensität des Lichts. Das Gegenteil der Transparenz ist die Lichtundurchlässigkeit oder Opazität O. Es gilt:

$$T = 1/O.$$

Bei völliger Transparenz beträgt der Wert T = 1, er sinkt bei völliger Schwärzung auf den Wert T = 0, während die Opazität unter gleichen Bedingungen von 1 bei völliger Transparenz auf unendlich steigt.

Als Maß für die Filmschwärzung ist die Opazität gut geeignet. Aufgrund der Tatsache, dass das menschliche Auge bei seiner Helligkeitswahrnehmung S nicht linear, sondern eher logarithmisch, weitgehend nach dem Fechnerschen Gesetz:

$$S = k \log L/L_0, \text{ mit } L_0 = 1 \text{ cd/m}^2,$$

auf die Reizintensität L reagiert, wird zur Kennzeichnung der Schwärzung jedoch der Dichtewert D verwendet, der logarithmisch mit der Opazität zusammenhängt. Es gilt:

$$D = \log O = \log 1/T.$$

Dabei gilt der Logarithmus zur Basiszahl 10, d. h. die doppelte Opazität oder halbe Transparenz ergibt sich bei einem Dichtewert $D = \log 2 = 0{,}3$. Zur Darstellung einer gleich abständigen Grautreppe wird ein Stufenkeil verwendet, dessen Dichtewerte je nach Anwendung linear um den Wert 0,1; 0,15 oder 0,3 ansteigen [229]. Aus praktischen Gründen wird bezüglich der grafischen Darstellung der Zusammenhänge auch bei der Belichtung wiederum mit dem Logarithmus, also log H gearbeitet.

Dichtewerte lassen sich auch zur Kennzeichnung von Neutraldichtefiltern verwenden die der farbneutralen Abschwächung des Lichts dienen. Ein Neutraldichtefilter mit D = 0,3 hat einen Transmissionsgrad von 0,5 und halbiert die Beleuchtungsstärke, was dem Schließen der Blende im Objektiv um eine Stufe entspricht. Da die Blende den wesentlichen Mechanismus zur Beeinflussung der Helligkeit der Abbildung darstellt, werden statt absoluter Belichtungswerte oft relative Blendenwerte angegeben.

5.3.2 Kennlinie und Kontrastumfang

Die Filmkennlinie, die auch als Schwärzungskurve oder Dichtekennlinie bezeichnet wird, stellt den Zusammenhang zwischen der Belichtung $\log E \cdot t$ und der Schwärzung des entwickelten Materials dar, die durch den Dichtewert D gegeben ist. Der Einfachheit halber wird hier zunächst nur die Schwarzweißdichte betrachtet, Farbfilmmaterialien haben separate Kennlinien mit ähnlichem Verlauf für die veränderliche Dichte der Farbstoffwolken.

Die Kennlinie verläuft insgesamt nicht linear, sondern weist eine charakteristische S-Form auf, die in Abbildung 5.9 für Negativmaterial dargestellt ist. Im linken Teil der Kurve wird deutlich, dass der Film auch ohne Belichtung

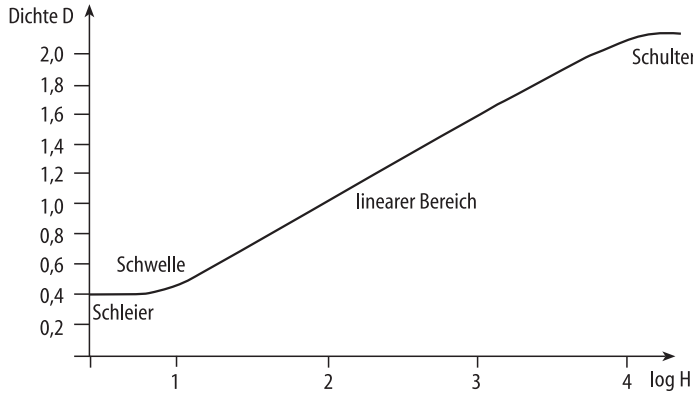


Abb. 5.9. Typische Kennlinie des Negativ-Filmmaterials

eine gewisse Dichte aufweist, die als Schleier bezeichnet wird. Diese Dichte resultiert daher, dass einerseits das Trägermaterial (engl. Base) nicht völlig transparent ist und andererseits die Silbersalze zu einem geringen Teil auch ohne Lichteinwirkung entwickelbar sind (engl.: Base plus fog bzw. D_{min} bei Farbfilm). Der Schleier ist daher konstant und wird kaum von der Entwicklung beeinflusst.

Bei steigender Belichtung gelangt man zur Schwelle, dem Bereich, ab dem die Schwärzung durch Lichteinwirkung beginnt. Im mittleren Teil der Kennlinie wird der Zusammenhang dann linear. Der Übergangsbereich, Schwelle oder auch Durchhang genannt, erstreckt sich bis zu einem Punkt, der um D = 0,1 über dem Schleier liegt.

Der lineare Teil der Kennlinie ist der eigentliche Aufnahmebereich, hier sollte der bildwichtige Kontrastumfang, d. h. die Hell-Dunkel-Differenz, liegen. Wenn das Dichtemaximum des entwickelbaren Materials D_{max} erreicht ist, wird weiterer Lichteinfall keine höhere Schwärzung bewirken, der Übergang zu diesem Bereich wird als Schulter bezeichnet.

Obwohl die Schwelle und die Schulter eigentlich ungeeignet für die Aufnahme sind, sind sie doch für die charakteristische Bildwirkung von Film von großer Bedeutung. Außerhalb des Bereichs, in dem der bildwichtige Kontrastumfang liegt, treten nämlich oft Spitzlichter auf, die sehr viel größere Leuchtdichten aufweisen als der mittlere Belichtungsbereich. Diese werden durch die Schulter sanft begrenzt, was eine erheblich andere Wirkung im Bild hat, als wenn eine abrupte Beschneidung aufträte. Hier unterscheidet sich Film deutlich von vielen der üblichen elektronischen Bildumsetzungsverfahren. Der Zusammenhang zwischen den Werten der Kennlinie und je einem Positiv- und Negativbild zeigt Abbildung 5.10.

Der fotografisch wesentliche Teil der Kennlinie ist der lineare Bereich. Abhängig vom verwendeten Filmmaterial kann dieser mehr oder weniger Stei-

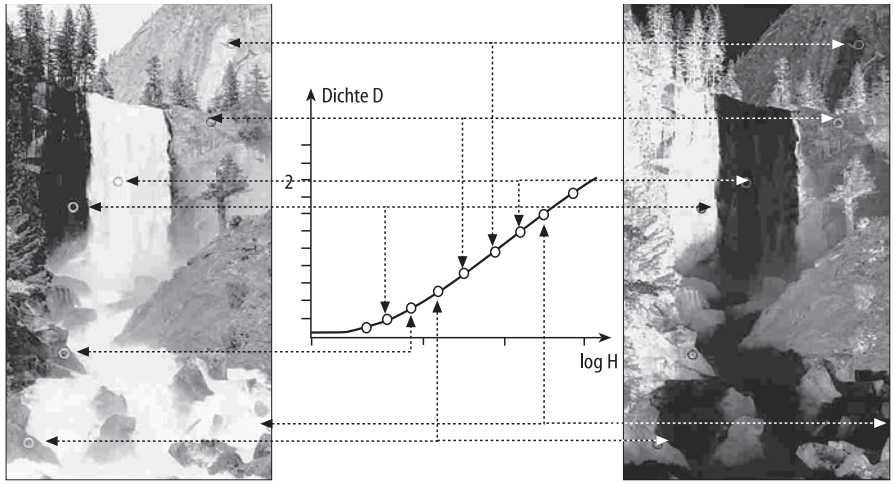


Abb. 5.10. Positiv- und Negativbild im Zusammenhang zur Aufnahmekennlinie [113]

gung aufweisen. Die Steigung wird als Gradation oder auch kurz als Gamma bezeichnet und ist als der Quotient zwischen Dichte- und Belichtungsänderung (log) definiert. Je größer die Steigung, desto größer ist der Szenenkontrast im Bild. Die Steigung wird mit Hilfe einer Tangente bestimmt. Die Beschränkung auf den wichtigsten, nämlich den geradlinigen Teil liefert den Gammawert:

$$\gamma = \Delta D / \Delta \log H.$$

Die Tangente ist so definiert, dass sie durch einen unteren Punkt geht, der einen Dichtewert aufweist, der um D = 0,1 über dem Schleier liegt. Ab diesem Punkt ist eine sichere Entwicklung gewährleistet. Der obere Schnittpunkt ist so definiert, dass er um 1,5 logarithmische Einheiten auf der Belichtungsachse höher liegt als der untere. Auf diese Weise ist ein Leuchtdichteumfang bzw. Objektkontrast von 1:32 definiert, ein Wert, der für den Leuchtdichteumfang durchschnittlicher Motive bei Studioausleuchtung festgelegt wurde. Bei Einbeziehung des Schulter- und Kniebereichs kann auf ähnliche Weise ein β-Wert definiert werden, der die mittlere Gradation angibt.

Die Steigung der γ-Kurve, auch bezeichnet als Steilheit des Materials, kann bei der Filmherstellung bestimmt werden. Nahe liegend ist zunächst, sowohl für das Positiv als auch für das Negativ ein Gamma von 1 zu wählen und auf diese Weise den Kontrastumfang der aufgenommenen Szene in den gleichen Kontrastumfang bei der Projektion umzusetzen. Aufgrund der geringen Leuchtdichte der Projektion, die nicht der von hellem Sonnenschein entspricht, sowie von Kontrastverringerungen durch Objektive und Streulichter wirkt das Filmbild allerdings zu matt, wenn eine »Über-alles-Kennlinie« mit $\gamma = 1$ verwendet wird. In der Praxis wird das System- oder Gesamt-Gamma daher auf einen Wert von ca. $\gamma = 1{,}7 \ldots 1{,}8$ angehoben. Dies geschieht durch Verwendung eines großen Gamma-Wertes beim Positivfilm.

Beim Positivfilm werden Gamma-Werte von ca. 3 benutzt. Da sich nun das Gesamt-Gamma aus dem Produkt der Einzelwerte mit

$$\gamma = \gamma_{neg} \cdot \gamma_{pos}$$

ergibt, folgt, dass der Negativfilm so hergestellt wird, dass er nur ca. $\gamma = 0{,}6$ aufweist. Der gesamte Belichtungsumfang soll nur von den Positiv- und Negativmaterialien abhängen. Intermediate- oder Duplikatmaterial dürfen den Kontrastumfang des Gesamtsystems nicht verändern, daher muss dieses Material einen Gamma-Wert von 1 aufweisen.

Die gezielte Verwendung eines geringen Kontrastumfangs auf der Aufnahmeseite bewirkt, dass der lineare Bereich der Kennlinie länger wird, denn die Steigung ist wegen des geringeren ΔD verringert. Damit lässt sich der oben angegebene Belichtungsumfang von $\log H = \log E \cdot t = 1{,}5$ mehrfach im geraden Teil der Kennlinie unterbringen oder auf höhere Werte ausdehnen (Abb. 5.11). Das heißt, dass der Normalbelichtungsumfang nicht absolut exakt in der Mitte des geraden Kennlinienteils liegen muss. Leichte Unter- oder Überbelichtungen können beim Kopierprozess verlustlos ausgeglichen werden, Belichtungsfehler sind tolerierbar. So bewirkt die Halbierung des Lichtstroms durch die kameraseitige Veränderung um einen Blendenwert eine Verschiebung um

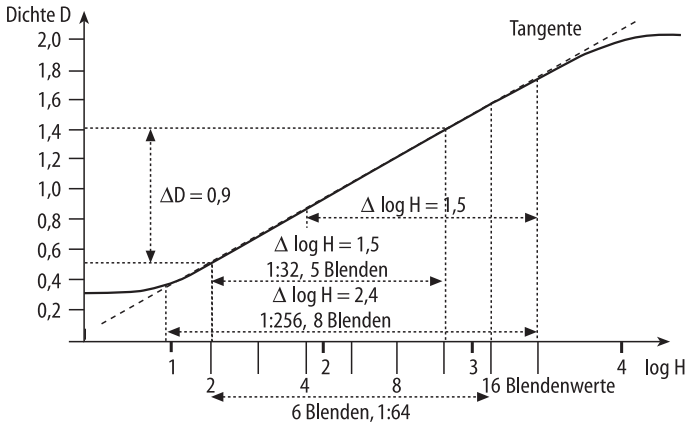


Abb. 5.11. Ermittlung des Gammawertes und der Belichtungsumfänge an der Kennlinie

0,3 Einheiten auf der log-H-Achse, was bei flacher Kennlinie leicht aufzufangen ist.

Eine Möglichkeit für die Beeinflussung der Kennlinie ist die Vorbelichtung. Dabei wird vor der eigentlichen Aufnahme durch diffuse Beleuchtung mit geringer Intensität eine Aufhellung der Schattenpartien bewirkt und die Kennlinie damit abgeflacht. Die gleiche Wirkung kann auch durch eine Nachbelichtung erreicht werden.

Da bei Umkehrfilmen der endgültige Kontrastumfang in einem Schritt entsteht, soll hier $\gamma = 1{,}8$ direkt vorliegen. Der Spielraum für die Belichtung ist damit erheblich geringer als beim Negativmaterial und es muss bei der Aufnahme sorgfältiger gearbeitet werden. Bereits Abweichungen um einen halben Blendenwert sind im Bild erkennbar.

Im praktischen Umgang wird der Kontrastumfang in der Szene eher über Logarithmen zur Basiszahl 2 statt über dekadische Logarithmen angegeben, da die Lichtmenge, die auf den Film fällt, durch die Blende im Objektiv der Kamera bestimmt wird. Deren Skalierung ist so gewählt, dass von Stufe zu Stufe die Lichtintensität um den Faktor 2 geändert wird. Der Kontrastumfang wird daher meistens in Blendenstufen angegeben. Ein Kontrast von 1:32 entspricht 5 Blenden, in der Praxis werden Kontrastumfänge von ca. 7 Blenden verwendet. Für den Mittelwert der von den Objekten in der Szene hervorgerufenen Gesamtremission wird ein Wert von 18% angegeben. Dieser Wert ist die Referenz für die sogenannte Bezugsblende, die in der Mitte des verwendeten Kontrastbereichs liegt. Diese Bezugsblende wird von einem Spot-Belichtungsmesser angegeben, der in der gegebenen Lichtsituation zur Ausmessung einer Graufläche mit dem Remissionswert 18% verwendet wird. Wird dabei z. B. der Blendenwert 5,6 bestimmt, so liegt ein Kontrastumfang von 6 Blenden zwischen den Werten 2 und 16 (Abb. 5.11). Der Kontrastumfang, angegeben in Blendenwerten, kann mit dem Spot-Belichtungsmesser bestimmt werden, indem er auf die hellsten und dunkelsten Stellen der Szenen gerichtet wird.

5.3.3 Farbfilmeigenschaften

Um Farben im Filmmaterial umsetzen zu können, werden lichtempfindliche Schichten verwendet, die wellenlängenselektiv reagieren. Bei Farbfilmmaterialien lassen sich damit im Vergleich zum Schwarzweißfilm statt einer drei Kennlinien für die drei Gelb-, Purpur- und Cyanschichten gewinnen. Die Kennlinien verlaufen prinzipiell genauso wie beim Schwarzweißfilm, können aber parallel verschoben sein (Abb. 5.12). Damit auch graue, d. h. unbunte Farben wiedergegeben werden können, müssten die drei Kennlinien idealerweise den gleichen Verlauf aufweisen. Die Parallelverschiebungen (Fehler 1. Ordnung) können beim Kopierprozess oder bei der Abtastung bzw. beim Colourmatching korrigiert werden. Kaum korrigierbare Fehler, als Fehler 2. Ordnung bezeichnet, treten dagegen auf, wenn die Kennlinien gekreuzt verlaufen. Man spricht von einem kippenden Farbstich, der in den Schattenpartien komplementär zu dem Stich im Bereich der Lichter ist. Die Ursache solcher Fehler liegt bei der Filmherstellung oder beim Kopierprozess.

Die Aufgabe der Farbschichten ist die Filterung des einfallenden Lichts. Im Idealfall sollte die Gelbfilterschicht im Blau-, die Magentaschicht im Grün- und die Cyanschicht im Rotbereich als Filter zur subtraktiven Mischung beitragen. Die jeweiligen Filterwirkungen sind jedoch unterschiedlich und weisen auch Dichtewerte in den Nachbarbereichen, sogenannte Nebendichten, auf. Die Summendichte ergibt sich aus der Gesamtheit der Haupt- und Nebendichten. Für die Gesamtwirkung muss ein Kompromiss gefunden werden, der vor allem bei unbunten Vorlagen keinen Farbstich zeigt. Das gelingt durch Unterdrückung unerwünschter Dichten mittels Masken, die während des Farbentwicklungsprozesses in den Emulsionen gebildet werden.

Aufgrund von Differenzen zwischen den Kennlinienverläufen von Positiv- und Negativmaterialien werden bei der messtechnischen Bestimmung der Dichten von Farbfilmen verschiedene Filtersysteme eingesetzt. Man unterscheidet daher sogenannte Status-M-Dichten für Negativ- und Intermediate-

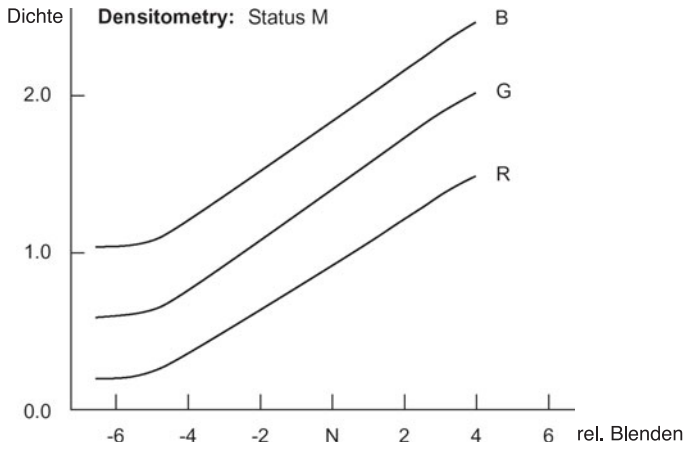


Abb. 5.12. Kennlinienverlauf des Farbnegativmaterials Kodak 5248 [113]

und Status-A-Dichten für Positivmaterial [151]. Mögliche Farbstiche hängen natürlich auch von der spektralen Verteilung des Lichts ab. Die Filmemulsionen werden entsprechend für den Spektralverlauf von Tageslicht oder von Kunstlichtquellen, d. h. Temperaturstrahlern (Tungsten Lamp), abgestimmt. Kunstlicht mit 3200 K zeichnet sich durch einen größeren Rotanteil aus als Sonnenlicht mit der Farbtemperatur 5600 K. Mit Hilfe von Konversionsfiltern, die den erhöhten Rotanteil unterdrücken, lassen sich Kunstlichtfilme für den Tageslichteinsatz verwenden. Das gilt auch umgekehrt, in beiden Fällen ist allerdings zu bedenken, dass der Filtereinsatz die Lichtintensität schwächt und so die effektive Filmempfindlichkeit reduziert.

5.3.4 Die Lichtempfindlichkeit

Die Lichtempfindlichkeit des Filmmaterials ist dadurch definiert, dass ermittelt wird, bei welcher Belichtung eine Dichte entsteht, die bei S/W-Filmen um D = 0,1 bzw. bei Farbfilmen um D = 0,2 über dem Schleier liegt [229]. Das mit 10 multiplizierte logarithmische Verhältnis dieses mit H_M bezeichneten Belichtungswertes zum Wert 1 lxs ergibt die DIN-Zahl:

$$\text{DIN-Zahl} = 10 \log (1 \text{ lxs}/H_M)$$

Die Definition ähnelt der Bildung von Pegelwerten in dB. Die doppelte Lichtempfindlichkeit wird damit durch die Steigerung um 3 DIN ausgedrückt. Zur Angabe eines linearen Empfindlichkeitswertes (Exposure Index, EI) dient der ASA-Wert. Er folgt aus der Beziehung:

$$\text{ASA-Wert} = 0{,}8 \text{ lxs}/H_M,$$

wobei H_M wieder die genannte Belichtung darstellt. Tabelle 5.2 zeigt das Verhältnis zwischen DIN- und ASA-Werten.

Empfindliche Filme erfordern bei gegebener Lichtintensität eine geringere Belichtungsdauer, man sagt umgangssprachlich: sie haben mehr speed. Die Empfindlichkeitsangaben beziehen sich auf definierte Entwicklungsbedingungen, denn auch über den Entwicklungsprozess kann die Empfindlichkeit des Materials gesteigert werden.

Eine derartige gezielte Steigerung wird als forcierte Entwicklung bezeichnet. Dabei wird durch die Verlängerung der Entwicklungszeit erreicht, dass auch Kristalle zu Silber gewandelt werden, die so wenig Lichtintensität ausgesetzt waren, dass sie bei gewöhnlicher Entwicklung zu den unterbelichteten Filmbereichen gehören. Die Forcierung wirkt gleichermaßen im Licht- und Schattenbereich, sie steigert die Dichte und bewirkt die Verschiebung der gesamten Kennlinie.

Tabelle 5.2. Vergleich von DIN- und ASA-Werten

DIN	12	15	18	21	24	27	30
ASA	12	25	50	100	200	400	800

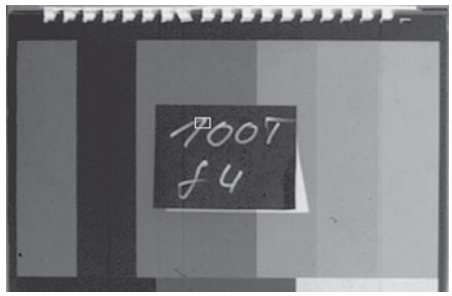
5.3.5 Das Filmkorn

Lichtempfindliche Filme lassen sich mit Hilfe großflächiger Silberkristalle gewinnen. Abgesehen von verschiedenen Weiterentwicklungen (T-Grain) lässt sich generell der Antagonismus zwischen kleinem Filmkorn und hoher Empfindlichkeit nicht aufheben. Große Kristalle werden als Filmkorn (Grain) besonders bei großen Flächen mit mittleren Grauwerten im Bild sichtbar und beeinflussen wesentlich den sogenannten Filmlook. Graue Flächen im Bild sind nicht homogen, sondern bestehen aus einer unregelmäßigen Verteilung der Silberkörner, die in Abhängigkeit vom Grauwert mehr oder weniger dicht ist (Abb. 5.13). Im Farbfilm gilt Ähnliches. An die Stelle der Silberteilchen, die ja beim Entwicklungsprozess ausgelöst werden, treten hier die Farbstoffwolken, was in Abbildung 5.2 sehr gut deutlich wird, wo anhand einer mikroskopischen Aufnahme einer Farbschicht geringer Dichte die Farbstoffwolken mit und ohne Silber gut erkennbar sind. Die Größenordnung der Durchmesser der Silberteilchen und Farbstoffwolken beträgt bei Negativmaterial ca. 5 μm. Abbildung 5.14 zeigt anhand verschiedener Kodak-Negativmaterialien, wie die Korngrößen mit der Lichtempfindlichkeit steigen [151]. Intermediate- und Print-Material kann feinkörniger sein, da hier die Empfindlichkeit nicht im Vordergrund steht. Die Körnigkeit ist in den einzelnen Farbschichten unterschiedlich. In der Blauschicht ist sie besonders ausgeprägt, denn sie muss aufgrund des geringen Blauanteils im Licht besonders bei Kunstlicht-Negativen empfindlicher sein als die anderen Schichten. Daher wird im Filmbereich für Stanztricks oft die Aufnahme vor Grün dem Blue-Screen-Verfahren vorgezogen.

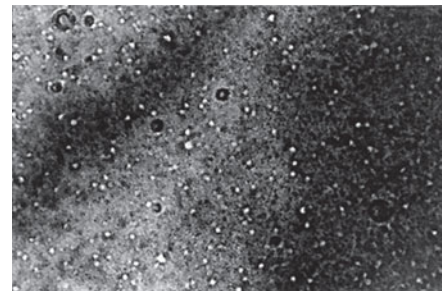
Die bei der Abtastung der Filmoberfläche und Wandlung der Dichtedifferenzen in ein elektrisches Signal entstehenden Unregelmäßigkeiten führen bei der Messung zur Granularität, die sich in elektrischer Form als Rauschsignal zeigt und in Abhängigkeit von der Korngröße verschiedene Amplituden aufweist. Über die Bildung der Wurzel aus dem Mittelwert über deren Quadrate lässt sich als Messwert die RMS-Körnigkeit gewinnen. Über den Logarithmus wird das Verhältnis dieses Wertes zum maximalen Nutzsignal als Signal to Noise Ratio S/N angegeben:

$$S/N = 20 \log (U_{max}/U_{RMS}).$$

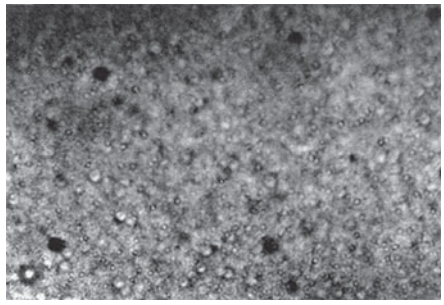
Abb. 5.13. Filmbildausschnitte bei verschiedenen Vergrößerungen [113]



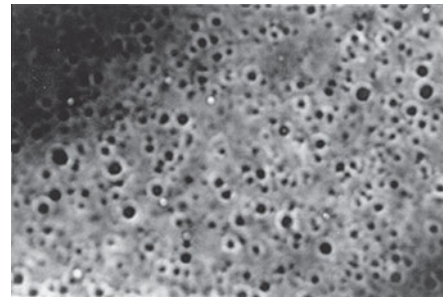
Gesamtbild mit Ausschnitt



Kodak 5248 mit 100 ASA



Kodak 5274 mit 200 ASA



Kodak 5279 mit 500 ASA

Abb. 5.14. Gleichmäßig vergrößerte Ausschnitte bei verschiedenen Filmempfindlichkeiten [133]

5.3.6 Das Auflösungsvermögen

Das Auflösungsvermögen beschreibt die Darstellbarkeit feiner Details in der Filmemulsion, die wiederum von der Körnigkeit des Materials abhängt. Die Bestimmung erfolgt anhand feiner Hell-Dunkel-Wechsel in Form von Linienpaaren. Je mehr Linienpaare auf einem Millimeter Breite untergebracht und unterschieden werden können (lp/mm), desto höher ist das Auflösungsvermögen.

Die visuelle Trennbarkeit der Linien ist dabei eine ungenaue Bestimmungsgröße, da bei geringen Abständen der Weißbereich zwischen zwei schwarzen Streifen grau erscheint. Dieser Umstand wird bei der Bestimmung der Modulationstransferfunktion MTF berücksichtigt. Aufgrund der Beziehung

$$MTF = m\,(HF) \,/\, m\,(LF)$$

gibt sie Auskunft über die allmähliche Kontrastabnahme bei Steigerung der Anzahl der Linienpaare pro Millimeter, indem sie die Hell-Dunkel-Differenz oder genauer den Modulationsgrad m nach der Beziehung

$$m = (s_{max} - s_{min}) \,/\, (s_{max} + s_{min})$$

bei hoher Liniendichte (HF) ins Verhältnis zum Modulationsgrad bei geringer Dichte (LF) setzt, bei der die Linien 100% getrennt erscheinen (Abb. 5.15).

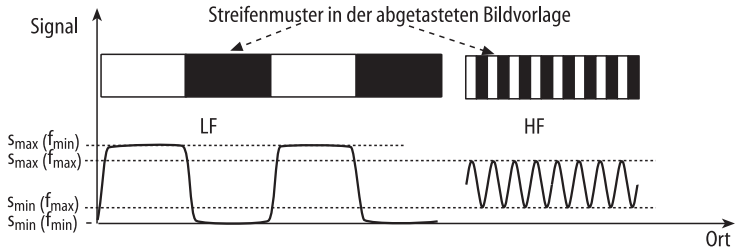


Abb. 5.15. Zur Definition der Modulationstransferfunktion

Die MTF eignet sich nicht nur zur Beschreibung des Auflösungsvermögens der Filmemulsion, sondern auch für weitere Elemente im Lichtweg oder elektronische Signalbearbeitungsstufen, bei denen dann HF und LF für High und Low Frequency stehen. Die Gesamt-MTF ergibt sich aus der Multiplikation aller Einzelwerte. Durchschnittlicher, nicht sehr hoch empfindlicher Negativfilm erreicht bei 120 lp/mm noch einen MTF-Wert von 50%. Abbildung 5.16 zeigt den für die RGB-Anteile verschiedenen MTF-Verlauf des Kodak-Negativmaterials 5248 und Tabelle 5.3 verschiedene Auflösungswerte für unterschiedliche Materialien jeweils bei einem MTF-Wert von 50%. Es sei hier noch einmal darauf hingewiesen, dass das Filmmaterial nicht das einzige auflösungsbegrenzende Element ist. Oft beeinträchtigt der MTF-Wert eines Objektivs das Bild stärker als die MTF des Filmmaterials.

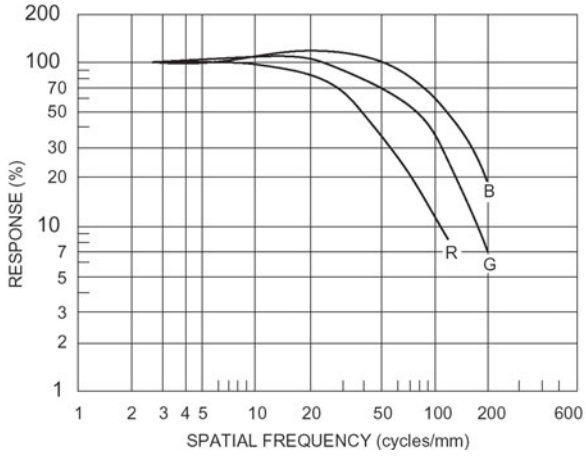


Abb. 5.16. Verlauf der Modulationstransferfunktionen des Negativ-Materials Kodak 5248 [113]

Tabelle 5.3. Vergleich der Auflösungswerte bei verschiedenen Filmmaterialien

Typ	5279	5274	5248	5246	5245
R	40	35	35	40	38 lp/mm
G	70	65	78	80	95 lp/mm
B	80	70	120	100	120 lp/mm

5.4 Filmentwicklung und -kopie

Vom ersten Drehtag bis zur Auslieferung der fertigen Filme ist das Kopierwerk ein wichtiger Partner beim Filmproduktionsprozess. Das Kopierwerk entwickelt die Negative und lagert und verwaltet sie. Sie werden hier nach der Schnittliste, die vom Schnittplatz kommt, geschnitten und geklebt. Weiterhin werden im Kopierwerk die Positive belichtet, wobei die Bestimmung der dafür erforderlichen Lichtmischung eine große Rolle spielt. Es können auch Bildformatänderungen und Verkleinerungen bzw. Vergrößerungen (Blow ups) und weitere Trickarbeiten vorgenommen werden. Schließlich werden auch die Positive entwickelt und für die Auslieferung konfektioniert.

5.4.1 Filmentwicklung

Zunächst ist es wichtig, dass das belichtete Filmnegativ möglichst schnell entwickelt wird, damit die empfindlichen latenten Bilder keinen Schaden nehmen. Das Negativmaterial wird daher täglich nach Drehschluss in schwarzer Folie und den sorgfältig beschrifteten Originaldosen zum Kopierwerk geschickt. Für die Negativentwicklung wird heute für Kodak-Material ein Prozess mit der Bezeichnung ECN-2 (Eastman Color Negativ) eingesetzt. Dabei wird der Film in einer Entwicklungsmaschine mittels Rollenantrieb so durch die verschiedenen Verarbeitungsstufen geführt, dass er dort über die jeweils erforderliche Dauer verbleibt. Zuerst wird in einem Vorbad in der Dauer von 10 s die Lichthofschutzschicht, die sich auf dem Negativ befindet, eingeweicht und dann in 5 s entfernt. Im eigentlichen Entwicklerbad bleibt der Film für 3 Minuten, bis in einem Stoppbad in 30 s die Entwicklung der Silberhalogenidkörner gestoppt und die Farbentwicklersubstanz aus dem Film gewaschen wird. Anschließend erfolgt eine 30-sekündige Wässerung, dann gelangt der Film für 3 Minuten in das Bleichbad, in dem das Silber, das sich beim Entwicklungsprozess neben den Farbstoffen gebildet hat, in Halogenidverbindungen verwandelt wird, die wiederum im Fixierprozess und der anschließenden Wässerung entfernt werden können. Im Fixierbad bleibt der Film für 2 Minuten, vor- und nachher wird er für eine bzw. zwei Minuten gewässert [128]. Der Prozess endet mit einer 10-sekündigen Stabilisierung und einer 5-minütigen Trocknung (Abb. 5.17). In ähnlicher Weise läuft auch der Positiv-Entwicklungsprozess ab, der bei Kodak mit ECP-2B bezeichnet wird.

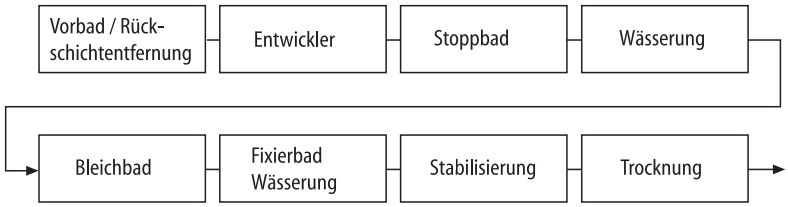


Abb. 5.17. Ablauf des Entwicklungsprozesses ECN-2

Der Entwicklungsprozess lässt sich auch verwenden, um Fehlbelichtungen auszugleichen, die über die Lichtbestimmung nicht erfassbar sind. Zum Ausgleich von Unterbelichtungen wird die forcierte Entwicklung verwendet, bei der die Verweildauer im Entwicklerbad verlängert wird. Auf diese Weise können Schwärzungsänderungen hervorgerufen werden, die einer Belichtungsänderung um mehrere Blendenstufen entsprechen. Da die Bildqualität durch diesen Vorgang gemindert wird, sollte die Änderung auf ca. eine Blendenstufe begrenzt bleiben, falls der Vorgang nicht bewusst als besonderer Effekt genutzt wird. Die Forcierung wird auch als Push-Prozess, das Umgekehrte als Pull-Prozess bezeichnet. Als zweite Möglichkeit ist die Vor- oder Nachbelichtung (Flashing) zu nennen, eine kurzzeitige ganzflächige Zusatzbelichtung, die eine Aufhellung von Schattenbereichen bewirkt. Eine weitere Bildbeeinflussungsmöglichkeit bei der Entwicklung stellt die Bleichbadüberbrückung dar, bei der das Silber zusammen mit dem Farbstoff in der Schicht verbleibt und zu einem dunkleren Bild mit gedämpften Farben führt.

5.4.2 Der Kopierprozess

Für die Umsetzung des Negativs in ein Positivbild (Printprozess) wird das Negativ durchleuchtet und mit dem transmittierten Licht das Positiv belichtet. Für diesen Prozess können beide Filme kontinuierlich oder schrittweise transportiert werden. Der Kopiervorgang kann weiterhin danach unterschieden werden, ob die Emulsionen der Filme beim Kopiervorgang in direktem Kontakt stehen oder ob eine optische Abbildung zwischengeschaltet wird. Wenn keine Bildgrößenänderungen erforderlich sind, wird in den meisten Fällen eine Kontaktkopie mit kontinuierlichem Lauf durchgeführt (Continuous Contact Printer), was den Film mehr schont als der schrittweise Transport und auch größere Geschwindigkeiten ermöglicht. Es werden Kopiergeschwindigkeiten von mehr als 200 m/min erreicht [229]. Das Prinzip einer Kontakt-Kopiermaschine wird in Abbildung 5.18 deutlich. Die Belichtung findet an der Kopiertrommel statt. Aufgrund ihrer Krümmung ist es erforderlich, dass der außen geführte Positivfilm einen minimal größeren Perforationslochabstand aufweist als der innen liegende Negativfilm. Bei der Erstellung von Endkopien muss die Kopiermaschine so konstruiert sein, dass gleichzeitig das Bild- und das Tonnegativ auf das Positiv abgebildet werden.

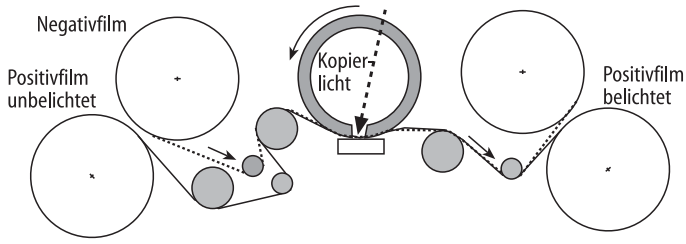


Abb. 5.18. Prinzip des Continous Contact Printers

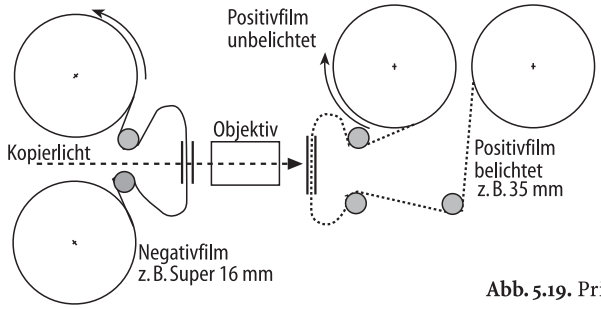


Abb. 5.19. Prinzip des Optical Step Printers

Die Verwendung eines Kopierprozesses mit zwischengeschalteter Abbildungsoptik ist erforderlich, wenn Bildgrößenanpassungen oder besondere Effekte erforderlich sind. Typische Fälle sind die Umsetzung (Blow up) von (Super-)16-Negativmaterial auf 35 mm-Positivfilm oder die Erzeugung eines anamorphotisch gestauchten Bildes. In den meisten Fällen wird bei der optischen Kopierung der schrittweise Transport verwendet, die Kopierleistung liegt mit ca. 6 m/min erheblich niedriger als bei den Massenkopiermaschinen mit kontinuierlichem Lauf. Abbildung 5.19 zeigt das Schema einer optischen Schrittkopiermaschine (Optical Step Printer). Es wird deutlich, dass das Licht aus dem Projektor über eine austauschbare Optik auf die Kamera geführt wird. Der gewünschte Effekt kann durch Abstands- und Abbildungsänderung hervorgerufen werden.

Zur Erzielung hoher Kopierqualitäten, insbesondere bei älterem, verkratztem Ausgangsmaterial, kann die Schrittkopiermaschine mit einer Nasskopiereinrichtung ausgestattet werden. Dabei liegen beide Filme in einem als Wetgate bezeichneten Kopierfenster, das mit einer Flüssigkeit gefüllt ist, die den gleichen Brechungsindex aufweist wie das Filmmaterial. Die von den Kratzern und Schrammen verursachten Unebenheiten werden durch die Flüssigkeit gefüllt und die schadhaften Stellen werden bei diesem Immersion-Print-Prozess nicht mehr abgebildet, da keine Differenzen im Brechungsgrad mehr vorhanden sind. Zum Zwecke der hochqualitativen Archivierung kann ein Film auch so verarbeitet werden, dass die drei RGB-Farbauszüge separat auf S/W-Material kopiert werden. Da auf diese Weise kein Ausbleichen der Farbstoffe zu befürchten ist, kann mit einer erheblich längeren Lebensdauer gerechnet werden.

5.4.3 Die Lichtbestimmung

Der in zwei Schritte aufgeteilte Filmproduktionsprozess, nämlich die Negativbelichtung am Drehort und die Positivbelichtung unter Laborbedingungen, hat zwei Vorteile: erstens, dass das Negativmaterial über die flache Gradationskurve so gestaltet werden kann, dass Fehlbelichtungen leicht ausgeglichen werden können, zweitens gibt es auch einen Spielraum für die Grauwerte und Farben. Am Drehort muss nicht und kann auch kaum darauf geachtet werden, dass eine Farbabgleichung zwischen den verschiedenen Szenen vorgenommen wird.

So ist es günstig, dass die Festlegung der Farbstimmung sowie die Abstimmung dieser Werte für die verschiedenen Szenen unter kontrollierten Bedingungen im Kopierwerk bei der Lichtbestimmung vorgenommen werden kann.

Der technische Prozess ist die Beeinflussung des Kopierlichts. Heute wird meist das additive Verfahren angewandt, bei dem das Licht einer 1000 W-Kunstlichtquelle mittels dichroitischer Filter in die Anteile Rot, Grün und Blau aufgespalten wird. Diese Anteile können über sogenannte Lichtventile in ihrer Intensität verändert werden. Die drei, ggf. veränderten, Anteile werden wieder zusammengeführt und für den Kopierprozess benutzt (Abb. 5.20). Die Lichtänderung ist exakt reproduzierbar, da sie stufenweise eingestellt wird, der Einstellumfang beträgt 50 Schritte [229]. Bei einem Positiv-Gamma von 1 entspricht ein Schritt einer Belichtungsänderung von 0,025 H, zwölf Schritte etwa einer Blendenstufe.

Die Farbanpassung erfordert, dass die Dichteänderungen beachtet werden, die sich aus der Änderung einzelner Anteile ergeben. Sie erfolgt mit Hilfe eines Color Analysers, d. h. das Negativ wird über eine Videokamera abgetastet, und das Bild wird auf einem Videomonitor dargestellt, der in besonderer Weise kalibriert ist, sodass die Farbwirkung weitgehend mit der übereinstimmt, die bei der Projektion des Positivs zu sehen sein wird. Allerdings ist aufgrund der Tatsache, dass das Videobild eine aktiv leuchtende Fläche darstellt, während auf der Leinwand reflektiertes Licht betrachtet wird, die Übereinstimmung auch bei bester Kalibrierung nie vollständig erreichbar, sodass die Handhabung durch erfahrene Lichtbestimmer erfolgen sollte. Diese Personen werden beim Angleichen der Szenen vom Kameramann unterstützt, denn dieser weiß, welche Teile ggf. absichtlich unterbelichtet oder farbstichig bleiben sollen. Die für jede Einstellung gewählten Farbwerte werden gespeichert, sodass beim eigentlichen Kopiervorgang die ausgewählten Werte automatisch aufgerufen werden können.

Zur Kalibrierung steht bei der Lichtbestimmung ein entwickelter Negativ-Kontrollfilm zur Verfügung, der Hell- und Dunkelfelder mit Dichten enthält, die 90% bzw. 2,5% Remission in der Szene entsprechen, weiterhin Abbildungen von Graustufen, Farbflächen und Hauttönen.

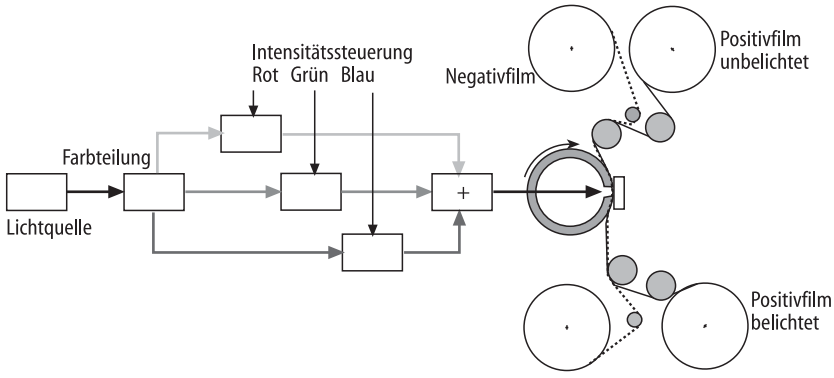


Abb. 5.20. Prinzip der additiven Lichtsteuerung

5.4.4 Blenden, Titel- und Trickarbeiten

Zu den Aufgaben des Filmkopierwerks gehört auch die Anfertigung von speziellen Filmsequenzen, die Titel, Überblendungen oder aufwendigere Tricks enthalten. Sie werden heute meist am Computer, früher aber am Tricktisch und an der Trickkopiermaschine durchgeführt. Der Tricktisch dient der Aufnahme flacher Vorlagen, wie z. B. Bilder oder Titel. In Abbildung 5.21 wird deutlich, dass die Kamera senkrecht steht und auf die Vorlagenfläche gerichtet ist. Die abzufilmende Vorlage kann von oben oder unten beleuchtet werden, sie darf somit auch transparent sein. Der Abstand der Kamera kann ebenso geändert werden wie die Position der Vorlage. Damit lässt sich die Konstruktion für die schrittweise Aufzeichnung von Einzelbildern einsetzen, um z. B. einen Rolltitel zu erzeugen, indem die Vorlage präzise motorisch gesteuert nach jeder Einzelbelichtung um eine definierte Strecke weiterbewegt wird.

Die Trickkopiermaschine (Optical Printer) ist eine weit aufwendigere Konstruktion, die mit sehr hoher mechanischer Präzision arbeitet. Hier wird auf einer optischen Bank eine Kamera montiert, die das Bild eines Projektors aufnimmt. Über die Bildkombination mit Hilfe von Strahlenteilern können auch mehrere Projektoren eingesetzt werden. In den optischen Weg können verschiedene Elemente, wie z. B. Wandermasken, eingesetzt werden, die das Bild in gewünschtem Sinne beeinflussen.

Im einfachsten Fall sind nur Auf- und Abblenden zum Schwarz zu erzeugen, dabei verschwindet das Bild durch eine logarithmisch gesteuerte Reduktion der Öffnung der Sektorenblende bei der Belichtung. Auf ähnliche Weise lassen sich auch Überblendungen von einer Sequenz zur anderen realisieren: hier wird für die erste Sequenz eine lineare Abblende erzeugt. Der Film wird dann um die Blendendauer zurückgespult und die zweite Szene wird mit einer entsprechenden Aufblende darüber kopiert. Unschärfeblenden können durch eine gezielte Verschiebung der Abbildungsoptik aus oder in den Schärfepunkt erzeugt werden.

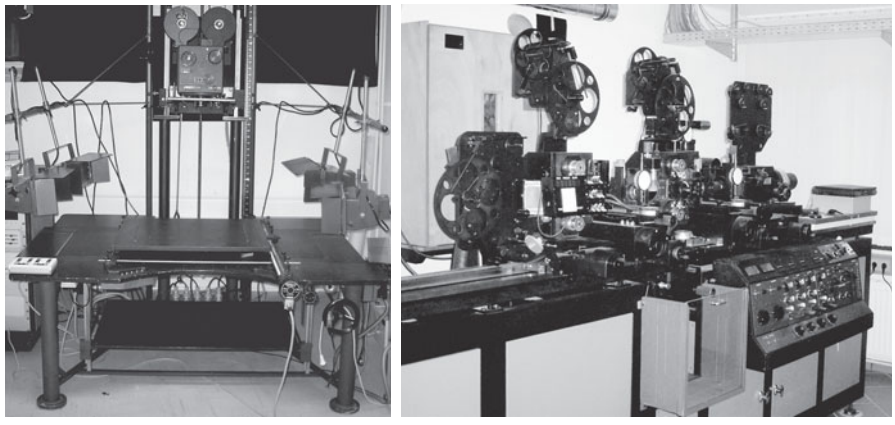


Abb. 5.21. Tricktisch und Trickkopiermaschine

5.5 Filmton

Der Filmton hat eine sehr große Bedeutung, die für die emotionale Wirkung besonders des fiktionalen Films kaum hoch genug eingeschätzt werden kann. Dementsprechend gab es seit dem Ende der Stummfilmzeit – die übrigens oft gar nicht so tonlos war, da wegen der angesprochenen Wirkung häufig Begleittöne z. B. in Form von Livemusik eingesetzt wurden – bereits früh ein großes Interesse am Tonfilm und im Laufe der Zeit viele tontechnische Entwicklungen, die mit dem Kino in Zusammenhang standen.

Allgemein ist festzustellen, dass es gegenüber der Rundfunk- und Fernsehtechnik für die Filmtechnik erforderlich ist, dass die Audioinformation nicht nur einfach übertragen, sondern auch gespeichert wird. Dazu wird sie zunächst durch Mikrofone in elektrische Signale gewandelt, die dann in Änderungen der mechanischen Auslenkung einer Nadel (Nadeltonverfahren), Änderungen der Intensität magnetischer Felder (Magnettonverfahren) oder Änderungen der Lichtintensität (Lichttonverfahren) umgesetzt und jeweils auf ein Medium aufgezeichnet werden, das sich am entsprechenden Aufnahmesystem vorbei bewegt. Heute wird für die Filmtechnik entweder das Magnetton- oder das Lichttonverfahren verwendet – jeweils in analoger und digitaler Form. Beide können auf Film aufgebracht werden und bilden auf demselben Streifen den kombinierten Magnet- oder Lichtton (COMMAG oder COMOPT). Aus Gründen der kostengünstigeren optischen Kopierbarkeit hat sich auf der Wiedergabeseite der Lichtton etabliert. Magnetton hatte hier nur in den 50er Jahren eine relativ große Bedeutung, als mit der Einführung der Breitbildformate auch eine intensivere Wirkung auf der Audioseite erreicht werden sollte. Man verwendete damals bereits Mehrkanalsysteme mit 4 Magnetspuren, die beim Cinemascope-Format beidseitig der beiden Perforationsreihen des 35-mm-Films liegen, und bis zu sechs Magnetspuren beim 70-mm-Film, der deshalb breiter ist als das 65-mm-Aufnahmematerial.

Auf der Aufnahmeseite wird der Ton getrennt verarbeitet, und es kommt das qualitativ hochwertigere Magnettonverfahren (SEPMAG) zum Einsatz – in analoger Form oder digital. Das magnetische Aufzeichnungsprinzip wird in Kap. 8 erörtert. Heute kommen bei Filmaufnahmen in den meisten Fällen Digitalgeräte zum Einsatz, früher wurden vor allem die portablen, analogen Geräte des Typs Nagra verwendete, mit denen in der Regel zwei Tonspuren aufgezeichnet werden können.

Im Bereich der Nachbearbeitung von Filmtönen sind erheblich mehr Spuren erforderlich, weil sich die Filmtonmischung aus den Dialogtönen, Umgebungsgeräuschen, Effekten und Filmmusik zusammensetzt, die zudem heute auch für eine mehrkanalige Wiedergabe abgemischt werden müssen. Da bei der Nachbearbeitung die Synchronität aller dieser Signale eine sehr große Rolle spielt, verwendete man für die analoge Audiopostproduktion im Filmbereich Perfoband (Cordband), d. h. perforierten Film von 16 mm, 17,5 mm oder 35 mm Breite, der mit 2 bis 6 Magnetspuren versehen ist. Aufgrund der Perforation kann hier die Synchronisation mehrerer sogenannter Perfoläufer mit dem Filmbild auf einfache Weise mechanisch sichergestellt werden. Auch im Bereich der Nachbearbeitung werden immer mehr Digitalgeräte verwen-

det. Die Aufzeichnung erfolgt hier auf Festspeichern oder auf Festplatten der Computertechnik. Die Digitaltechnik bietet die Möglichkeit, alle zu mischenden Töne zunächst zu speichern und für einen gleichzeitigen Zugriff bereitzuhalten. Sie stehen dann für den Tonschnitt mit einem Computer basierten Editingsystem, wie z. B. Avid Media Composer, zur Verfügung und können ohne weitere Wandlung der anschließenden Tonmischung mit einem Audiobearbeitungssystem, wie z. B. Pro Tools, übergeben werden. Dabei sollte vor allem dafür gesorgt werden, dass alle beteiligten Systeme mit denselben Datenformaten arbeiten können.

5.5.1 Das Lichttonverfahren

Beim Lichttonverfahren wird das elektrische Audiosignal in eine veränderliche Schwärzung des Filmmaterials umgesetzt. Dies kann einfach durch die Variation der Lichtintensität in Abhängigkeit von der Signalspannung geschehen. Diese Art wird als Intensitätsschrift bezeichnet und war die zuerst angewandte. Später kam dann die Transversal- oder Zackenschrift zum Einsatz, die den Vorteil bietet, dass keine Graustufen auftreten, sondern nur Schwarzweißkontraste, die einfacher zu reproduzieren sind als die Grauwerte der Intensitätsschrift. Bei der Zackenschrift wird in Abhängigkeit von der Signalspannung die vollständig geschwärzte Fläche mehr oder weniger groß, während der Restbereich transparent bleibt (Abb. 5.22).

Für die Aufnahme wird eine Lichttonkamera benutzt, die im einfachen Fall eine Schlitzblende (Spalt) enthält, die mechanisch mehr oder weniger stark abgedeckt wird (Abb. 5.22). Die Blende oder ein Spiegelsystem kann elektromechanisch angetrieben werden. Wenn die Abdeckung mittels eines Dreiecks geschieht, sind dabei nur kleine Auslenkungen erforderlich [228]. Auf diese Weise entsteht die sogenannte Doppelzackenschrift (bilateral), bei der die Änderung der Signalamplitude deutlich sichtbar ist. Für zweikanalige bzw. stereophonische Aufzeichnungen werden im Bereich der Monotonspur zwei schmale Audiospuren nach dem gleichen Prinzip untergebracht. Unter den verschiedenen Typen der Transversalschrift, wie Uni-, Bi- und Multilateralschrift, wird der dual bilaterale Typ heute am häufigsten verwendet (Abb. 5.23).

Das Lichttonverfahren bietet den großen Vorteil, dass Bild- und Toninformationen gleichzeitig auf optischem Wege kopiert werden können. Als zweiter

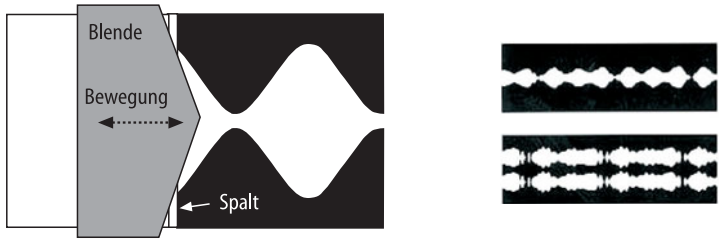


Abb. 5.22. Schlitzblende zur Erzeugung der einfachen und dualen Bilateralschrift

Vorteil ist die Tatsache zu nennen, dass der Lichtton eine höhere Dokumentenechtheit aufweist als der Magnetton, weil Letzterer erheblich einfacher gelöscht werden kann. Der große Nachteil des Lichttonverfahrens ist der schlechte Geräuschspannungsabstand, der unter anderem vom unregelmäßig verteilten Filmkorn hervorgerufen wird. Hinzu kommt die Gefahr des Auftretens des sogenannten Donnereffekts, der aber heute beim Kopierprozess weitgehend eliminiert werden kann. Der Donnereffekt resultiert aus nichtlinearen Verzerrungen, die dadurch entstehen, dass bei hohen Frequenzen die Auslenkungsspitzen aufgrund von Lichtdiffusion zu dicht beieinander liegen.

Die Weiterentwicklungen der Lichttonaufzeichnung führten zu einem besseren Störsignalabstand. Dies gelang durch die Nutzung eines Laserstrahls, der aufgrund seiner hohen Intensität die Verwendung von feinkörnigem Lichttonnegativmaterial erlaubt. Bei einer solchen Laser-Lichttonkamera wird der scharf gebündelte Strahl sägezahnförmig trägheitslos über eine akustooptische Ablenkeinheit mit einer Frequenz von 96 kHz ausgelenkt, wobei die Amplitude vom Audiosignal abhängt. Auf diese Weise entsteht die gewohnte Transversalschrift. Die Geräte wurden oft so gebaut, dass das Lichttonnegativ zusätzlich zu der Stereospur über ein Leuchtdioden-Array belichtet wird, um Audiosignale aufzuzeichnen, die nach dem im nächsten Abschnitt beschriebenen Dolby-Digital-Verfahren codiert sind.

Die analoge Lichttonspur befindet sich zwischen den Filmbildern und der Perforation (Abb. 5.23) [90]. Für die Wiedergabe wird sie mit konstanter Intensität über einen Spalt durchstrahlt. Die durch die Filmschwärzung veränderte Lichtintensität wird durch eine Fotozelle registriert, die eine entsprechende Spannung abgibt, die wiederum einer Audio-Verstärkeranlage zugeführt werden kann. Bei Stereospuren wird eine Doppeloptik mit zwei Fotozellen verwendet. Falls eine Stereospur von einem alten Monoabtastgerät erfasst wird, entsteht das Monosignal direkt durch die Summenbildung bei der Abtastung.

Ein Problem bei der Wiedergabe ist der gleichmäßige Lauf der Tonspur und der intermittierende Antrieb für den Bildtransport. Bild- und Tonoptik sind daher räumlich getrennt und werden über eine Filmschleifenbildung und träge Schwungmassen mechanisch entkoppelt. Die Lichttonabtastung erfolgt in einem genormten Abstand von 21 Bildern hinter dem Bildfenster, es existiert daher ein entsprechender Bild/Tonversatz, der bereits beim Kopiervorgang von Bild und Ton auf das gemeinsame Positiv berücksichtigt wird.

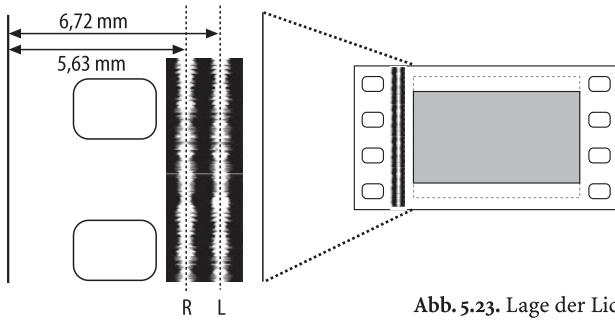


Abb. 5.23. Lage der Lichttonspuren auf dem Positiv

5.5.2 Mehrkanaltonverfahren

Mehrkanalsysteme dienen bei der Wiedergabe der besseren Einbeziehung der Zuschauer in das Kinoerlebnis. Die emotionale Wirkung wird einerseits durch die Wiedergabe mit hoher Qualität, d. h. in großer Lautstärke, mit hoher Dynamik und ausgedehntem Frequenzbereich sowie hohem Tiefbassanteil erreicht. Andererseits durch die Möglichkeit der Schallquellenortung und die Einbeziehung der Hörer in ein Umgebungsgeräuschfeld, das sie wie in der Natur auch von der Seite und von hinten erreicht.

Die menschliche Hörwahrnehmung hat eine Lokalisationsfähigkeit, die in der Ebene in Sichtrichtung am ausgeprägtesten ist. Die Richtungswahrnehmung beruht darauf, dass das Gehirn einerseits Pegel- und andererseits Laufzeitdifferenzen auswertet, die zwischen den beiden Ohren entstehen. Für die technische Umsetzung ist dabei interessant, dass das Gehör sich täuschen lässt und eine Schallquelle auch dort orten kann, wo sich keine Quelle befindet. Die Wahrnehmung derartiger Phantomschallquellen wird bei der Stereophonie ausgenutzt. Dabei wird ein Hörer, der im etwa gleichschenkligen Dreieck vor zwei gleichen Lautsprechern platziert wird, die Veränderung der horizontalen Lage einer Schallquelle wahrnehmen, wenn beide Lautsprecher dasselbe Signal mit unterschiedlichem Schallpegel abstrahlen.

Dieses einfache System ist für die Kinotonwiedergabe nicht geeignet, da das Gefühl, dass die Schallquelle bei gleichem Pegel in der Mitte liegt, nur an Orten erreicht wird, die mittig zwischen den Lautsprechern liegen. Im Kino ist das Auditorium aber über eine große Fläche verteilt. Damit auch an ungünstigen Sitzplätzen eine ausreichende Mittenortung – die ja gerade für die Filmdialoge von sehr großer Bedeutung ist – erreicht werden kann, wird das Stereosystem für den Kinoeinsatz durch einen Centerkanal erweitert, über den das Mittensignal mit Hilfe eines zentral hinter der perforierten Bildleinwand positionierten Lautsprechers abgestrahlt wird. Die Leinwand ist für diesen Zweck mit feinen Löchern versehen, durch die der Schall hindurchtreten kann.

Für die weitere Einbeziehung der Hörer in das Schallfeld wurde bereits früh mit Lautsprechern experimentiert, die sich hinter und seitlich von den Hörern befinden. Hierdurch soll meist keine direkte Ortung der Quellen, sondern nur eine Wiedergabe von Effektklängen oder eine Einhüllung durch Geräusche, die Ambience, erreicht werden. Signale mit einer solchen Funktion werden heute als Surroundsignale bezeichnet. Erste derartige Mehrkanalsysteme erlangten in den 50er Jahren Bedeutung, als durch die Einführung der Breitbildformate und der verbesserten Tonwiedergabe in den Kinos der Konkurrenz durch das Fernsehen Paroli geboten werden sollte. Das damals eingeführte Cinemascopesystem nutzte vier Kanäle, d. h. Left, Center, Right (L, C, R) und einen Effektkanal. Die größten Systeme, wie Cinerama, arbeiteten damals mit einem separaten 35-mm-Magnetband. Die Bildbreite war so groß, dass wegen der genannten Ortungsprobleme zwischen dem Center und den Seitenkanälen noch jeweils ein halblinker und halbrechter (HL, HR) Lautsprecher platziert wurde. Signale von einer sechsten und siebten Spur wurden von der rechten und linken Seitenwand abgestrahlt. In den 60er und 70er Jahren wurde dieser Aufwand nicht mehr betrieben und man kehrte vielfach zum Monoton zurück.

5.5.2.1 Dolby Stereo

Diese Situation änderte sich nachhaltig erst mit der Einführung des Dolby-Stereosystems in der Mitte der 70er Jahre. Der Erfolg des Konzepts der Dolby Laboratories beruht auf dem wegen der einfachen Kopierbarkeit ökonomischen Lichttonverfahren, mit dem ein hochwertiger Kinoton produziert wird. Dabei wurde erstens das Störabstandsproblem durch eine Kompandertechnik erheblich gemindert (Dolby Noise Reduction, Dolby NR). Ein zweiter Punkt für den Erfolg des Dolby-Systems war dann die konsequente Abwärtskompatibilität, die sich über die gesamte Systemfamilie vom Mono- bis zum Mehrkanal-Digitalton hinzieht [50].

Die maximale Systemdynamik, d. h. das geräteabhängige Verhältnis von leisestem zu lautestem Signal, wird vom Signal-Rauschabstand bestimmt, der beim Lichttonverfahren sehr gering ist. Der Signal-Rauschabstand kann verbessert werden, indem beim Audiosignal bei der Aufnahme eine frequenzbereichs- und pegelabhängige Signalanhebung vorgenommen und bei der Wiedergabe entsprechend rückgängig gemacht wird. Das Signal wird dazu aufnahmeseitig komprimiert (verringerte Dynamik) und wiedergabeseitig entsprechend expandiert (Abb. 5.24), wobei auch das Rauschen abgesenkt wird. Die Kombination von Kompression und entsprechender Expansion führte zu der Bezeichnung Kompander. Die zugehörigen Kennlinien müssen dabei standardisiert sein. Zunächst wurde das Dolby A-Verfahren entwickelt, das mit verschiedenen Dynamikeinschränkungen für unterschiedliche Frequenzbänder arbeitet und dementsprechend eine aufwendige Kalibrierung erfordert. Es folgten die vereinfachten Verfahren Dolby B und C, die aber nicht für den Filmton, sondern vor allem bei Heimcassettenrecordern verwendet wurden. Für die professionelle analoge Audioaufzeichnung entstand dann das Dolby SR-System (Spectral Recording), das sowohl beim Einsatz in analogen Tonstudios als auch im Kino eine hervorragende Qualität bietet.

Die Dolby-Rauschunterdrückung wurde für die Verwendung im Kino so geschickt mit einer mehrkanaligen Wiedergabe verbunden, dass bezüglich der Aufzeichnungsmedien nur zwei Kanäle zur Verfügung stehen müssen. Das analoge Verfahren nach diesem Prinzip wird mit Dolby Stereo bezeichnet. Es arbeitet mit zwei Lichttonspuren, die bei gemeinsamer Abtastung durch ein monophon arbeitendes Wiedergabegerät problemlos ein Monosignal ergeben.

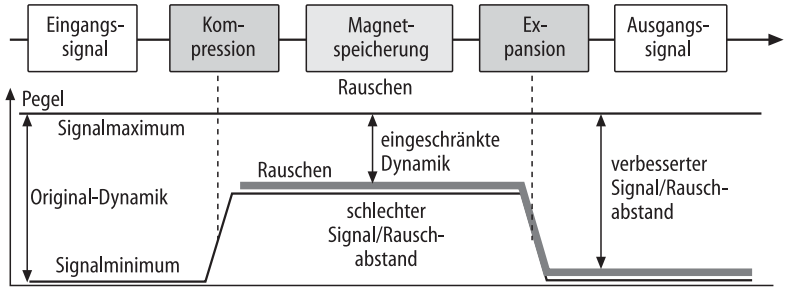


Abb. 5.24. Funktionsprinzip der Rauschunterdrückung

Dolby Stereo beschreibt aber keine gewöhnliche Stereophonie, sondern eine Matrixcodierung (Motion Picture Matrix), mit der in den beiden Audiospuren zusätzlich zu den Informationen für Links und Rechts (L und R) die Informationen für das Center- und das Surroundsignal verschlüsselt übertragen werden. Die vier Kanäle sind so gewählt, dass die Tonwiedergabesysteme aus den Zeiten der Cinemascope-Filme genutzt werden konnten.

Abbildung 5.25 zeigt die Dolby Stereo-Codierung. Die Informationen für die Lautsprecher links und rechts vorn werden unverändert auf die beiden Transportkanäle L_t und R_t gegeben, während das Centersignal zunächst um 3 dB abgesenkt wird und dann den Kanälen L_t und R_t in gleichem Maße zugesetzt wird. Das Surroundsignal wird ähnlich behandelt, allerdings wird es mit einer Phasenverschiebung von + 90° dem linken und von – 90° dem rechten Kanal zugemischt. Auf diese Weise bleibt auch das vierkanalige Signal monokompatibel, denn bei einer Summierung heben sich die phasenverschobenen Anteile auf. Im Dolby Stereo Decoder werden die Links- und Rechtsanteile wieder direkt den entsprechenden Wiedergabekanälen zugeordnet und der Centerkanal wird aus der Summenbildung C = L + R gewonnen. Das Surroundsignal ergibt sich aus der Differenz L – R. Damit die diesem Kanal zugehörigen Atmosphärengeräusche nicht direkt geortet werden können, wird das Surroundsignal durch ein Tiefpassfilter auf 7 kHz bandbegrenzt [90].

Die Variante von Dolby Stereo für den Heimbereich wird Dolby Surround genannt, hier ist eine Kombination mit dem Rauschunterdrückungsverfahren nach Typ B enthalten. Dolby Surround kann mit einem Surround Pro Logic Decoder betrieben werden, der das wesentliche Problem von Dolby Stereo, nämlich die schlechte Kanaltrennung, verringert. Der Decoder mindert vor allem das störende Übersprechen in den Surroundkanal und verbessert damit die Richtungswahrnehmung. Zunächst wird dazu das Surroundsignal gegenüber dem Hauptsignal um 20 ms verzögert, da nach dem Gesetz der ersten Wellenfront die Lokalisation vom zuerst eintreffenden Schallereignis bestimmt wird, das dann vorne liegt. Zusätzlich wird eine Schaltung eingesetzt, die die Aufgabe hat, das codierte Signal ständig auf eine dominante Schallquellenrichtung hin zu untersuchen. Mit Hilfe steuerbarer Verstärker werden entsprechend den Ergebnissen die Intensitäten der Kanäle so geregelt, dass eine Verstärkung der festgestellten Dominanz auftritt.

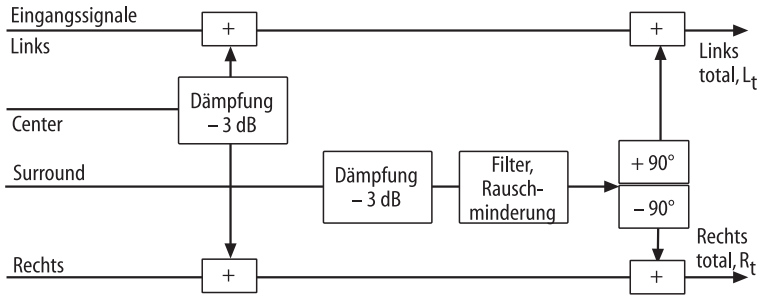


Abb. 5.25. Dolby Stereo-Codierung

5.5.2.2 Dolby Digital

Beim Übergang zur digitalen Tonaufzeichnung wurde bei Dolby festgelegt, dass weiterhin ein optisch kopierbares Aufzeichnungsverfahren verwendet werden sollte, das die Kompatibilität zum analogen System gewährleistet. Letzteres gelingt durch die Vorschrift, dass neben den neuen Digitalsignalen auf allen Filmkopien immer auch die zwei analogen Lichttonspuren verfügbar sein müssen.

Für die Speicherung der Dolby Digital-Daten werden die Flächen zwischen den Perforationslöchern belegt (Abb. 5.26). Sie reicht aus, um mit Laserdioden die robuste Aufzeichnung einer Matrix aus 76 x 76 Punkten zu gewährleisten, mit denen 5776 Bit dargestellt werden können. Da ein 35-mm-Film mit 24 Bildern pro Sekunde läuft, die jeweils mit vier Perforationslöchern versehen sind, folgt daraus eine aufzeichenbare Bitrate von 554,5 kbit/s. Davon werden im Kino 320 kbit/s genutzt. Diese Datenrate wird für fünf Audiokanäle und einen sechsten, auf 120 Hz bandbegrenzten Subbasskanal zur Verfügung gestellt.

Bereits die Datenrate eines einzigen Kanals überschreitet die oben genannte Bitrate, so dass die Gesamtinformation nur mit Hilfe eines Datenreduktionsverfahrens aufgezeichnet werden kann. Das bei Dolby verwendete Verfahren heißt AC3 und arbeitet nach den in Abschn. 3.8.2 beschriebenen Prinzipien. Die Datenreduktion ist so leistungsfähig, dass es möglich wurde, auch die Surroundbeschallung stereophonisch auszulegen. Bei Dolby Digital wird damit im Kino die sogenannte 5.1-Wiedergabe realisiert. Die fünf Kanäle stehen für die bekannten Frontsignale Left, Center und Right sowie für die zwei Surroundsignale Left und Right Surround (L_s und R_s) zur Verfügung. Die abgetrennte 1 bezeichnet den optional einsetzbaren Tiefbasskanal, der zur Verstärkung akustischer Effekte als Low Frequency Effects-Kanal (LFE) verwendet werden kann.

Die nächste Variante der Dolby-Familie wird mit Dolby Digital Surround EX bezeichnet. Es handelt sich dabei um eine zu Dolby Digital kompatible Erweiterung mit einem dritten Surroundkanal für die direkte Beschallung von hinten (Back Surround, B_s). Damit entsteht also ein 6.1-Format. Darauf entwickelte Dolby ein 7.1-System, bei dem auch der Back-Surroundkanal nochmal in zwei Teile aufgespalten wurde, d. h. neben LCR gibt es die vier Kanäle Left und Right Side Surround und Left und Right Rear Surround. Die neueste Entwicklung heißt Dolby Atmos und verwendet zusätzlich zwei Reihen von Lautsprechern die sich über den Köpfen der Zuschauer von der Leinwand zur Rückseite des Raums hinziehen und damit ein 3D-Audio-System darstellen (Abb. 5.27) [91].

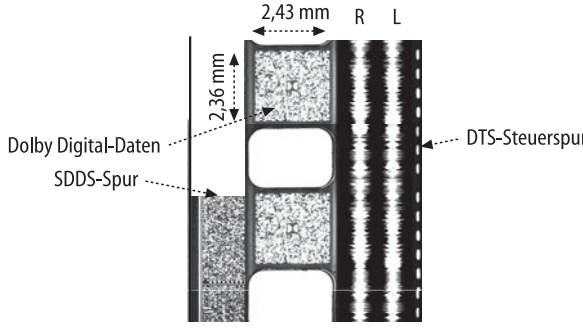


Abb. 5.26. Lagen der verschiedenen Digitaltonspuren

5.5.2.3 DTS und SDDS

Als Alternative zum Mehrkanalsystem von Dolby gibt es das Digital Theater System DTS, das mit einem Tonträgermaterial arbeitet, das vom Filmmedium getrennt ist. Bei DTS werden sechs Audiokanäle mit Hilfe eines Datenreduktionsverfahrens auf einer Doppel-CD-ROM gespeichert. Aufgrund der vergleichsweise hohen Speicherkapazität kann der Datenreduktionsfaktor dabei geringer sein als bei Dolby Digital. Auf dem Filmmaterial ist nur eine schmale, optisch lesbare Steuerspur für die Bild/Tonsynchronisation erforderlich, die zwischen der analogen Lichttonspur und dem Bildfeld Platz findet, sodass sie parallel zu einer Dolby Digital-Tonspur zur Verfügung steht (Abb. 5.26).

Als weitere Alternative ist das System Sony Dynamic Digital Sound (SDDS) mit der Datenreduktion ATRAC zu nennen. Hier stehen als 7.1-System acht Audiokanäle zur Verfügung, die auch die halbrechten und halblinken Frontlautsprecher versorgen können. Für diese Digitalinformation wird der Platz an den Filmrändern an beiden Seiten außerhalb der Perforation genutzt (Abb. 5.26).

Unter den vielen Begriffen für Mehrkanalsysteme taucht auch immer wieder die Abkürzung THX auf. Diese bezeichnet keine eigene Tonaufzeichnungsform, sondern die genaue Definition elektroakustischer Parameter einschließlich der Abhörbedingungen. Diese sind oft viel kritischer für die Qualität der Darbietung als die Toncodierungsarten, da verschiedene Räume die abgestrahlten Schallwellen in unterschiedlicher Form frequenzabhängig bedämpfen und reflektieren. Starke Reflexion steigern die Halligkeit des Raumes, die durch die sogenannte Nachhallzeit angegeben wird, d. h. durch den Zeitraum in dem der Schallpegel nach Abschalten der Quelle um 60 dB gesunken ist. Mit steigender Nachhallzeit sinkt die Sprachverständlichkeit. Daher schreibt das THX-System vor, dass die Nachhallzeit im Kino nicht mehr als 0,2 s betragen darf. Neben den raumakustischen Bedingungen sind bei THX weitere elektroakustische Parameter festgelegt, z. B. der Frequenzgang der Frontlautsprecher, die im Kino ein besonderes Problem darstellen, da sie hinter der perforierten Bildwand angebracht sind, die insbesondere die höheren Frequenzen dämpft.

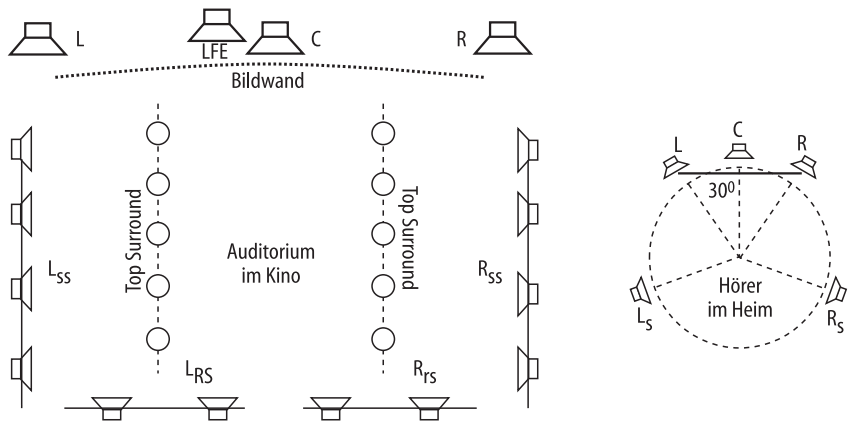


Abb. 5.27. Lautsprecheranordnungen für Dolby Digital und DTS im Kino und im Heimbereich

5.6 Der digitale Film

Filmbilder in Datenform wurden zunächst vor allem im Bereich der Postproduktion gebraucht, für Bildmanipulationen, die mit konventionellen Mitteln im Kopierwerk nicht herstellbar sind. Mit sinkenden Kosten für die digitalen Postproduktionsmittel wurden aufwendige Tricks bei immer mehr Filmen realisiert. Der Transfer des Filmbildes in die digitale Ebene erfolgt dabei mit Filmabtastern, der Rücktransfer der digitalen Daten mit Hilfe von Filmbelichtern. Um das ausbelichtete Material möglichst unauffällig wieder in den Film integrieren zu können, müssen die drei Schritte Abtastung, Bearbeitung und Belichtung mit hoher Auflösung und in optimaler Abstimmung aufeinander erfolgen.

Zu Beginn der Entwicklung entstanden geschlossene Systeme, die alle drei Bearbeitungsstadien integrierten, sodass die Daten bis zur Belichtung nicht ausgegeben werden mussten. Die bekanntesten Beispiele sind das Cineon-System von Kodak und Domino von Quantel, die nicht mehr vertrieben werden. Heute wird eher arbeitsteilig gearbeitet, d. h. Abtastung, Bearbeitung und Belichtung werden an separaten, für die jeweilige Aufgabe optimierten Systemen von Spezialisten vorgenommen.

Das Gesamtkonzept zum digitalen Film wird als Digital Intermediate (DI) oder Digital Lab bezeichnet, da es viele Prozesse abdeckt, für die sonst das Kopierwerk erforderlich ist. Im Vergleich zur klassischen Filmproduktionskette zeigte sich, dass beim DI weniger Filmduplikate erforderlich wurden, sodass ein Qualitätsgewinn entsteht. Darüber hinaus sind wesentlich mehr Bildmanipulationsmöglichkeiten gegeben als im klassischen Filmlabor.

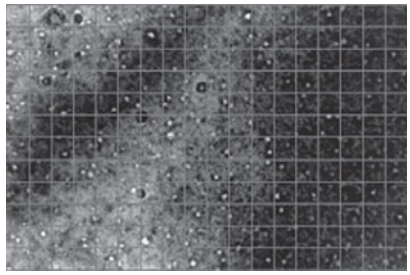
Alternativ zum Weg über das Filmnegativ können Bilder für den Kinobereich heute auch direkt elektronisch mit HD-Kameras aufgenommen werden, oder es werden am Computer generierte Bilder in den Film eingefügt [70]. Dabei spielen vermehrt nicht nur die ausgefeilten Trickmöglichkeiten eine Rolle, sondern auch ökonomische Erwägungen. Dies gilt auch hinsichtlich der mit `Digital Cinema` bezeichneten Form der Filmdistribution, bei der die Digitalbilder gar nicht mehr auf Film belichtet werden, sondern digital ins Kino gelangen und dort mittels elektronischer Projektoren gezeigt werden.

Damit wäre der gesamte Filmbereich digitalisiert. Ein Problem der rein digitalen Kette ist allerdings, dass sie nicht die Formatunabhängigkeit, Langzeitstabilität und internationale Austauschbarkeit gewährleistet, die bei Filmmaterial seit langem gegeben ist, sodass damit zu rechnen ist, dass der Film kurzfristig noch einige Bedeutung behalten wird.

Die wesentlichen Beurteilungskriterien für die Qualität, mit denen Filmbilder digital dargestellt werden können, sind das Auflösungsvermögen, der Kontrastumfang, und die Umsetzung von Grauwerten und Farben. Die Filmparameter werden bezüglich Negativfilmen betrachtet, da dieses Material abgetastet und auch bei der Ausbelichtung verwendet wird. Konkrete Beispiele beziehen sich hier auf die weit verbreiteten Kodak-Filme Eastman EXR 100T (5248), Kodak Vision 200T (5274) und Kodak Vision 500T (5279) mit den Empfindlichkeiten 100, 200 und 500 ASA. Für die Belichtung der Digitaldaten auf Film wird vorwiegend besonders feinkörniges und damit unempfindliches Material der Typen EXR 50D (5245) und Intermediate-Material des Typs 5244 benutzt.

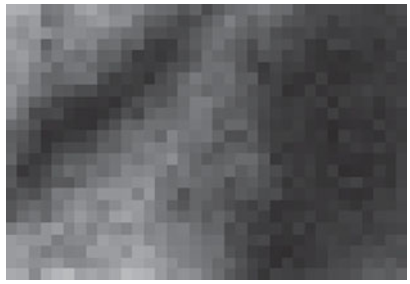
5.6.1 Die Bildauflösung

Bezüglich der Bildauflösung sollen zum Ersten im Digitalbereich so viele Bildpunkte zur Verfügung stehen, dass damit die Filmauflösung, gemessen in Linienpaaren pro Millimeter (lp/mm) oder als MTF, erreicht wird. Zweitens ist zu bedenken, dass der sogenannte Filmlook stark vom Filmkorn geprägt ist, das die besondere Eigenschaft der statistischen Verteilung aufweist (s. Abschn. 5.3.5). Beim Farbfilm treten an die Stelle des Filmkorns die Farbstoffwolken, die mit sinkender Filmempfindlichkeit kleiner werden. Die Größen betragen bei 500-ASA-Material ca. 8 µm, bei 200 ASA ca. 5 µm und bei 100 ASA ca. 3 µm. Um die Größenordnung der Farbstoffwolken zu erreichen, müssen bei einem Übergang zwischen Film und Digitaldaten die abgetasteten bzw. aufbelichteten Bildpunkte also in der Größenordung von 5 µm liegen, d. h., dass ein 35 mm Film mit full apertur von 24,9 mm x 18,7 mm durch 4980 x 3752 Bildpunkte repräsentiert werden muss. Bei 3-µm-Farbstoffwolken wären mehr als 8000 Bildpunkte horizontal erforderlich. Bei dieser Betrachtung ist zu bedenken, dass die Zerlegung des Bildes nach einem regelmäßigen Schema mit der Entstehung von Alias-Strukturen verbunden sein kann, die nur unter Beachtung des Abtasttheorems ausgeschlossen werden können. Nach dem Abtasttheorem müssen die Abtaststrukturen doppelt so fein sein wie die feinsten aufzulösenden Bildelemente. Abbildung 5.28 zeigt in den Simulationen einer 2k-, 4k- und 8k-Abtastung, die auf das vergrößerte Bild des 100 ASA-Materials angewandt wurden, dass die Farbstoffwolken nur bei 8k sichtbar sind und bei 4k-Auflösung verschwinden [151].

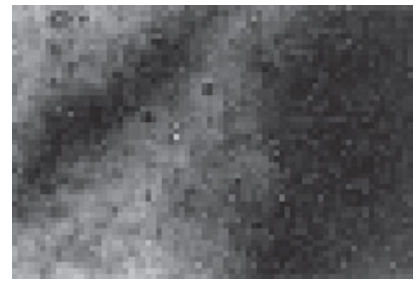


Kodak 5248 (100 ASA) aufgerastert

Kodak 5248 mit simulierter 2k-Auflösung



Kodak 5248 mit simulierter 4k-Auflösung



Kodak 5248 mit simulierter 8k-Auflösung

Abb. 5.28. Simulation der Abtastung von Filmmaterial mit verschiedenen Auflösungen [151]

Ohne Berücksichtigung der Farbstoffwolken kann die erforderliche Bildpunktanzahl in der digitalen Ebene auch mit Hilfe der Modulationstransferfunktion ermittelt werden. Nimmt man wieder den Kodak 5248-Film als Beispiel, so zeigt das Datenblatt (Abb. 5.16), dass die maximalen MTF-Werte bei Blau erreicht werden. Bei 50% MTF kann hier ein Auflösungswert von 120 lp/mm abgelesen werden. Um diese zu erreichen, sind unter Berücksichtigung eines Faktors zwei für das Abtasttheorem 2 x 24,9 mm x 120 lp/mm x 2/paar = 11 950 Pixel, d. h. eine 12k-Abtastung erforderlich. Bei Orientierung an der MTF für Grün mit 80 lp/mm genügt eine 8k-Abtastung, der gleiche Wert gilt für 500-ASA-Material Kodak 5279 bei Blau.

Die Feinheit der Bildauflösung hängt aber nicht nur vom Filmmaterial (Original Camera Negative, OCN), sondern von dem gesamten technischen System, insbesondere von der Kamera ab. Hier tritt vor allem die Modulationstransferfunktion der Objektivs hervor, die auch bei Verwendung von Prime Lenses mit $MTF_{ges} = MTF_{Film} \cdot MTF_{Kamera}$ einen Einfluss hat, der die Gesamtauflösung auf ca. 30 lp/mm bei $MTF_{ges} = 50\%$ senkt.

Falls der Abtastprozess, der ja selbst multiplikativ in die Gesamt-MTF eingeht, das Ergebnis nicht weiter verschlechtern soll, ist aus diesen Betrachtungen insgesamt zu folgern, dass der Film in der digitalen Ebene mit mindestens 4k, d. h. 4096 Bildpunkten in der Bildhorizontalen, vorliegen sollte. Bei 4k-Auflösung ergibt sich für ein 35 mm-Filmbild der maximalen Größe (full apertur, 24,576 mm x 18,67 mm) eine Anzahl von 3112 Bildpunkten vertikal [125]. Nach dem Rücktransfer der Digitaldaten auf Film sind schließlich weitere Auflösungsverluste durch den Kopierprozess und die Projektion zu erwarten, die nach Einschätzung der Fa. Cintel mit einer Auflösungsäquivalenz von 2k für die Kopie bzw. 1,5k für die Projektion gleichgesetzt werden können.

Mit dem letztgenannten Argument, aber vor allem auch aus Kostengründen, wird in der Praxis bisher selten mit 4k-Daten gearbeitet. In den meisten Fällen wird eine 2k-Auflösung der Filmbilder verwendet. Einem 35-mm-full-aperture-Bild sind horizontal 2048 Bildpunkte und vertikal 1556 oder auch 1536 Pixel zugeordnet. Unter Berücksichtigung der verschiedenen Filmbildformate werden diese Werte eingeschränkt. Für 1:1,85 selten auf 2048 x 1100 und meist auf 1998 x 1080 Pixel. Wenn ein HD-Bild mit 1920 x 1080 vorliegt, wird das Bild dafür auf eine Breite von 1998 Pixel skaliert und oben und unten beschnitten. Alternativ kann auch die Differenz zwischen 1998 und 1920 durch schwarze Ränder seitlich am Rand ausgeglichen werden. Weitere Größen sind 2048 x 1080 (Nutzung des HD-Formats ohne Rücksicht auf Filmformate) und 2048 x 858 (für Cinemascope bei 1:2,39). Um dem 4k-Format Rechnung zu tragen, werden die hier genannten Bildpunktzahlen alle verdoppelt.

Falls die Daten für die Filmbelichtung vorbereitet werden, geht man oft davon aus, dass die 2048 Bildpunkte der Fullframe-Breite zugeordnet sind, dann folgt für ein 22 mm breites Bild, das unter Berücksichtigung des Tonspurbereichs belichtet wird, dass nur 1833 Pixel horizontal und unter Berücksichtigung des Formats 1:1,85 (Standard 1,85) 991 Pixel vertikal genutzt werden. Diese Werte variieren je nach Ausbelichtungsbetrieb etwas. Es ist empfehlenswert, dort die gewünschten Werte zu erfragen, damit das Bild nicht noch einmal skaliert werden muss.

Tabelle 5.4. DPX-Codewortzuordnung

Grauwerte	D_s	Codewert	Grauwerte	D_s	Codewert
D-Min	0,00	95		1,00	593
	0,02	107	90%-Weiß	1,18	685
	0,05	122		1,36	777
2%-Schwarz	0,17	180		1,49	839
	0,30	244	D-Max	1,61	900
	0,49	341		1,73	961
18%-Grau	0,75	470		1,86	1023

5.6.2 Die Grauwertauflösung

Bei Negativmaterialien, die für die Aufnahme in der Kamera und zur Belichtung verwendet werden, tritt ein maximaler Dichteunterschied $\Delta D = 1{,}7$ auf. Um allen Eventualitäten gerecht zu werden, wird für die Betrachtung des digitalen Negativs oft eine Minimaldichte D_{min} von 0,2 und ein $\Delta D = 2$ festgelegt (Kodak). Die Digitalisierung, d. h. die Zuordnung von Abtastwerten zu einer begrenzten Anzahl von Grauwertklassen, erzeugt einen Quantisierungsfehler, für den das menschliche Auge besonders im dunkleren Graubereich empfindlich ist, da es dort ein höheres Differenzierungsvermögen hat als im hellen. Damit bei der Projektion des Positivs auch im Bereich des hohen Vermögens der Quantisierungsfehler unsichtbar bleibt, ist eine Auflösung eines jeden Abtastwertes in ca. 1000 Stufen bzw. eine Darstellung mit 10 Bit erforderlich. Abbildung 5.29 zeigt die Auswirkung unterschiedlicher Quantisierungen am Beispiel eines Graukeils, in den Grenzen der Druckauflösung des Buches.

Bei der Umsetzung eines Dichteumfangs von $\Delta D = 2$ ergibt sich aus 10 Bit eine Dichteabstufung von ca. 0,002 D pro Codewort. Man spricht in diesem Fall von 10 Bit logarithmischer Daten, da die Zahlen den logarithmisch gebildeten Dichteunterschieden des Negativfilms zugeordnet werden. Die log-Daten-Darstellung wurde als Format beim Cineon-System von Kodak verwendet und später für das DPX-Fileformat übernommen, das heute eine große Rolle bei der Filmdatenspeicherung spielt. Die Codewortzuordnung zeigt Tabelle 5.4.

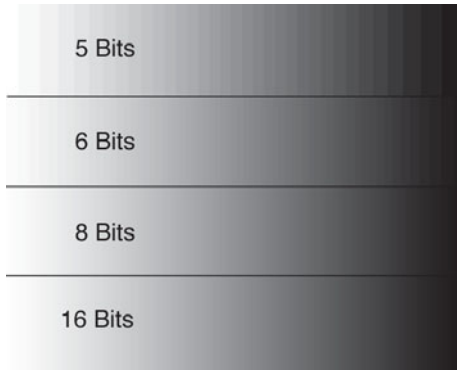


Abb. 5.29. Grauwertdarstellung mit verschiedenen Quantisierungen

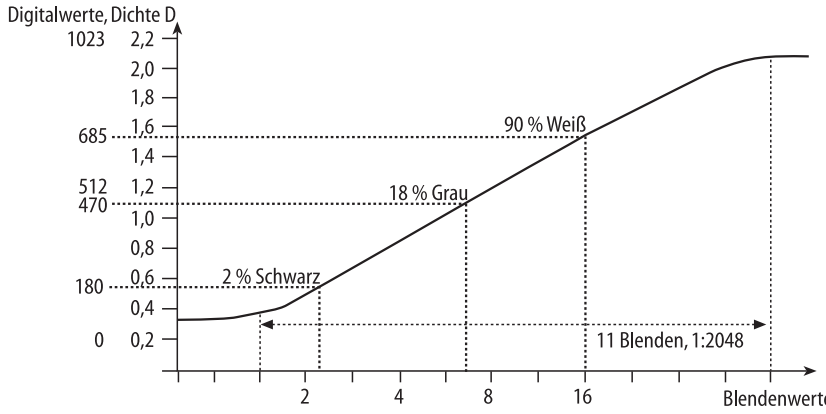


Abb. 5.30. Erfassung des Dichtebereichs beim Filmnegativ und Zuordnung zu Digitaldaten

Die Zuordnung der Codeworte gehorcht in etwa der Beziehung [26]:

$$\text{DPX-Wert} = (95 + 500 \cdot D_s),$$

wobei D_s für Dichte über dem Schleier (Density above base) steht und sich die Zahl 500 aus der Stufenbreite 0,002 D pro Codewort für $D_s = 1$ ergibt. Dem Filmschleier ist somit die Zahl 95, dem 2%-Schwarz die Zahl 180, dem 18%-Bezugsgrau die Zahl 470 und dem 90%-Weiß der Wert 685 zugeordnet. Bei einem Gamma von 0,6 entspricht eine Blendenstufe ungefähr 90 DPX-Werten.

In Abbildung 5.30 wird deutlich, dass für den dunkleren Graubereich wesentlich mehr Zahlen zur Verfügung stehen als im hellen, weil die Zuordnung anhand der flachen Kennlinie des Negativmaterials geschieht, die eine Steigung von ca. $\gamma = 0{,}6$ aufweist. Die Bezeichnung lineare Daten bezieht sich dagegen auf $\gamma = 1$, d. h. den gleichmäßigen Anstieg der Transparenz oder eines elektronischen Signals mit der Belichtung, so wie es bei einem elektronischen Bildwandler der Fall ist. 10-Bit-log-Daten stellen den Kontrastumfang also quasi komprimiert dar.

Zur Umrechnung zwischen log- und lin-Daten kann man von der oben genannten Dichtedifferenz $\Delta D = 2$ ausgehen und annehmen, dass sie aus der Belichtung eines Negativs mit $\gamma = 0{,}6$ herrührt (Abb. 5.30). Aus $\Delta D/\gamma = 3{,}33$ folgt dann ein Kontrastumfang von $10^{3{,}33} = 2154{:}1$, der mit $2^{11} = 2048$ wiederum etwa 11 Blendenstufen entspricht, während zu der Dichtedifferenz $\Delta D = 2$ ein Transparenzumfang von 100:1 gehört. Zur Bildung eines linearen Zusammenhangs zwischen Belichtung und Transparenz muss der Transparenzumfang danach um den Faktor 21,54 größer sein. Um diesem Faktor bei linearer Darstellung gerecht zu werden, müssen also 4 bis 5 Bit mehr verwendet werden. Da der Dichteunterschied der Rechengrundlage eher zu groß ist, reichen 14 Bit aus, um den Grauwertbereich der 10-Bit-log-Daten linear darzustellen. In der Praxis wird aber oft mit 16-bit-Daten gearbeitet, die dann der Computertechnik angepasst sind, z. B. im 16-Bit-TIFF-Format. Die weiteren 2 Bits werden als Headroom verwendet.

Der Unterschied zwischen derartig gebildeten log- und lin-Daten ist sehr bedeutend im Hinblick auf die Weiterverwertung des Digitalsignals. Wenn es im Endeffekt auf ein elektronisches Display gelangen soll, so ist die lineare Darstellung zu bevorzugen. Dann muss dieses Signal nur noch mit dem spezifischen Video-Gamma beaufschlagt werden, das lediglich zum Ausgleich der Nichtlinearität der Bildröhre erforderlich ist. Wenn dagegen das Digitalsignal wieder auf einen Film ausgegeben werden soll, so ist die Beibehaltung der log-Werte günstiger, da sich der gewünschte hohe Endkontrast der Positivkopie wie gewöhnlich direkt aufgrund des anschließenden Kopiervorgangs zum Positiv ergibt.

Eine Schwierigkeit bei der Beibehaltung der log-Daten stellt ihre Beeinflussung dar, wie sie z. B. für die Farbkorrektur unumgänglich ist. Für diesen Vorgang ist eine visuelle Kontrolle erforderlich, die mit Hilfe von Monitoren oder Projektoren erfolgt, die ihrerseits ihr spezielles Kontrastverhalten aufweisen. Das Displaysystem sollte daher so gestaltet sein, dass unter Berücksichtigung der immanenten Nichtlinearität der verwendeten Bildröhre oder eines anderen Displays eine Einstellung gewählt werden kann, die weitgehend der Gamma-Kurve der Filmprojektion entspricht. Eine ähnliche Schwierigkeit besteht, wenn Programme zur Bildmanipulation benutzt werden. In den meisten Fällen beruhen die Software-Routinen auf der Annahme einer linearen Helligkeitsverteilung, d. h. einer Verdopplung des Codewortwertes bei Helligkeitsverdopplung. Um Fehlberechnungen zu vermeiden muss das Film-Gamma rechnerisch ausgeglichen werden, dabei können auch die leicht unterschiedlichen Gamma-Werte der einzelnen Farbschichten berücksichtigt werden. Insgesamt stellt die Anpassung aller Bestandteile einen aufwendigen Kalibrationsprozess dar, der individuell für das jeweils verwendete Displaysystem und für das eingesetzte Filmmaterial durchgeführt werden muss.

Bei der hier vorgestellten Digitalisierung des Filmnegativs entsteht bei 10-Bit-log-Daten und 4k-RGB-Auflösung ein Datenumfang von 3 x 4096 x 3112 x 10 Bit = 45,6 MByte pro Bild. Bei 2k-Auflösung sind es noch 11,4 MByte, die sich bei linearer Darstellung mit 14 Bit auf ca. 16 MByte erhöhen. Bei 2k-Auflösung mit 14 Bit ist pro Minute eine Speicherkapazität von 23 GByte erforderlich. Wenn die diese reinen Bilddaten in Files verpackt werden, steigt der Datenumfang noch.

Um die Datenmenge zu reduzieren, kann die Theorie für ein digitales Positiv verwendet werden. Dabei wird von einem erheblich geringeren Dichteumfang ausgegangen, mit dem Argument, dass auch bei der Erstellung einer Positivkopie im Kopierwerk die Anpassung an die steile Kennlinie des Printfilms dazu führt, dass nur ein Ausschnitt aus der Negativkennlinie genutzt wird. In diesem Fall ist es bei der Quantisierung möglich, mit 8 Bit in linearer Darstellung auszukommen [171]. Die praktische Durchführung erfordert allerdings sowohl bei der Filmabtastung als auch bei der Ausbelichtung eine sehr exakte Anpassung an den gewünschten Dichtebereich und damit einen erheblichen Aufwand. Um sich für den weiteren Verarbeitungsprozess auch bei den Digitaldaten möglichst den Belichtungsspielraum zu erhalten, der beim Negativmaterial gegeben ist, sollte die oben angewandte Theorie des digitalen Negativs zugrunde gelegt werden und nicht die des Positivs.

5.6.3 Die Farbqualität

Das digitale Filmbild kann als Gemisch von RGB-Farbauszügen gewonnen werden. Bei der Filmabtastung wird das Negativ mit weißem Licht durchstrahlt und die Farbauszüge werden über RGB-Farbfilter erzeugt. Die Filter müssen dabei sehr steilflankig arbeiten, damit geringe spektrale Überlappungen auftreten. Abbildung 5.32 zeigt ein Beispiel für Filterkurven, die auf die Filter eines anschließenden Ausbelichtungsvorgangs abgestimmt sind.

Auch bei der Umsetzung der elektronischen Signale durch Filmbelichtung müssen die RGB-Auszüge separat aufgenommen werden. Idealerweise verwendet man dazu drei Laser, die jeweils monochromatisches Licht für die drei Grundfarben erzeugen. Die Alternative ist die Verwendung von weißem Licht und Filtern, wie es z. B. bei den CRT-Recordern geschieht. Die Filterkurven sollten möglichst gut auf die spektralen Eigenschaften des aufnehmenden Filmmaterials abgestimmt sein und mit den Filterkurven der Abtastung übereinstimmen.

Die Umsetzung von Farben ist insgesamt ein aufwendiger Prozess, der ein eigenes Farbmanagement erfordert. Die Anpassung der Farbräume von digitalen Medien und analogem Film gelingt nicht völlig. Der messtechnische Vergleich der beiden zugehörigen Farbräume in Abbildung 5.31 zeigt, dass Film ein größeres Gamut, also einen größeren Bereich darstellbarer Farben hat [216]. Dies kann zwar durch Verwendung höher gesättigter Primärfarben im Grünbereich angeglichen werden, so wie es beim Arrilaserbelichter geschieht, doch ist der Farbraum des Films nicht dreieckig und vor allem zeigen die Messpunkte in der Mitte von Abbildung 5.31, dass die Helligkeits- und Sättigungsabstufungen beim Film nicht so linear verlaufen wie im digitalen Raum der Videosignale.

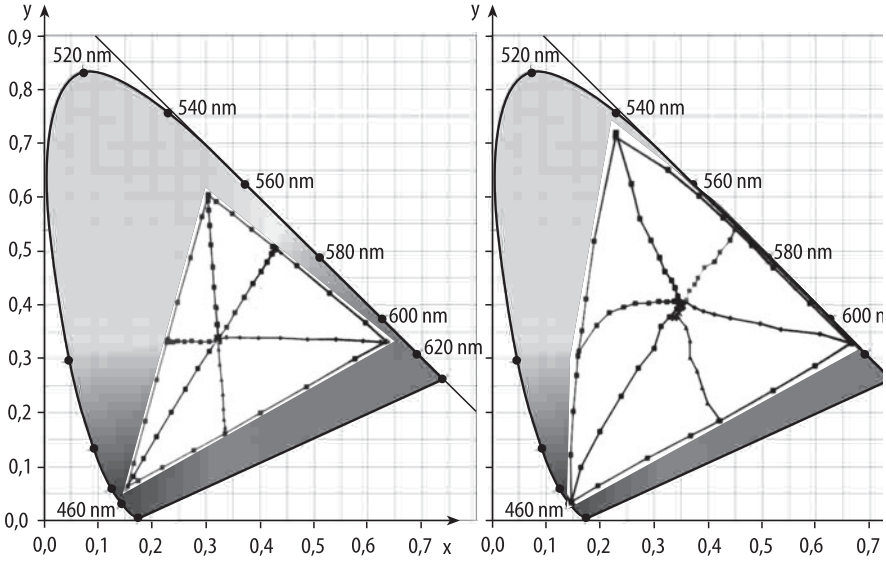


Abb. 5.31. Vergleich gemessener Farbräume von HD-Video (links) und Film (rechts) [216]

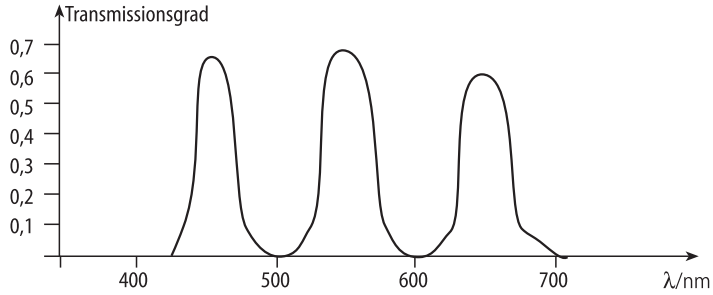


Abb. 5.32. Filterkurven

5.7 Filmabtaster

Filmabtaster dienen der Umsetzung des Filmbildes in elektronische Signale bzw. digitale Daten. Dazu wird der Film durchleuchtet und das Licht gelangt auf einen Bildwandler. Der Filmtransport lässt sich hinsichtlich intermittierenden und kontinuierlichen Laufs unterscheiden. Eine zweite Klassifikation ist möglich nach der ausgegebenen Signalqualität, also hinsichtlich der Unterscheidung in SDTV, HDTV oder Filmauflösung mit 2k oder 4k.

Für den Fernsehbereich erfolgte die Umsetzung in Standard- oder HD-Auflösung, mit Geräten, die Telecine genannt werden. Diese sind seit langer Zeit ein wichtiger Bildgeber, denn vor der Verfügbarkeit elektronischer Magnetbandaufzeichnungsverfahren war der Film das einzige Speichersystem für den Fernsehbereich, sodass die Umsetzung von Film in Videosignale eine große Rolle spielte. Da das Fernsehsystem die Bildfrequenz bestimmt, erfolgen die Filmproduktionen für den Fernsehbereich, meist auf 16 mm-Film, mit einer Filmgeschwindigkeit von 25 fps, sodass eine problemlose Umsetzung in 50 Halbbilder möglich wird.

Damals wie heute werden jedoch auch Kinofilme im Fernsehen gezeigt, die für eine Wiedergabe mit 24 fps konzipiert sind. Es muss also die Möglichkeit einer Anpassung zwischen der Filmgeschwindigkeit mit 24 Frames per Second (fps) und dem Videosignal mit 25 fps in Europa bzw. 30 fps in Amerika gegeben sein. In Europa ist die Anpassung unaufwendig: Der Film läuft bei der Abtastung einfach schneller, als er im Kino wiedergegeben wird. Die Abweichung von ca. 4,1 %, die für das Audiosignal mit einer Steigerung der Tonhöhe einhergeht, wird meist nur bei direktem Vergleich zwischen 24-fps- und 25-fps-Wiedergabe wahrgenommen. Bei der Filmabtastung für das NTSC-Format mit 30 fps kann der Film nicht einfach schneller laufen, da die Differenz zu 24 fps groß ist. Hier werden mit dem sogenannten NTSC-Pulldown-Verfahren pro Sekunde sechs Zusatzbilder in das Videomaterial eingearbeitet, d. h. nach zwei Filmbildern wird aus dem vorhergehenden Material ein zusätzliches Halbbild generiert.

Wenn auch für eine 25-fps-Abtastung die Filmgeschwindigkeit von 24 fps beibehalten werden soll, kann auf ähnliche Weise ein sogenanntes PAL-Pulldown vorgenommen werden, indem zu jedem 12. Filmbild ein Zusatzhalbbild

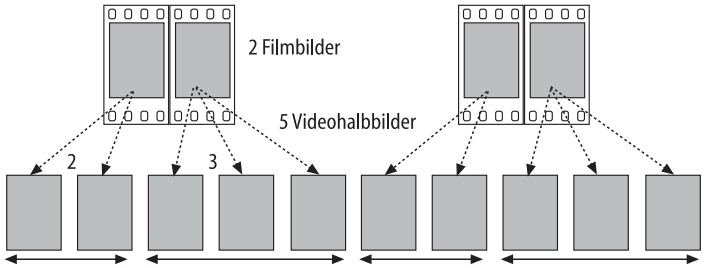


Abb. 5.33. Umsetzung von 24 fps in 30 fps nach dem NTSC-Pulldown-Verfahren

generiert wird. Eine Sekunde Videomaterial besteht dann insgesamt aus $2 \cdot (11 \cdot 2\ \text{HB} + 3\ \text{HB}) = 50$ Halbbildern (Abb. 5.33).

Die Filmabtaster für den Fernsehbereich wurden so weiterentwickelt, dass sie für die Erzeugung hoch aufgelöster Bilder, also für HDTV und für den digitalen Film geeignet sind. In der Regel sind diese, dann als Film Scanner bezeichneten Abtaster, in der Lage, sehr viele Signalformate parallel bereitzustellen. Eine besonders interessante Entwicklung ist diesbezüglich der Gedanke, dass bei einer Vision, die die vollständige digitale Archivierung von Filmen ins Auge fasst, der Aspekt der Formatunabhängigkeit immer bedeutsamer wird. In diesem Zusammenhang ist das Ziel, mit der Anfertigung einer sogenannten digitalen Filmkopie zunächst eine Abtastung mit möglichst hoher Auflösung zu gewinnen, die als Basis für alle Formate dient. Die höchste Unabhängigkeit wird erreicht, wenn ein digitales Faksimile des Films gewonnen wird, indem der Film unsynchronisiert durch den Abtaster läuft, wobei nicht nur der Bild-, sondern auch der Randbereich mit Tonspuren, Perforation etc. erfasst wird. Aus dem gewonnenen Datensatz können dann im Prinzip auch solche Daten- und Videoformate abgeleitet werden, die erst zukünftig definiert werden. Daher ist die digitale Filmkopie besonders für die Archivierung des wertvollen Filmmaterials von Bedeutung.

Angesichts der hier erwähnten Problematik mit der Vielfalt der Datenformate gibt es sogar auch einen gegenläufigen Trend, nämlich den Film zu nutzen, um Digitaldaten zu speichern, ähnlich, wie es beim Dolby-Digital-Lichtton geschieht. In diesem Zusammenhang geht man davon aus, dass der Film ein bewährtes, langzeitstabiles Speichersystem darstellt. Hier wäre die Abtastung zur Rückgewinnung der Daten erforderlich [23].

Moderne Filmabtaster eignen sich aufgrund der Austauschbarkeit der Bildfenster (Gates) sowohl zur Abtastung von 16- als auch 35 mm-Film. Sie erlauben den Formatwechsel zwischen 4:3 und 16:9 und enthalten eine automatische Steuerung eines angeschlossenen Videorecorders. Obwohl die Abtastung möglichst staubfrei erfolgt, kann mit einer Nassabtastung die Sichtbarkeit von Staub und Schrammen stark reduziert werden. Dabei wird unter der Verwendung von Bildfenstern (Wet-Gates) eine Flüssigkeit zur Filmbenetzung verwendet, die den Staub entfernt und die Schrammen füllt, wobei durch die Tatsache, dass die Flüssigkeit (Perchlorethylen) und das Filmmaterial den gleichen Brechungsindex aufweisen, die Streuung und Brechung des Lichts vermieden

wird. Die Alternative, das sogenannte »Digital Wet Gate«, beseitigt die Störungen auf elektronische Art.

In Filmabtastern können sowohl Original-Filmnegative als auch Duplikat-Positive verwendet werden, die den Vorteil aufweisen, keine Klebestellen zu haben, und bei denen Licht- und Schattenbereiche bereits filmtypisch komprimiert sind [145]. Insgesamt ist die Abtastung ein aufwendiger Prozess, denn der besondere Film-Charakter soll möglichst auch nach der Übertragung auf Video erhalten bleiben. Das Medium Film verhält sich anders als das Medium Video, vor allem ist der Kontrastumfang und das Auflösungsvermögen bei Film größer. Da die höchstentwickelten Filmabtaster mehrere Auflösungen parallel bereitstellen, erscheint eine Unterscheidung hinsichtlich der Bildauflösung ungeeignet. Daher wird hier eine Klassifikation nach den Abtastprinzipien vorgenommen, also bezüglich der Frage, ob die Abtastung bildpunktweise, zeilenweise oder bildweise arbeitet.

5.7.1 Filmabtastung bildpunktweise

Das Grundprinzip dieser Umsetzung ist das Flying-Spot-Verfahren. Es ist bereits sehr alt und eng mit der Lichtpunktabtastung nach Nipkow verwandt, bei der ein Lichtstrahl, durch die Nipkow-Scheibe gesteuert, den Film durchdringt und auf der anderen Seite die bildpunktabhängig variierende Lichtintensität mit einer Fotozelle in ein elektrisches Signal gewandelt wird.

Beim Flying-Spot-Prinzip tritt an die Stelle des Lichtstrahls der in einer Bildröhre (CRT) erzeugte Elektronenstrahl, der mittels elektromagnetischer Felder abgelenkt wird. Der Elektronenstrahl durchdringt aber nicht den Film, sondern fällt zunächst auf eine Leuchtschicht. Dort wo der Strahl auftrifft, entsteht ein heller Lichtpunkt, der über die Ablenkung des Elektronenstrahls dann ein gleichmäßig helles, unmoduliertes Raster erzeugt. Der einzelne Lichtpunkt wiederum ist die Lichtquelle, die den Film durchdringt. Durch die ortsabhängig verschiedenen Dichten wird das Licht in seiner Intensität verändert und gelangt schließlich zu einer Fotozelle, wo aus der Folge der einzelnen Lichtpunkte direkt ein serielles Signal entsteht (Abb. 5.34). In der Fotozelle werden durch die Energie der auftreffenden Lichtquanten Elektronen aus einem Metall gelöst (äußerer Fotoeffekt). Die negativen Elektronen werden von einem positiven Potenzial an der Anode angezogen, sodass ein Stromfluss entsteht, der

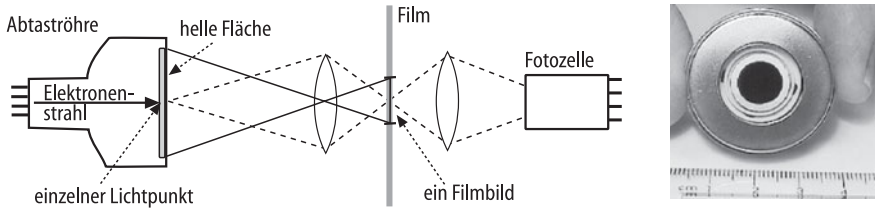


Abb. 5.34. Arbeitsprinzip des Flying Spot-Abtasters und Bild einer Fotodiode

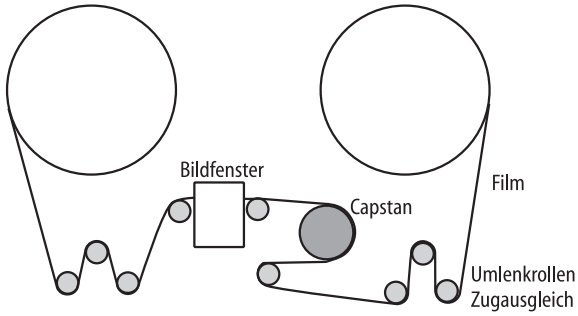


Abb. 5.35. Kontinuierlicher Filmtransport im Abtaster

der Anzahl der Photonen, d. h. der Lichtintensität, proportional ist. Alternativ kann der innere Fotoeffekt einer Fotodiode (Abb. 5.31) genutzt werden.

Der Filmtransport kann bei diesem Verfahren kontinuierlich und damit materialschonend durchgeführt werden. Der Antrieb wird mit Hilfe eines Capstans vorgenommen, bei dem der Film gegen eine sich gleichmäßig drehende Scheibe gedrückt wird, die mit einem Gummibelag versehen ist und den Film schlupffrei mitzieht, wobei die Berührung nur am Rand erfolgt (Abb. 5.35).

Im Markt für Flying-Spot-Abtaster tritt seit langer Zeit die Firma Cintel hervor, die den Abtastertyp C-Realitiy baut. Abbildung 5.36 zeigt den Abtaster und Abbildung 5.37 das Blockschaltbild. Er arbeitet mit einer gegen Magnetfelder abgeschirmten Bildröhre, die eine große Helligkeit erzeugt. Die Röhre ist mit einer dicken Frontplatte versehen, die die Bildung eines Lichthofes um den Abtastspot verhindert. Mit digitaler Steuerung wird eine hochpräzise Strahlablenkung erreicht. Das Licht der Abtaströhre gelangt über den Film zu einem farbselektiven Strahlteiler und von dort aus weiter zu drei Fotodioden [188].

Die Abtaststeuerung ist sehr flexibel, ein Formatwechsel zwischen 4:3 und 16:9 ist damit ebenso wenig ein Problem wie die Halbbildgewinnung. Die Abtastung kann an verschiedene Filmgeschwindigkeiten und Standbilder angepasst werden. Die digitale Steuerung ermöglicht auch den Ausgleich von Instabilitäten der Bildlage. Bildstandsfehler sind im Videobereich besonders

Abb. 5.36. Flying Spot-Abtaster C-Reality und Bild einer Röhre [188]

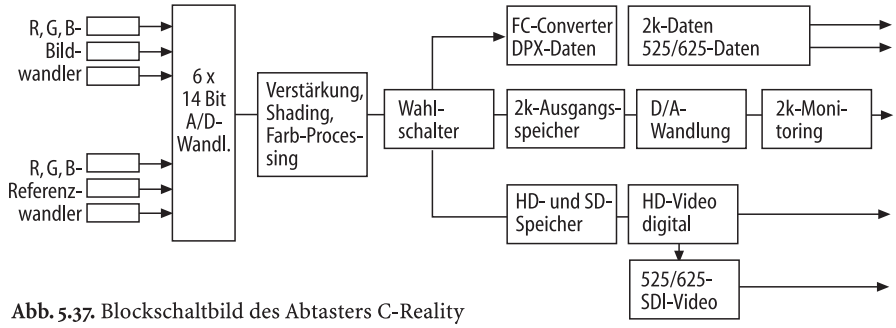


Abb. 5.37. Blockschaltbild des Abtasters C-Reality

kritisch, da das instabile, abgetastete Bild mit stabilen Videobildern gemischt oder überlagert werden kann (z. B. bei der Untertitelung). Die Bildstandsfehler werden unsichtbar, wenn das Abtastraster in gleicher Weise wie das Filmbild verschoben wird. Mit der digitalen Abtaststeuerung bei Flying-Spot-Abtastern kann dies erreicht werden, indem als Referenz für das Abtastraster die Filmperforation benutzt wird, die die gleiche Instabilität aufweist wie das Bild.

Auch die Signalverarbeitung geschieht auf digitaler Basis. Das vorverstärkte Signal wird mit hoher Auflösung A/D-gewandelt und in einen Datenspeicher eingelesen. Die Daten können bezüglich Farbe und Kontrast umfangreich verändert und als Analog- oder Digitalsignal ausgegeben werden. Der Typ C-Reality von Cintel bietet bei einer Digitalisierung mit 14 Bit eine maximale Auflösung von 4k. Bei 2k, d. h. wenn eine Zeile in 2048 Bildpunkte und das Bild in 1536 Zeilen (B/H = 4/3) aufgelöst wird, ermöglicht C-Reality eine Abtastgeschwindigkeit von 6 Bildern/s.

Ein Vorteil von Flying Spot gegenüber anderen Abtastverfahren ist, dass bei den vielfältigen Möglichkeiten der Bildformatänderung und -verzerrung, die aus Qualitätsgründen möglichst nahe an der Quelle, also im Abtaster selbst, vorgenommen werden, kein Auflösungsverlust auftritt, da die auf der Röhre abgetastete Fläche z. B. für einen Zoomeffekt einfach verkleinert werden kann.

Hersteller von Abtastern nach dem Flying-Spot-Prinzip nehmen für sich in Anspruch, dass diese Technik am besten den »Film-Look« erhält, da durch die veränderbare Verstärkung am Fotosensor die Erfassung des gesamten Kontrastbereichs des Filmmaterials erreicht werden kann, und die beste Anpassung an die Gradationskurve möglich wird. Damit soll die Umsetzung eher weicher und damit besser an das Medium Film angepasst sein. Als weiteres Qualitätsargument kann angeführt werden, dass Abtaster nach diesem Prinzip in der Lage sind, bei 2k- oder 4k-Auflösung die volle RGB-Auflösung, d. h. ohne unterschiedliche Auflösung im Luminanz- und Chrominanzbereich, zu bieten. Die Behauptung eines weicheren Bildeindrucks ohne die Härte, die oft mit Digitalsystemen verknüpft ist, kann durch das Argument gestützt werden, dass die Bildpunktbildung einem nicht ganz so starren Muster unterliegt wie bei Wandlern, die zeilen- oder bildweise arbeiten, und bei variablerem Muster weniger Alias-Störungen auftreten. Im gleichen Sinne wirkt auch die Form der Intensitätsverteilung über dem Bildpunkt selbst, die bei Flying Spot über einer runden Fläche gaußförmig ist.

5.7.2 Filmabtastung zeilenweise

Für diese Abtastungsart werden Halbleiterbildwandler in Form von CCD-Zeilen verwendet, deren Funktionsweise in Kap. 6 beschrieben ist. Bei den CCD-Abtastern wird zur Lichterzeugung eine leistungsstarke Halogen- oder Xenonlampe verwendet, deren Lichtspektrum mittels Filter an das Filmmaterial angepasst wird. Das Lichtbündel wird mit Zylinderlinsen so geformt, dass ein möglichst gleichmäßig ausgeleuchteter Lichtstreifen entsteht, der den Film in voller Breite durchdringen kann. Für die Abtastung verschiedener Bildformate muss das Licht unterschiedlich gebündelt werden, daher gibt es für die einzelnen Bildformate eigene Optikblöcke, die das Objektiv, die Kondensorlinsen und die Filmführung enthalten und als komplette Einheit ausgetauscht werden (Abb. 5.38 rechts).

Das durch den Film tretende Licht ist durch den Dichteverlauf moduliert und wird über ein Projektionsobjektiv auf ein Strahlteilerprisma abgebildet, das es in RGB-Anteile zerlegt (Abb. 5.38) [145]. Für jeden Farbauszug wird eine separate Bildwandlerzeile verwendet. Der Film wird kontinuierlich zwischen Zylinderlinse und CCD-Zeile geführt, der Antrieb wird auch hier meist mittels Capstan realisiert.

Der Bildwandler ist in eine lichtempfindliche und eine Speicherzeile aufgeteilt. Damit wird die Belichtungszeit unabhängig von der Auslesezeit. Die durch das Licht in der Zeile erzeugte Ladung wird während der H-Austastzeit schnell in die angrenzende CCD-Speicherzeile geschoben, die wie bei Flächenbildwandlern als Ausleseregister dient. Das H-Register kann so innerhalb der aktiven Zeilendauer entleert werden, während für die lichtempfindliche Zeile die gleiche oder eine kürzere Zeit zur Ladungsintegration zur Verfügung steht. Die CCD-Analogsignale werden dann A/D-gewandelt und gespeichert. Schließlich wird das Videosignal in digitaler oder analoger Form bereitgestellt. Dem Zeilensprungverfahren wird Rechnung getragen, indem das Signal halbbildrichtig aus dem Speicherbereich ausgelesen wird.

Bei CCD-Filmabtastern ist die Firma DFT/Thomson (ehem. Philips/BTS) marktführend, deren verbreitetster Abtastertyp (Abb. 5.39) die Bezeichnung Spirit Datacine trägt. Das Gerät erreicht eine Geschwindigkeit von bis zu 23 fps

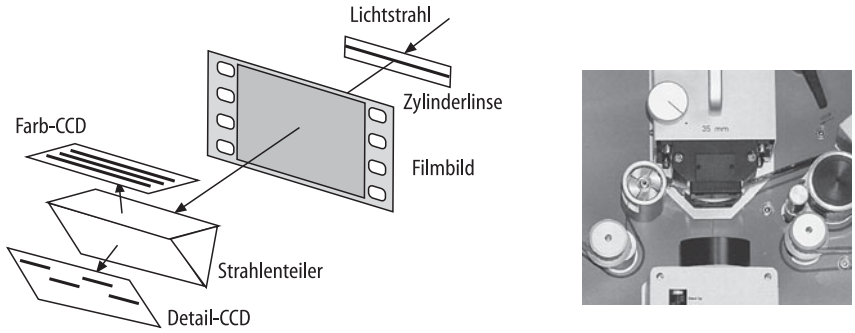


Abb. 5.38. Strahlverlauf und Bildfenster eines Halbleiterfilmabtasters [219]

und kann 35 mm-, 16 mm- und Super 8-Filme verarbeiten. Der Spirit Datacine-Abtaster arbeitet mit Bildwandlern, die von Kodak entwickelt wurden. Sie bestehen aus einem Detailsensor mit 1920 Bildpunkten sowie drei Farbsensoren mit je 960 Bildpunkten.

Die Weiterentwicklung, der Typ Spirit 4k, ist für 2k- und 4k-Abtastungen optimiert. Er hat zwei Detailsensoren mit 2048 und 4096 Pixels. Die Bildpunktgröße beim 2k-Sensor ist entsprechend größer und damit auch die Lichtempfindlichkeit. Entsprechend ist die Auswirkung auf die Scangeschwindigkeit. Im 2k-Modus werden 25 fps und im 4k-Modus 6 Bilder pro Sekunde erreicht [79].

Die Signale werden nach der Wandlung digitalisiert und in 14-Bit-Datenform im Bildspeicher abgelegt. Dabei stehen verschiedene Korrekturmöglichkeiten zur Verfügung. So können mit der Korrektur des sogenannten Fixed Pattern Noise (FPN) statische Ungleichmäßigkeiten der Lichtverteilung reduziert werden, indem eine Abtastung des Intensitätsprofils der Lichtquelle ohne Film vorgenommen wird. Für jeden Bildpunkt wird dann die Abweichung vom Mittel registriert und elektronisch korrigiert, sodass Unregelmäßigkeiten der Lichtverteilung ebenso ausgeglichen werden können wie Störungen einzelner CCD-Pixel. Weiterhin gibt es Geräte zur Reduzierung der Sichtbarkeit des Filmkorns, die nach dem Prinzip der Rauschminderung arbeiten, und eine elektronische Stabilisierung des Bildstands, der jedoch bereits weitgehend stabil ist, da wie beim Flying-Spot-Abtaster die Perforation mit abgetastet wird und damit als Korrekturmaß zur Verfügung steht.

Halbleiterfilmabtaster weisen gegenüber Flying-Spot-Abtastern vor allem die Vorteile der Halbleiter- gegenüber der Röhrentechnologie auf. Die Bildparameter sind sehr stabil und die erforderlichen Abgleichmaßnahmen minimal. Die CCD-Lebensdauer ist sehr hoch, während Röhren eher verschleißen. Filmabtaster nach dem CCD-Prinzip arbeiten vergleichsweise schnell werden daher vor allem dort benutzt, wo Wirtschaftlichkeit im Vordergrund steht.

Abb. 5.39. Halbleiterfilmabtaster Spirit Datacine [219]

5.7.3 Filmabtastung bildweise

Die bildweise Erfassung der Filmbilder ist ein nahe liegendes Prinzip, das bereits früh mit Röhrenbildwandlern (Speicherröhrenabtaster) verwirklicht wurde. Der Filmgeber funktioniert dabei wie ein Filmprojektor, transportiert also den Film schrittweise, während das Aufnahmesystem nach dem Prinzip einer Videokamera arbeitet, die das Filmbild während seines Stillstandes umsetzt. Aufgrund von Bildstandsproblemen wurde dieses Prinzip lange Zeit nicht mehr benutzt und Abtastverfahren mit kontinuierlichem Filmtransport bevorzugt. In neuerer Zeit wird das Verfahren jedoch wieder aufgegriffen, z. B. vom Projektorhersteller Kinoton, für den das Prinzip aufgrund der dort verfügbaren Technologie hochpräziser Schrittschaltwerke nahe liegend ist. Die Film Transfer Machine der Firma konvertiert 35-mm-Filmformate in Echtzeit in ein SDTV- oder ein HDTV-Signal mit bis zu 850 Linien Auflösung.

Als weiterer Hersteller baut die Firma Arri einen Filmabtaster mit der Bezeichnung Arriscan nach dem Prinzip der bildweisen Abtastung (Abb. 5.40). Auch er nutzt also den nur bei diesem Verfahren gegebenen Effekt der hohen Belichtungsdauer, die ja bei der punkt- oder zeilenweisen Abtastung erheblich kürzer ist. Das Gerät verarbeitet 16-mm-, Super 16 und 35-mm-Film (2-, 3- und 4-perf) und profitiert von der Erfahrung beim Bau äußerst präziser Filmtransportmechanismen, die die Firma als Hersteller hochwertiger Kameras hat.

Bei Filmscanner Arriscan werden viele neue Technologien eingesetzt. Zunächst wird ein besonders präziser Bildstand (pin registered) erzielt. Als Lichtquelle kommt dann statt der üblichen Xenon-Lampe ein Array aus Leuchtdioden (LED) zum Einsatz, die den Film im fixierten Zustand durchleuchten. Bei diesen Bauelementen gab es seit Beginn des Jahrhunderts einen großen Technologiesprung. Sie stellen heute sehr stabile und langlebige Lichtquellen mit hoher Lichtausbeute dar. Das bedeutet hier, dass ihre Lichtstärke kurze

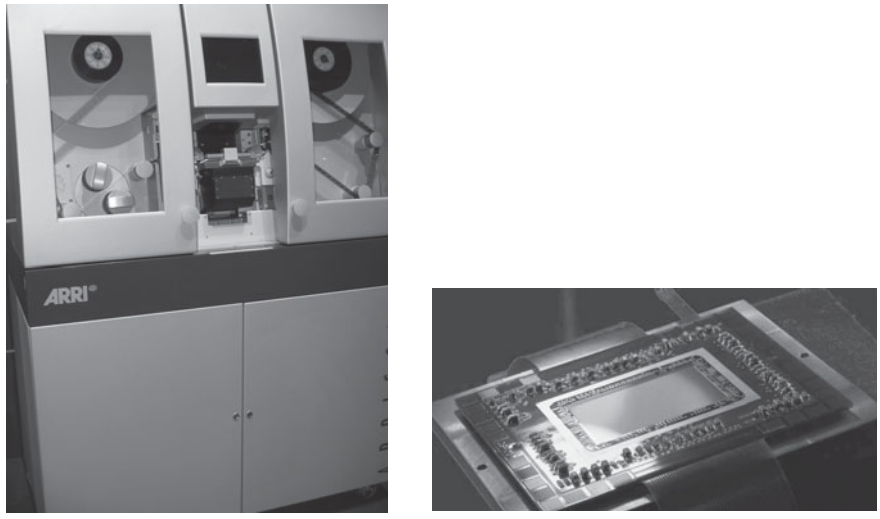


Abb. 5.40. Filmscanner Arriscan und zugehöriger CMOS-Bildwandler [9]

Beleuchtungszeiten ermöglicht und trotzdem die Wärmebelastung des Films sehr gering gehalten werden kann.

Der Film wird mit Lichtblitzen aus rot-, grün- und blauleuchtenden LED durchstrahlt. Mit Hilfe einer Infrarot-LED kann ein zusätzliches Bild generiert werden. Dieses kann als Alpha-Datenkanal mitgeführt werden und ermöglicht eine einfache Erkennung bzw. Korrektur von Bildfehlern wie Schrammen etc.

Als Bildwandler wird wie in der Arri-D20-Kamera ein CMOS-Flächensensor eingesetzt (Abb. 5.40). Er hat eine Fläche von 25 mm x 19 mm, sodass ein 35 mm-Filmbild ohne Zwischenabbildung umgesetzt werden kann. Aufgrund mehrerer paralleler Ausgänge kann der CMOS-Sensor sehr viel schneller ausgelesen werden als ein CCD-Sensor.

Der Chip bietet eine 3k-Auflösung. Die Daten können direkt ausgegeben oder auf 2k-Daten konvertiert werden. Mittels piezogetriebenem Versatz des Sensors kann die Auflösung auf 6k gesteigert werden. Dazu wird der Sensor in vier Schritten horizontal und vertikal um die Breite der aktiven Sensorfläche verschoben (microscanning), sodass auch Licht erfasst wird, das ohne Verschiebung auf CMOS-Flächenteile fällt, die von Bussystemen und Transistoren belegt sind. Sowohl im 2k- als auch im 4k-Modus stehen damit an der Quelle mehr Daten zur Verfügung, sodass durch Subsampling weitgehend Alias-freie 2k- oder 4k-Daten erzeugt werden können. Der Abtastvorgang kann für jedes Bild zweimal mit einer um den Faktor 10 verschiedenen Lichtintensität vorgenommen werden. Die jeweils gewonnenen 12-Bit-Daten werden dann elektronisch kombiniert, sodass ein besonders großer Kontrastbereich erfasst und gleichzeitig das Rauschen minimiert wird [79].

Die Signalverarbeitung erfolgt mit 16 bit. Mit diesen Parametern kann ein 35-mm-Film in 2k-Auflösung mit 1 fps und in 4k-Auflösung mit 0,25 fps abgetastet werden. Mit einer optionalen Speed-Erweiterung werden 4 fps bei 2k- und 1 fps bei 4k-Auflösung erreicht. Die Datenübertragung erfolgt über Ethernet-Verbindungen und als Datenformate werden Cineon/DPX-10-bit-log, DPX-16-bit und Tiff-16-bit genutzt.

5.7.4 Gradations- und Farbkorrektur

Die Filmabtastung ist ein aufwendiger Prozess, bei dem die unterschiedlichen Medien Film und Video aneinander angepasst werden müssen. Allein schon wegen seiner flachen Gradationskurve kann das bei der Abtastung verwendete Filmnegativ nicht unkorrigiert verarbeitet werden. Der Prozess der Filmabtastung muss aber nicht nur in Hinblick auf die Reduktion des Kontrastumfanges, sondern auch wegen der Anpassung der Farbwerte kontrolliert werden.

Die Farbkorrektur ist unumgänglich, da verschiedene Filmmaterialien unterschiedliche Farbwerte hervorrufen. Außerdem ändert sich z. B. die Farbtemperatur des Sonnenlichts tageszeitabhängig und damit ggf. von Szene zu Szene. Das gilt natürlich nicht nur bei der Abtastung sondern auch für die Farbkorrektur im gewöhnlichen Filmprozess, wo bei der Umsetzung des Negativs in das Positiv die Gradation und die Farbwerte angepasst werden, indem das in RGB-Farbauszüge geteilte Kopierlicht kanalselektiv unterschiedlich stark

gedämpft wird. Im Vergleich dazu bietet der Digital-Intermediate-Prozess besondere Vorteile, denn die Farbkorrektur findet anhand digitalisierter Filmdaten statt, sodass die Korrekturen in der digitalen Ebene über diese sogenannte primäre Farbkorrektur weit hinausgehen und wesentlich diffiziler sein können.

Auch bei der elektronischen Korrektur geht es im Kern immer darum, die RGB-Werte in gegenseitigem Verhältnis anzupassen und so z. B. durch gleichmäßige Anhebung in allen Kanälen eine Sättigungsminderung zu erzielen oder auch einen besonderen „Look" zu erzeugen. Die Anpassung von Gradation und Farbigkeit sollte möglichst an der Signalquelle, also z. B. am Filmabtaster vorgenommen werden, dort, wo noch der gesamte Informationsgehalt verfügbar ist. Nach der Umsetzung ist zu erwarten, dass der Informationsgehalt in vielerlei Hinsicht reduziert ist. Die Hersteller von Filmabtastern bieten daher für ihre Systeme eine primäre Farbkorrekturmöglichkeit an.

Die Kontrolle und Veränderung von Helligkeits- und Farbwerten geschieht entweder schnell im Synchronmodus, d. h. während des kontinuierlichen Durchlaufs von Hand, oder, genauer, im Programmmodus, bei dem die Werte für jede Szene einzeln individuell korrigiert und gespeichert werden. Beim eigentlichen Abtastprozess stellen sich dann im Programmmodus anhand von Timecode-Daten die entsprechenden Werte automatisch ein.

Bei hohen Ansprüchen werden die Standard-Farbkorrektureinrichtungen durch aufwendige, digital arbeitende separate Geräte wie Pandora Pogle oder Da Vinci 2k ersetzt. Das sind echtzeitfähige Bildverarbeitungssysteme für verschiedene Auflösungen bis zu 4k, die HD-Daten in unterschiedlichen Formaten verarbeiten können, sodass eine HD-Tape-to-Tape-Korrektur ebenso möglich ist wie eine Farbkorrektur auf der Basis von linearen oder logarithmischen Filmdatenformaten. Eine direkte Anbindung an verbreitete Filmscanner wie Spirit Datacine oder C-Reality ist gewährleistet. Derartige Systeme weisen eine sehr hohe Farbselektivität auf und bieten eine sekundäre Korrektur, mit der nicht nur das gesamte Bild beeinflusst wird, sondern eine bestimmte Farbe einzeln herausgegriffen und verändert werden kann. Zur einfacheren, intuitiven Bedienung stehen hier Joysticks oder Kugeln zur Verfügung, mit denen die Farborte und die Verstärkung für den Bereich der Lichter und der Schatten einzeln eingestellt werden können (Abb. 5.41).

Die als Grading bezeichnete Anpassung zwischen den Medien Film und Video ist kritisch, weil das menschliche Auge ein sehr schlechtes Instrument zur Farbbeurteilung darstellt. Der Mensch verbindet aus seiner Erfahrung Objekte mit ihm bekannten Farben und passt sich Farbgemischen an. Daher sollte das Grading von erfahrenen Coloristen durchgeführt werden, die durch ein gutes Umfeld in die Lage versetzt werden, möglichst objektiv arbeiten zu können. Dazu gehört zunächst, dass der Arbeitsraum eine neutrale Beleuchtung und keine extremen Farben oder Streulichter aufweisen sollte. Neutrale Referenzlichtquellen sind wünschenswert (Abb. 5.41). Dann ist ein möglichst gutes Display erforderlich, auf der Basis eines Monitors, der nicht mit der Gamma-Einstellung für Videosysteme, sondern mit einer Gradation arbeitet, die weitgehend dem Ausgabemedium, wie z. B. Film, angepasst ist. Noch besser ist die Verwendung einer Großbildprojektion, die im gleichen Sinne bezüglich aller dabei relevanten Teile kalibriert ist (s. u.).

Abb. 5.41. Arbeitsplatz und Bediengeräte zur Farbkorrektur

Das Farbmanagement ist ein zentraler Prozess für das gesamte Digital Intermediate (DI) [2]. Es ermöglicht die Anpassung zwischen den Ausgabemedien mittels des sogenannten Gamut-Mapping. Im einfachsten Fall kann nur eine Farbortverschiebung und eine Anpassung des Weißpunktes vorgenommen werden. Farben, die nach der Umsetzung in den neuen Farbraum nicht mehr dargestellt werden können, werden dabei meist auf die Grenzwerte des neuen Gamuts gesetzt, sodass bestimmte Farben beschnitten werden. Alternativ kann eine Verschiebung der Primärfarben vorgenommen werden, sodass alle Farben in ihrem Charakter erhalten bleiben und nur entsättigt erscheinen.

Optimierte Mapping-Verfahren mit Farbmanagementsystemen wie Truelight oder Lustre nutzen Look-Up-Tabellen (LUT), in denen für einen Satz von Zahlen, die die Farbwerte in einem System beschreiben, ein veränderter Satz von Parametern gespeichert ist [162]. Man unterscheidet ein- und dreidimensionale LUT. Erstere haben für jeden Wert in einem Farbkanal einen zugeordneten Wert in demselben Kanal. 3D-LUT verkoppeln dagegen die Kanäle und halten für jede Kombination von Eingangstripeln eigene Ausgangstripel bereit. Da die hier entstehende Anzahl von Einträgen (bis zu 10^9 bei 10 Bit) unpraktikabel ist, beschränkt man sich auf eine Auswahl, die z. B. bei 33^3, also ca. 36000, liegt [79].

Mit Hilfe der LUT kann die Vorschau im DI so kalibriert werden, dass die Bildwirkung des Ausgabemediums während der Bearbeitung des Materials weitgehend vorweg genommen werden kann [192]. Die optimale Kalibrierung benötigt LUT-Profile für alle farbrelevanten Aspekte: für die Ausgabe auf Film ein Negativprofil und ein Positivprofil zur Simulation der Kopierung auf Negativ und Positiv, hier muss mit dem Kopierwerk kooperiert werden. Darüber hinaus ein Lichtquellenprofil zur Simulation der Lichtquelle und das Displayprofil zur Anpassung des verwendeten digitalen Projektors oder Displays.

5.8 Filmbelichtung

Die Filmbelichtung dient dem Transfer der digitalen Bilddaten in die Filmebene. Der Prozess wird auch als Filmaufzeichnung (FAZ) oder Filmrecording bezeichnet. In der Regel wird auf Intermediate-Negativfilm belichtet, von dem die Positivkopien gewonnen werden. Neuerdings gibt es Bestrebungen, auf Positivmaterial zu belichten und damit direkt eine Vorführkopie zu erzeugen, was allerdings nur für wenige Kopien kostengünstiger ist.

Das nächstliegende Verfahren zur Filmbelichtung ist das schrittweise Abfilmen eines hochauflösenden Monitors. Dazu wird das Monitorbild formatfüllend auf den Film abgebildet und für einen automatischen Ablauf muss nur die Filmbewegung in der Kamera mit der Frequenz synchronisiert werden, mit der neue Bilder auf dem Monitor erscheinen. Um zu sehr guten Ergebnissen zu kommen, ist dieses Verfahren jedoch ungeeignet. Die Helligkeit gewöhnlicher Monitore ist so gering, dass sehr lange Belichtungszeiten verwendet werden müssen, und die Bildqualität wird durch die Schattenmaske in der Kathodenstrahlröhre und ihren relativ geringen Farbraum oder durch das Pixelmuster eines LC-Displays herabgesetzt.

5.8.1 CRT-Belichter

Diese Probleme könnten dadurch gelöst werden, dass der Film über drei hintereinander gelegene S/W-Bildröhren oder monochrome LCD geführt wird, deren Licht den Film nacheinander erreicht. Die Röhren benötigten dann keine Schattenmaske, sie müssten mit den RGB-Signalen separat versorgt werden und mit entsprechenden Leuchtstoffen oder Filtern versehen sein, sodass jede Röhre für die Belichtung eines Farbauszugs verwendet würde. Die dreifache Belichtung würde auch das Problem der Belichtungsdauer mindern.

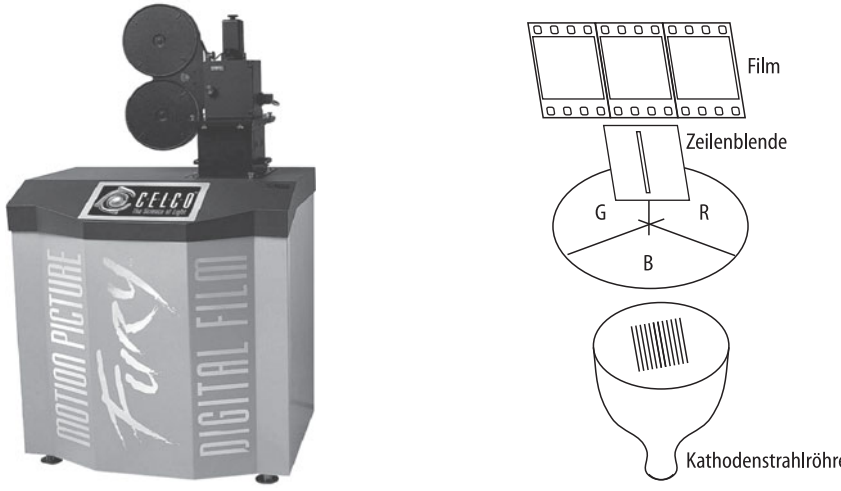


Abb. 5.42. Kathodenstrahlbelichter und Funktionsprinzip

Das beschriebene Verfahren bringt jedoch das Problem der Rasterdeckung der Farbauszüge mit sich, sodass es sicherer ist, es so zu modifizieren, dass während der Belichtung eines Bildes die RGB-Signale zeitlich nacheinander auf nur einer S/W-Röhre dargestellt werden, wobei über ein Filterrad jeweils das zugehörige Farbfilter zwischen Röhre und Film positioniert wird. Nach diesem Prinzip arbeiten die modernen Kathodenstrahlbelichter. Bekannte Systeme sind Nitro von Celco oder die Belichter der Solitaire-Cine-Reihe, wie z. B. Cine III von Management Graphics. Um bei diesen Geräten auch Streulichtprobleme zu minimieren wird mit Hilfe einer bewegten Zeilenblende die Belichtung der RGB-Anteile zeilenweise durchgeführt (Abb. 5.42 rechts).

Das Gerät Nitro HD der Firma Celco eignet sich nach Herstellerangaben zur Belichtung aller Filmformate von 16 mm- bis zu 65 mm-Material. Es ist für Intermediate-Film ebenso geeignet wie für Negativfilme. Das Magazin fasst eine Filmlänge von bis zu 600 m. Es werden alle gängigen Fileformate und alle Auflösungen zwischen Video über HD-Video bis zu 4k und optional sogar bis zu 8k verarbeitet. Der Belichtungsspot hat bei der Belichtung von 35 mm-Film nur einen Durchmesser von 6 µm. Der maximal erzeugte Dichteumfang beträgt $\Delta D = 1{,}5$. Als Belichtungsdauer für HD-Signale wird weniger als 5 s pro Bild angegeben. Das Nachfolgegerät Fury (Abb. 5.42 links) soll eine Geschwindigkeit von 1,3 s pro Bild bei Belichtung von 2k-Daten erreichen.

5.8.2 Laserbelichter

Bereits früh wurde erkannt, dass das theoretisch optimale Filmaufzeichnungsverfahren die Belichtung mit Laserstrahlen ist. Es stehen separate Laser für die drei Farben RGB zur Verfügung, die aufgrund ihrer sehr hohen Intensität das Potential einer sehr schnellen Belichtung bieten und eine hohe Dichte erzeugen können. Die erzeugten Farben sind sehr gesättigt (Abb. 5.43), sodass ein erheblich größerer Farbraum als bei Farbmonitoren erfasst wird. Darüber hinaus entstehen beim Laserbelichter keine Probleme mit der optischen Abbildung und mit Streulicht, da der Film punktweise belichtet wird.

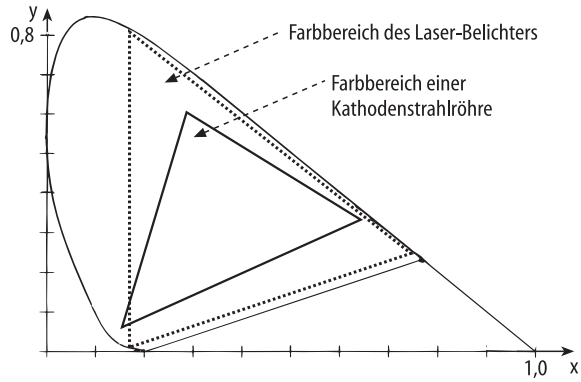


Abb. 5.43. Vergleich der mit Lasern und Kathodenstrahlröhren erreichbaren Farbräume

Der Einsatz der Lasertechnologie war zunächst mit der Schwierigkeit verbunden, dass nur Gaslaser verfügbar waren, die Stabilitätsprobleme aufwiesen und einen sehr hohen Energiebedarf hatten. Ein Beispiel für einen Belichter mit Gaslaser ist der Lightning-Recorder, den die Firma Kodak alternativ zum Solitaire Cine III in ihrem Cineon-System verwendete. Die Weiterentwicklung führte zum Laserbelichter der Fa. Arri (Arrilaser). Dabei ist der wesentliche Punkt, dass Festkörper- statt Gaslaser benutzt werden, sodass ein kompaktes Gerät entsteht, das nicht in klimatisierten Räumen untergebracht werden muss (Abb. 5.44 links). Drei Laser erzeugen Lichtstrahlen in den Farben RGB, die separat über Intensitätsmodulatoren geführt und schließlich zu einem Strahl vereinigt werden. Die Intensitätsänderung in den Modulatoren wird beim Arrilaser mittels des Beugungseffektes für das Licht erreicht. Es passiert dazu einen Kristall, der mit einer akustischen Welle angeregt wird, deren Amplitude die Beugungswirkung und damit die Lichtintensität steuert.

Der vereinigte Gesamtstrahl wird dann zu einem Scanner-Modul geführt, das ein mit 6000 U/min rotierendes Ablenkprisma enthält, und schließlich auf den Film abgebildet. Während der Belichtung wird der Film mit konstanter Geschwindigkeit quer zum Strahl bewegt, sodass das Bild zeilenweise aufgezeichnet wird (Abb. 5.44 rechts). Die hohe Strahlintensität führt dazu, dass für ein Bild mit 4096 x 3112 Bildpunkten auch bei Verwendung von unempfindlichem Intermediate-Film nur eine Belichtungszeit von ca. 3,2 s erforderlich ist. Bei 2k-Auflösung kann diese Zeit halbiert werden. Der Gesamtzyklus für die Bearbeitung eines Bildes verlängert sich durch die Start- und Bremszeiten sowie die Rücktransportzeit um ca. 2 Sekunden [210].

In den meisten Fällen wird auf 35 mm- Intermediate-Material vom Typ Kodak 5244 belichtet. 16 mm-Film kann mit dem Arrilaser nicht bearbeitet werden, zukünftig soll jedoch 65-mm-Material verwendbar sein. Neben der hohen Geschwindigkeit zeichnet sich der Arrilaser vor allem dadurch aus, dass er einen großen Farbumfang hat, Dichtedifferenzen von mehr als $\Delta D = 2$ erreicht und keine Streulichtartefakte aufweist, die bei Kathodenstrahlbelichtern nie ganz ausgeschlossen werden können.

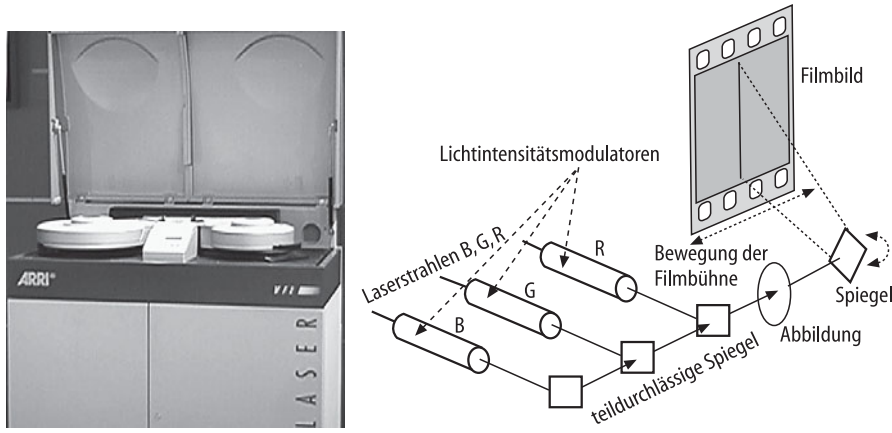


Abb. 5.44. Filmbelichter Arrilaser mit Filmspulen und Funktionsprinzip

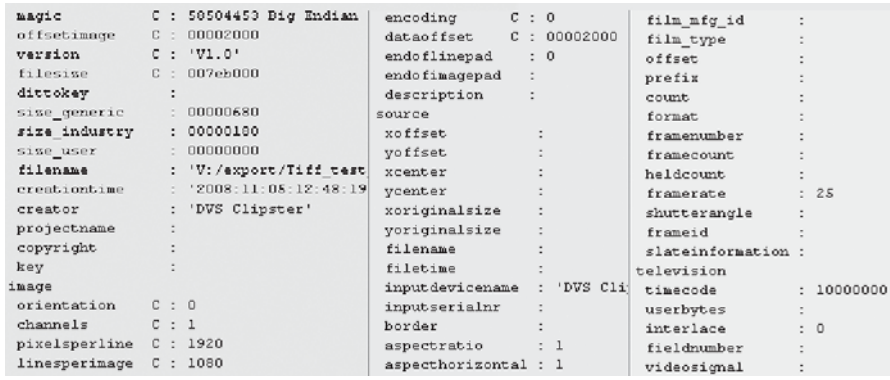


Abb. 5.45. Datenbereiche im DPX-Header

5.9 Filmdatenspeicherung

Die Speichersysteme für die Filmdaten stammen in den meisten Fällen aus der HD-Video- oder Computertechnik, wo die Daten in verschiedenen Fileformaten gespeichert werden. Fileformate für den Filmbereich unterscheiden sich von Standardfileformaten. Sie enthalten spezifische Informationen, die z. B. beschreiben, ob die Werte als lin- oder log-Daten zu interpretieren sind. Weiterhin existieren Angaben über die Bildhöhe und -breite, die Bits pro Pixel und die Bildorientierung (X-Y-Origin). In den Datenstrom können Timecode- und Keycodedaten sowie ein niedrig aufgelöstes Bildduplikat als schnell aufrufbares Beispielbild aufgenommen werden.

Eine der ersten Filedefinitionen entstand im Zusammenhang mit dem früh entwickelten Cineonsystem von Kodak. Die Standardendung trägt die Bezeichnung „.fido“. Hier wird ein File Information Header genutzt, der u. a. die Ordnung der Bytes angibt, das Erstellungsdatum und die Gesamtgröße der Files [96]. Weiter existiert ein Image Information Header u. a. zur Angabe der Bildorientierung und der Darstellung als RGB, YCrCb etc. Aus dem Cineon-Format wurde als standardisiertes Fileformat das Digital Moving Picture Exchange Format abgeleitet (Abb. 5.45). Es trägt die Abkürzung DPX, die auch für die Filekennung „.dpx“ verwendet wird. Hier gibt es File-, Image- und Orientation-Header sowie den Industry-Specific-Header mit den Teilen Film und Television Info. Das Format arbeitet mit Datenwörtern von 32 bit Länge, die ausreichen, um die 3 RGB-Komponenten einer 10-bit-log- Darstellung aufzunehmen, wobei dann 2 Bits ungenutzt bleiben. Tabelle 5.5 zeigt den dabei entstehenden Speicherbedarf für verschiedene Auflösungen. DPX unterstützt die verlustlose Datenreduktion auf Basis des Run Length Encoding (RLE).

Tabelle 5.5. Speicherbedarf von DPX-Files bei verschiedenen Auflösungen

Auflösung	SD	1k	HD	2k	4k
1 Sek	38,4 MB	76,8 MB	197 MB	300 MB	1200 MB
1 Min	2,3 GB	4,6 GB	12 GB	18 GB	72 GB

Tabelle 5.6. Datenraten bei verschiedenen Auflösungen

Format	Auflösung	Bildrate	Farbcodierung	Datenrate
SD	720 x 576	25 fps	YCrCb 4:2:2, 10 bit	26 MB/s
HD	1920 x 1080	25 fps	YCrCb 4:2:2, 10 bit	130 MB/s
HD	1920 x 1080	25 fps	RGB 4:4:4, 10 bit	194 MB/s
Film	2048 x 1080	24 fps	RGB 4:4:4, 10 bit	199 MB/s
Film	4096 x 3112	24 fps	RGB 4:4:4, 10 bit	1147 MB/s

Allgemein werden auch die Fileformate aus dem Gebiet der Computergrafik wie z. B. TIFF, BMP oder Pict verwendet. TIFF-Daten werden meist mit 3 x 16 bit codiert und als Bildsequenz gespeichert. Tabelle 5.6 zeigt eine Übersicht über die beim Datentransfer entstehenden Datenraten in MByte/s. Für den Echtzeitbetrieb müssen die Schnittstellen mindestens diese Werte erreichen.

Für den Austausch der Daten steht mit HD-SDI eine echtzeitfähige Schnittstelle für ein HD-Komponentensignal zur Verfügung, die sich im Parallelbetrieb mittels Dual-SDI auch für RGB-Daten nutzen lässt. Bei HD-SDI-Daten mit 10 Bit muss festgelegt werden, wie diese genutzt werden sollen. Die am Datenaustausch beteiligten Partner müssen sich einigen, ob sie den SDI-Legal Range (64 ... 940) oder den Open Range (0 ... 1023) nutzen wollen, also den Headroom leer lassen oder ihn mit den Daten ausnutzen (Fullrange).

Die Verwendung von Datenreduktion ist – abgesehen von der Distribution im digitalen Kino – bisher wenig verbreitet. Für Filmapplikationen kommt meist das Wavelet-Transformation (s. Kap. 3) zum Einsatz. Dabei gilt schon seit längerem die Erkenntnis, dass für die Qualität der datenreduzierten Bilder besonders die Ausgangsqualität entscheidend ist. In diesem Zusammenhang zeigt die Firma Cintel anhand von Filmmaterial, Beispiele, die es günstig erscheinen lassen, an der Quelle eine Auflösung von 4k bereitzustellen, da dann eine Reduktion um den Faktor 52 eine bessere Bildqualität liefert als eine Reduktion um den Faktor 30 bei 2k-Auflösung (Abb. 5.46).

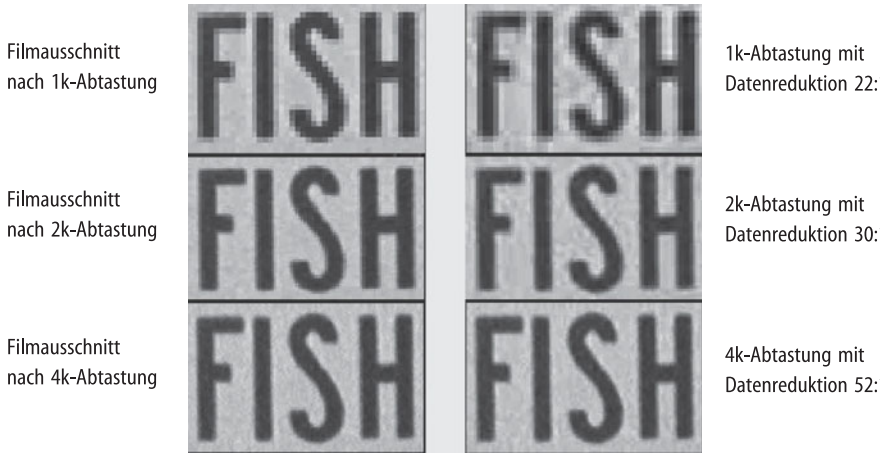


Abb. 5.46. Wirkung der Datenreduktion bei Vorlagen verschiedener Auflösung [214]

6 Bildaufnahmesysteme

Bildaufnahmesysteme dienen der Umsetzung der Lichtintensitätsverteilung innerhalb einer optischen Abbildung in ein speicherbares, elektronisches oder chemisches Abbild. Zu den Bildaufnahmesystemen oder auch Bildgebern gehören die elektronische Kamera, die Filmkamera und der Filmabtaster. Flimabtaster sind bereits in Kap. 5 behandelt worden, wo auch die chemischen Prozesse zur Bildentstehung beim Film beschrieben sind. In diesem Kapitel geht es daher um die Umsetzung einer Lichtmodulation in die Modulation eines elektrischen Signals sowie den Aufbau der Kameras mit ihren optischen und elektrischen Bestandteilen.

Bildwandlerelemente sind kleine lichtempfindliche Gebiete, die einzeln, als Anordnung mehrerer in einer Zeile oder einer Fläche eingesetzt werden. Bildpunkt- und Zeilenwandler werden bei Filmabtastern verwendet und sind in Kap. 5 beschrieben. Im Gegensatz zu den Abtastern kann bei Kameras aber nicht mit externem, den Film durchdringenden, sondern nur mit dem Licht gearbeitet werden, das von den aufzunehmenden Objekt reflektiert wird. Dazu ist eine flächige Anordnung vieler lichtempfindlicher Elemente erforderlich. Jedes Element entspricht dann einem Bildpunkt (Picture Element, Pixel) und hat die Aufgabe, eine elektrische Ladung in Abhängigkeit von der Lichtintensität zu verändern. Wie beim Film spielt neben der Intensität auch die Belichtungszeit eine Rolle, je länger das Licht wirken kann, desto größer ist die veränderte Ladungsmenge und damit die Lichtempfindlichkeit des Bildwandlers.

Es gibt Bildwandler mit Röhren-, CCD- und CMOS-Technologie. Die Röhrentechnologie war bis in die 80er Jahre hinein dominant. Obwohl CMOS- und CCD-Halbleiterbildwandler bereits in den 60er und 70er Jahren entwickelt wurden, konnten sie die Röhren erst spät ablösen, da die Fertigungstechnologien noch nicht die erforderlichen sehr kleinen Halbleiterstrukturen ermöglichten. Dieser Umstand war besonders für die CMOS-Technologie hinderlich, sodass nach den Röhren zunächst vor allem CCD-Chips marktrelevant waren. Nachdem die Halbleiterchips immer weiter verkleinert werden konnten, wurden ab ca. 2005 auch im professionellen Broadcastbereich vermehrt CMOS-Sensoren verwendet.

Die Anforderungen an die Chipentwicklung sind hoch. Heute soll mit 2/3"- oder sogar 1/3"-Bildwandlern eine HD-Auflösung von 1920 x 1080 Pixel erreicht werden, d. h., dass auf einer Bildwandlerbreite von weniger als 10 mm etwa 2000 Pixel untergebracht werden müssen. Die Bildpunkte haben also eine Bildpunktbreite von 3 ... 4 µm. Die Verkleinerung der Bildpunkte soll trotzdem nur wenig zu Lasten der Lichtempfindlichkeit gehen.

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4_6, © Springer-Verlag Berlin Heidelberg 2013

6.1 Bildwandler

Die Grundlage der Umsetzung von Lichtintensität in ein elektrisches Signal ist durch die Lichtenergie gegeben, die Ladungsträger aus einem Metall lösen oder in einem Halbleitermaterial bilden kann (äußerer und innerer Photoeffekt). Die Nutzung des äußeren Effekts war nur zu Beginn der Entwicklung relevant, heute wird, bei CCD- und CMOS-Chips und auch bei Röhren, nur der innere Photoeffekt genutzt. In Halbleitermaterialien, wie Silizium, existiert ein Leitungsband, in das die Ladungsträger mit Hilfe der Licht- bzw. Photonenenergie gehoben werden. Die dafür erforderliche Photonenenergie E ist dabei vom Material und der Temperatur abhängig. Für Silizium bei Zimmertemperatur beträgt sie 1,12 eV, woraus mit der Beziehung $E = h \cdot f = hc/\lambda$, die angibt, dass die Photonenenergie mit der Frequenz f steigt, folgt, dass eine bestimmte maximale Grenzwellenlänge für diesen Photonenabsorptionsprozess erforderlich ist, die im angegebenen Fall bei 1,1 μm liegt (h bezeichnet das Plancksche Wirkungsquantum). Sichtbares Licht hat kürzere Wellenlängen und kann somit die Photonenabsorption auslösen. Die Ladungsträger im Halbleiter erhöhen die Leitfähigkeit bzw. senken den Widerstand mit steigender Beleuchtungsstärke.

Halbleitermaterialien können p- und n-dotiert sein, die Kombination führt zu einer Diode. An der Grenzschicht der zwei Materialien entsteht eine ladungsträgerfreie Sperrschicht, die einen Isolator bildet, sodass die Diode eine Kapazität aufweist. Eine Fotodiode oder ein anderer Halbleiterbildwandler ist so gebaut, dass diese Schicht von Licht erreicht wird und abhängig von der Photonenintensität Ladungsträger, also Elektronen und Löcher, gebildet werden. Der Zusammenhang zwischen Photonenzahl und gebildeter Ladungsträgermenge ist linear.

Die Eindringtiefe der Photonen in das Halbleitermaterial hängt von ihrer Energie und damit von der Wellenlänge ab. Kurzwelliges Licht wird an der Oberfläche absorbiert, langwelliges dringt tiefer in das Material ein, da die hohe Energie der kurzwelligen Strahlung dazu führt, dass schnell viele der verfügbaren Energieniveaus besetzt sind. Die Anzahl der besetzbaren Energiezustände nimmt mit der Eindringtiefe ins Halbleitermaterial ab, sodass langwelliges Licht einen weiteren Weg im Halbleiter zurückgelegt als kurzwelliges. Die Werte betragen ca. 0,2 μm bei 400 nm Wellenlänge (Blau), 2,3 μm bei 500 nm (Grün) und 8,5 μm bei 700 nm (Rot). Dieser Effekt wird als Möglichkeit zur Farbseparation, also zur Gewinnung der Farbauszüge mit einem einzelnen Bildwandler untersucht (Foveon-Sensor, Abb. 6.1).

Abbildung 6.2 zeigt am Beispiel eines MOS-Kondensators (Metal Oxid Semiconductor), der in CCD-Chips verwendet wird, wie die Ladungen durch Photo-

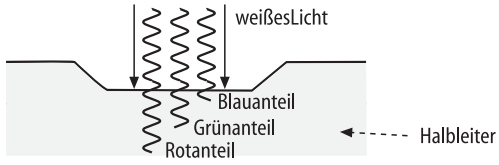


Abb. 6.1. Wellenlängenabhängige Eindringtiefe der Photonen

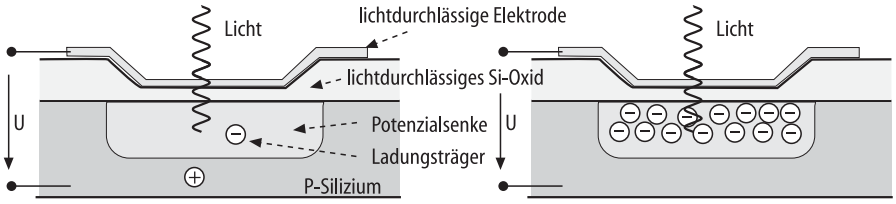


Abb. 6.2. Bildung von Ladungsträgern durch Lichteinwirkung

nenabsorption erzeugt und wie die Ladungsmenge durch Anlegen einer Spannung in der dadurch entstehenden Potenzialsenke gesammelt wird. Bei hoher Lichtintensität kann sich soviel Ladung bilden, dass die Aufnahmekapazität überschritten wird. Wie bei einem Fass, das überläuft, wird in diesem Falle die Ladung das Gebiet des Pixels verlassen und die Ladung der Nachbargebiete verfälschen. Der Effekt wird als Blooming bezeichnet. In alten Röhrenwandlern trat er oft deutlich zutage, in CCD-Chips werden aufwendige Maßnahmen zur Verhinderung des Überlaufs ergriffen. Hier zeigt sich ein Vorteil der CMOS-Technologie, bei der jedes Pixel mit einem eigenen Transistor versehen ist, sodass überschüssige Ladungen einfach abgeführt (gelöscht) werden können.

Abbildung 6.3 zeigt, dass die Abgrenzungen und Strukturen von Röhren-, CCD- und CMOS-Bildwandlern sehr unterschiedlich sind. Die Röhre hat eine homogene Fläche, die erst über die Abtastung mittels eines Elektronenstrahles in Zeilen aufgerastert wird. Beim CCD wird mit Hilfe von Steuerspannungen eine horizontale und vertikale Bildpunktstruktur gebildet. Die gleiche Struktur gibt es auch bei den CMOS-Chips, hier sind aber über die auf dem Chip integrierten Transistoren die Pixel klarer voneinander abgegrenzt.

Die wesentlichen Beurteilungsparameter sind bei der Bildwandler-Chips die gleichen wie bei der Filmabtastung, nämlich die Empfindlichkeit, die Bildauflösung bzw. die MTF und der Dynamikbereich, d. h. der maximal verarbeitbare Szenenkontrast. Letzterer ist nach oben hin durch die Maximalladung im Pixel und nach unten durch das unvermeidliche Rauschen bestimmt.

Insgesamt ist im Überblick festzuhalten, dass der Bildwandlungsprozess für den einzelnen Bildpunkt bei den drei Bildwandlertypen recht ähnlich ist, dass sie sich aber in dem Grad der Abgrenzung einzelner Pixel und vor allem in der Auslesung der einzelnen Bildpunktinformationen aus der Bildwandlerfläche unterscheiden.

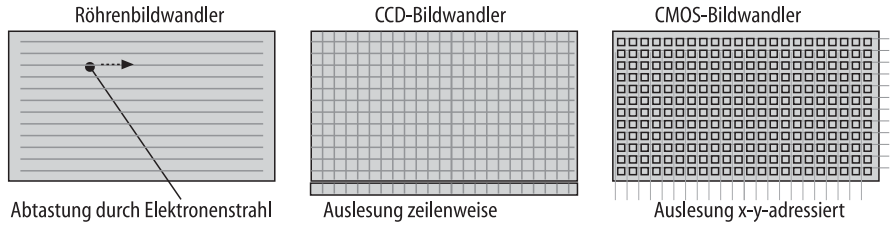


Abb. 6.3. Bildstrukturen und Auslesemechanismen bei Röhren-, CCD- und CMOS-Wandlern

6.1.1 Röhrenbildwandler

Der erste elektronische Bildwandler, das Ikonoskop, stand um 1936 zur Verfügung. Bei diesem Gerät wird durch eine fotoelektrische Schicht auf Basis des äußeren Fotoeffektes die Aufladung von Bildpunkt-Kondensatoren erreicht, die durch einen Elektronenstrahl ausgelesen werden. Dieser Bildwandler wurde in den 50er Jahren zum Vidikon weiterentwickelt, welches einen erheblich einfacheren Aufbau aufweist, da es mit einer Halbleiterschicht arbeitet.

6.1.1.1 Das Vidikon

Das Vidikon besteht aus einem Röhrenkolben, der evakuiert ist, damit die Elektronenbewegung nicht von Luftmolekülen gestört wird (Abb. 6.4). Die vordere Glasplatte ist fotoempfindlich beschichtet und stellt das Target dar. Am anderen Ende des Röhrenkolbens befindet sich das Elektronenstrahlerzeugungssystem. Ein Heizdraht liefert die Energie zur Emission von Elektronen aus einer Metallkathode. Die Elektronen werden als Strom der Größenordnung 1 μA in Richtung auf ein positives Potenzial beschleunigt und passieren dabei ein Elektrodensystem, welches sie so zu einem Strahl bündelt, dass der Strahldurchmesser beim Auftreffen ca. 10 μm beträgt. Außerhalb des Röhrenkolbens befinden sich Spulen, die ein veränderliches magnetisches Feld erzeugen, das dazu dient, den Elektronenstrahl fernsehnormgerecht abzulenken und über das Target zu führen. Der Elektronenstrahl tastet zeilenweise die Ladungszustände aller Punkte auf der Fläche ab, sodass insgesamt ein serielles Signal entsteht.

Die Röhrenkolben der genannten Bildwandlertypen sind zylindrisch aufgebaut (Abb. 6.5) und bezüglich des Durchmessers standardisiert. Standardwerte sind 1" (25,4 mm), 2/3" und 1/2". Die Länge der Röhre liegt etwa bei 10–20 cm.

Nicht die gesamte Röhrendiagonale kann bis zum Rand als Bildfenster genutzt werden. Für die aktive Bilddiagonale einer 1"-Röhre wurde der Wert 16 mm festgelegt, bei 2/3" und 1/2" gelten proportional kleinere Werte. Abbildung 6.6 zeigt eine Übersicht über die Standardwerte und die genutzten Bildwandlerflächen für die Formate 4:3 und 16:9. Bei beiden Formaten sind die Werte für die Diagonalen gleich und die Seitenlängen dafür entsprechend verändert.

Die Bezeichnungen werden nun auch für CCD- und CMOS-Sensoren übernommen, eine Kamera mit 2/3“-Sensor hat entsprechend eine Bildwandlerdia-

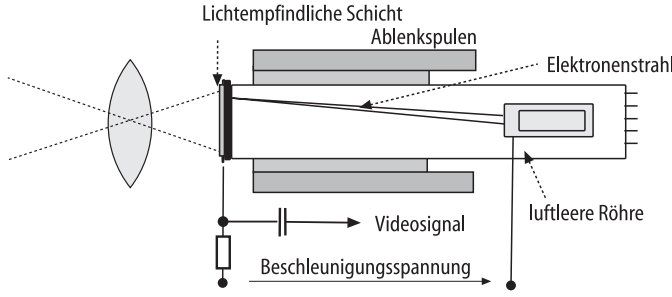


Abb. 6.4. Aufbau des Vidikons

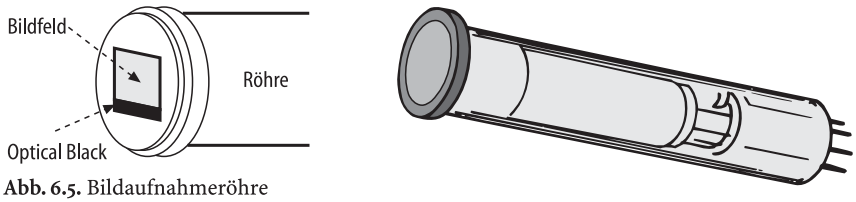


Abb. 6.5. Bildaufnahmeröhre

gonale von 11 mm. Beim 16/9-Format hat die Fläche dann einer Größe von 9,59 mm x 5,39 mm.

Moderne Kameras für den Fernsehbereich haben maximal 2/3"-Wandler. Im Heimanwenderbereich werden Sensorgrößen bis hinab zu 1/4“ verwendet. Beim Vergleich zum Film fällt auf, dass Filmbilder wesentlich größer sind, was erheblichen Einfluss auf die Schärfentiefe und die Erzielung eines Filmlook hat. Daher werden elektronische Kameras für den Bereich der Digital Cinematography möglichst mit großen Bildwandlern ausgerüstet, die eine Fläche haben, die der Filmbildflächengröße entspricht. Hier wird meist das Super-35-mm-Format als Maßstab genommen, mit einer von Fläche von ca. 24,8 mm x 18,6 mm.

Durch die Verwendung von relativ preiswerten, digitalen Spiegelreflex-Fotokameras (Digital Single Lens Reflex, DSLR) kommen weitere Bildgrößen ins Spiel, wie z. B. der sogenannte Vollformat-Sensor. Diese Bezeichnung bezieht sich auf das Kleinbild-Format 36 mm x 24 mm. Hier beträgt die Bilddiagonale 43 mm und ist damit erheblich größer als beim 35 mm-Film, mit einem entsprechend anderen Bildeindruck. Die DSLR-Kameras haben darüber hinaus noch erheblich mehr verschiedene Größen, einige sind in Abbildung 6.6 aufgeführt. Der 4/3“-Sensor hat auch die Bezeichnung MFT-Sensor (Micro Four Thirds).

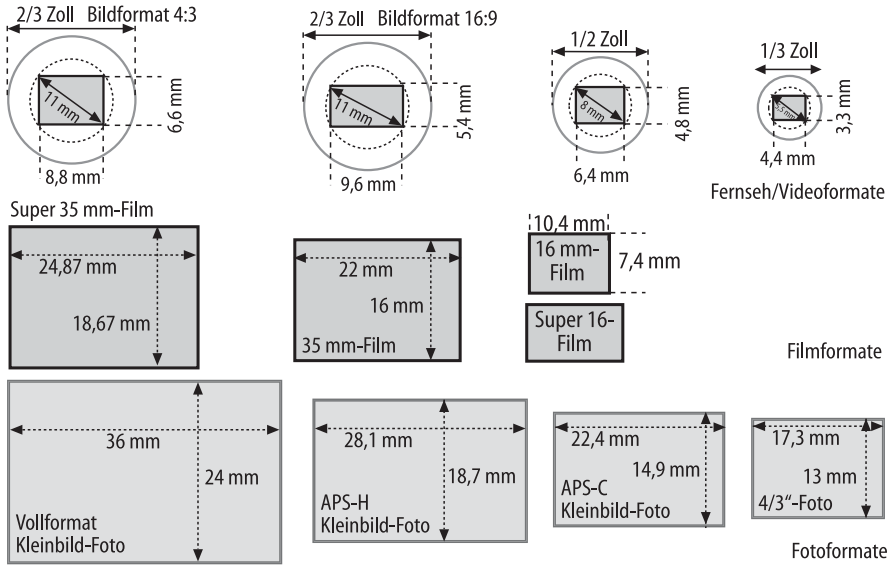


Abb. 6.6. Bildwandlerflächen im Maßstab 1:1

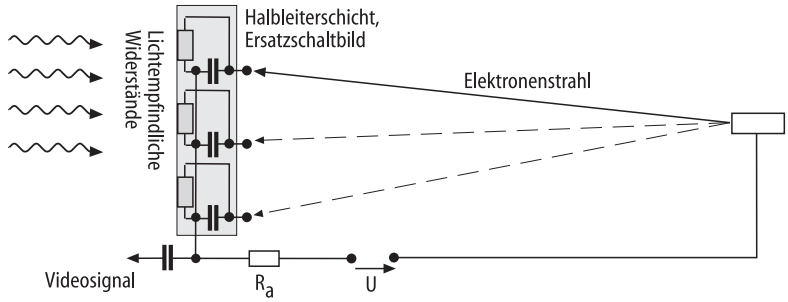


Abb. 6.7. Ersatzschaltbild der lichtempfindlichen Schicht und das Röhrenfunktionsprinzip

6.1.1.2 Die lichtempfindliche Schicht

Das Target besteht aus einer transparenten Zinnoxidschicht die als Signalelektrode dient und einer zweiten, lichtempfindlichen Schicht aus Selen oder Antimontrisulfid. Die Wirkung des Lichtes auf diese Schicht lässt sich am einfachsten verstehen, wenn man sie sich in Halbleiterpunkte zerlegt denkt und die kapazitive sowie ladungsträgerbildende Wirkung von Fotodioden betrachtet. So kann man sich vorstellen, dass jeder Bildpunkt aus einer Parallelschaltung eines Kondensators mit einem lichtempfindlichen Widerstand besteht. Gemäß der Zunahme von Ladungsträgern mit der Lichtintensität sinkt der Widerstand mit der Beleuchtungsstärke. Der Kondensator wird durch den Elektronenstrahl aufgeladen. Bei Dunkelheit ist im Idealfall der Widerstand so groß, dass sich der Kondensator bis zur nächsten Aufladung nicht entlädt. Bei hohen Lichtintensitäten wird dagegen der Widerstandswert klein, sodass der Kondensator über den Widerstand entladen wird. Der Entladungszustand jedes Kondensators repräsentiert damit die Helligkeit des entsprechenden Bildpunktes (Abb. 6.7). Der Elektronenstrahl hat nun die Aufgabe, die Bildpunkte fernsehnormgerecht abzutasten und die Kondensatoren jedes Bildpunktes wieder aufzuladen. Wenn der Strahl nach einer Vollbilddauer wieder die gleiche Stelle abtastet, ist durch das Maß der Nachladung des Bildpunktkondensators bestimmt, welche Helligkeit an dem betreffenden Bildpunkt vorgelegen hat. Zwischen den Abtastungen entleeren sich die Kondensatoren fortwährend, damit wird zeitlich über die Lichtstärke integriert und somit eine große Empfindlichkeit erreicht. Der Elektronenstromkreis wird über die Zinnoxidschicht und den Arbeitswiderstand R_a geschlossen. An R_a fällt in Abhängigkeit vom Nachladestrom mehr oder weniger Spannung ab (Abb. 6.7).

Die Speicherschicht verhält sich nicht ideal, die Kondensatoren entladen sich auch bei völliger Dunkelheit ein wenig, sodass beim Nachladen der sogenannte Dunkelstrom fließt, der temperaturabhängig ist. Zur Korrektur des daraus resultierenden Signalfehlers wird am Rand des Targets eine Schwarzmaske angebracht (Optical Black, Abb. 6.5). Das Signal aus diesem Abtastbereich entspricht dem Dunkelstrom und dient damit als Korrekturreferenz. Auf diese Weise wird gleichzeitig auch die Temperaturabhängigkeit berücksichtigt.

Die Ortsauflösung des Röhrenbildwandlers wird wesentlich durch den Durchmesser des Elektronenstrahls am Ort des Targets sowie durch die Ver-

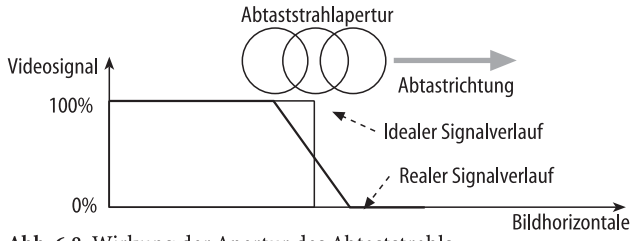


Abb. 6.8. Wirkung der Apertur des Abtaststrahls

teilung der Elektronen innerhalb des Strahles bestimmt, daraus ergibt sich die sogenannte effektive Apertur. Die Apertur begrenzt die höchste Signalfrequenz, wirkt also als Tiefpassfilter. Hohe Signalfrequenzen treten an scharfen Kanten auf. Abbildung 6.8 zeigt, wie ein solches Signal durch die Abtastung mit großer Apertur verformt wird. Mit Hilfe einer Aperturkorrektur durch Anhebung hoher Frequenzen wird der Aperturfehler in der Kamera kompensiert. Kameras mit Röhren spielen in der modernen Medienpraxis keine Rolle mehr, der Begriff Aperturkorrektur wird aber zuweilen auch bei der Signalverarbeitung für CCD- und CMOS-Kameras benutzt.

6.1.2 CCD-Bildwandler

Diese Bildwandlertypen waren die ersten, die die Röhrentechnologie ablösen konnten. Das brachte erhebliche Vorteile, denn die hier genutzten Halbleiter sind mechanisch robust, benötigen nur kleine Spannungen und Leistungen, haben ein geringes Gewicht und sind unempfindlich gegen Überbelichtung. Ein CCD-Chip (Charge Coupled Device) ist ein Schieberegister, das auch als Eimerkettenspeicher bezeichnet wird. Im CCD wird elektrische Ladung wie Wasser in einer Eimerkette von einer Speicherzelle zur nächsten transportiert.

Eine CCD-Speicherzelle wird durch einen MOS-Kondensator gebildet. MOS steht für Metall Oxid Semiconductor, Abbildung 6.2 zeigt den Aufbau. Eine Elektrode ist durch die Oxidschicht auf dem Halbleitermaterial vom Halbleiter getrennt. Durch eine an der Elektrode angelegten Spannung von z. B. + 10 V wird im Halbleiter eine Potenzialsenke gebildet, in der sich Elektronen sammeln können, während die positiven Löcher abgestoßen und in das Innere des Halbleiters gedrängt werden [142].

Durch Aneinanderreihung von MOS-Kondensatoren kann nun der Ladungstransfer bewirkt werden, der dem Bauelement seinen Namen gab. Dies geschieht durch die Veränderung der Spannungsniveaus an den Speicherzellen, z. B. mit dem 4-Phasenbetrieb, der in Abbildung 6.9 dargestellt ist. Jeder Speicherzelle sind hier vier Taktelektroden zugeordnet. Eines davon hat eine Spannung von z. B. 0 V, während die anderen drei auf + 10 V liegen, Unter diesen bildet sich die Potenzialsenke, während das 0 V-Potenzial, die Speicherbereiche oder Bildpunkte voneinander abgrenzt. Die Ladung kann nun auf einfache Weise verschoben werden, indem ein zweites Potenzial auf 0 V gesetzt und der Ladungssammelbereich verkleinert wird. Anschließend wird die Spannung an

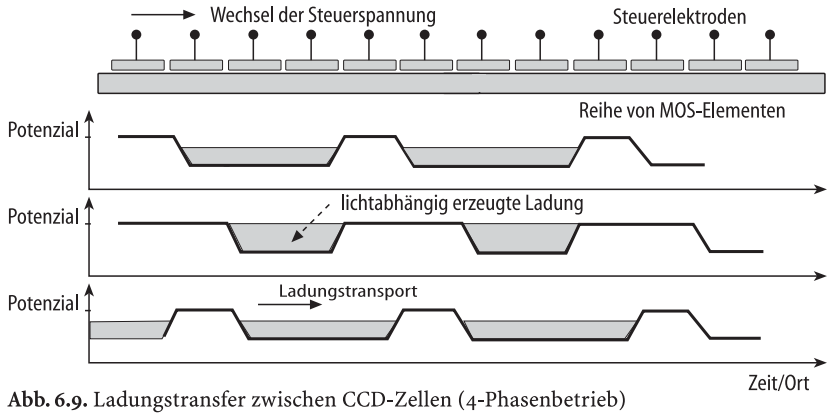


Abb. 6.9. Ladungstransfer zwischen CCD-Zellen (4-Phasenbetrieb)

der Elektrode, die in der Richtung des Transports liegt und die bisher die Zellen trennt, auf 10 V gelegt, sodass sich der Speicherbereich der Ladung um die Breite einer Elektrode verschiebt. Durch zyklische Wiederholung dieses Vorgangs kann schließlich die Ladung durch den gesamten Halbleiter geschoben werden (Abb. 6.9). Dabei ist zu bemerken, dass ein CCD ein analoges Bauelement ist. Die Spannungen an den Elektroden, die den Transport der Ladung bewirken, nehmen zwar nur wenige diskrete Zustände an, die lichtabhängig gesammelte Ladung ist aber nicht wertdiskret sondern analoger Art.

Die eindimensionale Anordnung der Zellen führt zu einer CCD-Zeile (Abb. 6.10). Sie kann zur Bildwandlung in Filmabtastern eingesetzt werden. Hier wird der Film zeilenweise abgetastet, indem er durchleuchtet und kontinuierlich vor der CCD-Zeile vorbei geführt wird (s. Abschn 5.7.2). Dabei können die CCD mit separaten Fotodioden kombiniert sein und selbst, lichtdicht abgedeckt, nur für den Ladungstransport zuständig sein. Alternativ kann die Lichtempfindlichkeit der CCD-Zelle selbst ausgenutzt werden.

Zum Einsatz in Kameras sind Bildwandlerflächen erforderlich, eine CCD-Bildwandlerfläche wird aus CCD-Zeilen zusammengesetzt. Durch die Spannungen an den MOS-Zellen wird das Bild bei der Wandlung sowohl horizontal als auch vertikal in Bildpunkte (Pixel) zerlegt. Da der Ladungstransfer auf der Fläche vertikal (zeilenweise) geschieht, kann die horizontale Trennung spaltenweise fest vorgenommen werden, indem Streifen aus hochdotiertem p-Silizium als Channel Stop benutzt werden.

Je mehr Pixels vorhanden sind, desto besser ist die Auflösung bzw. die Qualität des Bildwandlers. Die Anzahl der CCD-Zeilen sollte mindestens der aktiven Zeilenanzahl der Videonorm entsprechen. Die Pixelanzahl pro Zeile folgt aus

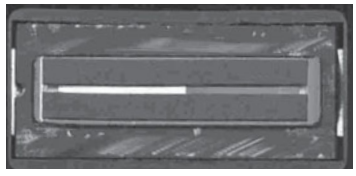


Abb. 6.10. CCD-Zeile [32]

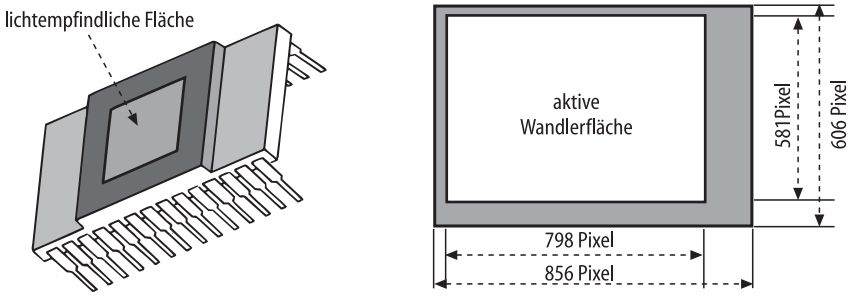


Abb. 6.11. CCD-Chip im Gehäuse und Beispiel für die Bildpunktanzahl auf der Wandlerfläche

dem Seitenverhältnis und der Berücksichtigung des Kellfaktors. Eine Abschätzung der aufgrund von Alias auftretenden Störkomponenten ergibt aber, dass bei SD-Auflösung für die aktive Zeile 800 Pixel erforderlich sind, wenn die zur Videobandbreite 5 MHz gehörigen Pixel (s. Kap. 2) störungsfrei wiedergegeben werden sollen. Ein CCD-Wandler für ein 4:3-SD-Bild sollte also mehr als 800 H x 580 V Pixel enthalten [195]. HD-Sensoren für TV-Anwendungen weisen aus ökonomischen Gründen in den meisten Fällen nur die erforderliche aktive Bildpunktzahl von 1920 x 1080 Pixels auf.

Nicht alle CCD-Zellen werden zur eigentlichen Bildwandlung genutzt. Auch CCD weisen einen Dunkelstrom auf, der wie bei Röhren mit Hilfe von zusätzlichen, lichtdicht abgedeckten Pixels korrigiert wird (Optical Black).

Die CCD-Chips werden in mit einem Fenster versehenen Gehäusen untergebracht, wie sie gewöhnlich in der Halbleiterindustrie verwendet werden (Abb. 6.11). Für die Bildwandlerabmessungen werden die Werte der aktiven Wandlerflächen von Röhrenbildwandlern übernommen (Abb. 6.6 und 6.12). Das gilt auch für die Bezeichnung, d. h. ein Wandler mit einer Bilddiagonale von 11 mm wird als ein 2/3“-CCD-Bildwandler bezeichnet. Es werden Wandlergrößen bis hinab zu 1/4“ gefertigt, in Fernseh-Studiokameras kommen meist 2/3“-CCD zum Einsatz. Die Pixelanzahl muss also auf einer Fläche von ca. 0,5 cm² untergebracht werden, der Anspruch an die Fertigungstechnologie ist dementsprechend hoch. Durch Fehler beim Fertigungsprozess kommt es vor, dass einzelne Sensorelemente ausfallen. Darum werden die CCD geprüft und selektiert. CCD mit Fehlstellen werden für preiswerte Kameras verwandt. Beim Auslesen der Fehlstellen wird das Videosignal verfälscht, es entsteht der Flaw-Effekt, der mit Hilfe elektronischer Maßnahmen verdeckt werden kann (s. Abschn. 6.2.3).

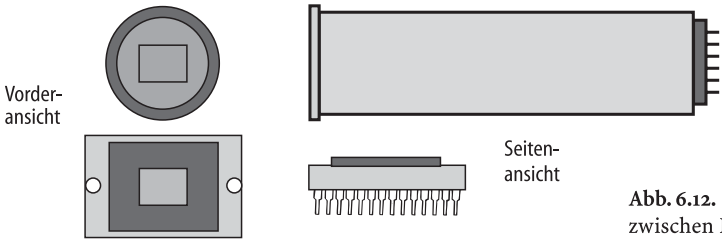


Abb. 6.12. Größenvergleich zwischen Röhre und CCD

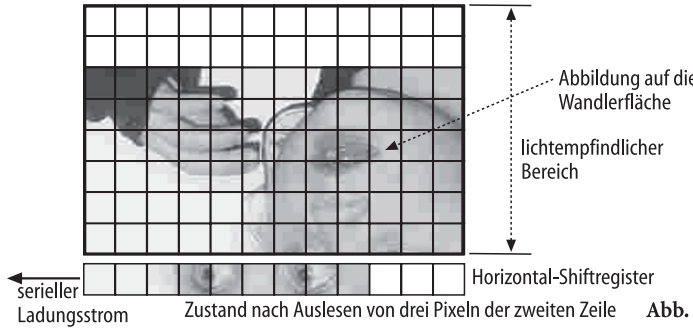


Abb. 6.13. Full-Frame-CCD

6.1.2.1 Full-Frame-CCD

Das Ladungsbild der Bildpunkte repräsentiert die optische Abbildung. Die Ladungen der Zellen müssen ausgelesen und zu einem seriellen Signal geformt werden. Das gelingt am einfachsten beim Full-Frame-CCD, bei dem unterhalb der lichtempfindlichen Fläche eine einzelne CCD-Zeile positioniert wird (Abb. 6.13). Zum Auslesen der Information wird die Ladung vertikal zeilenweise in dieses horizontale Ausleseregister geschoben. Wenn dieses dann wiederum vor jedem Zeilentransfer durch eine horizontale Ladungsverschiebung geleert wird, entsteht das serielle Signal. Full-Frame-CCD werden in Fotoapparaten verwendet. Nach der Ladungssammlung wird der Chip hier lichtdicht abgedeckt und die CCD-Zellen können zum Ladungstransport verwendet werden. Im Videobereich stehen aber für Ladungssammlung und -transport nur ca. 20 ms zur Verfügung, sodass, zur Erzielung einer hohen Empfindlichkeit, die Ladung in den Zellen möglichst lange aufintegriert werden muss. Daher werden Ladungssammel- und Auslesevorgang getrennt. Die Auslesung der Ladung geschieht nicht direkt, sondern es wird ein Speicherbereich genutzt. Die verschiedenen CCD-Typen werden im Videobereich daher hinsichtlich der Anordnung ihrer Speicherbereiche in FT-, IT- und FIT-Typen unterschieden.

6.1.2.2 Frame Transfer (FT)

Ein nach dem Frame Transfer-Prinzip aufgebauter CCD-Chip besteht aus einem lichtempfindlichen Sensorteil, einem lichtdicht abgedeckten Speicherbereich und einer einzelnen CCD-Zeile, die als horizontales Ausleseregister dient (Abb. 6.14). Die durch das Licht erzeugte Ladung wird während der aktiven (Halb-) Bilddauer gesammelt und während der Vertikalaustastung in den Speicherbereich verschoben, dabei dienen die lichtempfindlichen CCD-Zellen auch als Ladungstransportzellen. Damit die Ladungen während des Transports nicht verfälscht werden, muss die Sensorfläche in dieser Zeit lichtdicht abgedeckt werden. Die Abdeckung wird über eine rotierende mechanische Flügelblende erreicht (Abb. 6.14), die die maximale Belichtungszeit auf ca. 16 ms pro Halbbild begrenzt. Während im Sensorteil die Ladungen des nächsten Bildes gesammelt werden, kann der Speicherbereich der Fernsehnorm entsprechend zeilenweise ausgelesen werden. Dazu wird jeweils die unterste Zeile des Speicherbereichs während der horizontalen Austastlücke in das H-Ausleseregister übernommen. Das Register wird während der aktiven Zeilendauer geleert,

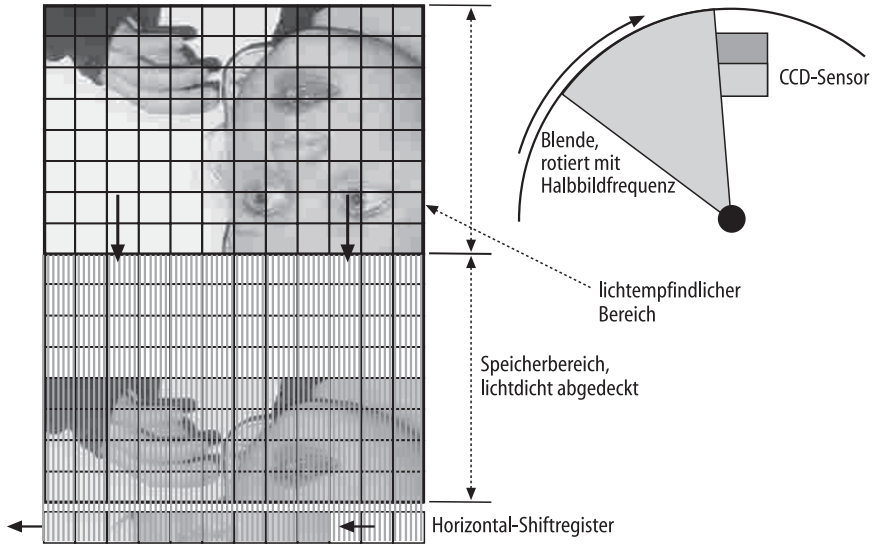


Abb. 6.14. Frame Transfer Prinzip

sodass ein serieller Ladungsstrom entsteht. Wenn der FT-Bildwandler nur für Videosysteme mit Zeilensprung genutzt werden soll, kann die CCD-Zeilenzahl halbiert werden. Der örtliche Arbeitsbereich der Zeile wird dann spannungsgesteuert für jedes Halbbild nach oben oder unten versetzt. Das geschieht durch Verschiebung der Potenzialwälle, die die Zeilen trennen. Der Speicherbereich muss dann nur die Zeilen eines Halbbildes aufnehmen.

Zur Verhinderung des Überlaufens der Potentialtöpfe bei Überbelichtung wird durch eine hohe Spannung von etwa + 20V dafür gesorgt, dass die überschüssigen Ladungen in das Substrat, d. h. in die Tiefe des Chips abfließen. Zur Verbesserung der Sensoreigenschaften werden auch CCD entwickelt, die von hinten belichtet werden können, damit die Taktelektroden nicht stören.

Das FT-Prinzip hat gegenüber anderen den Vorteil, dass die Pixeldichte sehr groß ist, womit sich eine hohe Empfindlichkeit und ein gutes Auflösungsvermögen ergibt. Der wesentliche Nachteil ist die Notwendigkeit der Verwendung einer mechanischen Blende, allerdings kann diese auch durch einen vorgesetzten Shutter auf Flüssigkristallbasis realisiert werden (LC-Shutter). Beispiele für den Einsatz von 2/3“-FT-Bildwandlern sind die HD-Kamera LDK 8000 und die SD-Kamera LDK 20 der Firma Thomson. Der Sensor hat bei letzterer 1000 x 594 Pixel, die bei MTF = 70% eine Auflösung von 800 Linien ermöglichen.

6.1.2.3 Interline Transfer (IT)

Beim IT-Bildwandler liegt der Speicherbereich nicht unterhalb, sondern in der Bildwandlerfläche. Neben den lichtempfindlichen Sensorelementen befinden sich die lichtgeschützten Speicherzellen, der Speicherbereich besteht also aus vertikalen Spalten. Während der Dauer eines Halbbildes wird die Ladung in den Sensorelementen gesammelt. In der vertikalen Austastlücke wird sie dann innerhalb von weniger als einer Mikrosekunde in den Speicherbereich gescho-

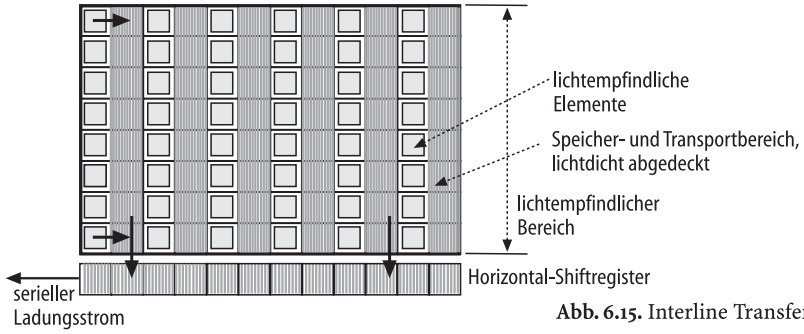


Abb. 6.15. Interline Transfer Prinzip

ben, so schnell, dass eine Lichtabdeckung während dieses Transportvorganges nicht erforderlich ist. Anschließend beginnt die Integrationszeit für das nächste (Halb-) Bild. Während dieser Zeit werden die Speicherspalten wie beim FT-Prinzip mit Hilfe des horizontalen Transportregisters fernsehnormgerecht ausgelesen (Abb. 6.15).

Ein IT-CCD arbeitet ohne mechanische Blende und bietet den Vorteil, dass wegen der schnellen Ladungsübernahme in den Speicherbereich die Integrationsdauer hoch ist. Nachteil dieses Prinzips ist die geringe Pixeldichte. Der Füllfaktor, also der Anteil der für das Licht wirksamen Fläche, liegt bei 30 ... 50%. Hinzu kommt die Anfälligkeit für den sogenannten Smear-Effekt. Der Begriff »Smear« beschreibt ein Verschmieren von Bildpunkten, das vor allem bei IT-CCD auftritt. Ursache ist der Ladungstransport in den Speicherspalten, während gleichzeitig die benachbarten Sensorelemente dem Licht ausgesetzt sind. Die Abmessungen der Lichtabdeckung für die Speicherspalten sind sehr klein, es lässt sich nicht völlig vermeiden, dass Licht auch unter die Abdekkung gelangt. Liegt beispielsweise ein sehr heller Bildpunkt in einer dunklen Umgebung, so bewirkt das helle Licht eine Ladungsvermehrung für alle Ladungen, die durch die benachbarte Speicherzelle transportiert werden. Da die transportierten Ladungsmengen aus dunklen Bildbereichen sehr gering sind, ist die Verfälschung sehr auffällig. Sie erscheint als senkrecht verlaufender Strich, der durch den hellen Bildpunkt läuft, der die Störung hervorruft (Abb. 6.16). Um diesen Nachteil gegenüber FT-CCD zu vermindern, wird von vielen Herstellern die On-Chip-Lens-Technik (OCL) verwandt. Dabei wird über jedem Sensorelement eine eigene kleine Linse angebracht, die auch das Licht auf

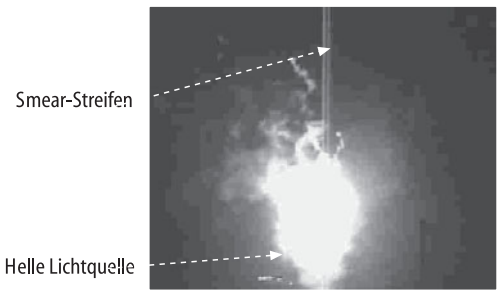


Abb. 6.16. Vertical Smear Effect

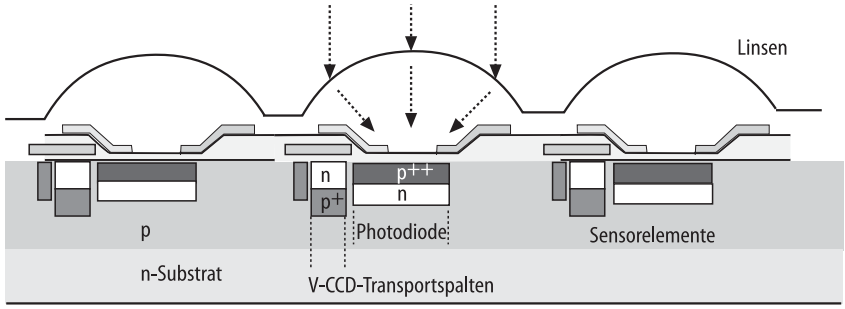


Abb. 6.17. IT-Sensoraufbau mit On Chip-Lens

die Sensoren bündelt, das sonst auf die Speicherspalten fällt und damit verlorengeht. Die damit bewirkte Lichtkonzentration trägt auch zur Reduzierung des Smear-Effekts bei, da das Licht von den Speicherspalten weg gelenkt wird (Abb. 6.17). Bei FT-CCD tritt der Smear-Effekt nicht auf, da der Sensorbereich während des Ladungstransportes abgedeckt wird.

Die Verhinderung des Überlaufens der Potenzialwälle bei Überbelichtung geschieht beim IT-CCD, indem mittels einer hohen Spannung, die gegenüber dem p-dotierten Bereich von + 15 V auf ca. + 45 V gesteigert wird, die Ladungen in das n-Substrat abgeleitet werden (Drain). Abbildung 6.17 zeigt die dafür verwendete Struktur des IT-Sensors. Der lichtempfindliche Bereich ist hier als Photodiode ausgeführt, damit keine Lichtabsorption in einer darüberliegenden Elektrode stattfindet. Die oberste Schicht der Diode ist sehr stark mit Löchern dotiert (Hole Accumulation Diode, HAD), womit die Sensoreigenschaften, wie die Verringerung von Dunkelstromstörungen verbessert werden können.

6.1.2.4 Frame-Interline-Transfer (FIT)

Das Frame Interline Transfer-Prinzip kombiniert die IT- und die FT-Technik (Abb. 6.18). Der Speicherbereich besteht wie beim IT-CCD aus Speicherspalten,

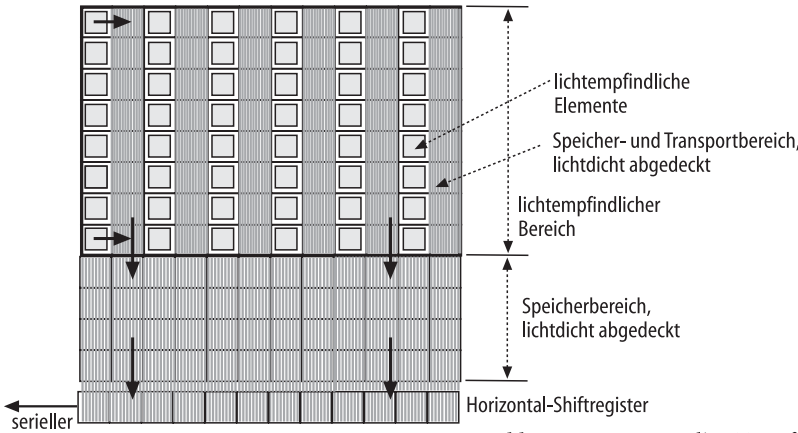


Abb. 6.18. Frame Interline Transfer Prinzip

außerdem ist ein FT-ähnlicher Speicherbereich unterhalb des Sensorbereichs vorhanden. Die Ladung aus den Sensorelementen wird wie beim IT-CCD zunächst in die lichtdicht abgedeckten Spalten geschoben. Anschließend wird sie dann sehr viel schneller als beim IT-CCD in den Speicherbereich geleitet, von wo aus sie normgerecht ausgelesen werden kann. Durch die Kombination des IT- und FT-Prinzips wird auch eine Kombination der Vorteile erreicht: Einerseits ist keine mechanische Blende erforderlich, andererseits ist der Smear-Effekt stark reduziert, denn dieser ist von der Transportgeschwindigkeit in den Speicherspalten abhängig. Nachteil des FIT-CCD ist der hohe technische Aufwand und die gegenüber dem FT-Prinzip verringerte Pixeldichte.

6.1.2.5 Auslese-Modi bei IT und FIT

Die Gewinnung von Halbbildern kann bei IT- und FIT-Chips auf verschiedene Arten erfolgen. Im Frame-Reset-Mode (Super Vertical Mode) wird die Ladung jeder zweiten Zeile ausgewertet, je Halbbild also die geraden oder ungeraden Zeilen. Die Ladung der jeweils ungenutzten Zeilen wird vernichtet. Im Field-Integration-Mode (Field Read Mode) werden dagegen zwei Zeilen paarweise zusammengefasst, wobei die Paarbildung halbbildweise wechselt (Abb. 6.19). Dieser Modus ist die Standardbetriebsart und bietet den Vorteil einer verdoppelten Empfindlichkeit, allerdings auf Kosten der Vertikalauflösung.

Als dritte Möglichkeit steht manchmal der Frame Read Mode zur Verfügung. Hier wird, wie beim Frame Reset Mode, jede Zeile einzeln ausgelesen, die damit verbundene geringere Lichtempfindlichkeit gegenüber dem Field-Integration-Mode wird durch eine Verdopplung der Belichtungszeit ausgeglichen. Die Halbbilder werden paarweise überlappend zusammengefasst. Trotz der begrifflichen Assoziation entspricht dieser Auslesemodus nicht dem Verhalten einer Filmkamera, denn dort wird das Bild nur über die Dauer eines Halbbildes gewonnen. Zwar verringert sich auch beim Frame Read Mode die Bewegungsauflösung, das Bild wird aber auch unschärfer, da es zwei Bewegungsphasen enthält (Abb. 6.20). Ein filmtypischer Bewegungseindruck kann nur mit Hilfe des Ausleseprinzips beim Pro-Scan-Mode erreicht werden.

Pro-Scan steht für einen Auslese-Modus der progressiv und nicht mit Zeilensprung arbeitet. Dieser Modus ist vor allem im Zusammenhang mit der digitalen Cinematographie interessant, die grundsätzlich auf Vollbildern basiert. Zur Erzielung eines filmähnlichen Bildeindrucks wird dieser Modus in letzter Zeit auch in semiprofessionelle Kameras implementiert. Einer der ersten Chips die progressiv ausgelesen werden konnten arbeitete nach dem FIT-Prinzip und trug die Bezeichnung M-FIT (Multiple Frame Interline Transfer).

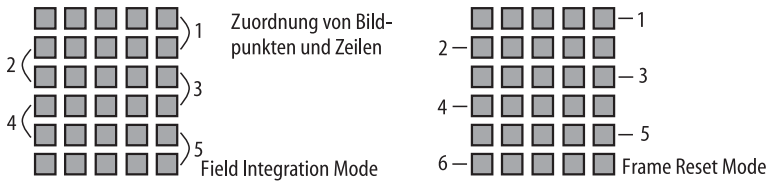


Abb. 6.19. Methoden zur Halbbildgewinnung bei IT- und FIT-CCD

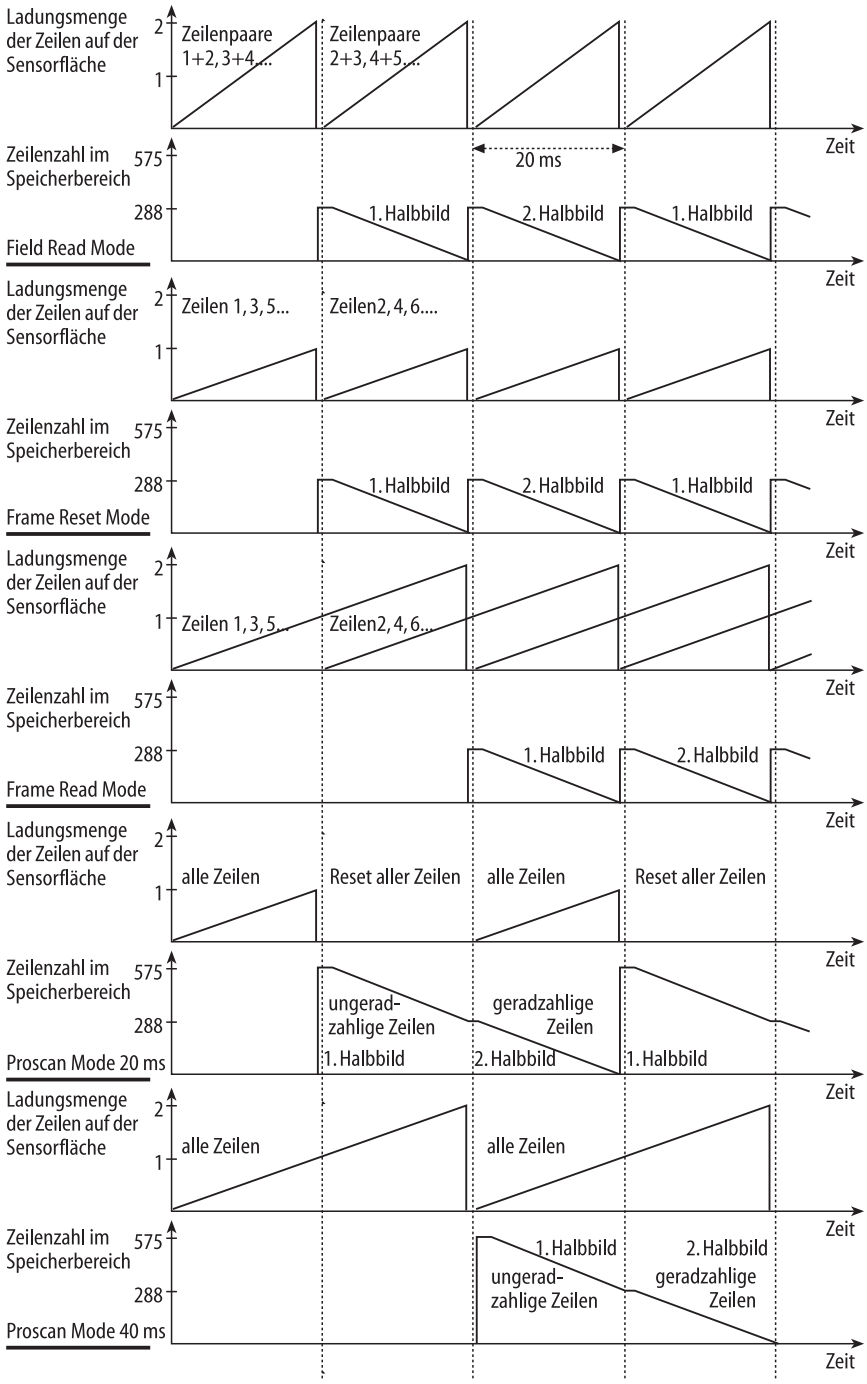


Abb. 6.20. Vergleich aller Auslese-Modi bei IT- und FIT-CCD

Bei diesem und anderen Pro-Scan-Sensoren ist die Speicherfläche unterhalb des bildaktiven Bereichs verdoppelt, sodass nicht nur die Zeilen eines Halbbildes, sondern alle 576 (bzw. 1080 oder mehr bei HD-Kameras) aktiven Zeilen zur gleichen Zeit abgelegt werden können. Damit eine Bewegungsauflösung erzielt werden kann, die der von Filmkameras bei 25 fps entspricht, wird der Bildwandler im 20-ms-Modus so gesteuert, dass die Ladung nur während der Dauer eines Halbbildes gesammelt wird und während der folgenden V-Austastlücke zunächst die ungeraden und direkt danach dann die geraden Zeilen in den Speicherbereich geschoben werden. In der folgenden Halbbildperiode wird, wie üblich, das erste Halbbild ausgegeben, die während dieser Zeit im Bildbereich entstehenden Ladungen werden abgeführt. Anschließend wird das zweite Halbbild ausgelesen, während die Ladungen für das nächste Vollbild gesammelt werden. Mit diesem 20 ms-Modus wird erreicht, dass die Bewegungsauflösung nicht auf zwei verschiedenen Halbbildern und 50 Hz, sondern auf einem Vollbild und 25 Hz beruht (Progressiv Scan). Die Ladungsbildung für ein Bild entsteht in allen Zeilen gleichzeitig. Alternativ, aber nicht filmähnlich kann im 40 ms-Modus die Ladung über die Vollbilddauer gewonnen werden. Für die videotechnische Umsetzung werden erst anschließend, wie bei der Filmabtastung, zwei Halbbilder gewonnen.

Eine Pro-Scan-Kamera erzeugt im 20-ms-Modus also eine Bewegungsauflösung, die der einer Filmkamera entspricht, denn auch dort entsteht ein Vollbild mit 24 oder 25 Hz [236]. Diese ist schlechter als im gewöhnlichen Videosystem, aber gerade das soll erreicht werden, denn damit kommt man dem vielzitierten Filmlook näher, der sich vor allem auch durch eine besondere, in gewissem Maße realitätsfremde Bildanmutung auszeichnet, die sich neben der geringen Bewegungsauflösung in der unregelmäßigen Verteilung der Filmkorns und auch im nicht perfekten Bildstand niederschlägt (s. Kap. 5). Abbildung 6.20 zeigt den Vergleich aller Auslesemodi [236].

6.1.3 CMOS-Bildwandler

MOS-Bildwandler wurden noch vor den CCD-Wandlern entwickelt, doch dominierten Letztere im praktischen Einsatz aufgrund des Umstandes, dass CCD-Wandler kleinere Bildpunkte ermöglichten. Die MOS-Technologie wurde zu CMOS (komplementär) weiterentwickelt und ist die Grundlage moderner Computerschaltkreise. Sie ermöglicht heute Strukturgrößen, die mit denen bei CCD-Sensoren konkurrieren können.

Der grundlegende Aufbau basiert auf einem hochintegrierten CMOS-Schaltkreis mit einer Vielzahl von Transistoren. Einigen von ihnen werden lichtempfindliche Bereiche zugeordnet, in denen sich, entsprechend der auffallenden Intensität, Ladungsträger bilden, wobei sowohl die MOS-Struktur mit einem Photogate als auch mit Photodioden genutzt werden können [227].

Die Ladungen können nun direkt auf dem CMOS-Chip verarbeitet werden, was gegenüber CCD mit vielen prinzipiellen Vorteilen verbunden ist. Einer davon ist der Zugriff auf die Ladung der einzelnen Bildpunkte, der bei CMOS durch die direkte Adressierung des Bildpunktes mittels einer xy-Matrix erfol-

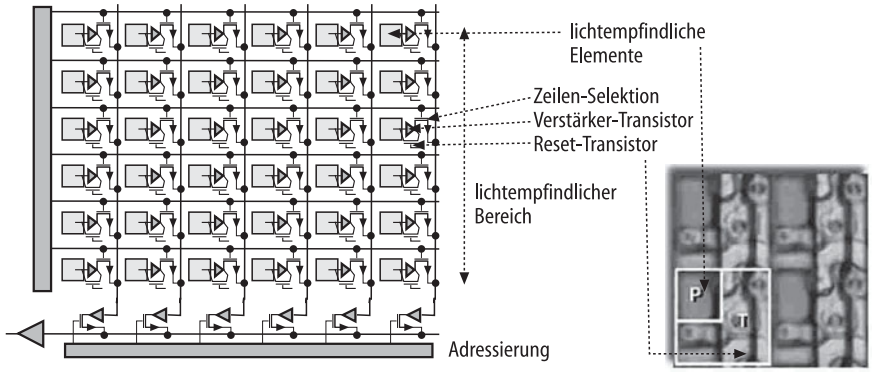


Abb. 6.21. CMOS-Funktionsprinzip und Ausschnitt aus der Sensorfläche

gen kann. Die Transistoren auf dem Chip schalten die adressierte Information auf die Leitungen (Abb. 6.21), sie müssen nicht durch den ganzen Chip geführt werden. Damit ist eine einfache Auslesung wahlweise im interlaced- oder progressiv-Mode möglich. Darüber hinaus kann auch nur ein Teilbereich des Wandlers, die sogenannte Region of Interest (ROI), ausgegeben werden.

Zu Beginn der Entwicklung war jedem Bildpunkt nur ein Transistor zugeordnet (Passive Pixel Sensor, PPS), moderne CMOS-Sensoren arbeiten dagegen mit „Aktiven Pixel Sensoren“ (APS). Hier werden mindestens drei Transistoren (3T-Pixel) eingesetzt, die zur Adressierung, Verstärkung und zur Ableitung der Ladung dienen (Reset). Bei Verwendung von vier oder fünf Transistoren pro Pixel ist eine aufwendige Signalverarbeitung direkt auf dem Chip möglich, von der Rauschreduktion bis hin zur A/D-Konversion, sodass Sensoren gebaut werden können, die direkt ein Digitalsignal ausgeben.

Die direkte Verstärkung im Bildpunkt führt zu einer Umsetzung der gesammelten Ladung in einen Spannungswert, der anschließend weiterverarbeitet wird. Von Vorteil ist dabei, dass die Verstärkung an dem Ort stattfindet, an dem noch der größte Signal/Rauschabstand vorliegt. Im Gegensatz dazu muss beim CCD die Ladung durch den gesamten Sensor geschoben werden und wird erst am Ausgang in eine Spannung umgesetzt [168]. In Verbindung mit niederohmigen Buselektroden ergibt sich dadurch bei CMOS-Chips ein Leistungsbedarf im mW-Bereich, während er bei CCD viel höher liegt. Dieser Umstand ist wichtig, weil damit die Erwärmung und die damit verbundene Dunkelstromproblematik bei CMOS eine geringere Rolle spielt.

Die Verarbeitung des Signals direkt am Bildpunkt bedeutet auch, dass hier die Ladungswechselfrequenz relativ gering ist, während beim CCD-Ausgang mit Pixelraten von mehr als 10 Mpix/s gerechnet werden muss. Daher ist es relativ einfach möglich, CMOS-Sensoren für schnelle Bildwechsel auszulegen.

Neben der hohen möglichen Bildrate ergibt sich ein besonderer Vorteil der CMOS-Technologie im Hinblick auf Bildwandler, die für die professionelle Videotechnik und digitale Cinematografie geeignet sind. Hier ist von großem Interesse, den verarbeitbaren Szenenkontrast über die 50 ... 60 dB hinaus zu steigern, die mit CCD-Wandlern erreichbar sind. Durch teillinearares Auslesen

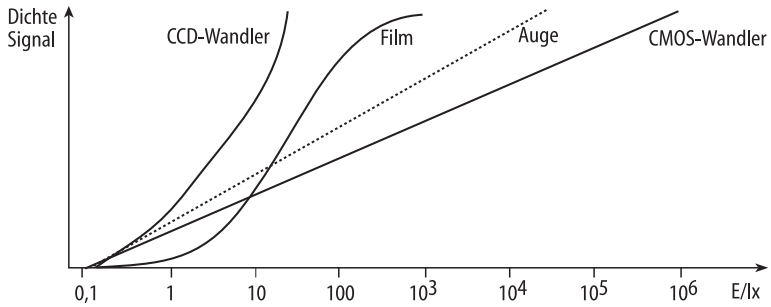


Abb. 6.22. Vergleich der erfassbaren Szenenkontraste

der Informationen der einzelnen Bildpunkte kann bei CMOS ein Szenenkontrast von über 90 dB erreicht werden. (Abb. 6.22)

Neben den vielen Vorteilen von CMOS-Sensoren gibt es auch Nachteile. Diese resultieren vor allem daraus, dass die Transistoren eine große Fläche auf dem Chip beanspruchen. In Abbildung 6.21 sind sie rechts durch der Buchstaben „T“ gekennzeichnet. Die aktive Sensorfläche wird daher verhältnismäßig klein, d. h. der Füllfaktor sinkt. Abhilfe kann, wie bei IT-CCD, durch die On-Chip-Lens-Technologie geschaffen werden. Ein zweites Problem ergibt sich durch die große Zahl der Transistoren, die in ihrer Vielzahl nicht völlig gleich gestaltet werden können. Unvermeidlich folgt daraus, dass die Ladungen in jedem Pixel etwas unterschiedlich behandelt werden. Der störende Effekt wird als Rauschen bezeichnet, da er unregelmäßiger Natur ist. Die genauere Bezeichnung lautet Fixed Pattern Noise (FPN), da es aus der örtlichen Verteilung resultiert. Das FPN darf nicht mit dem unten dargestellten, gewöhnlichen, zeitabhängigen Rauschen verwechselt werden. Das FPN war lange ein Hinderungsgrund für die CMOS-Entwicklung. Heute kann es durch Filterprozesse, die direkt auf dem Sensor stattfindet, erheblich gemindert werden [115].

Ein im Bild sichtbarer Fehler, der bei CCD nicht auftritt, ergibt sich bei CMOS-Sensoren aus der zeilenweisen Adressierung beim Auslesen der Bildwandlerfläche. Während CCD-Sensoren immer einen Speicherbereich beinhalten, der gewährleistet, dass der Ladungssammel- vom Auslesevorgang getrennt ist, ist dies bei CMOS nicht selbstverständlich. Um hier zu gewährleisten, dass alle ausgegebenen Ladungen aus demselben Zeitbereich stammen, müssen die Ladungen auch zwischengespeichert werden. Diese als Synchronous oder Global Shutter bezeichnete Funktion ist mit einem hohen Signalverarbeitungsaufwand sowie vielen Transistoren auf dem Chip verbunden. Ein Sensor nach dieser Bauart, die also den Lichtsammel- vom Auslesevorgang trennt, wird unter dem Namen FT-CMOS entwickelt. Die meisten Kameras arbeiten bisher jedoch mit dem sogenannten Rolling Shutter, d. h., dass eine Zeile oder Spalte ausgelesen wird, während die Nachbarzeile gerade Ladung sammelt. Die Zeilen stammen damit aus zeitlich unterschiedlichen Belichtungsphasen, was bei hohen Geschwindigkeiten im Motiv oder schnellen Kameraschwenks Bildverzerrungen bewirken kann. Horizontale Bewegungen führen dazu, dass vertikale Strukturen schräg abgebildet werden. Bei schnellen vertikalen Bewegungen können

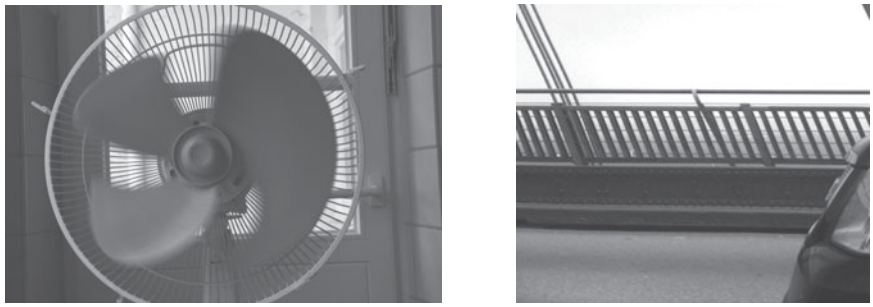


Abb. 6.23. Bildartefakte bei CMOS-Kameras [41]

Bildteile in die Länge gezogen werden, wenn sie sich in Richtung des Auslesevorgangs bewegen. Abbildung 6.23 zeigt diesen Effekt an der verzerrten Darstellung eines Ventilatorblatts, das links gegen und rechts mit der Bewegung rotiert [41]. Im rechten Teil der Abbildung ist die verzerrte Darstellung eines Brückengeländers bei schneller Vorbeifahrt zu sehen. Wie stark der Effekt auftritt, hängt von der Bauweise der Kamera ab. Vollständig kann er auch dadurch kompensiert werden, dass eine mechanische Flügelblende den Sensor nach jedem Bild abdeckt, sodass hier ein Auslesezeitraum zur Verfügung steht, in dem keine Ladung gesammelt wird. Viele Hersteller versuchen, den Fehler dadurch zu minimieren, dass der Auslesevorgang möglichst schnell vorgenommen wird.

6.1.4 Plenoptischer Bildwandler

Für diese Form der Bildumsetzung wird auch der Begriff Lichtfeldkamera verwendet. Die Besonderheit besteht darin, dass anstelle einer zweidimensionalen Abbildung ein mehrdimensionales Lichtfeld verarbeitet wird, wobei insbesondere die Richtung erfasst wird, aus der ein auf den Sensor fallender Lichtstrahl kommt. Das mehrdimensionale Lichtfeld kann durch eine plenoptische Funktion theoretisch beschrieben werden. Die Informationen aus dem Lichtfeld werden auf eine Vielzahl von Bildpunkten gegeben (Abb. 6.24). Sie können dazu benutzt werden, in einer Abbildung nachträglich die Schärfe zu bestimmen oder zu verlagern, daher braucht auch kein Fokussiervorgang abgewartet zu

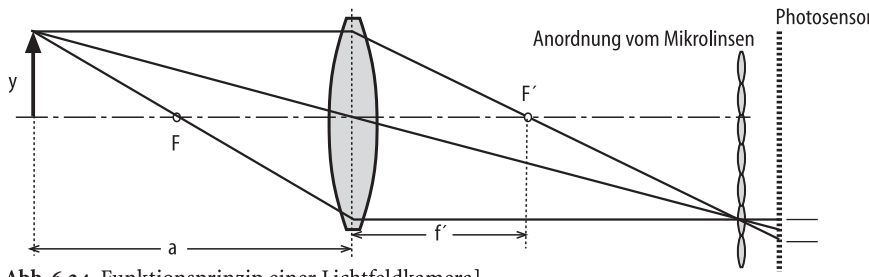


Abb. 6.24. Funktionsprinzip einer Lichtfeldkamera]

werden. Interessant ist die Kamera auch für stereoskopische 3D-Anwendungen, weil sich aus dem Lichtfeld auch die erforderlichen verschiedenen Ansichten eines Objekts gewinnen lassen.

Eine einfache Konstruktion zur Realisierung besteht darin, zwischen der Hauptlinse und dem Bildwandler ein Array von Mikrolinsen zu platzieren. Die Lichtstrahlen, die von einem Objektpunkt ausgehen, werden dadurch nicht in einen Sensorpunkt, sondern auf eine Fläche auf dem Sensor verteilt, in der die verschiedenen Strahlrichtungen repräsentiert sind. Die Ansprüche an die Auflösung des Bildwandlers sind entsprechend hoch, bzw. muss bei gewöhnlicher Bildpunktanzahl mit einer Verringerung der effektiven Auflösung der Bilddarstellung gerechnet werden [77].

6.1.5 Sensorempfindlichkeit

Die Empfindlichkeit von CMOS- und CCD-Sensoren ist von der Quantenausbeute (Ladung pro Lichtquant) bestimmt. Einen Einfluss hat auch der Aufbau des Sensors. FT-CCD z. B. sind aufgrund der größeren Pixel prinzipiell empfindlicher als IT- und FIT-CCD, allerdings wird ein Teil des Lichts durch die Taktelektroden absorbiert, die bei FT-Chips über den lichtempfindlichen Bereichen und bei IT- und FIT-CCD auf den Speicherspalten verlaufen.

Die für die Videopraxis relevante Angabe der Sensorempfindlichkeit beruht auf der Angabe der Blendenzahl, die erforderlich ist, um bei gegebener Beleuchtungsstärke E den vollen Videosignalpegel zu erreichen. Der Bezugswert ist die Beleuchtungsstärke in der Szene. Zu Messzwecken wird eine weiße Fläche mit einem Remissionsgrad von 89,9 % mit Kunstlicht der Farbtemperatur 3200 K und einer Beleuchtungsstärke von 2000 lx beleuchtet. Bei modernen Videokameras wird der Videopegel von 100 % unter den genannten Bedingungen bei ca. Blende k = 11 erreicht.

Durch zusätzliche Verstärkung kann bewirkt werden, dass der maximale Videopegel bereits bei geringerer Beleuchtungsstärke erzielt wird, allerdings wird dabei der Rauschen mit verstärkt und es vermindert sich der Signal-Rauschabstand. Bei gängigen Kameras kann eine Maximalverstärkung von 18 dB zugeschaltet werden, der Signal-Rauschabstand ist dann um genau denselben Wert reduziert. Mit der oben angegebenen Empfindlichkeit ist bei F = 2 und einer Verstärkung von 18 dB dann nur noch eine Beleuchtungsstärke von 7,8 lx erforderlich. Für Amateurkameras wird oft mit noch kleineren Minimalbeleuchtungsstärken geworben. Derartige Angaben stellen keinen Vergleichsmaßstab dar, denn sie beziehen eine Verstärkung ein, die nicht angegeben wird.

Im Zuge der Entwicklung der Digital Cinematography ist der Wunsch gestiegen, wie mit konventionellen Filmkameras zu arbeiten. Daher geben viele Hersteller die Empfindlichkeit auch in ISO(ASA)-Werten an sowie die Möglichkeit vor, an der Kamera verschiedene Empfindlichkeiten zu wählen. Praktisch bedeutet die Wahl einer erhöhten Empfindlichkeit die Zuschaltung einer Verstärkung. Eine Empfindlichkeitsverdopplung von z. B 400 auf 800 ISO entspricht einer Verstärkung um 6 dB. Um den gleichen Wert wird auch das Rauschen angehoben. Die damit verbundene Signalverschlechterung ist aber

nicht immer wahrnehmbar, da viele derartige Kameras intern einen großen Signal/Rauschabstand aufweisen (s. Abb. 6.28). Die Empfindlichkeitsangabe in ISO-Werten folgt mit der Belichtungsdauer t pro Bild aus der Beziehung

$$\text{ISO-Wert} = k^2 \cdot 226 \text{ lxs} / (R \cdot E \cdot t).$$

Diese lässt sich aus folgender Abschätzung herleiten: Aus der Filmtechnik ist der ISO/ASA-Wert für Farbbilder definiert als 1,41 lxs/H_M. Dabei ist $H_M = E_M \cdot t$ die minimale Belichtung, die gerade zur Filmschwärzung führt. Es wird angenommen, dass sie um 5 Blenden (Faktor 32) unter der Maximalbeleuchtung liegt, die für den oben genannten Fall zur Videosignalvollaussteuerung führt. Unter der weiteren Annahme, dass die Beleuchtungsstärke E_s auf der Sensorfläche A_2 mit einer eingeschränkten Linsentransmission τ behaftet ist, folgt mit $E_s = \tau \cdot \Phi/A_2$ die Beziehung $H_M = 1/32 \cdot \tau \cdot \Phi/A_2 \cdot t$. Im folgenden Abschnitt wird gezeigt, dass der Lichtstrom über die Gleichung $\Phi = 1/4 \cdot R \cdot E \cdot A_2/k^2$ aus der Beleuchtungsstärke E und dem Remissionsgrad R folgt, sodass sich, unter der konservativen Annahme einer Linsentransmission von $\tau = 0{,}80$, die genannte Beziehung ergibt.

Wenn ein Hersteller z. B. angibt, dass eine Kamera unter den oben genannten Messbedingungen bei 2000 lx die Vollaussteuerung bei Blende F 10 erreicht, so folgt daraus ein Empfindlichkeitswert von ca. 600 ISO, wenn die Kamera auf Filmbedingungen mit 24 fps und eine Äquivalenz mit einer 180°-Blende, d. h. eine Belichtungsdauer von 1/48 s, eingestellt wird.

6.1.6 Rauschen und Dynamik

Jedes Bildwandlersignal ist mit Rauschen verbunden. Die Ursache dafür ist bei Röhrenkameras vor allem die erste Verstärkerstufe, beim CCD- oder CMOS-Sensor ist es der Bildwandler selbst. Die Photonen-Absorption ist ein Zufallsprozess und führt zum sogenannten Schrotrauschen. Weiterhin tritt das thermische Rauschen auf, Dunkelstromrauschen, das Fixed Pattern Noise und weitere Anteile, die in Abbildung 6.25 dargestellt sind.

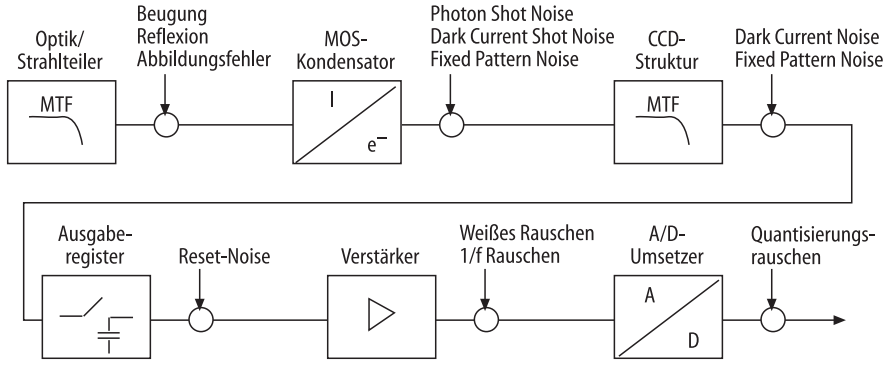


Abb. 6.25. Rausch- und Störeinflüsse beim Bildwandler

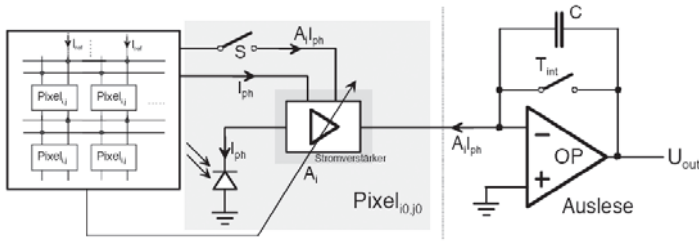


Abb. 6.26. Ladungsbewertung zur Steigerung des Szenenkontrasts im CMOS-Sensor [196]

Das Rauschen ist ein wichtiger Parameter für die Bildqualität. Ein verrauschtes Bild weist unregelmäßige kleine Flecken auf, es wirkt körnig. Das Rauschen ist daher ein wichtiger Parameter für die Dynamik bzw. den vom Bildwandler verarbeitbaren Szenenkontrast. Die Dynamik entspricht dem Aussteuerbereich, der nach oben hin durch die Maximalladung im Bildpunkt (full-well capacity, saturation signal) und nach unten durch das Rauschen (noise floor, rms-noise) begrenzt ist. Die Definition des Dynamic Range (DR) lautet:

DR = 20 log (saturation signal / rms-noise).

Die Schwelle der visuellen Erkennbarkeit von Rauschstörungen liegt bei ca. S/N = 42 dB. Moderne CCD weisen insgesamt einen linearen Signal-Rauschabstand von mehr als 55 dB auf, sodass noch eine Reserve zur Verfügung steht, die ein gewisses Maß an elektronischer Verstärkung zulässt [115]. Neben dem hier genannten linearen oder unbewerteten wird manchmal auch ein bewerteter Störabstand angegeben, der die verringerte Empfindlichkeit des menschlichen Auges bezüglich hoher Ortsfrequenzen berücksichtigt, sodass die bewerteten Störabstände um ungefähr 8 dB höher liegen als die linearen.

Mit der CMOS-Technologie werden zunächst ähnliche Störabstandswerte wie beim üblichen CCD erreicht, doch gibt es hier die Möglichkeit der modifizierten Auslesung. Während CCD-Sensoren linear arbeiten und nur die Ladung des gesamten Chips beeinflusst werden kann, nämlich ggf. mit einer Ableitung der Ladung in das Substrat, ist dieser Vorgang bei CMOS pixelselektiv möglich, da jeder Bildpunkt über einen eigenen Reset-Transistor verfügt. Durch teillineares Auslesen, also einer pixelselektiven Bewertung der Ladung (lokaladaptive Pixel) können hier sehr helle Bildbereiche ebenso erfasst werden wie dunkle [196]. Abbildung 6.26 zeigt, wie dabei die Signale aus der Photodiode in Abhängigkeit von den Signalen der Nachbarpixel geregelt werden. Das Ergebnis ist, dass ein sehr großer Szenenkontrast in der Bildvorlage umgesetzt werden kann. Existierende Wandler erlauben die Verarbeitung eines Szenenkontrasts von 120 dB der die Fähigkeit des Auges übersteigt (Abb. 6.22 und 6.27). Abbildung 6.27 zeigt, dass die im linken Bild (CCD) überstrahlten Bereiche im rechten (CMOS) noch differenziert werden können.

Die Erfassung dieser extrem hohen Szenenkontraste ist vor allem für wissenschaftliche Zwecke erforderlich. Für Video/Broadcastanwendungen reichen die oben genannten 54 dB aus. Im Rahmen der Digital Cinematography möchte man sich jedoch den Gegebenheiten des analogen Films nähern. Dieser ist

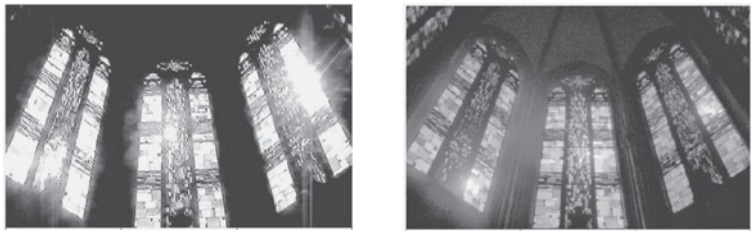


Abb. 6.27. Vergleich des Belichtungsverhaltens von CCD (links) und CMOS-Wandler [111]

durch die Erfassung von Szenenkontrasten bis zu 12 Blenden gekennzeichnet. Das erlaubt einen umfangreichen Spielraum bei der Belichtung am Set und auch bei der Lichtbestimmung, dem Grading, in der Postproduktion, denn im Gegensatz zur Broadcastfall wird am Filmset die Signalaussteuerung bzw. die Bildqualität meist nicht direkt bei der Aufnahme endgültig festgelegt.

Ein Szenenkontrast von 12 Blenden entspricht einem Dynamikbereich von 72 dB, jede Blendenstufe entspricht einem Faktor 2 bzw. 6 dB. Dieser Wert kann mit modernen Bildwandlern bereits erreicht werden, ohne auf das für CMOS-Sensoren mögliche teillineare Auslesen zurückgreifen zu müssen. So erreicht man bei guten Sensoren z. B. eine full-well-Kapazität von 80 000 Elektronenladungen und kann das Rauschen so weit reduzieren, dass es nur noch 17 Ladungen erzeugt: DR = 20 log (80 000 / 17) = 73,5 dB.

Diese Betrachtung beruht auf einer linearen Zuordnung der Lichtwerte (Szenenkontrast) zu den Signalwerten (Dynamikbereich). Während Filmmaterial aufgrund der flachen Gradation von $\gamma = 0{,}6$ bei Negativfilm den Szenenkontrast komprimiert in den Schwärzungs- bzw. Signalbereich umsetzen kann, ist bei elektronischen Kameras diese nichtlineare Zuordnung nur mit Hilfe der nachgeschalteten Signalverarbeitung möglich. Der Bildwandler selbst arbeitet linear und muss daher die DR = 72 dB erreichen, wenn tatsächlich ein Szenenkontrast von 12 Blenden erfasst werden soll (Abb. 6.28) [220].

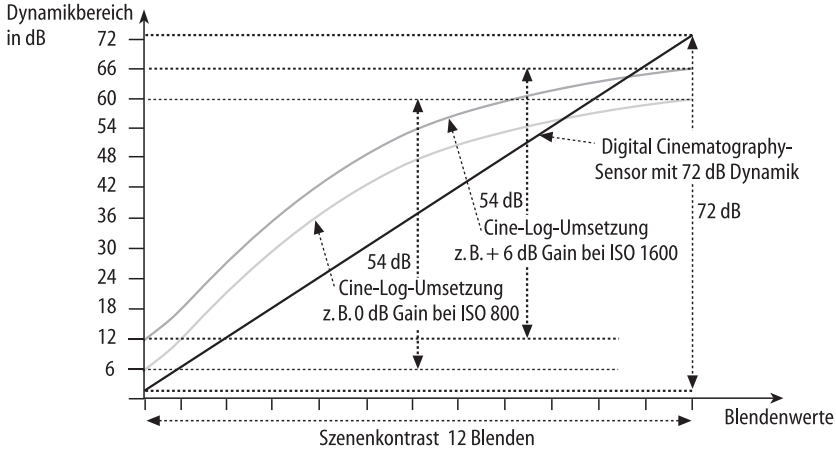


Abb. 6.28. Szenenkontrast und Dynamik

Dabei reicht es weiterhin, für das Endprodukt nach der Postproduktion einen kleineren Dynamikumfang von z. B. 54 dB bereitzustellen. Entsprechend ergibt sich also einen Belichtungsspielraum von 18 dB oder 3 Blenden, der in der Postproduktion genutzt werden kann oder bei der Aufnahme. Deshalb geben Hersteller z. B. an, dass eine (Über-)Belichtung bis zu 800 % mit ihrer Kamera möglich sei. Oder eine Steigerung der Empfindlichkeit von ISO 640 auf 2560, ohne dass sich dabei der SNR von 54 dB am Signalausgang ändert [220].

6.1.7 Global Shutter

Der Begriff steht für eine Veränderungsmöglichkeit des Belichtungszeit auf elektronischem Wege, abgeleitet von der Verschlusszeit einer mechanischen Blende. Sie betrifft das ganze Bild und wird auch Electronic Shutter genannt. Durch verkürzte Belichtungszeiten lassen sich schnelle Bewegungsabläufe ohne Unschärfen darstellen. Bewegt sich z. B. ein Tennisball in 20 ms von links nach rechts durch das Bildfeld, so erscheint normalerweise seine Position über die ganze Bildbreite verschmiert, da er sich am Anfang der Ladungssammelzeit am linken und am Ende der Zeit am rechten Bildrand befindet. Durch den elektronischen Shutter (Abb. 6.29) wird eine Verkürzung der Integrationszeit bewirkt, und der Ball erscheint nur an einer Stelle im Bild und wird so deutlicher abgebildet. Dazu wird die im ersten Teil der Gesamtintegrationsdauer gesammelte Ladung abgeführt und nur die Ladung genutzt, die im restlichen Teil der Halbbildperiode auftritt. Zur Abfuhr der überschüssigen Ladung wird beim CCD-Sensor die Ladungsabfuhr mittels des Drain benutzt der auch zur Abfuhr überschüssiger Ladung verwendet wird.

Die normale Belichtungszeit von 1/50 s pro Halbbild wird durch den Shutter auf 1/100 s, 1/500 s, 1/1000 s etc. verringert. Moderne Kameras erlauben auch eine stufenlose Einstellung der Belichtungszeit. Damit lässt sich die Kamera an beliebige Frequenzen anpassen, was z. B. von Bedeutung ist, wenn man das Bild eines Computermonitors, der mit einer Bildwechselfrequenz von 75 Hz arbeitet, ohne störende Interferenzstreifen aufnehmen will (Clear Scan). Bei Einsatz des Shutters sinkt die Empfindlichkeit proportional zur verminderten

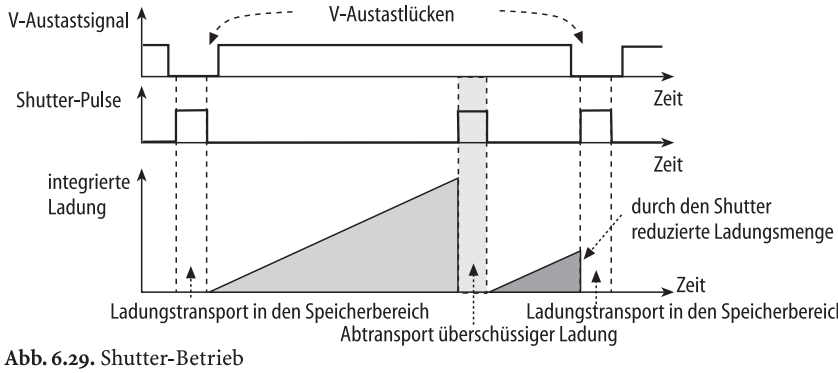


Abb. 6.29. Shutter-Betrieb

Ladungsintegrationszeit. Zum Ausgleich muss die Beleuchtungsstärke oder die Blende verändert werden.

Anders herum betrachtet, kann der Shutter auch zur Reduktion einer zu großen Lichtintensität genutzt werden, z. B. um mit relativ weit geöffneter Blende zu arbeiten und so eine geringere Schärfentiefe zu erzielen. Dabei ist allerdings für die Bildästhetik die damit verbundene, oben genannte relativ große Bewegungsschärfe zu berücksichtigen. Wenn es darum geht, mit einer elektronischen Kamera möglichst genau den Filmlook zu imitieren, so ist bei einer Bildrate von 24 fps der Shutter so zu wählen, dass die Belichtungszeit 1/48 s beträgt, um dem Bildeindruck nahe zu kommen, der gewöhnlich mit einem Hellsektor von 180° bei der Filmkamera erreicht wird.

6.1.8 Defokussierung

Im Unterschied zu Röhren- ist das Bild bei CCD- und CMOS-Bildwandlern auch in horizontaler Richtung in diskrete Punkte zerlegt. Damit ist eine räumliche regelmäßige Strukturierung in zwei Dimensionen vorgegeben die als räumliche Abtastung aufgefasst werden kann. Die räumlichen Strukturen in dem abgetasteten Bild können nun mit der Abtaststruktur selbst interferieren, Wenn dabei das Abtasttheorem nicht beachtet wird, d. h. dass das Abtastmuster doppelt so fein sein muss, wie die feinsten Strukturen in abgetasteten Bild, kann es zu Störmustern kommen, die als Aliasing bezeichnet werden. Die strenge Erfüllung des Abtasttheorems erfordert die Beschränkung der Bildauflösung, sodass nur Bildstrukturen den Sensor erreichen, die halb so gut aufgelöst sind wie es die Wandlerauflösung selbst vorgibt. Praktisch geht man nicht ganz so streng vor, da sehr feine Muster selten lange statisch im Bild bleiben.

Mögliche Aliasanteile werden vermindert, indem das aufzunehmende Bild mit einem vor dem Bildwandler platzierten optischen Tiefpassfilter defokussiert wird, das aus mehreren doppelbrechender Quarzplatten unterschiedlicher Orientierung besteht und damit ein breitbandiges Filterverhalten aufweist. Die möglichen Interferenzen (Alias) sind bei CMOS-, IT- und FIT-Sensoren wegen des geringeren Füllfaktors stärker ausgeprägt als bei FT-CCD. Bei Röhrenbildwandlern brauchte die Defokussierung nur bezüglich der Vertikalen vorgenommen zu werden. Hier wird die Öffnung des Abtaststrahls genutzt, die so eingestellt wird, dass er mit den Nachbarzeilen überlappt.

Das Defokussierungsfilter muss der Auflösung bzw. der Bildpunktanzahl des Bildwandlers angepasst sein. Diesbezüglich gibt es Probleme bei Geräten, die mit unterschiedlichen Auflösungen arbeiten. Besonders auffällig sind diese bei den sogenannten DSLR-Kameras (Digital Single Lens Reflex), die als hochauflösende Fotoapparate konstruiert sind. Sie werden im Videomode gerne verwendet, um preiswert einen Filmlook mit geringer Schärfentiefe zu erzeugen. Die Defokussierung bezieht sich hier aber auf den Fotobereich mit einer Auflösung von z. B. 20 MPixels anstelle der für die HD-Auflösung erforderlichen ca. 2 MPixels. Die Begrenzung der Auflösung im Motiv passt damit für hoch aufgelöste Fotos, sie ist für HD-Anwendungen aber deutlich zu gering. Daher stellt Aliasing ein typisches Problem für Filmdrehs mit DSLR-Kameras dar.

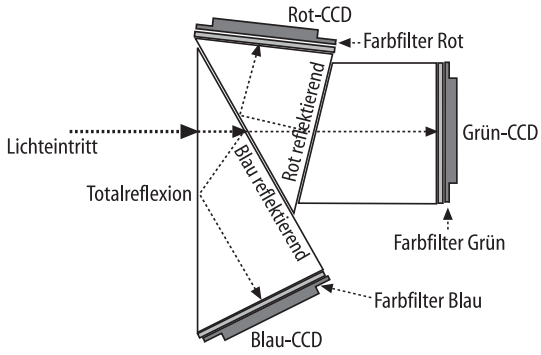


Abb. 6.30. Strahlteilerprisma

6.1.9 Farbbildwandler

Für die Gewinnung eines Farbbildes müssen aus der vorliegenden Abbildung drei Farbauszüge bezüglich der Farbwertanteile für Rot, Grün und Blau gewonnen werden. Das Licht wird durch Filter in die entsprechenden Spektralanteile aufgeteilt, die dann zu drei oder nur zu einem Bildwandler gelangen. Zukünftig könnte die Farbseparation ggf. auch über die wellenlängenabhängige Eindringtiefe in die Halbleitermaterialien vorgenommen werden (s. Abschn. 6.1)

Um das Licht drei separaten Wandlern zuführen zu können, muss eine räumlich getrennte Farbseparierung vorgenommen werden. Zu Beginn der Farbfernsehentwicklung wurden für diese Aufgabe teildurchlässige Spiegel verwendet, die mit verschiedenen dichroitischen Schichten (s. Abschn. 6.2.1) bedampft sind, sodass sie bestimmte Frequenzbereiche transmittieren und andere re-

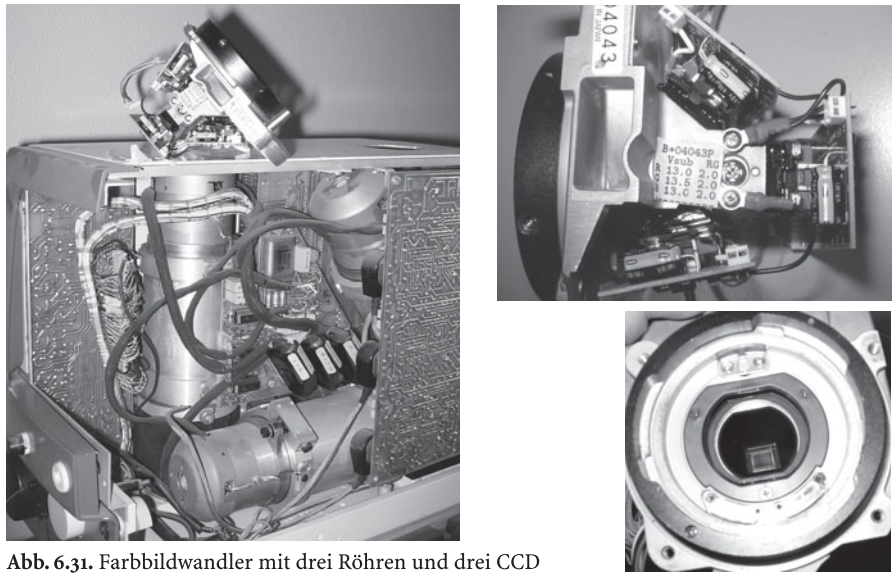


Abb. 6.31. Farbbildwandler mit drei Röhren und drei CCD

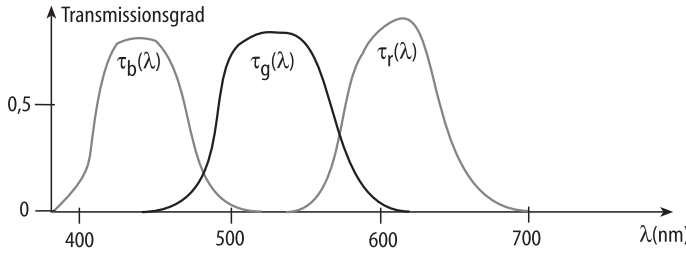


Abb. 6.32. Spektralcharakteristik des Strahlteilers

flektieren. Heute kommt zur Lichtstrahlteilung im professionellen Bereich fast ausschließlich ein fest verbundener Prismensatz zum Einsatz (Abb. 6.30).

Halbleitersensoren sind klein und leicht und können fest mit dem Prismenblock verbunden werden. Die Anordnung kann dabei mit großer Genauigkeit vorgenommen werden, sodass sich keine Rasterdeckungsprobleme ergeben. Röhren sind für eine derartige feste Verbindung zu groß, sie können sich leicht dejustieren. Wegen dieser Probleme wurden auch Vier-Röhren-Kameras entwickelt, bei denen ein eigener Kanal zur Gewinnung des Leuchtdichtesignals vorhanden ist (YRGB- bzw. WRB-Kamera). Die Fotos in Abbildung 6.31 verdeutlichen die Größenverhältnisse am Beispiel einer Drei-Röhren-Kamera. Zum Vergleich mit modernen Bildwandlern ist ein CCD-Prismenblock oben aufgelegt. Daneben ist dieser in der Seitenansicht dargestellt wobei sich zeigt, dass es sich um drei Chips handelt, die mitsamt ihrer Platinen am Prismenblock befestigt sind. Das dritte Bild zeigt den Blick auf die Chips durch die Objektivfassung.

Die Einzelprismen sind zur Erzielung des Filtereffektes mit dichroitischen Filter-Schichten versehen. Das Licht tritt in das erste Prisma des Prismenblocks ein und trifft dann auf die schräge Kante (Abb. 6.30). Deren Oberfläche ist mit einer blaureflektierenden Schicht versehen, und über eine Totalreflexion wird der reflektierte Anteil zum CCD für Blau gespiegelt. Die Grün- und Rotanteile des Lichtes werden an der schrägen Fläche des ersten Prismas transmittiert und treten in das zweite Prisma ein. Die schräg zum Strahl stehende Fläche des zweiten Prisma ist mit einer rotreflektierenden Schicht versehen, sodass der entsprechende Anteil des Lichts zum Rot-Sensor gespiegelt wird. Schließlich tritt der übrig gebliebene Grünanteil in das dritte Prisma ein und fällt auf den Bildwandler für Grün. Um eine korrekte Abbildung für alle Spektralanteile gewährleisten zu können, ist es wichtig, dass, abgesehen von einer Feinjustierung, die drei Strahlwege gleich lang sind, daraus resultiert auch die Bauweise des Prismas.

Die verwendeten dichroitischen Schichten arbeiten nicht als ideale Filter. Da Blau zuerst ausgefiltert wird, finden sich vor allem Blauanteile auf den Wandlern für Grün und Rot. Um die unerwünschten Anteile zu absorbieren, werden zusätzlich vor jedem Bildwandler Farbkorrekturfilter angebracht. Abbildung 6.32 zeigt die Filterwirkung des gesamten Strahlteilers.

Von großer Bedeutung ist die Positionierung der drei Bildwandler relativ zueinander (Rasterdeckung). Wenn die Pixel für die drei Farbwerte nicht exakt

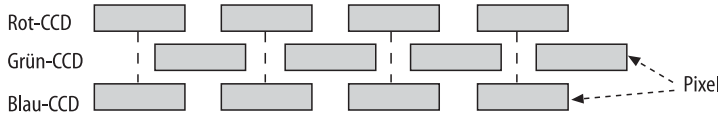


Abb. 6.33. Spatial Offset bei CCD

übereinanderliegen, ergeben sich unscharfe Schwarz-Weiß-Übergänge. Eine besonders genaue Positionierung ist bei HD- und anderen hoch auflösenden Kameras erforderlich. In vielen Fällen wird aber die Soll-Lage des Grün-Sensors durch Offsetmontage bewusst versetzt, sodass seine Lage gegenüber den R- und B-Sensoren um eine halbe Pixelbreite seitlich verschoben ist (Abb. 6.33). Damit wird für das Helligkeitssignal eine scheinbare Erhöhung der Pixelanzahl und so eine Verringerung der Alias-Anteile bzw. eine erhöhte Auflösung in dem für die Schärfewahrnehmung dominanten Luminanzsignal erreicht, da dieses dann aus den räumlich versetzten Anteilen von einerseits G sowie andererseits R und B gebildet wird.

Bei Kameras mit nur einem Bildwandler kommen für die Farbseparation Mosaikfilter zum Einsatz, bei denen sich vor jedem Sensorpixel ein Filter für die wellenlängenabhängige Transmission der einzelnen Farbanteile R, G oder B befindet. Die Anordnung folgt nach einem regelmäßigen Muster. Abbildung 6.34 zeigt das oft verwendete Bayer-Pattern mit den Filtern für RGB. Hier liegt der Grün-Anteil doppelt so häufig vor wie Blau und Rot, da dieser am stärksten in dem für den Schärfeindruck wichtigen SW-Anteil auftritt. Da für jeden Bildpunkt nur die Information für eine einzelne Farbe vorliegt, müssen die fehlenden aus den Nachbarpixeln entnommen werden. Die Darstellung des farbigen Gesamtbildes ist damit erst nach einem Rechenprozess möglich, dem sogenannten Debayering, das herstellerabhängig mit unterschiedlicher Qualität vorgenommen werden kann. Bei Kameras, die ein RAW-Format anbieten, die also die Daten weitgehend unverarbeitet ausgeben, kann dieser Prozess in die Postproduktion verlagert werden.

Single-Sensor-Kameras (SLS) mit Mosaikfilter kommen also mit nur einem Wandler aus. Sie wurden früher vor allem aus Kostengründen gebaut und vorwiegend im Amateurbereich verwendet. Neuerdings gewinnt die Farbseparation mittels Bayer-Pattern aber auch im High-End-Bereich an Bedeutung und zwar im Kontext der Digital Cinematography. Hier ist das Ziel, sich mit den elektronischen Kameras weitgehend der gewöhnlichen Filmkameras anzunähern (s. Abschn. 6.4). Das bedeutet insbesondere, dass die im Filmbereich verwendeten Objektive verwendet werden sollen, nicht zuletzt auch deshalb, weil sie eine erheblich finanzielle Investition darstellen. Abgesehen von ggf. unterschiedlichen Bildwandlergrößen, sind aber die Objektive für Strahlteiler-Kameras anders als Filmobjektive, da bei ihrer Berechnung der Lichtweg durch das Strahlteilerprisma berücksichtigt werden muss. Falls also die Filmobjektive in der Digital Cinematography verwendet werden sollen, sind einzelne flache Bildwandler erforderlich, die an die Stelle des gewöhnlichen Films gesetzt werden können. Unabhängig ob für CMOS- oder CCD-Technologie, wird für diesen Fall meist die Farbseparation mit dem Bayer-Muster verwendet.

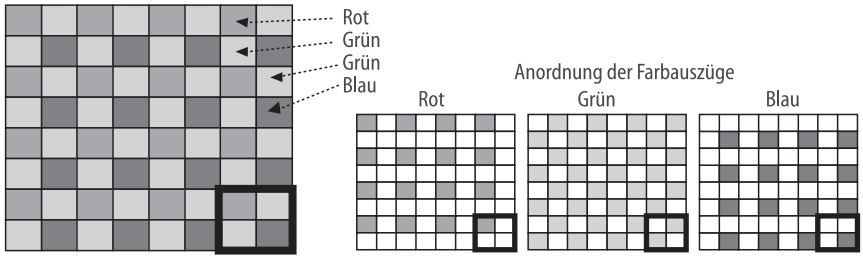


Abb. 6.34. Bayer Pattern als Filter für 1-Chip-CCD- oder CMOS-Sensoren

Gegenüber Kameras mit drei Bildwandlern ist bei der Farbbildgewinnung mit nur einem Sensor bei den meisten Kameras die Farbauflösung verringert. Dieser Nachteil kann jedoch umgangen werden, wenn die Sensorauflösung viermal so groß ist wie die Bildauflösung. Bei vierfacher HD-Auflösung mit 3840 x 2160 Pixels können dann die Bildpunkte z. B. so zugeordnet werden, dass ein Quadrupel genau einem von 1920 x 1080 resultierenden Bildpunkt entspricht (Abb. 6.34, rechts unten). Hier stehen dann die drei Farbinformationen gleichzeitig zur Verfügung, wobei der Grünanteil, räumlich versetzt, gleich zweimal vorhanden ist und so ein mögliches Aliasing für den Y-Kanal, der im Wesentlichen vom Grünauszug bestimmt ist, verringert wird [220]. Ein rechnerisches Debayering ist in diesem Falle nicht mehr erforderlich.

6.1.10 Bildformatwechsel

Während SD-Kameras früher nur für das Bildseitenverhältnis 4:3 ausgelegt waren und HD-Kameras eigentlich nur 16:9 zu liefern bräuchten, ist es oft erwünscht, dass eine Kamera Bilder in beiden Formaten ausgeben kann. Zur Erfüllung dieses Anspruchs stehen verschiedene Verfahren auf optischer oder elektronischer Basis zur Verfügung.

Die einfachste Formatumschaltung auf elektronische Art lässt sich durch den Einsatz eines 16:9-Bildwandlers erreichen, bei dem die Seitenbereiche für die Erzeugung eines 4:3-Formates einfach beschnitten werden, d. h. die Seitenbereiche des Sensors bleiben ungenutzt (Abb. 6.35). Bei unverändertem Objek-

Abb. 6.35. Beschneidung des 16:9-Bildes

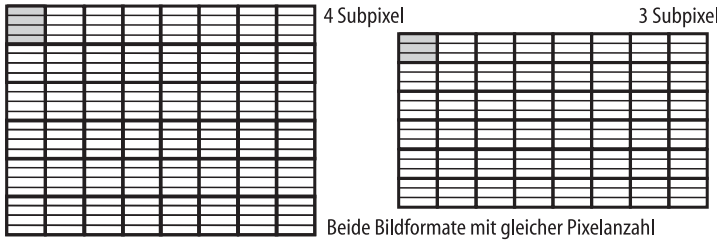


Abb. 6.36. Dynamic Pixel Management

tiv wird durch das verkleinerte Format der Bildwinkel kleiner, bzw. die kleinste Brennweite des Objektivs scheint größer zu sein. Dieser Umstand kann bei teuren Objektiven durch sogenannte Minuskonverter korrigiert werden. Bei dieser Art des Formatwechsels, die auch am häufigsten eingesetzt wird, sollten Sensoren mit möglichst hoher Pixelanzahl verwendet werden, damit die Auflösung bei der 4:3-Abtastung nicht zu stark eingeschränkt wird.

Auf optische Weise gelingt die Umstellung einer 4:3-Aufnahmeeinheit auf 16:9 durch die Nutzung einer anamorphotischen Abbildung, ähnlich wie sie beim Cinemaskop-Filmformat verwandt wird. Ein Anamorphot verzerrt das Bild derart, dass es horizontal komprimiert wird, ein 16:9-Bild lässt sich so auf das 4:3-Format reduzieren. Auf der Wiedergabeseite wird das Bild im 16:9-Format wiedergegeben, indem es elektronisch horizontal auseinandergezogen wird. Dieses Prinzip ist mit horizontalem Auflösungsverlust verbunden, es wird als 16:9-Vollformat bezeichnet (s. Kap. 7.4).

Eine Lösung zum Formatwechsel, die keine Einschränkungen mit sich bringt, ist der Austausch oder die Umschaltung zweier kompletter Prismenblöcke, die jeweils mit den Sensoren für 4:3 und 16:9 bestückt sind. Diese Lösung ist aufwendig und teuer und eignet sich vor allem für Studiokameras.

Beim Einsatz von CCD nach dem FT-Prinzip lässt sich ein Formatwechsel auch über ein sogenanntes Dynamic Pixel Management (DPM) erreichen. Dies ist eine rein elektronische Lösung, die die volle Auflösung des Bildwandlers in beiden Formaten beibehält. Bei DPM werden die Bildpunkte in der Vertikalen quasi in vier Subpixels unterteilt. Das Ausgangsformat ist 4:3. Für dieses Seitenverhältnis werden alle vier Subpixels zu einem Bildpunkt zusammengefasst. Für das Format 16:9 wird die abgetastete Vertikale um ein Viertel reduziert, gleichzeitig werden aber nur noch drei Subpixels einem Bildpunkt zugeordnet, sodass sich in beiden Fällen exakt die gleiche Auflösung ergibt (Abb. 6.36). Bei diesem Verfahren kann der Formatwechsel auf Knopfdruck erfolgen, es sind keine optischen oder mechanischen Änderungen erforderlich. Ein Nachteil ist die um 1/4 verringerte Empfindlichkeit beim 16:9-Betrieb [21].

Auf ähnliche Weise kann bei FT-Chips auch im HD-Bereich eine Umschaltung der Bildauflösung realisiert werden. Wenn hier im HD-Modus mit 1080 Zeilen für jeden Bildpunkt vier Subpixels zusammen gefasst werden, müssen für das Format mit 720 aktiven Zeilen dann sechs Subpixels einem resultierenden Bildpunkt zugeordnet werden. Aus der Zuordnung von drei Subpixels kann dann z. B. auch ein Format mit 1440 Zeilen gewonnen werden.

Abb. 6.37. Diffuse und gerichtete Reflexion

6.2 Kameraoptik

6.2.1 Grundlagen der Optik

Für die Umsetzung des Bildes in ein elektrisches Signal muss ein Gegenstand auf eine kleine zweidimensionale Fläche abgebildet werden. Die Abbildung geschieht mit Linsensystemen, die mit Hilfe der geometrischen Optik beschrieben werden können. Diese Art der Beschreibung ist zulässig, wenn ein Lichtbündel auf seinem Weg nur auf Hindernisse trifft, die groß gegenüber der Wellenlänge sind. Bei Abwesenheit von Hindernissen breitet sich danach das Licht als Strahl geradlinig aus. An Hindernissen (Übergang in ein Medium mit anderem Brechungsindex) wird das Licht reflektiert und gebrochen.

Für die Reflexion an ebenen Flächen (Unebenheiten kleiner als die Wellenlänge) gilt, dass der einfallende Strahl mit dem Lot auf der reflektierenden Fläche den gleichen Winkel bildet wie der ausfallende Strahl: Einfallwinkel gleich Ausfallwinkel. An unebenen Hindernissen wird das Licht gestreut, d. h. diffus nach allen Seiten reflektiert (Abb. 6.37). Ideal diffuse Reflexion wird als Remission bezeichnet. Mehr oder weniger von Objekten reflektiertes Licht bestimmt deren Kontrast mit der Umgebung und macht so das Objekt sichtbar. Die Brechung betrifft den Teil der Welle, der in das Medium eindringt (Abb. 6.38). Hier ist die Ausbreitungsgeschwindigkeit verändert, ein schräg einfallender Lichtstrahl verlässt daher seine Richtung, der Lichtweg erscheint »gebrochen«. Das Brechungsgesetz besagt, dass das Verhältnis der Sinus von Ein- und Ausfallwinkel gleich dem Brechungsindex n ist. Es gilt:

$$\sin \varepsilon / \sin \varepsilon' = n.$$

Beim Übergang in ein dichteres Medium wird der Strahl zum Einfallslot hin gebrochen, beim Übergang in ein dünneres Medium vom Lot weg. Dabei fin-

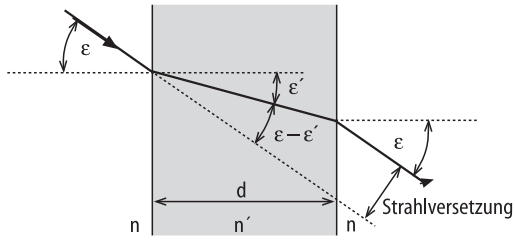


Abb. 6.38. Brechung an einer planparallelen Platte

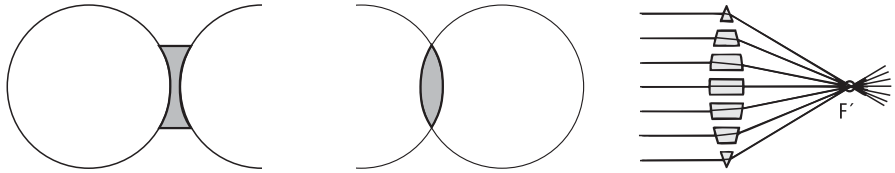


Abb. 6.39. Entstehung sphärischer Linsen und Brechung an einer dünnen Linse

det ab einem bestimmten Winkel überhaupt keine Brechung mehr statt – alles Licht wird dann reflektiert (Totalreflexion). Die Brechung bewirkt beim Strahldurchgang durch eine planparallele Platte eine Strahlversetzung.

6.2.1.1 Optische Abbildung

Die Abbildung geschieht mit Linsen, durchsichtigen Körpern, die von zwei sphärischen Flächen begrenzt werden (Abb. 6.39). Konkave Linsen, die in der Mitte dünner sind als am Rand, zerstreuen das Licht; konvexe Linsen, in der Mitte dicker als am Rand, bündeln das Licht. Trifft der Lichtstrahl auf die Mitte dieser sog. Sammellinse und liegt er parallel zum Lot, so wird er nicht gebrochen, dieser Weg entspricht der optischen Achse. Der Einfallswinkel und damit die Strahlbrechung steigt, je weiter außen der Strahl die Linse durchdringt (Abb. 6.39). Bezugsebene für Linsenrechnungen ist bei sog. dünnen Linsen die Mittelebene. Der Abstand zwischen Mittelebene und dem Treffpunkt von Strahlen (Brennpunkt), die auf der anderen Linsenseite parallel zur optischen Achse verlaufen, heißt Brennweite f (Abb. 6.40). Der reziproke Wert der Brennweite ist die Brechkraft D = 1/f, mit der Einheit Dioptrie (1 dpt = 1/m).

Lichtstrahlen, die genau durch die Linsenmitte gehen, sind Zentralstrahlen, die ihre Richtung nicht verändern. Mit Hilfe von Zentral- und Parallelstrahl lassen sich Bilder von Objekten einfach konstruieren. Dazu werden nur Zentral- und Parallelstrahl herausgesucht und deren Lichtweg verfolgt. Der Zentralstrahl bleibt, abgesehen von einer kleinen Strahlversetzung, unverändert, während der Parallelstrahl zum Brennstrahl wird. Der Ort, an dem sich die beiden hinter der Linse treffen, ist der Bildpunkt. Hier schneiden sich bei Abwesenheit von Abbildungsfehlern auch alle anderen Strahlen, die vom gleichen Objektpunkt ausgehen. Unter bestimmten Bedingungen treffen sich die Strah-

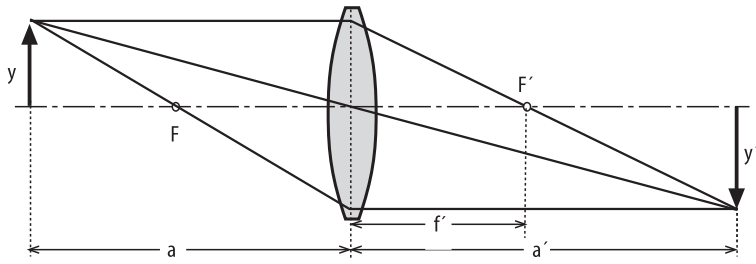


Abb. 6.40. Abbildung mit einer dünnen Linse

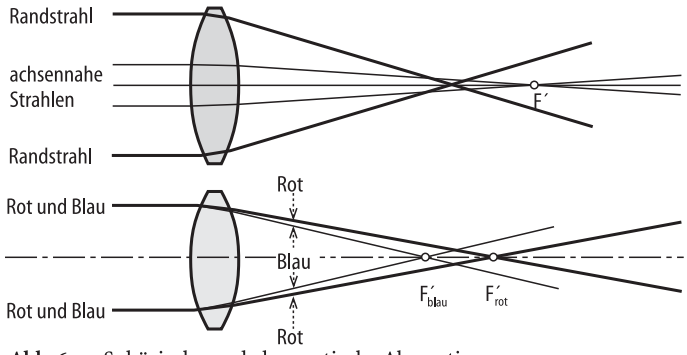


Abb. 6.41. Sphärische und chromatische Aberration

len nie, da sie divergieren. Dann entstehen virtuelle Bilder mit einer scheinbaren Lage, die sich aus der rückwärtigen Verlängerung dieser Strahlen ergibt.

Zur Berechnung von Bildlage und -größe wird ein Strahlverlauf von links nach rechts betrachtet und die Größen links der Bezugsebene negativ gezählt. Der Abstand zwischen Bezugsebene und Objekt heißt Gegenstandsweite a, zwischen Bezugsebene und Bild liegt die Bildweite a‘. Die bildseitige Brennweite heißt f‘. Befindet sich vor und hinter der Linse das gleiche Medium, so gilt: – f = f‘. Die Objektgröße wird mit y und die Bildgröße mit y‘ bezeichnet. Für Sammellinsen folgt daraus die Abbildungsgleichung [97]:

$$1/f' = -1/a + 1/a'.$$

Für das Abbildungs- bzw. Größenverhältnis b gilt:

$$b = y'/y = a'/a.$$

Beispiel: Eine 1,80 m große Person, die in 3,6 m Abstand vor der Linse steht, soll so auf einen Bildwandler abgebildet werden, dass sie eine Bildwandlerhöhe von 1 cm genau ausfüllt. Es gilt: $y'/y = 1/180 = a'/a$, die Bildweite beträgt $a' = 20$ mm. Die erforderliche Brennweite der Linse beträgt $f' = 19{,}9$ mm. Bei großen Bildweiten, relativ zu f‘, kann die Näherung $a' = f'$ verwendet werden.

Linsen rufen Abbildungsfehler hervor, die umso stärker in Erscheinung treten, je weiter die Lichtstrahlen von der optischen Achse entfernt sind. Wichtige Fehler sind die sphärische Aberration, durch die sich die Parallelstrahlen nicht genau in einem Punkt treffen da die Brechung am Rand der Linse anders ist als in der Mitte und die chromatische Aberration, hervorgerufen durch einen frequenzabhängigen Brechungsindex (Abb. 6.41). In der Praxis werden Abbildungen mit Linsenkombinationen (Objektiven) erzeugt, damit sich die Linsenfehler weitgehend kompensieren. Mit Linsenkombinationen kann auch ein Zoomobjektiv mit variabler Brennweite gebaut werden (s. Abschn. 6.2.2).

Aufgrund der Vielzahl der im Objektiv enthaltenen Linsen ist die Berechnung von Abbildungsverhältnissen schwierig, sie kann aber durch die Einführung von Hauptebenen vereinfacht werden. Dabei wird der Weg eines Lichtstrahls nicht über alle Brechungen verfolgt, sondern so vereinfacht, dass er scheinbar nur aus einer einzelnen Brechung an der Hauptebene resultiert, wie

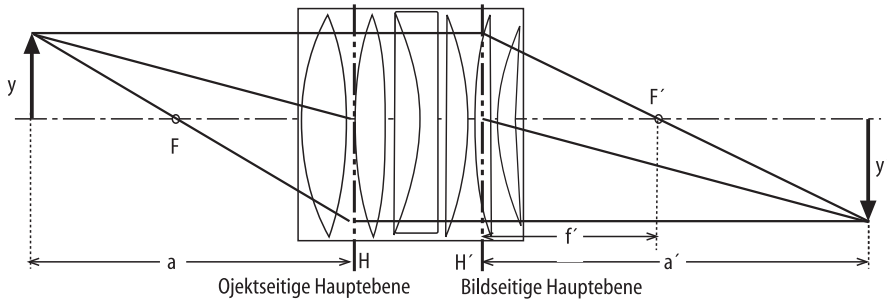


Abb. 6.42. Bildkonstruktion bei optischen Systemen mittels Hauptebenen

es bei einer idealen, dünnen Linse der Fall ist. Es gibt dann die jeweils die beiden objektseitigen und bildseitigen Hauptebenen und die zugehörigen Brennweiten, mit denen ein komplexes Linsensystem vollständig beschrieben werden kann (Abb. 6.42). Wenn bei Objektiven also von Brennweiten die Rede ist, ist diese in Bezug zur Hauptebene zu verstehen, wobei sich die Hauptebene auch außerhalb des Linsensystems befinden kann.

6.2.1.2 Linsenvergütung

Hochwertige Objektive enthalten sehr viele Linsen. An jeder wird ein Teil des Lichtes (ca. 5%) reflektiert und in der Summe kommt es zu nicht akzeptablen Transmissionsverlusten. Außerdem entsteht störendes Streulicht. Zur Reduzierung dieser Verluste werden die Oberflächen entspiegelt, dabei wird die Interferenzfähigkeit des Lichts ausgenutzt.

Wenn zwei Lichtwellen ohne Phasenverschiebung gegeneinander und mit gleicher Amplitude und Frequenz überlagert werden, tritt konstruktive Interferenz auf. Wenn dagegen die Phasenlage um 180° differiert, wird die Welle mit destruktiver Interferenz ausgelöscht (Abb. 6.43). Zur Entspiegelung wird auf die Linsenoberfläche eine aus Metallverbindungen bestehende Schicht aufgedampft, deren Dicke d einem Viertel Wellenlänge λ dividiert durch den Brechungsindex n_s der Schicht entspricht. Mit $n_g > n_s$ gilt: $d = \lambda/(4n_s)$.

Die am Glas mit Brechungsindex n_g reflektierte Welle durchläuft die Schicht zweimal und ist dann gegenüber der an der Metallschicht reflektierten Welle um die halbe Wellenlänge verschoben, sodass destruktive Interferenz auftritt.

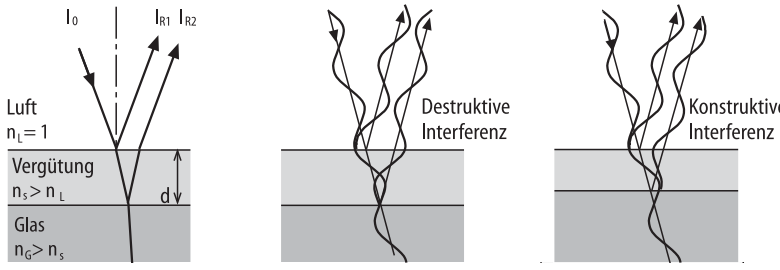


Abb. 6.43. Interferenz an dünnen Schichten

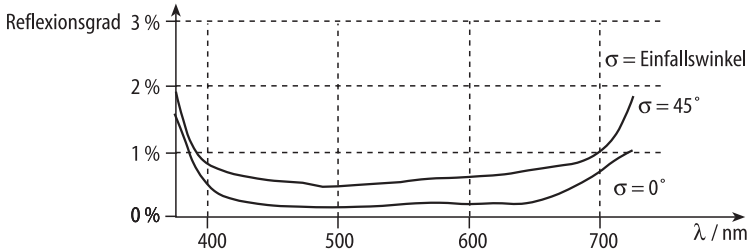


Abb. 6.44. Reflexionsverhalten bei Entspiegelung mit mehreren dünnen Schichten

und keine Lichtenergie mehr reflektiert wird. Mit anderen Schichtdicken kann auf gleiche Weise auch verspiegelt werden. Um den ganzen Wellenlängenbereich des sichtbaren Lichts zu erfassen, müssen mehrere Schichten verwendet werden (Abb. 6.44). Man kann mit diesen Schichten also frequenzselektiv reflektieren und Farbfilter herstellen. In Strahlteilern von Kameras werden sie z. B. als sogenannte dichroitische Schichten eingesetzt.

Objektive unterscheiden sich in ihrer Abbildungsgüte und Lichtstärke, und im Preis. Sehr hoch vergütete Linsen mit minimalen Abbildungsfehlern werden als Prime Lenses bezeichnet. Die verwendeten Objektive sind nicht kamera- oder herstellerspezifisch, sie können bei Nutzung desselben Abbildungsmaßstabes ausgetauscht werden, wenn sie den passenden Befestigungsmechanismus aufweisen. Objektive aus dem Filmbereich können aber nicht problemlos für Videokameras verwendet werden, allein schon aufgrund der Tatsache, dass bei der Objektivberechnung für Videokameras der durch den Strahlteiler verlängerte Lichtweg zum Bildwandler berücksichtigt werden muss.

6.2.1.3 Die Blende

Die Blende mit ihrem veränderlichen Durchmesser begrenzt den nutzbaren Bereich der Linse, damit ist der auf den Bildwandler fallende Lichtstrom regulierbar. Die Blende bewirkt auch eine Verminderung der Abbildungsfehler, denn Randstrahlen, bei denen die durch die Linse erzeugten Abbildungsfehler besonders groß sind, werden ausgeblendet.

Kennzahl für die Objektive ist die Lichtstärke d_{max}/f', das Verhältnis von größtem nutzbaren Objektivdurchmesser d_{max} zur Brennweite f'. Das Verhältnis zwischen der Brennweite f' und dem tatsächlich verwendeten Durchmesser d heißt Blendenzahl k (oder F), es gilt: $k = f'/d$. Die Blendenzahl ist von der Einstellung der Blende (engl.: F-Stop) im Objektiv abhängig. Alternativ findet man auch die Bezeichnung T-Stop. Dieser Wert bezieht sich nicht auf das angegebene berechnete Verhältnis, sondern auf lichttechnisch gemessene Werte [12]. Die T-Stop-Kalibrierung ist somit genauer, da die Transmissionsverluste im Objektiv berücksichtigt sind. Auf dem Blendeneinstellring sind die Blendenzahlen angegeben. Die Einstellung k = 4 besagt, dass die Brennweite viermal so groß ist wie der Pupillendurchmesser. Die Blendenzahlwerte sind in Stufen angegeben, die sich um den Faktor $\sqrt{2}$ unterscheiden (1,4; 2; 2,8; 4; 5,6; 8; 11; 16), da sich dabei die Pupillenfläche und damit die Lichtintensität bei der Steigerung um einen Stufenwert genau halbiert. Die Blendengröße bestimmt

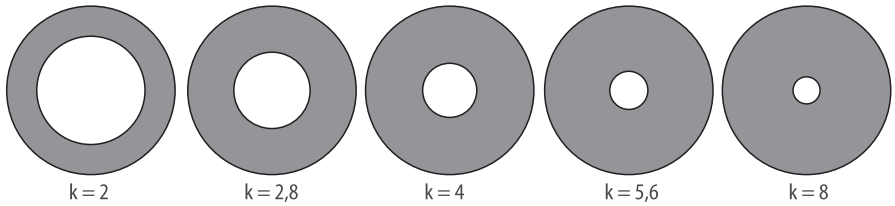


Abb. 6.45. Blendengrößen

wesentlich die Schärfentiefe, sie steigt mit kleinerer Blende bzw. größerer Blendenzahl (Abb. 6.45). Bei sehr kleinen Öffnungen können Bildbeeinträchtigungen durch Beugungseffekte auftreten, die mit dem wellentheoretischen Modell der Lichtausbreitung erklärt werden. Dagegen tritt bei stark geöffneter Blende, d. h. kleinen Blendenzahlen, der Effekt auf, dass die Lichtintensität zum Bildrand hin abnimmt außerdem wird die sphärische Aberration stärker sichtbar. Die beste Bildqualität wird bei mittleren Blendenzahlen, z. B. bei k = 2,8 erreicht.

Bei der Berechnung des durch die Blendenöffnung fallenden Lichtstroms gilt: $\Phi = I \cdot \Omega = L \cdot A_1 \cdot \Omega$. Die Linsenöffnung $d^2 \cdot \pi/4$ entspricht der wirksamen Leuchtfläche A_1 und für den zugehörigen Raumwinkel gilt: $\Omega = d^2 \cdot \pi/4 \cdot 1/a^2$ (Abb. 6.46 unten). Da die leuchtende Fläche A_1 über das Quadrat des Abbildungsverhältnisses a´/a mit der Bildwandlerfläche A_2 verknüpft ist gilt $A_1/A_2 = a^2/a'^2$ und daraus folgt $\Phi = L \cdot A_2 \cdot \pi/4 \cdot d^2/f'^2$, wobei der Einfachheit halber eine große Gegenstandsweite und damit f´= a´ vorausgesetzt wurde. Nun sei angenommen, dass die Leuchtdichte L über eine ideal diffus streuende Fläche mit dem Remissionsgrad R aus der Beleuchtungsstärke E erzeugt wird, sodass gilt: $L = R \cdot E/\pi$, dann folgt mit $d^2/f'^2 = 1/k^2$ die im vorigen Abschnitt zur Berechnung der Kameraempfindlichkeit genutzte Beziehung $\Phi = 1/4 \cdot R \cdot E \cdot A_2/k^2$.

Die Blende ist ein relativ frei wählbarer Parameter. Die Bezugsblende wird für den mittleren Grauwert der Szene eingestellt, der mit 18 % Remission angenommen wird. Neben der mittleren Blende interessieren auch die gemessenen Blendenwerte, die an den hellsten bzw. dunkelsten Stellen im Bild auftauchen. Falls die Differenz bei Video 5 bis 6 Blendenwerte und bei Film 6 bis 8 Blendenstufen übersteigt, ist evtl. der Kontrast zu vermindern, indem Schatten aufgehellt werden. Als Alternative lassen sich zu starke Lichtquellen dämpfen, indem z. B. ein hell beschienenes Fensters durch Abkleben mit Filterfolie verdunkelt wird. Um die Kamera selbst abschatten zu können, wird auf dem Objektiv oft ein Kompendium befestigt, das mit Flügeltoren ausgestattet ist, die die Entstehung von Lichtreflexen vermeiden helfen. Das Kompendium dient auch als Halter von Vorsatzfiltern und wird dann Mattebox genannt.

6.2.1.4 Schärfentiefe

Bei einer exakten Abbildung werden nur Objektpunkte in der Gegenstandsweite als Punkte auf die Bildwandlerebene abgebildet. Für Objekte, die in der Tiefe, also in Richtung der optischen Achse, davor und dahinter liegen erscheint das Bild als kleiner Fleck. Je größer der Fleck, desto unschärfer ist das Bild. Unschärfen sind tolerierbar, solange die Unschärfebereiche kleiner sind als

der das Wandler-Auflösungsvermögen bestimmende Bereich u‘. Gegenstände innerhalb eines Bereichs vor und hinter der optimalen Gegenstandsweite a, zwischen a_v und a_h, werden also noch scharf abgebildet – dies ist der Schärfentiefebereich (eng.: depth of field, Abb. 6.46). Die vorderen und hinteren Schärfenbereichsgrenzen können mit folgender Gleichung berechnet werden [97]:

$a_{h,v} = a /(1 \pm u‘ k (a + f‘)/f‘^2)$ genähert gilt:

$1/a_{h,v} = 1/a \pm u‘ k/f‘^2$.

Beispiel: In welchem Bereich kann sich die 1,80 m große Person aus dem in genannten Beispiel (Gegenstandsweite a = 3,6 m, Wandlerhöhe 1 cm, f‘ = 19,9 mm) bewegen, ohne den Schärfebereich zu verlassen, wenn die Bildhöhe in 1000 Zeilen aufgelöst ist und die Blende auf a) k = 4 und b) k = 8 eingestellt ist? Ergebnis: u‘ = 1 cm/1000, a) a_v = 2,6 m, a_h = 5,6 m, b) a_v = 2,1 m, a_h = 13 m.

Die Schärfentiefe hängt bei konstanter Brennweite also von der eingestellten Blendengröße ab. Eine kleinere Blendenöffnung (größere Blendenzahl k) ergibt eine größere Schärfentiefe. Andererseits gilt bei fest eingestellter Blende: je größer die Brennweite, desto kleiner der Schärfentiefebereich. Das Fokussieren auf ein Objekt (Scharfstellen) ist deshalb im Telebereich am einfachsten.

Die Brennweite, die eine Normalabbildung hervorruft, hängt wiederum mit der Bildwandlergröße zusammen. Aus den Betrachtungen im nächsten Abschnitt lassen sich die Brennweiten für die dort definierte Normalabbildung errechnen, die in Tabelle 6.1 dargestellt sind. Es wird deutlich, dass die Normalbrennweiten in Videosystemen am geringsten sind. Für eine Normalabbildung auf einen 35 mm-Film ergibt sich eine viel größere Brennweite.

Daraus folgt, dass Videokameras und der 16-mm-Film im Vergleich zum 35-mm-Film bei gleicher Blende eine viel größere Schärfentiefe bzw. eine wesentlich geringere selektive Schärfe aufweisen. Nur mit großen Bildflächen ist es daher möglich, den sogenannten Filmlook zu erzeugen, der von geringer

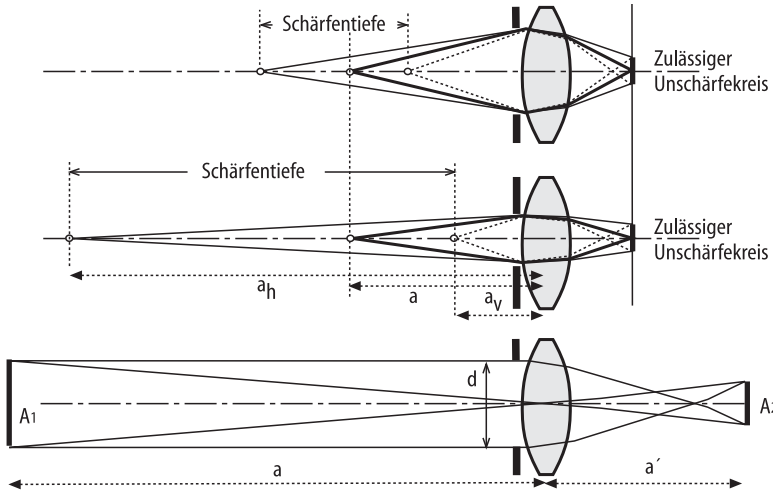


Abb. 6.46. Veränderung der Schärfentiefe in Abhängigkeit von der Blende

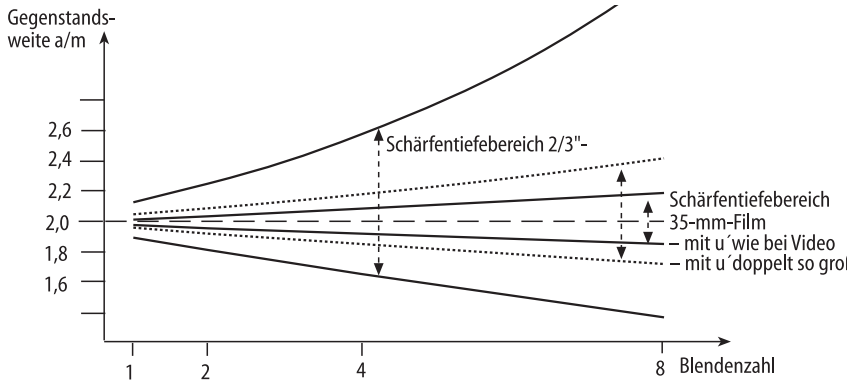


Abb. 6.47. Vergleich der Schärfentiefebereiche von 35 mm-Film und 2/3"-Sensor bei a = 2 m

Schärfentiefe geprägt ist. Der dominante, weil quadratische, Einfluss der Bildgröße bzw. Brennweite auf die Schärfentiefe lässt sich über die Blende in den meisten Fällen nicht ausgleichen. Abbildung 6.47 zeigt dies am Beispiel des Schärfentiefebereich von 35 mm-Film im Vergleich zu dem eines 2/3"-Videobildsensors bei einer Gegenstandsweite a = 2 m dargestellt und zwar unter der Annahme, dass der Unschärfebereich u' = 5 µm bei Film a) gleich und b) wegen der größeren Fläche doppelt so groß ist wie bei (HD-)Video. Darin wird deutlich, dass an der Videokamera ein Blendenwert von ca. 2 eingestellt werden müsste, um die gleiche Schärfentiefe zu erreichen, die der 35-mm-Film bei Blende 8 erzielt.

6.2.1.5 Miniaturisierungseffekt

Eine geringe Schärfentiefe kann als gestalterisches Mittel benutzt werden, um z. B. Objekte im Vordergrund von einem unscharfen Hintergrund abzuheben und so die Aufmerksamkeit der Zuschauer zu lenken.

Eine besondere gestalterische Form bekommt sie beim Einsatz als Miniaturisierungseffekt, der im Bezug auf historische Bilddarstellungen auch als

Abb. 6.48. Miniaturisierungseffekt durch gezielte unscharfe Abbildung [67]

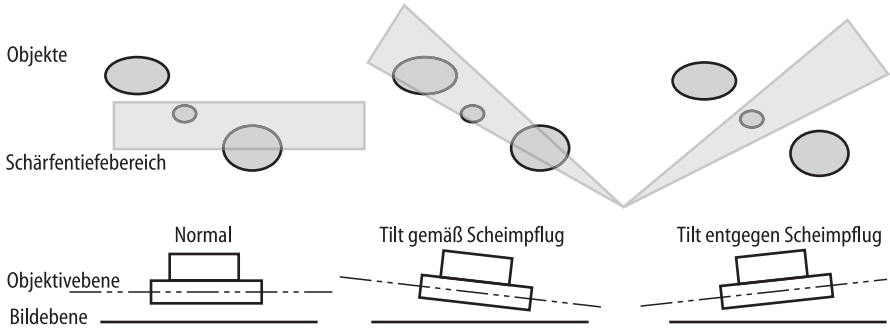


Abb. 6.49. Veränderung der Schärfenbereiche durch Tilt-Objektive

Dioramaeffekt bezeichnet wird. Hier wird eine Wahrnehmungstäuschung ausgenutzt, indem Abbildungen von großen Objekten in relativ großem Abstand, möglichst in Aufsicht erzeugt werden. Wenn in solchen Bildern große Bereiche unscharf dargestellt sind, erscheint es den Menschen so, als schauen sie auf eine Modellszene mit nachgebauten, verkleinerten Objekten, denn die große Unschärfe im Sichtfeld kennen sie aus dem natürlichen Sehen nur bei Annäherung an sehr kleine Gegenstände, während der Anblick großer Objekte aus großer Entfernung nie mit Unschärfen verknüpft ist. Der Wahrnehmungskonflikt und die Interpretation als Modell wird durch intensive Färbung und hohe Kontraste in den scharfen Bildbereichen verstärkt (Abb. 6.48).

Die Erzeugung der Unschärfen gelingt wegen der großen Abstände zwischen Kamera und Objekt mit gewöhnlichen Mitteln meist nur unzureichend. Daher kommen hier sogenannte Shift/Tilt-Objektive zum Einsatz, die es erlauben, das Objektiv relativ zur Bildwandlerebene zu verschieben und zu verschwenken. Für die Erzielung der Unschärfe ist die Tilt-Funktion wichtig, denn sie führt dazu, dass die Bildebene und der Bildsensor nicht mehr parallel zueinander liegen. Das Schärfentiefegebiet wird verkippt und keilförmig (Abb. 6.49). Das kann ausgenutzt werden, um mehrere Objekte mit einer bestimmten Tiefenstaffelung möglichst scharf darzustellen. Theodor Scheimpflug definierte dafür die sogenannte Scheimpflugsche Regel, die besagt, dass sich für eine möglichst scharfe Abbildung die Bild-, Objektiv- und Schärfeebene in einer gemeinsamen Geraden schneiden. Für den Miniaturisierungseffekt wird diese Regel nun gezielt missachtet, um die gewünschten Unschärfebereiche im Bild zu erzeugen (Anti-Scheimpflug).

Die erwünschten Unschärfen im Bild können auch nachträglich mit Bildbearbeitungsprogrammen erzeugt werden, doch wirken optisch erzeugte Unschärfen meist überzeugender. [67]

6.2.1.6 Normalabbildung und Perspektive

In der Praxis sind die abzubildenden Gegenstände oft weit von der Linse entfernt, und damit ist die Gegenstandsweite viel größer als die Brennweite der Linse. Für diesen Fall erfolgt aus der Abbildungsgleichung, dass die Bildweite sich nur sehr wenig von der Brennweite unterscheidet.

Tabelle 6.1. Normalbrennweiten für verschiedene Bildformate

Format	1/2"-CCD	2/3"-CCD	16 mm-Fim	Super 16	35 mm-Film	Super 35
Bilddiagonale	8 mm	11 mm	12,7 mm	14,4 mm	25,7 mm	31,1 mm
Normalbrennweite	9,7 mm	13,2 mm	15,4 mm	17,3 mm	31,0 mm	37,6 mm

Die Brennweite wird oft als Orientierung für die Bildwirkung angegeben, entscheidender ist diesbezüglich jedoch der Bildwinkel. Abgeleitet vom deutlichen Sehwinkel des menschlichen Sehfeldes wird ein Normal-Bildwinkel angegeben, der ca. 45° bezüglich der Bilddiagonalen beträgt. Beim Bildseitenverhältnis 4:3 folgt daraus ein vertikaler Bildwinkel von ca. 28° und ein horizontaler von 36°. Aus diesen Werten kann die zugehörige Normalbrennweite abgeleitet werden (Tabelle 6.1). Als relative Orientierung ist sie nur bei einem festen Bezug zu einer definierten Bildwandlerfläche tauglich, die Sensorgröße bestimmt die Normalbrennweite.

Bei größerer Brennweite als normal spricht man von einem Teleobjektiv, das einen kleinen Bildwinkel hervorruft, und umgekehrt vom Weitwinkelobjektiv. Weitwinkelobjektive erzeugen eine scheinbare Vergrößerung und Teleobjektive entsprechend eine scheinbare Verkürzung zwischen den in der Tiefe gestaffelten Gegenständen in der Szene (Abb. 6.50).

Für den Austausch von Objektiven folgt daraus, dass ein Objektiv, das z. B. auf einem 2/3"-Sensor eine Normalabbildung erzeugt, einen kleineren Bildwinkel hervorruft und damit ein Teleobjektiv darstellt, wenn es an einer Kamera mit 1/2"-Sensor eingesetzt wird (Abb. 6.51). Die Abhängigkeit der Normalbrenn-

Abb. 6.50. Vergleich der Bildwirkung von Normal- (oben) mit Weitwinkel- und Teleabbildungen

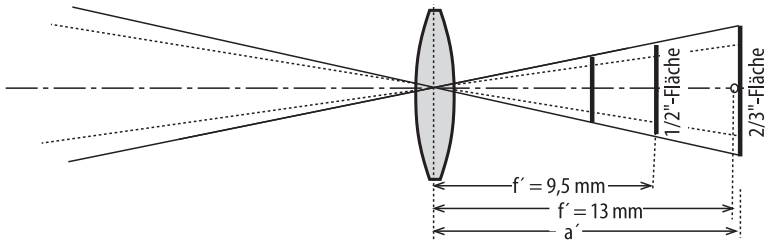


Abb. 6.51. Normalbrennweiten bei verschiedenen Wandlergrößen

weite von der Wandlergröße zeigt sich auch darin, dass Fotografen mit einer Kleinbildkamera (Bildfeldgröße 36 mm x 24 mm) eine Normalbrennweite von 50 mm verwenden, während sie bei 35 mm Film ca. 32 mm beträgt.

Sowohl beim Kleinbild als auch beim Kinobild wird ein Betrachtungsabstand verwendet, der 2 bis 3 mal der Bildhöhe H entspricht. Damit ergibt sich ein Betrachtungswinkel, der gleich dem Aufnahmewinkel ist und damit ist eine perspektivisch richtige Umsetzung erreicht. Beim Videosystem wird diese nur bei HDTV erzielt, da bei SDTV der Betrachtungsabstand aufgrund der geringen Auflösung ca. 6 H beträgt. Fernsehen mit Standardauflösung ist mit größerer Distanz zum Medium verbunden. Durch die Verwendung von Großaufnahmen kann man sich der richtigen Perspektive nähern, dann müssen Brennweiten vom ca. dreifachen der Normalbrennweite verwendet werden.

Da im Videobereich mit verschiedenen Bildwandlergrößen gearbeitet wird, hat sich beim Videoobjektiv keine allgemein gültige Normalbrennweite herauskristallisiert. In professionellen Kameras wird oft ein 2/3"-Chip mit einer Sensordiagonale von 11 mm eingesetzt. Daraus folgt für die Normalabbildung beim 2/3"-Wandler eine Bildweite von 13,2 mm.

Die scharfe Abbildung eines Gegenstandes in 2 m Entfernung von der Linsenebene erfordert mit a' = 13,2 mm und 1/f' = 1/(2000 mm) + 1/(13,2 mm) eine Brennweite f' = 13,1 mm. Praktisch rechnet man mit f' = 13 mm für eine Normalabbildung auf einen 2/3"-Sensor. Tabelle 6.1 zeigt weitere Werte.

Um ein Objekt mit einer anderen Gegenstandsweite scharf abzubilden, muss beim Fokussierungsvorgang (Scharfstellen) der Abstand zwischen Linse und Bildwandler, der sogenannte Auszug, verändert werden.

Da moderne Videokameras i. d. R. mit Zoomobjektiven bestückt sind, können die Brennweite und der Bildwinkel hier sehr leicht verändert werden. Die menschliche Wahrnehmung ermöglicht aber keinen derartigen Zoomvorgang. Um eine andere Perspektive zu gewinnen, muss sich der Mensch an einen anderen Ort begeben. Der Wechsel der Brennweite widerspricht damit der menschlichen Wahrnehmung und wird aus diesem Grund von vielen Regisseuren und Kameraleuten gemieden. Sie verwenden für Szenen gleichen Typs gleiche Festbrennweiten und wahren damit die visuelle Kontinuität. Das bedeutet, dass zur Annäherung an ein Objekt anstelle der Zoom- eine Kamerafahrt eingesetzt wird. Die Alternative ist, den Standort beizubehalten und den Bildwinkel zu verkleinern, dann verändert sich jedoch die Bildwirkung erheblich, weil sich die Relationen von Vorder- und Hintergrundobjekten ver-

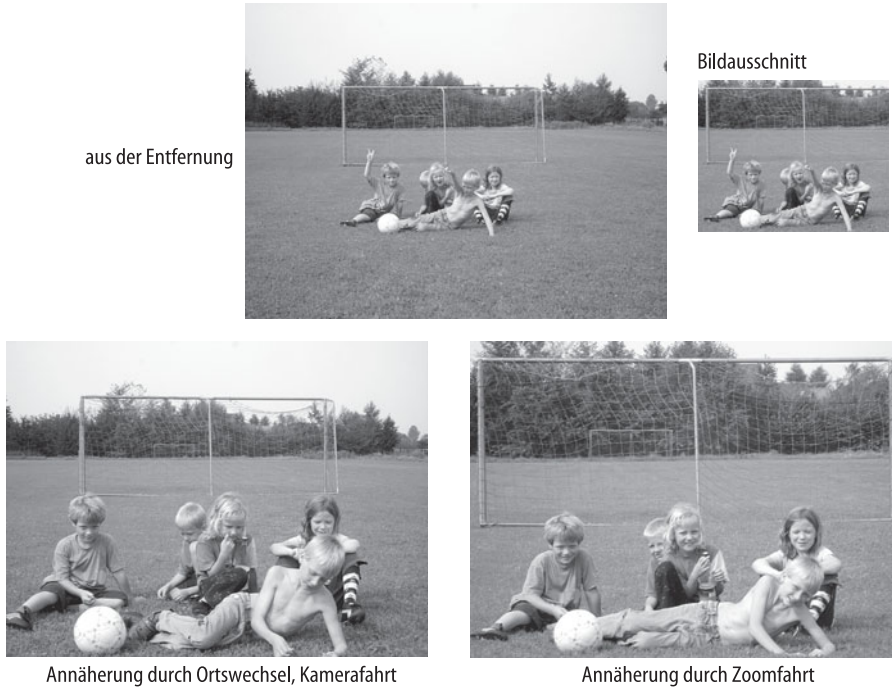


Abb. 6.52. Kamera- und Zoomfahrt im Vergleich

schieben. Im Telebereich scheinen die Abstände verkürzt, das Bild wird flach. Abbildung 6.52 zeigt den Vergleich der Annäherungen an ein Objekt mittels Kamera- und Zoomfahrt bei gleicher Vordergrundbreite.

Die Abbildung macht auch deutlich, dass die Annäherung mittels Zoom wie die Vergrößerung eines Bildausschnitts wirkt. Das bedeutet im Umkehrschluss wiederum, dass ein Zoomfahrt auch dadurch realisiert werden kann, dass einfach ein Sensorausschnitt per Bildverarbeitung in der Kamera vergrößert wird. Dieser Vorgang ist mit einem Verlust an Detailauflösung verbunden. Diese, als Digitalzoom o. ä., bezeichnete Funktion ist nicht bezüglich der Bildwirkung, aber bezüglich der Bildqualität dem optischen Zoom unterlegen.

Für die Positionierung eines Objekts innerhalb des Bildwinkels haben sich einige Begriffe herausgebildet, die an der Abbildung des Menschen, insbesondere des Gesichts, orientiert sind. Die Begriffe müssen den Kameraleuten vertraut sein, damit die Kommandos aus der Regie, wie z. B. »3 verengen bis groß« oder »1 nah, 4 aufziehen bis total«, richtig ausgeführt werden: Eine Großaufnahme zeigt den menschlichen Kopf und den oberen Schulterbereich (Abb. 6.53). Dieses Close Up (CU) kann bis zum Very und Extrem Close Up (sehr, extrem groß) gesteigert werden, indem nur noch Augen und Mund oder der Augenbereich erfasst werden.

Eine Nahaufnahme umfasst auch die Brust und wird als Bust Shot oder Medium Close Up bezeichnet. Bei größeren Abständen zum Objekt folgen die

Extrem weit, very long shot

Weit, long shot

Total, full length shot

Halbtotal, medium long shot

Amerikanisch, tight shot

Halbnah, mid shot

Nah, medium close up

Groß, close up

Ganz groß, big close up

Sehr groß, very close up

Extrem groß, extreme close up

Abb. 6.53. Einstellungsgrößen

halbnahe Einstellung, die die Person bis zur Hüfte (Mid Shot) erfasst, weiterhin die amerikanische Einstellung bis zu den Oberschenkeln (Thight Shot), die Halbtotale bis zu den Knien und schließlich die Totale, die die gesamte Person erfasst (Full Length Shot). Größere Abstände werden als weit und extrem weit bezeichnet (Long und Very Long Shot) [47]. Abbildung 6.53 zeigt die Einstellungsgrößen in der Übersicht.

Weitere Kamerakommandos sind weitgehend selbsterklärend, wie z. B.: »Achtung 3, nächste 2, 4 schwenk hoch, 1 unscharf werden, 2 Gast mittig einsetzen oder 3 über Schulter«. Um schnelle Reaktionen zu ermöglichen, sollte dafür gesorgt werden, dass die Nummer der angesprochenen Kamera immer zuerst genannt wird.

6.2.2 Das optische System der Kamera

Das Objektiv ist das wesentliche optische Element in der Kamera das zur Abbildung der Gegenstände auf den Bildwandler dient. Es besteht im einfachsten Fall aus einer Linse, in der Praxis aus einer aufwendigen Linsenkombination.

6.2.2.1 Objektive mit Festbrennweiten

Zu Beginn der Kameratechnik enthielten die Kameras Objektive mit fester Brennweite, teilweise waren mehrere auf einen Revolverkopf montiert, um einen Wechsel der Brennweite zu ermöglichen. Derartige Objektive bieten zwei Verstellmöglichkeiten, und zwar für den Auszug (zur Fokussierung) und für die Blende. Die Bedienelemente sind meist ringförmig angebracht. Über einen Zahnkranz kann zusätzlich ein Follow-Focus-Rad angebracht werden, das mittels einer Übersetzung eine exakte Schärfeverlagerung ermöglicht (Abb. 6.54).

Auch wenn Zoomobjektive, die eine stufenlose Variation der Abbildungsgröße ermöglichen, weit verbreitet sind, werden auch heute noch Festbrennweiten verwendet. Dies geschieht hauptsächlich in Einsatzgebieten wie der Spielfilmproduktion; also dort, wo eine möglichst optimale Bildqualität erzielt werden soll. Der Begriff höchste Qualität bezieht sich auf die Minimierung von Abbildungsfehlern, auf die Abbildungsschärfe und die Lichtstärke. Diese Eigenschaften können bei Festbrennweiten leichter erreicht werden, da sie weniger Linsen enthalten die aufeinander abgestimmt werden müssen. Diese können zudem auch asphärische Linsen sein, die kaum sphärische Aberration aufweisen.

Bei Filmproduktionen werden daher oft Sätze von Festbrennweiten angemietet, z. B. 16, 25, 32, 75 und 100 mm zur Abdeckung von zwei Weitwinkel-, einer Normal- und zwei Tele-Abbildungen auf ein 35-mm-Filmformat. Bekannte Hersteller von hochwertigen, sogenannten Prime Lenses sind die Firmen Zeiss und Cooke. Die Zeiss-Objektive der Master Prime-Klasse erreichen eine sehr gute Lichtstärke, d. h. eine kleinste Blendenzahl von T = 1,3, preisgünstigere Objektive der Compact Prime-Klasse haben den Wert T = 2,1.

Grundsätzlich müssen die Objektive an die zu belichtende Sensorfläche angepasst sein. Wenn nicht passende Objektive verwendet werden, ist zum Ersten damit zu rechnen dass die Sensorfläche nicht richtig ausgeleuchtet wird. Das ist bei der Belichtung kleiner Flächen mit großen Objektiven unkritisch und qualitativ wegen der Tatsache, dass die Randstrahlen den Sensor nicht

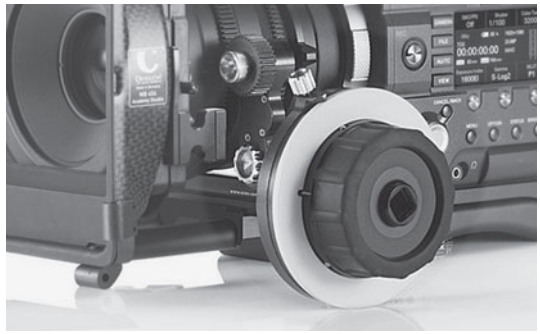


Abb. 6.54. Follow Focus als Schärfezieheinrichtung [208]

erreichen, sogar gut. Umgekehrt führt die Verwendung kleiner Objektive an großen Sensoren aber dazu, dass die Ecken der Sensoren abgeschattet werden und somit ein Randlichtabfall auftritt, der manchmal auch als Vignettierung bezeichnet wird. Zum Zweiten ändert sich auf jeden Fall der Bildwinkel. Wenn z. B. ein Foto-Objektiv, das für ein Kleinbildformat von 36 mm x 24 mm gebaut wurde, für eine Bildwandlerfläche verwendet werden soll, die 35-mm-Film entspricht, also für eine Bildbreite von 24 mm, so ist die Brennweite um den Faktor 1,5 zu groß und das Objektiv, das am Fotoapparat ein Normalobjektiv ist, ruft an dem kleineren Sensor eine Telewirkung hervor. Dieser Brennweitenverlängerungsfaktor wird Format- oder Crop-Faktor genannt. Der Begriff resultiert daher, dass durch den zu großen Bildkreis, den das Fotoobjektiv auszeichnet, das Bild auf dem kleinen Sensor beschnitten wird.

6.2.2.2 Das Zoomobjektiv

Bei Fernseh- und Videoproduktionen soll eine Perspektiv- bzw. Brennweitenänderung aus Zeitgründen meist ohne Umbau vorgenommen werden können. Daher sind Videokameras in den meisten Fällen mit Zoomobjektiven bestückt, die zusätzlich zu Fokus und Blende eine stufenlose Brennweitenvariation ermöglichen. Hier kann ein Wechsel von Weitwinkel zu Tele auch bei laufender Kamera vorgenommen werden, womit eine Zoomfahrt möglich wird.

Ein Zoomobjektiv enthält eine Vielzahl von Linsen (Abb. 6.55), die in vier Gruppen, die Fokussiergruppe, den Variator, den Kompensator mit Blende und

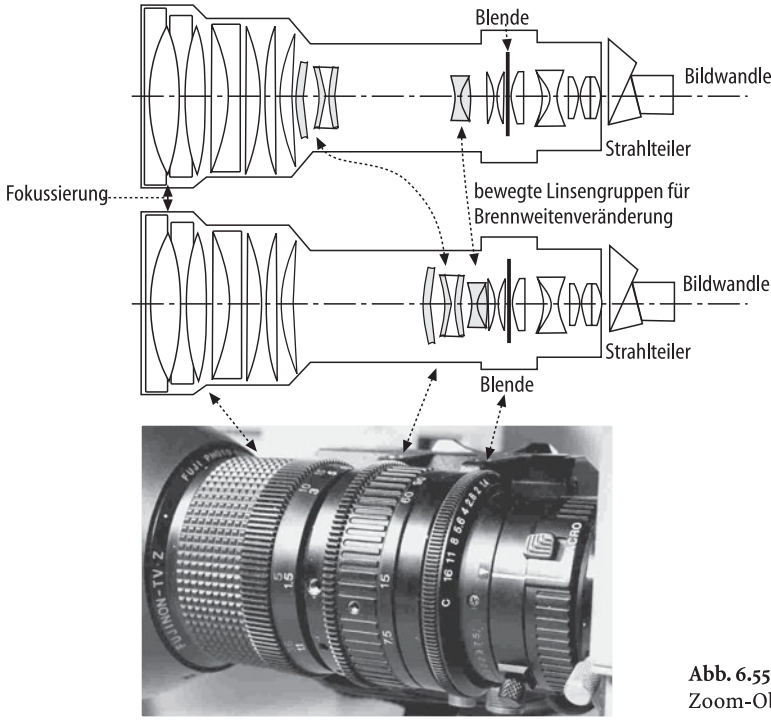


Abb. 6.55. Aufbau des Zoom-Objektivs

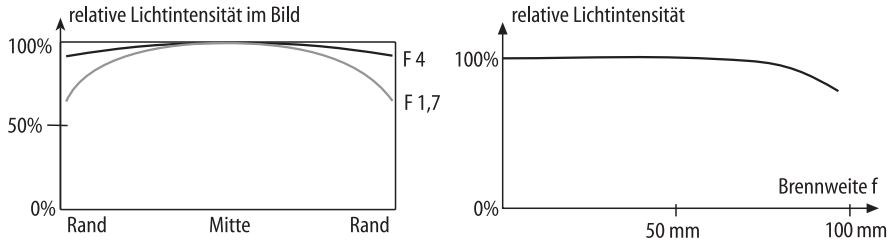


Abb. 6.56. Variation der Lichtintensität in Abhängigkeit von Blende und Zoom

die Relaisgruppe eingeteilt sind. Die Fokussiergruppe bestimmt die Schärfe der Abbildung. Man unterscheidet frontfokussierte Objektive, bei denen die Frontlinse bewegt wird, und innenfokussierte Objektive, bei denen sie feststeht. Innenfokussierte Objektive sind aufwendiger gebaut, weisen geringere Abbildungsfehler auf und sind besser gegen Schmutz geschützt. Hier können Filter direkt an der Linsenfassung montiert werden, ohne dass sie verdreht werden.

Der Variator ist eine bewegliche Linsengruppe, die zur Veränderung der Brennweite dient. Für die Bewegung des Variators gibt es am Objektiv einen dritten Einstellring, der entweder manuell oder über einen mit der Zoomwippe gesteuerten Motor angetrieben wird. Der elektrische Antrieb ermöglicht weiche Zoomfahrten, oft mit variabel einstellbaren Geschwindigkeiten. Das Bild des Variators entsteht auf dem ortsfesten Kompensator, der die Blende enthält. Die Relaisgruppe dient schließlich zur Abbildung auf den Bildwandler.

Die variable Brennweite erleichtert den Scharfstellvorgang durch Einstellung der maximalen Teleposition. Im Telebereich ist die Schärfentiefe besonders klein, damit werden Abweichungen vom Schärfepunkt deutlich sichtbar. Die Schärfe wird daher möglichst immer im Telebereich eingestellt, danach wird der gewünschte Bildausschnitt gewählt. Einige Objektivhersteller bieten in diesem Zusammenhang Typen mit Assisted Internal Focus (AIF) an. Bei diesen Typen wird auf Knopfdruck automatisch der Telebereich eingestellt. Die eigentliche Fokussierung und die Bildausschnittswahl werden weiterhin manuell vorgenommen.

Wichtige Merkmale sind bei Zoomobjektiven, neben der Minimalapertur, die geringste mögliche Brennweite (Anfangsbrennweite) und der Zoomfaktor, beide werden häufig zusammen angegeben. Ein typischer Wert für universell einsetzbare Objektive für 2/3“-Bildwandlern ist 15 x 8, d. h. die minimale Brennweite beträgt 8 mm, die maximale 120 mm. Der maximale diagonale Bildwinkel ergibt sich aus der Relation zur Bildwandlerdiagonale von 11 mm und beträgt hier 69°. Er erzeugt eine Weitwinkelabbildung. Der minimale Bildwinkel beträgt dann 3,1° und erzeugt eine Teleabbildung. Man kann nicht immer davon ausgehen, dass Zoomobjektive auch den Weitwinkelbereich abdecken. Aus Kostengründen beschränken sich Hersteller in einigen Fällen auf den Telebereich. Einige Objektive erlauben das Einbringen einer weiteren Linsengruppe, des Extenders, in den Strahlengang. Dadurch kann dieser Effekt verstärkt werden, da sich die Brennweite dann um einen festen Faktor (z. B. 2x) vergrößert. Der Extender verringert die Lichtstärke.

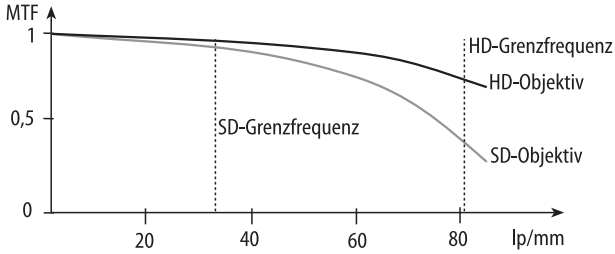


Abb. 6.57. Auflösungseigenschaften von SD- und HD-Objektiven

Zoomobjektive weisen oft einen Helligkeitsabfall am Bildrand auf, der umso stärker ist, je weiter die Blende geöffnet wird (Abb. 6.56). Des Weiteren tritt der Effekt auf, dass bei langen Brennweiten die Lichtstärke abnimmt. Dieser Effekt wird F-Drop oder Ramping genannt und führt zu einer Bildhelligkeitsveränderung beim Zoomen (Abb. 6.56). Der Effekt kann durch große Frontlinsen gemindert werden.

Zoom- oder Varioobjektive begrenzen auch die feinsten abbildbaren Strukturen des Bildes. Im Allgemeinen kann aber der Auflösungsverlust durch das Objektiv gegenüber der relativ schlechten Auflösung eines SD-Videobildwandlers vernachlässigt werden. Bei Objektiven für den HD-Einsatz werden die Schwächen der Objektive sehr viel deutlicher, Unschärfen fallen bei großer HD-Bilddarstellung viel mehr auf. Das Scharfstellen ist bei HD- erheblich schwieriger, als bei SD-Produktionen, nicht zuletzt, weil die Schärfe auf den kleinen Suchern schlecht beurteilt werden kann. Wenn gute Displays zur Verfügung stehen, ist die ungewollte Abweichung vom Fokus deutlich zu erkennen ebenso wie ein eventuelles Nachziehen der Schärfe im Live-Betrieb.

HD-Objektive sind i. d. R. aufwendiger gebaut und bieten insbesondere bei hohen Raumfrequenzen bessere Werte für die Modulationstransferfunktion (MTF), die das Maß für die Auflösung darstellt (Abb. 6.57) [202].

Die Objektive für den professionellen Videobereich kommen in den meisten Fällen von Canon oder Fujinon. Sie können so gebaut werden, dass sie nicht viel schwerer als 1 kg sind. Sie werden dann meist ohne weitere Hilfsmittel mit einem Bajonettverschluss, an der Kamera befestigt (Abb. 6.58). Lichtstarke,

Abb. 6.58. Objektive für portablen und Studioeinsatz [37]

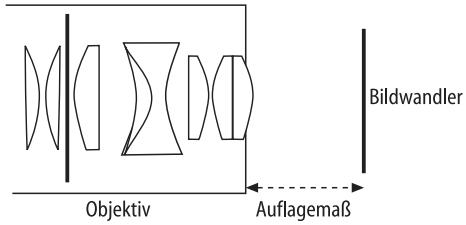


Abb. 6.59. Auflagemaß

sogenannte Boxobjektive mit großem Variationsbereich (bis 100-fach) können dagegen sehr groß und bis zu 40 kg schwer werden, sodass eher die Kamera am Objektiv befestigt wird. Hier sind zusätzliche Stützen erforderlich. Der Preis eines Studioobjektivs mit Vollservobedienung beträgt ca. 80 000 Euro. Wenn die Objektive manuell bedient werden, ergibt sich manchmal ein Problem durch die Tatsache, dass bei Videoobjektiven der Drehwinkel für eine gewünschte Veränderung erheblich kleiner ist als bei Objektiven für Filmkameras, sodass z. B. Schärfeverlagerungen erschwert werden. Follow-Focus-Systeme bieten auch hier Abhilfe.

6.2.2.3 Das Auflagemaß

Dieser Begriff (engl. Flange Focal Distance) hat große Relevanz für alle Kameras mit wechselbarem Objektiv. Er bezeichnet den Abstand zwischen der Objektivfläche, die an die Kamera angesetzt wird, und dem Bildsensor (Abb. 6.59). Für eine scharfe Abbildung, auch bei jeder Zoomstellung, ist die Einhaltung eines genau definierten Abstandes zwischen der letzten Objektivlinse und dem Bildwandler erforderlich. Kleinste Abweichungen schmälern die Qualität der Abbildung. Deshalb wird dieser Wert für Filmproduktionen mit Festbrennweiten vom Verleiher vermessen und genau kalibriert.

Bei Zoomobjektiven für den Videobereich wird das Auflagemaß von den Bildtechnikern direkt am Produktionsort eingestellt. Durch Stöße etc. können sich Abweichungen ergeben. Das Auflagemaß ist an der Kamera nach Lösen einer Arretierschraube korrigierbar. Der zugehörige Ring befindet sich am hinteren Ende des Objektivs. Zur Kontrolle und Einstellung des Auflagemaßes wird zunächst in Tele-Stellung auf ein kontrastreiches Motiv fokussiert und dann im Weitwinkelbereich mit Hilfe eines guten Monitors die Schärfe überprüft. Muss das Auflagemaß geändert werden, so sind nach jeder Verstellung die beiden Schritte zu wiederholen, bis sich ein Optimum ergibt.

6.2.2.4 Objektivbefestigungen

Das Auflagemaß ist auch ein wichtiges Merkmal für die verschiedenen, herstellerabhängigen Objektivbefestigungsarten (engl: lens mount). Im Bereich der klassischen Filmproduktion hat sich der PL-Mount (positive lock), ein Bajonettverschluss des Kameraherstellers Arri etabliert der das Objektiv mit einer kurzen Drehung an der Kamera verriegelt. Er hat ein Auflagemaß von 52 mm bei 54 mm Durchmesser. Für professionelle 2/3“-Video- und Fernsehkameras kommt als Quasi-Standard der Bajonett-B4-Mount der Firma Sony zum Ein-

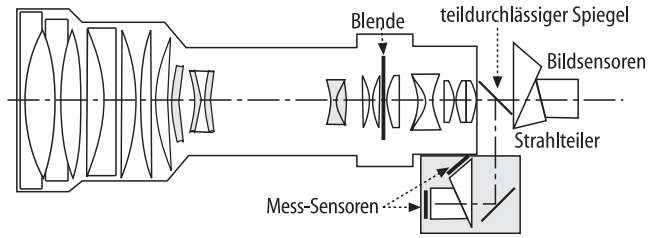


Abb. 6.60. Zur Funktion eines Autofocus-Systems

satz. Hier beträgt das Auflagemaß 48 mm. Die Kennzeichnung C-Mount steht für Einschraubgewinde an Industrie- und Überwachungskameras. Das Auflagemaß hat hier den Wert 17,526 mm, der Durchmesser beträgt 1“ und ist nicht kompatibel zum ähnlichen CS-Mount, dessen Auflagemaß um 5 mm kürzer ist.

Seit digitale Videokameras mit großen Bildsensoren zur Verfügung stehen wird aus Kostengründen vermehrt versucht, hochwertige Fotoobjektive an diesen Kameras zu betreiben. Diese kommen z. B. vom Hersteller Canon und sind entsprechend z. B. mit dem Canon EF-Mount versehen. Das Auflagemaß beträgt für den EF-Mount 44 mm und hat einen ähnlichen Wert wie der Nikon F-Mount mit 46,5 mm. Wenn diese Objektive nicht direkt befestigt werden können, gibt es die Möglichkeit, einen Adaptierungsring zum Anschluss zu verwenden. In diesem Zusammenhang ist es günstig, wenn das Auflagemaß für die Kamera klein ist, damit Platz für den Adapter bleibt. Dies ist z. B. bei dem E-Mount der Firma Sony der Fall ist. Hier beträgt das Auflagemaß nur 18 mm und erlaubt den Anschluss sehr vieler Objektive. Bei der Adaptierung ist zu beachten, dass sich der Formatfaktor der Abbildungen ändern kann.

6.2.2.5 Autofocus

Systeme zur automatischen Scharfstellung wurden zunächst nur für Heimanwenderzwecke benutzt, heute werden sie auch in professionellen Systemen eingesetzt. Aufgrund der oben genannten Schwierigkeiten beim Fokussieren ist es besonders im HD-Bereich von Vorteil, eine Fokussierhilfe zu verwenden. Sie muss nicht ständig den Fokus nachregeln, sondern kann auch punktuell, auf Knopfdruck aktiviert werden.

Eine Möglichkeit, Autofokussysteme zu realisieren, besteht in der Nutzung von Kontrastinformationen. Feine Strukturen im Bild bieten bei optimaler Fokussierung die größten Kontraste. Die Kontrastmessung kann z. B. mit zwei Bildsensoren durchgeführt werden, die unabhängig vom eigentlichen Bildwandler vor und hinter der Fokussierebene liegen. Bei optimaler Fokussierung erzeugen die beiden Sensoren gleiche Signale, bei Abweichung vom Schärfepunkte lässt sich aus der Differenz ermitteln, wie der Fokussierpunkte verschoben werden muss. Für dieses Verfahren werden ca. 20% des Lichts auf die Mess-Sensoren gelenkt. Bei Fujinon wird das Verfahren „Precision Focus“ genannt (Abb. 6.60) [62].

Als Alternative kann die Phasenvergleichsmethode verwendet werden. Auch hier wir ein Teil des Lichts zu einem Sensor gelenkt, allerdings wird es in zwei

Teile geteilt und mit einem örtlichen Unterschied durch eine Linse auf einen Zeilensensor fokussiert. Der Abstand der Fokuspunkte ergibt dabei dann die Information über einen korrekten Fokus, bzw. darüber, ob zu nah zu fern fokussiert ist [37].

6.2.2.6 Filter

Bevor das Licht aus dem Objektiv auf den Bildwandler trifft, durchläuft es meist ein Infrarotsperrfilter, das dazu dient, die Sensoren in diesem Spektralbereich unempfindlich zu machen, dann ggf. ein wechselbares Filter im Konversionsfilterrad und schließlich das optische Tiefpassfilter.

Auf dem Filterrad professioneller Videokameras sind meist verschiedene Filter fest eingebaut, die durch Drehung in den Strahlengang eingebracht werden können (Abb. 6.61). Es sind Konversionsfilter vorhanden, die zur Anpassung der üblicherweise für eine Farbtemperatur von 3200 K (Kunstlicht) ausgelegten Sensoren an eine andere Farbtemperatur dienen (z. B. für Tageslicht mit 5600 K). Weiterhin stehen Neutraldichtefilter zur Verfügung, mit denen die Lichtintensität geschwächt werden kann (ND). Es gibt auch kombinierte Farb- und ND-Filter.

Hinter dem Filterrad befindet sich schließlich das bereits beschriebene optische Tiefpassfilter, das eine definierte Begrenzung der Abbildungsschärfe bewirkt, um Aliasstrukturen zu verhindern, die durch die örtlich diskrete Zerlegung des Bildes im CCD- oder CMOS-Sensor bewirkt werden.

Filter haben eine wichtige Funktion für die Bildgestaltung und werden auch als Vorsatzfilter vor dem Objektiv verwendet. Die wichtigsten Filter für den Bereich der Bildaufnahme dienen der Lichtabschwächung und der Konversion der Farbtemperatur. Die reine Dämpfung der Lichtintensität wird über eine verminderte Transparenz mit entsprechender optischer Dichte erreicht. Sie soll für alle Wellenlängen des sichtbaren Lichts denselben Wert aufweisen. Derartige Filter werden daher als Neutraldichtefilter bezeichnet (ND). Wie beim Filmmaterial ist die Dichte über den Logarithmus des Kehrwertes der Transmission definiert d. h. die Halbierung der Lichtintensität entspricht einer Transmission von 50 %, einer Blendenstufe oder einer Dichteänderung vom Wert 0,3.

Wenn der Dichtewert nach einer Seite des Filters über einen gewissen Bereich gleichmäßig ansteigt, spricht man von einem Verlaufsfilter. Er wird so

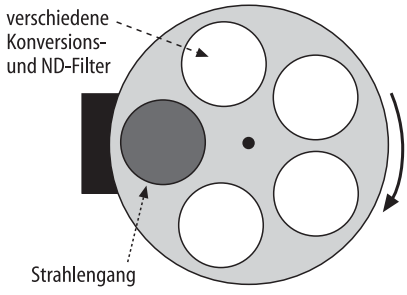


Abb. 6.61. Filterrad hinter dem Objektiv

eingesetzt, dass z. B. bei einer Landschaftsaufnahme die größere Dichte parallel zum Horizont auf dem Himmel liegt. Damit kann die Helligkeit in diesem Bereich so weit herabgesetzt werden, dass im Himmel wieder Strukturen durch Wolken etc. erkennbar werden.

Auch außen angebrachte Konversionsfilter dienen dazu, Bildwandler oder Filmmaterialien, die für eine bestimmte Farbtemperatur des eingesetzten Lichts sensibilisiert wurden, mit Lichtquellen verwenden zu können, die eine andere Farbtemperatur aufweisen. Ein typischer Fall ist die Verwendung von Bildsensoren oder Filmmaterialien, die für Kunstlicht mit 3200 K sensibilisiert sind, mit denen aber Tageslichtaufnahmen bei einer Farbtemperatur von z. B. 5500 K gemacht werden sollen. Für diesen Zweck geeignete Konversionsfilter werden über den Mired-Verschiebungswert gekennzeichnet. Das bedeutet, dass die Farbtemperaturen im Nenner eines Bruches stehen, dessen Zähler den Wert 10^6 aufweist. Der Verschiebungswert ergibt sich aus der Differenz zwischen den Brüchen mit den verschiedenen Farbtemperaturen, die dann in Mired angegeben werden (s. Kap. 2). Die angegebene Anpassung des Tageslichtspektrums an das Kunstlicht erfordert bei dem genannten Fall einen Konversionswert von $10^6/3200 - 10^6/5500 = +131$ mrd mit Hilfe eines gelb-orange gefärbten Filters. Der umgekehrte Fall, d. h. die Konversion von Kunstlicht, erfordert – 131 mrd mit einem blau gefärbten Filter. Da Blaufilter eine viel größere Dichte aufweisen als orangefarbene, werden Filmmaterialien und auch Videokameras filterlos meist für Kunstlicht kalibriert.

Weitere Filter werden eingesetzt, um bestimmte Effekte zu erzeugen, wie z. B. Farbeffekte, eine Kontrastminderung, Sterneffekte oder den Weichzeichnereffekt. Ein Filtertyp, der nicht direkt in diese Klasse fällt, ist das Polarisationsfilter. Dieses wird in der täglichen Praxis verwendet, wenn starke Reflexionen von Metallen oder Glasflächen vermindert werden sollen [155]. Da das von diesen Gegenständen reflektierte Licht meist stark polarisiert ist, d. h., dass die Schwingungsebene der Lichtwelle eine sehr dominante Richtung hat, kann mit einem Filter gearbeitet werden, das nur Licht einer definierten Schwingungsebene transmittiert. Zur Minderung der Reflexe muss dann das Polarisationsfilter einfach so gedreht werden, dass die Transmissionsschwingungsebene senkrecht zur Schwingungsebene des Lichts steht und die Reflexion gemindert wird (Abb. 6.62).

Abb. 6.62. Reflexion an einer Scheibe, aufgenommen ohne (links) und mit Polarisationsfilter

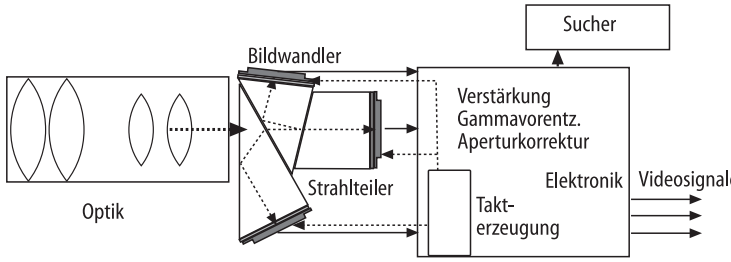


Abb. 6.63. Aufbau der Videokamera

6.3 Die Videokamera

Die Video- und Fernsehkamera ist die klassische elektronische Kamera. Sie wird hier unterschieden von der Digital Cinematography Kamera, die erst im aktuellen Jahrhundert als Ersatz für die Filmkamera entwickelt wurde und die in Abschnitt 6.4 dargestellt wird. Ein wesentlicher Unterschied der beiden Klassen besteht in der Verwendung verschiedener Bildsensorgrößen. Bei typischen Videokameras werden relativ kleine Sensoren bis zur Größe 2/3“ verwendet und damit eine vergleichsweise große Schärfentiefe erreicht, während die Digital-Cinematography-Kameras große Sensoren nutzen, die kleine Schärfentiefebereiche hervorrufen. In letzterem Bereich gibt es zudem das Ziel, wie bei klassischem Film, auch mit hohen Kontrasten und variablen Bildraten arbeiten zu können, während bei Ersterem ein Vorteil darin besteht, dass die große Schärfentiefe die Produktion vereinfacht.

Die hier betrachteten Fernseh- und Videokameras enthalten als Farbkameras ein optisches System, das im Wesentlichen ein Objektiv und bei 3-Chip-Kameras auch den Prismenblock umfasst, die eigentlichen Bildwandler und schließlich verschiedene elektronische Schaltungen zur normgerechten Signalerzeugung, Signalveränderung und Signalwandlung (Abb. 6.63).

Videokameras sind im Hinblick auf die verschiedenen Einsatzgebiete unterschiedlich gebaut [123]. Zunächst kann eine Unterscheidung in SD- und HD-Kameras vorgenommen werden. Bezüglich des Aussehens und der dem praktischen Einsatz angepassten Bauform ist diese jedoch weitgehend nebensächlich. Daher ist folgende Differenzierung gebräuchlich: Zunächst gibt es die Studiokamera, die ohne Rücksicht auf Gewicht und Größe vor allem für beste Signalqualität und komfortable Bedienung konzipiert ist. Da im Studio der Mehrkamerabetrieb dominiert, arbeiten diese im Verbund und werden als Systemkameras bezeichnet.

Professionelle Kameras für den Außeneinsatz werden meist als EB-Kameras (Elektronische Berichterstattung, engl.: Electronic News Gathering, ENG) mit angeschlossener Aufzeichnungseinheit ausgeführt. Am unteren Ende der Qualitätsskala rangieren die semiprofessionellen und die Amateurkameras, die vor allem preiswert, sehr leicht und mit vielen Automatikfunktionen ausgestattet sind. Der Kameraeinsatz in den genannten Gebieten überschneidet sich in vielen Fällen.

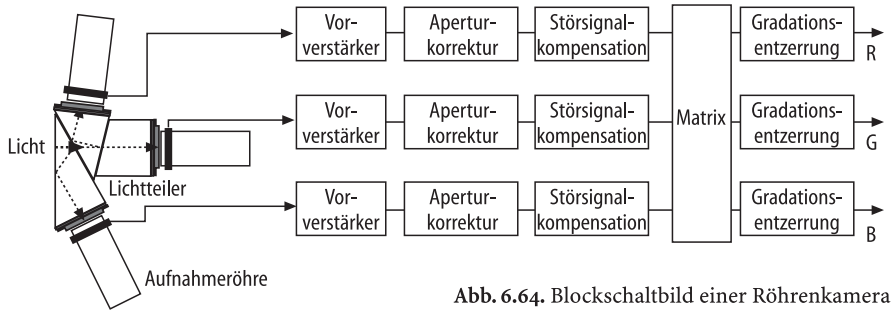


Abb. 6.64. Blockschaltbild einer Röhrenkamera

6.3.1 Das elektronische System der Kamera

Das elektronische System der Videokameras umfasst Baugruppen zur normgerechten Synchronsignalerzeugung, zur Erzeugung von Steuerspannungen für die Röhren-, CCD- oder CMOS-Sensoren. Es umfasst Steuerungssysteme für das Objektiv, verschiedene Automatikfunktionen und schließlich Baugruppen zur Beeinflussung des Videosignals selbst. Wichtig sind hier die Gamma-Vorentzerrung und die Aperturkorrektur. Abbildung 6.64 zeigt als Muster das Blockschaltbild einer Röhrenkamera.

6.3.1.1 Die Ansteuerung des Bildwandlers

Der Bildwandler wird gemäß der durch den Synchronsignalgenerator (SSG) vorgegebenen Zeitstruktur gesteuert. Bei Röhrenkameras müssen dazu, ähnlich wie in jedem Fernsehgerät, sägezahnförmige Ablenkspannungen oder -ströme für die H- und die V-Ablenkung erzeugt werden. Die Sägezahnform bewirkt eine langsame Strahlführung während der aktiven Bilddauer und einen schnellen Strahlrücksprung in der Austastlücke. Aufwendige Korrekturmaßnahmen sorgen für einen linearen Abtastverlauf.

CCD- und CMOS-Sensoren bieten gegenüber Röhren u. a. den großen Vorteil, dass die Geometrieparameter nicht justiert zu werden brauchen. Aufgrund der festen Verbindung der Bildwandler mit dem Strahlteiler können sich die Chips nicht gegeneinander verschieben, außerdem ist der Auslesevorgang durch einen hochstabilen Takt bestimmt. Die Steuerung des Halbleiterbildwandlers ist dagegen umfangreich, so müssen z. B. beim CCD die Steuerspannungen für die Transfergates und viele verschiedene CCD-Register in schneller Folge und sehr exakt bereitgestellt werden. Die erforderlichen elektronischen Schaltkreise können aber in einem Chip integriert werden, sodass sich ein geringer Schaltungsaufwand ergibt.

Der Synchronsignalgenerator erzeugt die Zeitstruktur der gewünschten Videonorm. Ein kompletter SSG kann als integrierte Schaltung realisiert werden, die die H- und V-Synchronsignale, die Austastsignale, den PAL-Kennimpuls und den Farbhilfsträger generiert. Bei professionellen Kameras ist es wichtig, dass der SSG im sogenannten Gen Lock-Betrieb von einem Studiotakt fremdsynchronisiert werden kann, dass er also im Master-Slave-Verbund als Slave agieren kann. Im Studiobetrieb wird damit erreicht, dass alle Kameras die

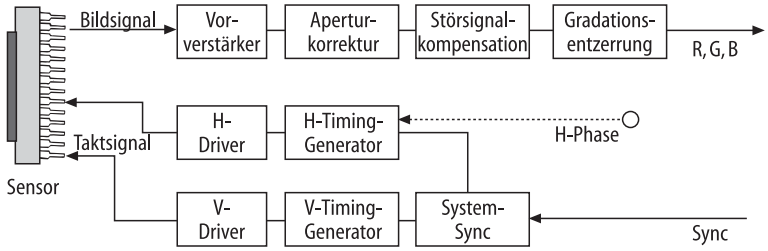


Abb. 6.65. CCD-Ansteuerung

Bilder im gleichen Rhythmus liefern, sodass sie problemlos gemischt werden können. Zur optimalen Anpassung gibt es an der Kamera Einstellmöglichkeiten für die Horizontal- und Farbträgerphase (H-Phase und SC-Phase Abb. 6.65). Die Möglichkeit zum Gen-Lock-Betrieb ist ein wichtiger Unterschied zwischen professionellen und Amateur-Kameras.

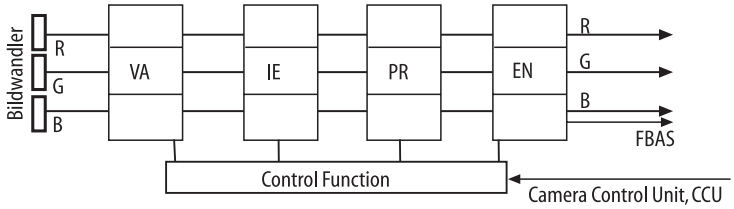
6.3.1.2 Der Videosignalweg

Am Arbeitswiderstand eines Röhrenbildwandlers liegt bereits das Videosignal an. Je heller ein abgetasteter Bildpunkt war, desto größer ist hier die Spannung. Ein möglichst rauscharmer Verstärker bringt dann das Signal auf den Videonormpegel.

Im Gegensatz zum Röhrenbildwandler wird am Ausgang des im CCD enthaltenen horizontalen Schieberegisters die Ladung der einzelnen Bildpunkte nicht kontinuierlich, sondern schrittweise ausgegeben. Der Schrittakt hängt von der Anzahl der horizontalen Bildpunkte ab. Enthält der CCD in jeder Zeile z. B. 800 Pixel, und werden die zugehörigen Ladungen aus dem Schieberegister unter Berücksichtigung einer seitlichen Schwarzmaske in ca. 53 µs herausgeschoben, so beträgt die Taktfrequenz f = 800/(53 µs) ≈ 15,1 MHz. Zur Beseitigung von Störungen, die bei dem zeitlich diskretisierten Auslesevorgang auftreten können, werden die aus dem Schieberegister kommenden Signalanteile zunächst in einer Sample & Hold-Schaltung mit der genannten Taktfrequenz abgetastet. In diese Schaltung kann die Flaw-Korrektur einbezogen werden, d. h. die elektronische Minimierung der durch fehlerhafte Pixel bewirkten Bildfehler. Hierzu werden die Positionen der Fehlstellen registriert und in einem Speicher (ROM) dauerhaft abgelegt. Immer wenn die Ladung einer Fehlstelle am Ausgang der S & H-Stufe erscheint, wird ROM-gesteuert statt des aktuellen Wertes der Wert des vorangegangenen Nachbarpixels ausgelesen. Nach einer Glättung und Verstärkung entsteht am CCD und auch am CMOS-Sensor ein noch nicht normgerechtes Videosignal, denn der Zusammenhang zwischen dem Lichtstrom und der erzeugten Ladung ist weitgehend linear.

6.3.1.3 Signalbearbeitungsstufen

Eine Übersicht über die Signalbearbeitungsstufen ist in den Abbildungen 6.66 und 6.67 dargestellt. Nach der Vorverstärkung kann zum Ausgleich unzureichender Beleuchtungsstärken am Verstärker der Verstärkungsgrad in festen Stufen (meist in Schritten von 1 dB oder 3 dB, meist bis maximal 18 dB) er-



VA = Video Amplifier, Einstellung von: Gain, Knie, Flarekompensation
IE = Image Enhancement Board für: Crispening und Detailing
PR = Processing Board, Einstellung von Schwarz- und Weißwert, Gamma
EN = Encoder, Zusammensetzung von Bild- und Sync-Signalen, FBAS-Erzeugung

Abb. 6.66. Signalbearbeitungsstufen in der Kamera

höht werden (Gain Up). Dabei vermindert sich der Signal-Rauschabstand um den gleichen dB-Wert. Es folgen die Stufen für die Aperturkorrektur, für die Bildverbesserung (Image Enhancement), für die Festlegung der Bildparameter wie Gamma, Weiß- und Schwarzwert, und schließlich die Matrix und die Codierungsstufe für die gewünschte Ausgabesignalart.

Direkt nach der Vorverstärkung wird eine erste Korrektur des Signals zur Beseitigung von Shading und Flare vorgenommen. Die Begriffe Black- und White-Shading beschreiben eine trotz gleichmäßig beleuchteter Bildvorlage ungleichmäßige Helligkeitsverteilung, die im Schwarzbereich durch verschieden große Dunkelströme einzelner Pixel und im Weißbereich durch den Helligkeitsabfall am Bildrand bewirkt wird. Diese Ungleichmäßigkeiten im Signal können elektronisch korrigiert werden. Flare bezeichnet eine ungleichmäßige Aufhellung des Bildfeldes durch das im Objektiv erzeugte Streulicht. Auch das Streulicht kann lokalisiert und kompensiert werden. Anschließend folgt das Image Enhancement mit der Aperturkorrektur. Die genannten Stufen sind hier der Verdeutlichung halber anhand einer analogen Kamera dargestellt, sie werden heute in Digitalkameras in sehr viel kleinerer Bauform realisiert.

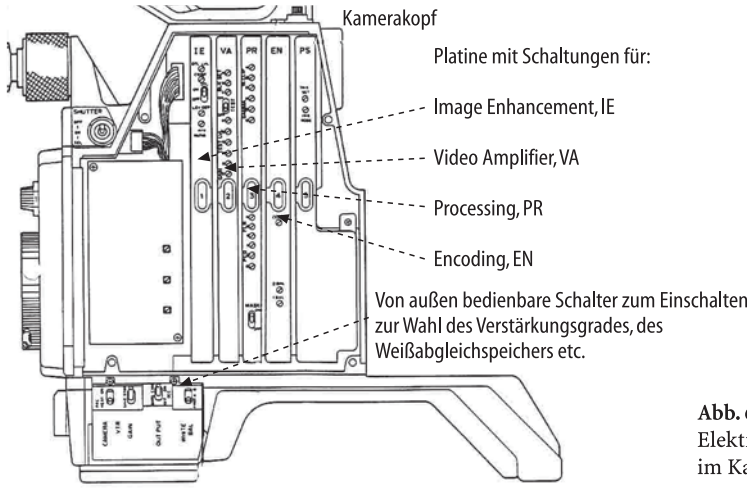


Abb. 6.67. Elektronik-Platinen im Kamerakopf [121]

6.3.1.4 Aperturkorrektur

Dieser Begriff bezieht sich auf den Durchmesser der Öffnung (Apertur) des abtastenden Elektronenstrahls bei Röhrenbildwandlern. Eine große Strahlapertur bewirkt eine Unschärfe im Bild. Aus Gründen der Aliasvermeidung ist eine gewisse Unschärfe erwünscht, die Apertur darf nicht zu klein werden. Der Begriff Aperturkorrektur wird nicht nur bei Röhrenbildwandlern, sondern auch bei CCD und CMOS benutzt, auch hier haben die Pixel eine definierte aktive Fläche. Allgemein wird der Begriff für den elektronischen Ausgleich dieser Unschärfe, gebraucht, die zusätzlich durch das optische Tiefpassfilter bewirkt wird. Elektronische Maßnahmen zur Steigerung des Schärfeeindrucks im Bild werden auch als Konturkorrektur, Detailing oder Crispening bezeichnet.

Ein Verlust der Schärfewirkung fällt besonders an scharfen Bildkanten auf. Abbildung 6.8 zeigt, wie dort Signale durch den Abtaststrahl abgeschrägt werden. Als Ausgleich kann mit einer Crispening-Schaltung eine Erhöhung der Kantenschärfe erreicht werden, indem das Videosignal zweifach differenziert und dieses Signal dann invertiert dem Originalsignal hinzugefügt wird. Nach einem ähnlichen Prinzip, jedoch rauschärmer, arbeitet die in Abbildung 6.68 erklärte Detailing-Schaltung (H-Aperturkorrektur).

Generell wird der Rauschpegel im Signal durch die Anwendung der Aperturkorrektur erhöht. Insbesondere bei der Abbildung von Gesichtern bringt das Detailing Probleme, da die Haut bei hoch eingestelltem Wert sehr flekkig erscheint. Hier muss nach subjektiven Kriterien auf Kosten der Schärfe des übrigen Bildes ein kleiner Detailing-Level gewählt werden. Zur Umgehung dieses Problems enthalten einige Kameras sogenannte Skin-Detailing-Schaltungen. Dabei wird der Hautton der aufgenommenen Person in der Kamera gespeichert und jedesmal, wenn dieser Hauttonwert auftritt, wird automatisch das Detailing herabgesetzt. Während der Abtastung der anderen Bildbereiche bleibt ein hoher Detailing-Wert wirksam.

6.3.1.5 Die Kamerakennlinie

Die Kennlinie beschreibt den Zusammenhang zwischen dem auf den Bildwandler treffenden Lichtstrom und dem Videosignalwert. Die Kennlinie ist durch den Schwarz- und den Weißwert begrenzt und verläuft aufgrund der Gamma-Vorentzerrung nichtlinear. Abbildung 6.69 zeigt den Verlauf einer Kamerakennlinie links bei linearer Teilung der Koordinatenachsen für die Signal-

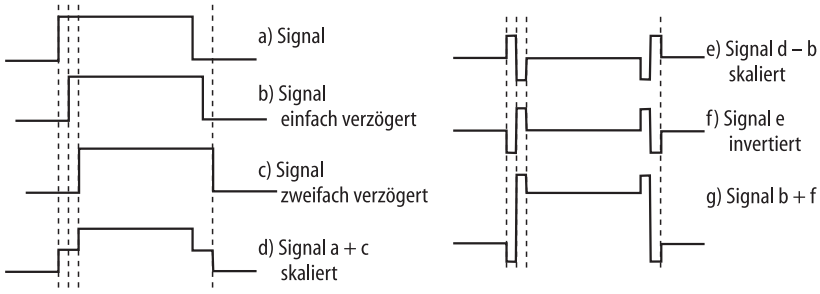


Abb. 6.68. H-Aperturkorrektur

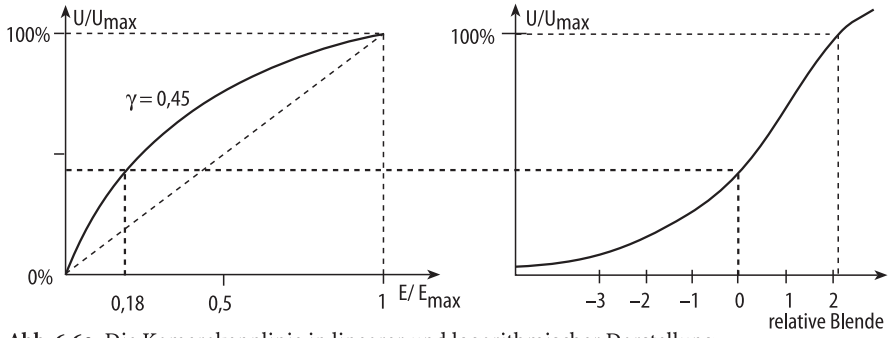


Abb. 6.69. Die Kamerakennlinie in linearer und logarithmischer Darstellung

spannung U (y-Achse) und die Beleuchtungsstärke E (x-Achse) und rechts in halblogarithmischer Darstellung, bei der die x-Achse logarithmisch eingeteilt ist und damit Blendenstufen entspricht, die empfindungsgemäß gleichabständig sind und damit besser dem visuellen Eindruck gerecht werden.

Die Gamma-Vorentzerrung bestimmt den nichtlinearen Verlauf der Kamerakennlinie, der nicht vom Bildwandler herrührt, sondern bewusst herbeigeführt wird, um den nichtlinearen Zusammenhang zwischen Videosignalpegel und entstehender Leuchtdichte bei der früher zur Wiedergabe eingesetzten Kathodenstrahlröhre zu kompensieren. Die Nichtlinearität wird mit Hilfe des Exponenten γ angegeben. Auf der Aufnahmeseite gilt:

$$U/U_m = (E/E_m)^{\gamma_A},$$

wobei U/U_m das Verhältnis der Videosignalspannung zum Maximalwert und E/E_m das Verhältnis der zugehörigen Beleuchtungsstärken ist. Der Gamma-Wert der Wiedergaberöhre ist mit $\gamma_W = 2{,}2$ theoretisch vorgegeben, um eine lineare Gesamtübertragung zu erhalten, muss auf der Aufnahmeseite also der Kehrwert $\gamma_A = 1/2{,}2 \approx 0{,}45$ eingestellt werden. Die Grauwertabstufungen im Bild lassen sich durch die Variation des Gamma-Wertes ändern. Professionelle Kameras erlauben darüber hinaus eine Einstellung des Blackgamma oder Blackstretch (Abb. 6.70), also eine Veränderung des Kennlinienverlaufs nur für die dunklen Bildpartien. Insgesamt lässt sich bei einer guten Kamera die Kennlinie zur Anpassung an besondere Beleuchtungssituationen vielfältig verändern.

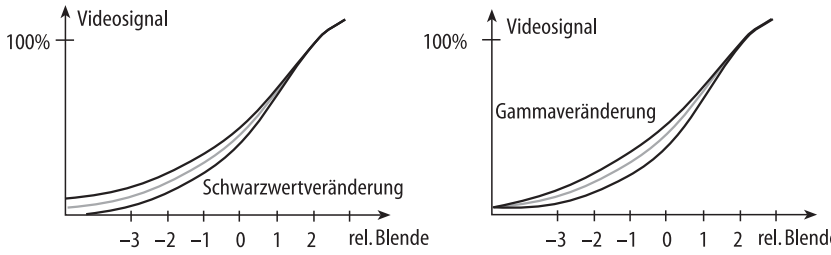


Abb. 6.70. Beispiele für Veränderungen der Kamerakennlinie

6.3.1.6 Schwarzklemmung, Weißbegrenzung und Aussteuerung

Das Videosignal ist bei Dunkelheit nicht gleich Null, es tritt ein Dunkelstrom auf, der mit Hilfe des Signals aus dem lichtdicht abgedeckten Bereich des Sensors kompensiert wird. Dazu wird eine Schwarzwertklemmung vorgenommen, wobei das Signal aus dem abgedeckten Bereich als Schwarzbezugswert dient.

Zwischen Austast- und Schwarzwert wird für das Luminanzsignal eine Schwarzabhebung (Pedestal) von normalerweise 0 ... 3 % des Maximalwertes eingestellt. Wenn der Schwarzwert zu hoch eingestellt wird, ist die Dynamik eingeschränkt, und das Bild wirkt flau (Abb. 6.71). Der Pedestal kann absichtlich etwas angehoben werden, um die Unterscheidbarkeit von Details, die in dunklen Bildbereichen liegen, zu verbessern. Das spielt vor allem im Studiobetrieb eine Rolle, da die Kameraparameter dort während der Aufnahme ständig verändert werden können. Das Luminanzsignal darf dagegen den sogenannten Ultra-Schwarzbereich unterhalb des Austastwertes nicht erreichen, in diesem Bereich dürfen nur FBAS-Chrominanzwerte auftreten.

Der Weißwert ist der Luminanz-Videosignalwert bei 0,7 V (100%). Signalspitzen dürfen nur geringfügig größer sein, denn sie werden durch das Clip-

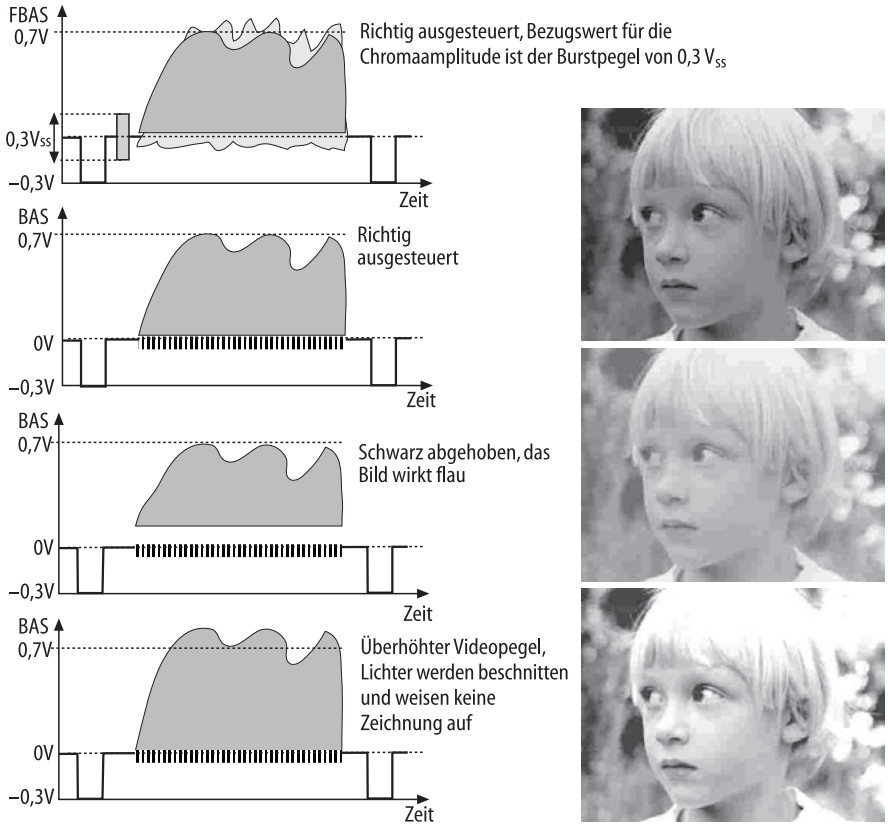


Abb. 6.71. Schwarz- und Weißwerteinstellung

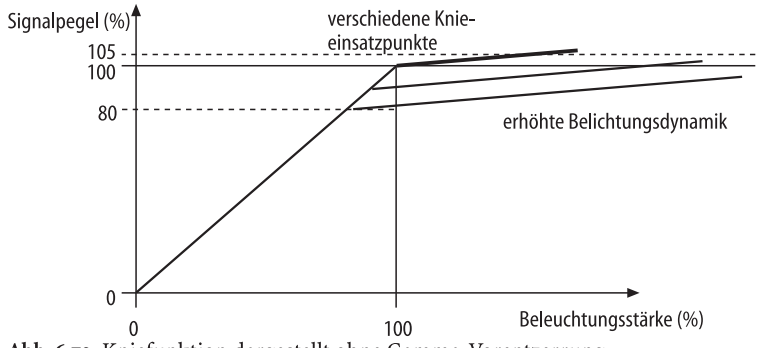


Abb. 6.72. Kniefunktion dargestellt ohne Gamma-Vorentzerrung

ping elektronisch beschnitten. Wenn große Anteile der Luminanzamplitude abgeschnitten werden, gehen die Abstufungen (Zeichnung) in den Lichtern verloren (Abb. 6.71). Der Clippingwert wird minimal größer als 100% gewählt. Bei Beiträgen, die für die Fernsehausstrahlung vorgesehen sind, sollte der Wert nicht größer als 103% sein, denn vor der Ausstrahlung wird das Signal nochmals begrenzt. Der Grad der durch das Clipping bewirkten Beschneidung der Spitzlichter wird im Studio nach subjektiven Kriterien gewählt. Zusammen mit der Pedestal-Einstellung wird das Signal immer so eingestellt, dass die wichtigen Bildteile genügend Zeichnung aufweisen.

Damit die Lichter im Weißbereich nicht abrupt abgeschnitten werden und noch Unterschiede erkennbar bleiben, d. h. ausreichend Zeichnung aufweisen, kann die Kennlinie im Signalbereich um 100% abgeflacht werden (Abb. 6.72). Durch diese Kniefunktion lässt sich eine Steigerung des Kontrastumfangs von mehr als einer Blendenstufe im Bereich der Lichter erreichen. Der Kniepunkt, d. h. der Einsatzpunkt der Abflachung, kann etwa zwischen 80% und 110% eingestellt werden. Eine Autokniefunktion bewirkt eine selbsttätige Verschiebung dieses Punktes in Abhängigkeit vom Motivkontrast. Diese Funktion erlaubt die Verarbeitung von Pegeln, die bis zu 600% der Normalaussteuerung betragen, die Kameraelektronik muss entsprechend hoch aussteuerbar sein.

6.3.1.7 Zebra

Der Einsatz des Signalclipping kann nur mit Hilfe von Messgeräten, z. B. einem Waveformmonitor, sicher festgestellt werden. Im Studiobetrieb sind diese Geräte vorhanden. Beim EB-Betrieb, also im Reportageeinsatz, muss das Bild allein anhand des Kamerasuchers beurteilt werden, wobei Signalübersteuerungen schwer erkennbar sind. Um sie dennoch deutlich werden zu lassen, kann bei vielen Kameras das Zebra aktiviert werden. Die Zebra-Funktion bewirkt, dass das Bild im Kamerasucher an den Stellen, an denen ein bestimmter Signalpegel vorliegt, ein oszillierendes Muster aus Punkten oder Streifen aufweist. Ein 100%-Zebra dient vor allem zur Übersteuerungskontrolle. Es kann als Einstellhilfe für optimale Belichtung nur dann verwendet werden, wenn im Bild eine Stelle mit hohem Weißanteil auftaucht, eine Stelle also, an der das Videosignal tatsächlich 100% erreichen soll. Wenn es eher um optimale Aussteuerung als

Zebramuster bei 70% Videopegel

Zebramuster bei 100% Videopegel

Abb. 6.73. Zebrafunktion

um Übersteuerung geht wird das 70%-Zebra eingesetzt. Das 70%-Zebra liegt in einem Signalpegelbereich, der bei der Abbildung des Hauttons hellhäutiger Menschen auftritt. Wenn die Blende so gewählt wird, dass das Zebra auf den hellsten Hautpartien erscheint, so werden diese gut abgebildet (Abb. 6.73).

Die Zebra-Funktion kann mit Hilfe der „Gain up"-Methode auch für andere Signalbereiche verwendet werden. So kann z. B. ein Signalpegel von 50% gefunden werden, indem die Verstärkung (Gain) um + 6 dB angehoben wird. Die Blende wird dann so eingestellt, dass das 70%-Zebra an den Stellen erscheint, die 50% Pegel aufweisen sollen und dann wird die Verstärkung wieder abgeschaltet. Dieser Vorgang führt zu einer Signalverdopplung, die sich aufgrund der γ-Vorentzerrung nur zu 70% auswirkt.

6.3.1.8 Linear Matrix

Mit dem Begriff Matrix ist bezüglich der Kamera nicht die Schaltung zur Umsetzung von Komponentensignalen in RGB-Signale gemeint. Zwar wird auch in der Kamera-Matrix eine gegenseitige Signalverkopplung zwischen den RGB-Kanälen vorgenommen, sie dient jedoch der Korrektur der Farbwiedergabe. Diese Korrektur ist erforderlich, um die negativen Anteile (Abb. 6.74) der Spektralwertkurven der standardisierten Bildschirm-Primärvalenzen zu berücksichtigen. Die Spektralwertkurven des Aufnahme-Bildwandlers haben keine negativen Anteile. Sie bilden also ein virtuelles Primärvalenzsystem, das durch eine lineare Transformation in jedes andere, also auch das Primärvalenzsystem des Bildschirms überführt werden kann. Diese lineare Transformation wird in

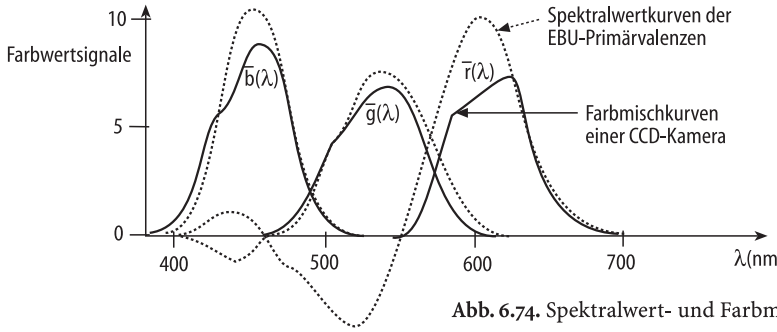


Abb. 6.74. Spektralwert- und Farbmischkurven

elektronischer Form mit Hilfe der Matrix vorgenommen. Mit der Matrix ist es also quasi möglich, Farbmischkurven mit negativen Anteilen zu erzielen [136]. Da die Matrix die farbliche Feinabstufung und damit den Charakter der Hauttonwiedergabe beeinflusst, gibt es dafür unterschiedliche Vorgaben, z. B. der EBU oder von Sendeanstalten wie BBC, ARD/ZDF, RAI etc.

6.3.2 Die Digitalkamera

Der Begriff Digitalkamera bezieht sich nicht auf die Bildwandlung, sondern auf die Signalverarbeitung. Die Ladungen aus dem CCD- oder CMOS-Sensor liegen am Bildpunkt in analoger Form vor. Das Signal wird vorverstärkt und zunächst in einem Preprocessing-Modul verarbeitet, das für die Änderung der Verstärkung, des Shading und des Weißabgleichs digital gesteuert wird.

Anschließend wird das Signal mit einem Analog-Digital-Umetzer hoher Qualität digitalisiert und durch Rechenprozesse verändert. Dabei werden vor allem Gamma- und Knie-Parameter und das Detailing auf digitaler Basis festgelegt. Hinzu kommen der Schwarzwert und die Matrix, die auf digitaler Ebene für alle drei Kanäle sehr exakt eingestellt werden kann. Für ein hochwertiges Videosignal ist am Ausgang der Signalbearbeitungsstufen eine 10 Bit-Auflösung der RGB-Anteile erforderlich. Bei der A/D-Umsetzung reicht die Quantisierung mit 10 Bit jedoch nicht aus, hier sollte mit mindestens 14 Bit-Auflösung gearbeitet werden. Außerdem ist auch beim digitalen Processing ist ein zusätzlicher Spielraum für die Signalanpassung erforderlich (Abb. 6.75). Um mit der Kniefunktion z. B. ein Signal von bis zu 600% Überpegel verarbeiten zu können sind 2,5 Bit sind erforderlich. Der Bereich der Gammaeinstellung erfordert 2 Bit, zwei weitere sind erforderlich, um genügend Spielraum zur Einstellung des Schwarzwertes zu erhalten. Idealerweise sollte eine interne Verarbeitung mit 18 Bit möglich sein [144].

Wenn das Signal am Ausgang des Kamerakopfes mit 10-Bit-Auflösung vorliegt, so entsteht, je nach Bildauflösung, schon bei hochwertigen SD-Kameras ein Datenstrom von dreimal ca. 170 Mbit/s, die zusammen mit weiteren Signalen zur Basisstation (CCU) übertragen werden müssen. Diese Datenrate ist für die Digitalübertragung über Triaxkabel zu hoch. Daher werden die Signale zunächst D/A-gewandelt und analog übertragen. Zukünftig wird durch den Einsatz mehrwertiger Modulationsverfahren, wie z. B. 16-QAM (s. Kap. 4.5.2), eine digitale Übertragung möglich sein. Aktuell wird vermehrt die digitale Übertragung per Glasfaser genutzt.

Ein wesentlicher Vorteil der digitalen Signalverarbeitung ist die Möglichkeit, auf die elektronischen Parameter und die Kennlinie sehr umfangreich Einfluss

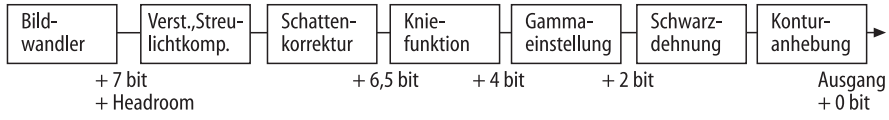


Abb. 6.75. Für die Verarbeitung in einer Digitalkamera erforderliche Bitzahl

nehmen zu können, die gewählten Einstellungen zu speichern und vor allem exakt reproduzieren zu können. Verschiedene Kamera-Set-Ups können auf austauschbaren Karten gespeichert werden, sodass sehr umfangreiche Einstellungen auf Knopfdruck aufgerufen werden können. Darüber hinaus erfordern Digitalkameras wenig Abgleich und behalten die eingestellten Parameter unverändert bei.

6.3.3 Der Weißabgleich

Wie bereits in Abschn. 2.3 beschrieben, hat das Licht verschiedener Lichtquellen unterschiedliche spektrale Zusammensetzungen. Temperaturstrahler können mit Hilfe der Farbtemperatur charakterisiert werden. Sonnenlicht hat z. B. eine Farbtemperatur von ca. 5800 K, im für den Menschen sichtbaren Spektralbereich sind hier die Frequenzen etwa gleich intensiv vertreten. Künstliche Temperaturstrahler weisen dagegen eine Farbtemperatur von ca. 3200 K auf, rote Spektralanteile sind hier intensiver als blaue.

Das menschliche Auge passt sich dem Frequenzgemisch an und empfindet die Farbe eines Blattes Papier als Weiß, egal ob es mit Sonnenlicht oder Kunstlicht mit 3200 K beleuchtet wird. Nur wenn beide Lichtarten gleichzeitig wirksam sind, wird der Unterschied sichtbar. Die Kamera muss elektronisch an veränderte Frequenzgemische angepasst werden. Dies geschieht mit dem Unbuntabgleich in Bildweiß und -schwarz, umgangssprachlich Weiß- und Schwarzabgleich genannt, für jede veränderte Beleuchtungssituation erneut. Beim Weißabgleich wird die Kamera auf eine weiße Fläche gerichtet, die von der jeweiligen Lichtart beleuchtet wird. Während des Abgleichs werden die Verstärkungen in den R- und B-Kanälen der Kamera so eingestellt, dass alle drei Kanäle gleiche Signalpegel aufweisen (Abb. 6.76). Im Studiobetrieb wird der Abgleich über eine Kontrolleinheit vorgenommen, wobei mit dem Vektorskop (s. Kap. 2) sichergestellt wird, dass das Signal keine Buntanteile mehr enthält. Bei professionellen Kameras im EB-Einsatz wird der Abgleich auf Knopfdruck automatisch vorgenommen (Automatic White Control, AWC). Dabei reicht es aus, die weiße Referenzfläche nur auf einem kleinen Teil der Bildwandlerfläche

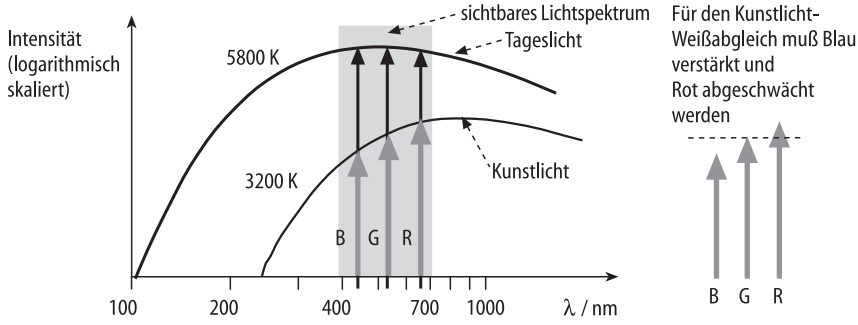


Abb. 6.76. Weißabgleich

abzubilden. Meistens können mehrere Abgleichparameter gespeichert werden, z. B. ein Satz für Tageslicht und einer für Kunstlicht, sodass bei einem Wechsel zwischen Innen- und Außenaufnahme statt eines erneuten Weißabgleichs nur der zugehörige Speicher aufgerufen werden muss.

Moderne Kameras können sehr große Farbbereiche zu Unbunt abgleichen. Trotzdem sollten vorher die Konversionsfilter im Filterrad auf die entsprechende Farbtemperatur eingestellt werden damit genug Spielraum für die elektronische Regelung erhalten bleibt. Nach dem Weißabgleich darf das Filterrad und die Beleuchtung nicht mehr verändert werden. Entsprechend dem Weißabgleich wird auch der Schwarzabgleich (Automatic Black Control, ABC) bei geschlossener Blende so vorgenommen, dass die Signale der drei Farbkanäle im Schwarz gleiche Pegelwerte aufweisen. Diese Einstellung ist unabhängig von der Beleuchtungsart.

6.3.4 Automatikfunktionen

Automatikfunktionen werden vor allem in Amateurkameras eingesetzt, um ungeübten Personen die Bedienung zu erleichtern. In diesen Bereich gehören vor allem der fortwährend durchgeführte Weißabgleich (Auto Tracing Whitebalance, ATW), die automatische Blendensteuerung (Auto Iris) und die automatische Scharfstellung (Autofocus). Automatikfunktionen haben im professionellen Bereich eine geringere Bedeutung. Die Automatikparameter beziehen sich zwangsläufig auf Standardaufnahmesituationen und bergen immer die Gefahr der Fehleinstellung. Auch bei Amateurkameras sollte die Automatik daher abschaltbar sein. Im Zuge der technischen Entwicklung arbeiten die Kontrolleinheiten für diese Funktionen allerdings zunehmend differenzierter, sodass auch in professionellen Kameras, insbesondere bei denen, die für den EB-Bereich konzipiert sind, immer mehr Automatikfunktionen integriert werden.

Mit dem automatischen Weißabgleich werden die Verstärkungen der Farbkanäle ständig an neue Umgebungslichtsituationen angepasst. In der Nähe des Objektivs wird ein Sensor angebracht, der z. B. Fotodioden enthält, denen das Licht über Rot-, Grün- und Blau-Filter zugeführt wird. Helle Bildbereiche werden als Weiß interpretiert und diesbezüglich wird mit Hilfe der aus den Sensoren gewonnenen Spannungen die Verstärkung der Blau- und Rot-Kanäle verändert. Der automatische Weißabgleich macht die Wiedergabe ausgeprägter Farbstimmungen (Abendrot) unmöglich.

Die automatische Blendensteuerung ist auch in vielen professionellen Kameras enthalten. Eine elektronische Schaltung registriert den Videosignalpegel und steuert die Blendenöffnung so, dass immer ein mittlerer Signalpegel erhalten bleibt und gleichzeitig wenig Übersteuerung auftritt. Nachteil dieser Automatik ist ein in bestimmten Situationen auftretender unnatürlich schwankender Helligkeitseindruck, z. B. wenn eine Person die abgebildete helle Szene durchquert und sie damit kurzzeitig abdunkelt. Bei Änderung der Blende verändert sich auch immer die Schärfentiefe. Häufig kann die automatische Blendenfunktion neben der festen Einstellung auch nur kurzzeitig auf Tastendruck aufgerufen werden. Damit erreicht man schnell eine relativ optimale

Blendeneinstellung, die dann nur wenig modifiziert beibehalten werden kann. Die automatische Blendensteuerung funktioniert nur solange, bis die Blende vollständig geöffnet ist. Um auch noch bei niedrigeren Beleuchtungsstärken arbeiten zu können, kann bei einigen Kameras eine automatische Verstärkungserhöhung eingeschaltet werden (Automatic Gain Control).

Die Fokussierung der Kamera ist schwierig zu automatisieren, denn auch die beste Elektronik kann nicht feststellen, ob der Vorder- oder der Hintergrund in einer Szene scharf abgebildet werden soll. Trotzdem ist die Automatik weit verbreitet. Amateurkameras enthalten in der Regel immer eine Autofokus-Funktion. Auch in professionellen Kameras sind sie vermehrt vorhanden, besonders bei HD-Kameras, bei denen der Fokus schwierig zu finden ist. Ihre Funktion ist in Abschnitt 6.2.2 beschrieben.

6.3.5 Bildstabilisierungssysteme

Amateurkameras sind im Laufe der Entwicklung immer kleiner und leichter geworden, sie können ohne Anstrengung mit der Hand geführt werden. Mit dieser Art der Führung ist aber der Nachteil verbunden, dass die Kamera sehr schlecht ruhig gehalten werden kann. Zur Erzeugung von nicht verwackelten Bildern ist ein Stativ unerlässlich, womit allerdings die Flexibilität der Kamera wieder eingeschränkt ist. Um die Verwackelung zu reduzieren, wurden für Amateurkameras Bildstabilisierungssysteme entwickelt, die auch im professionellen Bereich Einzug gehalten haben. Man unterscheidet Stabilisatoren bezüglich des Funktionsprinzips in elektronisch und optisch arbeitende Typen [182].

6.3.5.1 Elektronische Bildstabilisatoren

Diese Art der Stabilisierung wird auch als Digital Image Stabilizer (DIS) bezeichnet. Zunächst muss die Verwackelungsbewegung erfasst werden. Dazu werden die Helligkeitswerte von mehr als 100 repräsentativen Bildpunkten gespeichert und mit den Werten der Punkte im nächsten Halbbild verglichen. Bei Übereinstimmung liegt keine Verwackelung vor.

Falls die Werte für die Referenzpunkte dagegen nicht gleich sind, wird festgestellt, wo die Referenzpunkte mit dem neuen Bild in Übereinstimmung gebracht werden können. Daraus werden dann Bewegungsvektoren bestimmt, die die Lageveränderung durch die Verwackelung beschreiben. Diese Art der

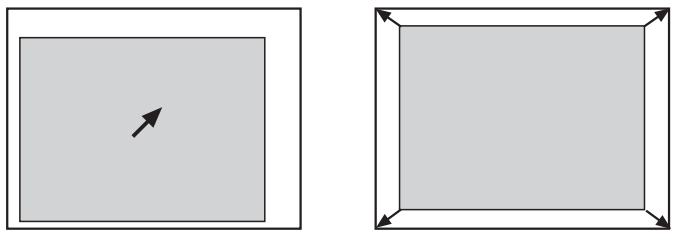


Abb. 6.77. Digital Image Stabilizer Prinzip

Bewegungsdetektion erfordert ausreichende Helligkeit. Eine andere, lichtunabhängige Methode zur Erfassung der Verwackelung beruht auf der Nutzung von zwei piezoelektrischen Beschleunigungsaufnehmern für die horizontale und vertikale Bewegungsrichtung. In jedem Fall muss eine elektronische Schaltung dafür eingesetzt werden, die Verwackelungsbewegung von einer gewollten Bewegung (Kameraschwenk) zu unterscheiden und ggf. die Stabilisierung danach auszurichten.

Nachdem die Bewegungsvektoren festgelegt sind, wird das Bild korrigiert. Dazu wird beim DIS-Verfahren dafür gesorgt, dass die Sensorfläche größer als der Bildausschnitt ist. Es entsteht ein Rand um die aktive Bildfläche herum, der ca. 15% der Sensorfläche umfasst (Abb. 6.77). Damit kann der aktive Bildausschnitt, durch die ermittelten Bewegungsvektoren gesteuert, so verschoben werden, dass die Verwackelung ausgeglichen wird. Einige Systeme erzeugen den Rand um die aktive Bildfläche durch einen elektronischen Verkleinerungsprozess. Nach der Korrektur wird hier das Bild wieder elektronisch vergrößert.

6.3.5.2 Optische Bildstabilisatoren

Stabilisatoren auf optischer Basis (Steady Shot) arbeiten mit mechanischen Mitteln. Sie bieten den Vorteil, dass keine Bildverfälschung durch einen elektronischen Zoom-Prozess auftreten kann. Es kann die gesamte Bildwandlerfläche genutzt werden. Außerdem funktioniert das System unabhängig von der Lichtintensität. Die Erschütterungsdetektion erfolgt meist mit den bereits genannten piezoelektrischen Beschleunigungssensoren.

Zur Korrektur werden zwei geneigte Glasplatten in den Strahlengang gebracht. Die Glasplatten sind durch einen hermetisch abgeschlossenen, beweglichen Balgen verbunden, der mit Silikonöl gefüllt ist, welches den gleichen Brechungsindex wie Glas aufweist (Abb. 6.78). Die Gläser sind auf Achsen gestützt, sodass eine horizontale und vertikale Verdrehbarkeit um ca. ± 3° ermöglicht wird. Die beiden Gläser bilden damit ein veränderliches, ein aktives Prisma. Zur Bildstabilisierung wird eine von den Bewegungsvektoren bestimmte Spannung erzeugt, die zum Antrieb der an den Achsen der Gläser angebrachten Stellmotoren dient. Die Achsen werden dann so verdreht, dass das im Objektiv hinter der Frontlinse platzierte aktive Prisma die Verwackelung durch Ausrichtung der optischen Achse ausgleicht.

Für den professionellen Bereich gibt es optische Stabilisatoren in Sandwich-Bausweis, die für 2/3"-Kameras zwischen Objektiv und Kamera montiert werden können. Sie verändern den Brennweitenbereich.

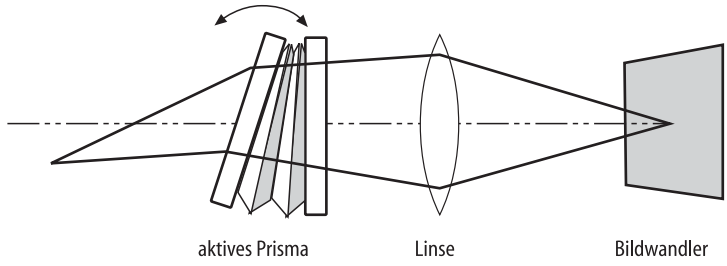


Abb. 6.78. Aktives Prisma zur optischen Bildstabilisierung

6.3.6 Der Kamerasucher

Kamerasucher (engl.: finder) dienen zur optischen Kontrolle des aufgenommenen Bildes. Zur Bildwiedergabe werden hauptsächlich hochauflösende kleine S/W-Monitore eingesetzt, deren Funktionsprinzip in Kap. 6 beschrieben wird. Neuerdings kommen vermehrt auch Farbmonitore auf LCD- oder OLED-Basis zum Einsatz, die Bildauflösung ist für genaues Arbeiten, insbesondere im HD-Bereich, jedoch oft unzureichend. Nach Möglichkeit wird das Farbbild mit einem separaten Farbmonitor kontrolliert. Im Sucher wird das Bild im »Under-Scan«-Modus, d. h. unbeschnitten, wiedergegeben.

Die Suchereinheit ist bei Studiokameras relativ groß (z. B. 12,7 cm (5“) Bilddiagonale) und kann aus einiger Entfernung betrachtet werden (Abb. 6.79 rechts). Es besteht die Möglichkeit, zwischen der Wiedergabe des Kamerabildes und eines weiteren Signals zu wählen. Das ist z. B. im Studioeinsatz mit mehreren Kameras hilfreich, da zur Orientierung auf das gerade am Bildmischer ausgewählte (geschnittene) Bild umgeschaltet werden kann.

Ob eine Kamera »geschnitten«, oder »auf Sendung« ist, wird durch eine Markierung deutlich gemacht, die in den Sucher eingeblendet wird. Die gleiche Markierung wird auch mit einem Rotlicht (Tally) nach Außen angezeigt. Diese Anzeige kann abgeschaltet werden. In den Sucher können weitere Informationen eingeblendet werden wie z. B. eine Sicherheitszonen-Anzeige (Safe-Area), die angibt, welcher Ausschnitt des dargestellten Bildes beim Empfänger sicher wiedergegeben wird. In den Sucher werden auch Informationen eingeblendet, die einen ggf. angeschlossenen Recorder betreffen. Neben der Aufnahmeanzeige finden sich z. B. Anzeigen zur Audiopegelkontrolle, der eingestellten Blende und den Shutter, Warnanzeigen über ein sich näherndes Bandende, Feuchtigkeit im Recorder oder den Ladezustand des Akkus. Als Einstellmöglichkeiten sind am Sucher Steller für die Bildhelligkeit, den Kontrast und das sogenannte Peaking vorhanden, d. h. für eine elektronische Anhebung der Konturen allein des Sucher-Bildes, die das Scharfstellen erleichtert.

An EB- EFP- und Amateurkameras werden sehr kleine, leichte Sucher verwendet (z. B. 1,5“ Bilddiagonale), deren Bildschirm über ein Okular betrachtet werden muss (Abb. 6.79 links). Mittels eines Stellringes kann das Okular an

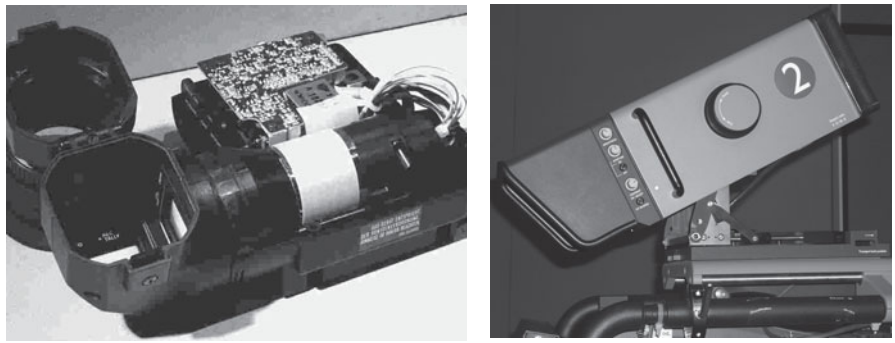


Abb. 6.79. Kleiner Kamerasucher, geöffnet, und großer Kameramonitor

das Auge des Betrachters angepasst werden. Für die Einstellmöglichkeiten gilt das Gleiche wie bei der Studiokamera. Bei den meisten Amateurkameras können allerdings Helligkeit, Kontrast und Peaking nicht eingestellt werden. Im Amateurbereich stehen kleine, ausklappbare LCD-Monitore, allein oder zusätzlich zum Sucher, zur Verfügung. Darüber hinaus gibt es heute auch eine Vielzahl kleiner LCD- oder OLED-Monitore von Drittherstellern, die auch höhere Bildqualitätsansprüche erfüllen.

6.3.7 Die Studiokamera

Die Studiokamera ist mit wenig Rücksicht auf Gewicht und Energiebedarf für höchste Signalqualität konzipiert. Bis in die 90er Jahre hinein basierte sie auf Röhren und war entsprechend groß. Mit dem Einzug von CCD-Sensoren konnte sie immer kompakter gebaut werden und ist heute in vielen Fällen so konstruiert, dass sie auf der Schulter getragen oder in einen Large Lens Adapter gesetzt werden kann. In letzterem Fall können die großen Studioobjektive verwendet werden. Die Abbildungen 6.80 und 6.81 ermöglichen den Vergleich zwischen einer Röhren- und einer modernen CCD-Studiokamera.

Studiokameras haben meist hochwertige Objektive mit großem Zoombereich. Sie sind mit großen Suchern und einer Kommunikationseinrichtung mit der Regie (Intercom) ausgestattet. Das Bedienpersonal (Kamera-Operator) arbeitet mit einer Hör-Sprechgarnitur (Headset) und konzentriert sich entsprechend den Regieanweisungen nur auf die Kameraführung, die Wahl des Bildausschnitts mittels des Zoomobjektivs und die Fokussierung. Damit die Hände an den Führungsgriffen verbleiben können, sind für Zoom und Schärfe

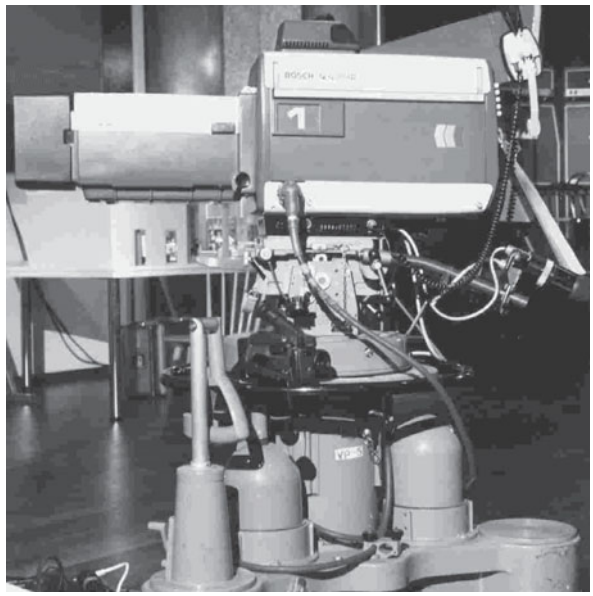


Abb. 6.80. Röhrenkamera auf Pumpstativ

Abb. 6.81. Studiokamera

Fernbedienelemente an den Griffen angebracht (Abb. 6.81 und Abb. 6.82). Hier kann meistens auch die Zoomgeschwindigkeit eingestellt und mittels einer Taste das Sucherbild zwischen dem Kamera- und einem Alternativsignal aus der Regie umgeschaltet werden.

Studiokameras werden auch als Systemkameras bezeichnet, da sie für Mehrkameraproduktionen in einem Kameraverbund ausgelegt sind. Im Studiobetrieb werden die Kameras bei den meisten Produktionen so eingesetzt, dass sie jeweils eine Nummer erhalten. Bei Kamerakommandos aus der Regie ist es nützlich, wenn diese Nummer zuerst genannt wird, damit das Bedienpersonal

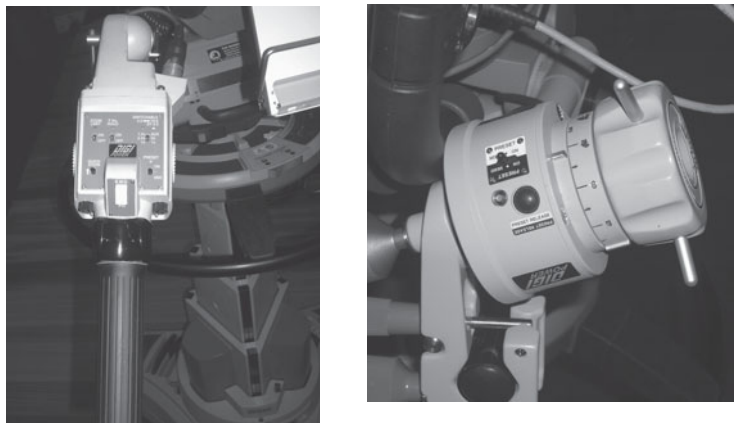


Abb. 6.82. Bedienelemente für Zoom und Fokus

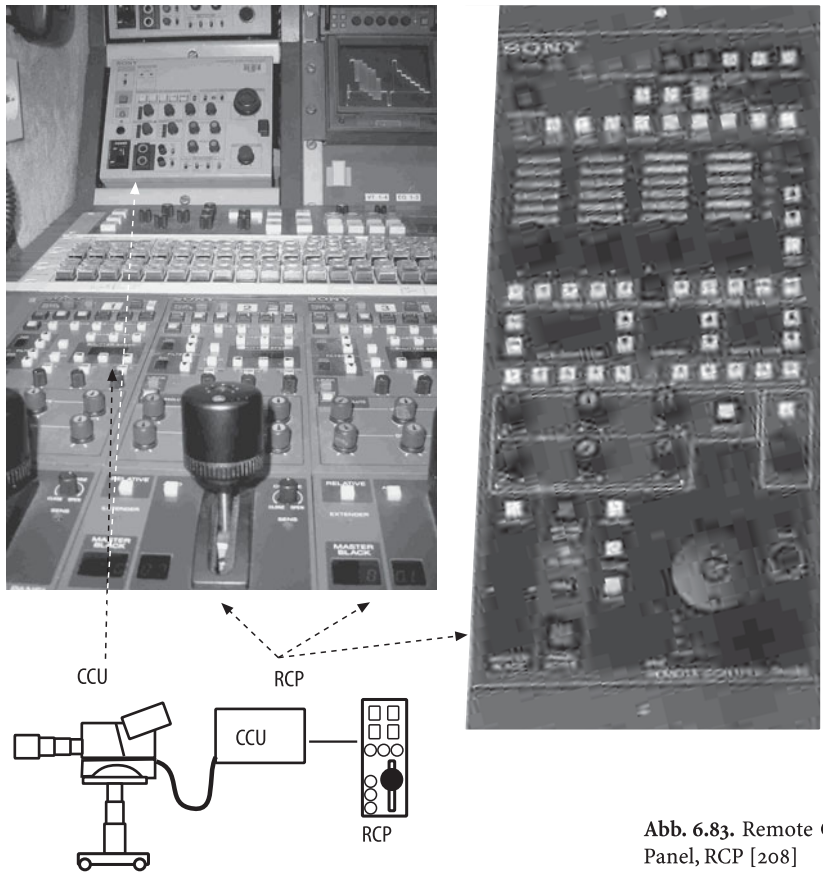


Abb. 6.83. Remote Control Panel, RCP [208]

zunächst weiß, wer angesprochen wird. Dann folgt eine sehr knappe Anweisung, damit auch weitere Kameras schnell angesprochen werden können. Dies spielt besonders im Live-Betrieb eine große Rolle, da alle Beteiligten unter starker Anspannung arbeiten.

Die Einstellung der Blende und aller bildwichtigen Parameter wird aus der Bildtechnik ferngesteuert. Die Fernbedienung geschieht mit einem separaten Kamerakontrollgerät (Remote (Operational) Control Panel, RCP (OCP)), das die Kameraelektronik in der von der Kamera getrennten Basisstation (Camera Control Unit, CCU) steuert. Am RCP können die Blende, der Schwarzwert, die Farbmischung etc. eingestellt werden. Abbildung 6.83 zeigt links die Verbindung von Kamera, CCU und der Bedieneinheit und rechts ein Remote Control Panel (RCP). Die wichtigsten Bildparameter lassen sich mit dem Hebel unten rechts einstellen. Dessen Vor- Rückbewegung steuert die Blende, die Drehung am Knopf den Schwarzwert und der Druck auf den Hebel bewirkt, dass das zugehörige Kamerabild auf dem Bildkontrollmonitor geschaltet wird.

Die Kamera ist mit nur einem Kabel mit der CCU verbunden, über das alle Signale, die Kommunikation, die Stromversorgung usw. geführt werden.

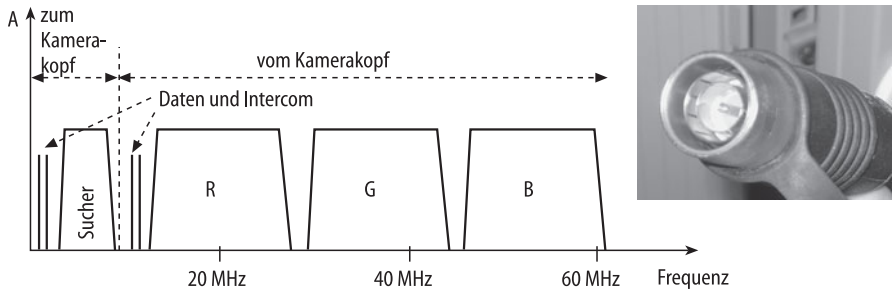


Abb. 6.84. Frequenzmultiplex und Steckverbindung im Triax-System

Für die Verbindung zwischen Kamerakopf und CCU kamen früher Multicore-Kabel zum Einsatz, bei denen die Signale einzeln geführt werden und die entsprechend viele Adern enthalten. Das ist aus Kostengründen auch heute noch manchmal der Fall. Die Kabel sind aufwendig, groß und teuer und die Maximallänge ist auf ca. 300 m beschränkt. Um flexiblere Verbindungen zu erhalten, mit denen Entfernungen bis zu 2000 m überbrückt werden können, wurden die heute meist verwendeten Triax-Kabel entwickelt, bei denen alle Signale zwischen Kamerakopf und CCU mit nur einer, doppelt abgeschirmten, Leitung (Triax) übertragen werden können. Es werden dazu herstellerspezifische Zweiweg-Multiplex-System verwendet (Trägerfrequenzverfahren). Signale zum Kamerakopf hin (Versorgungsspannung, Intercom, Returnwege für Sucher und Teleprompter) werden z. B. im Bereich 0 ... 8 MHz und Signale von der Kamera zur CCU (RGB-Video, Audio, etc.) im Bereich 10 ... 60 MHz übertragen (Abb. 6.84 und 6.85). Bei anderen Systemen wird das Videosignal zum Teil auch als Komponentensignal im Frequenzmultiplex übertragen, z. B. mit einer Bandbreite von 10 MHz für das Luminanzsignal und von 6 MHz für die Farbdifferenzsignale. Moderne Systeme, insbesondere im HD-Bereich, ermöglichen auch die Nutzung von Glasfaser-Lichtwellenleitern anstelle der Triax-Leitung. Dies wird insbesondere dann erforderlich, wenn mit HD-Signalen der Qualität 1080p/50 gearbeitet werden soll

Jede Kamera erfordert ein eigenes Triaxsystem, eine eigene CCU und RCP. Dieser Gerätesatz wird zusammen mit dem Sucher, der Fernbedien- und Intercomausstattung etc. als Kamerazug bezeichnet. Bei vielen Systemen können die

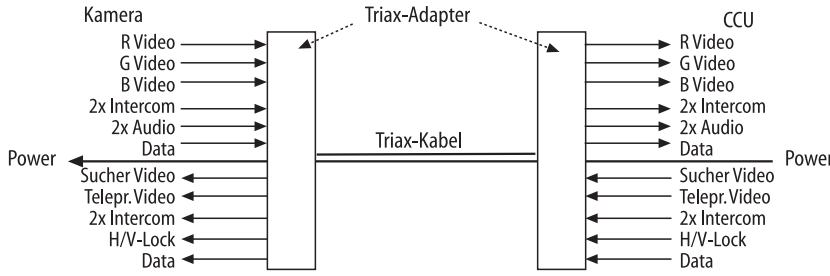


Abb. 6.85. Beispiel für Signalarten im Triax-System

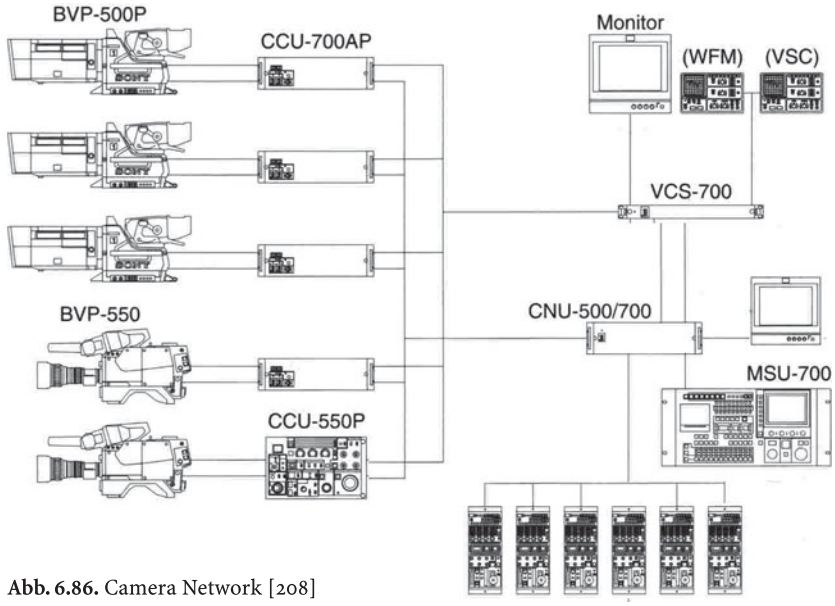


Abb. 6.86. Camera Network [208]

Kamerazüge zu Einheiten vernetzt werden. Das bietet den Vorteil, dass mittels einer Master System Unit (MSU) von einem übergeordneten Bildingenieursplatz aus auf die Parameter aller Kameras im Netz zugegriffen werden kann, was z. B. das Angleichen der Kameras vor der Produktion erheblich erleichtert. Die Steuersignalverteilung geschieht dann mit Hilfe der Camera Network Unit (CNU). Für die durch den Hebel ausgelöste Umschaltung der Videosignale, die zur Kontrolle zum Video- und Waveformmonitor gelangen, dient eine Videokreuzschiene (VCS) (Abb. 6.86).

Ein Beispiel für eine Röhren-Studiokamera ist die in den 80er Jahren verbreitete 3-Röhren Kamera von Bosch (Abb. 6.80). Heute werden vor allem CCD-Kameras eingesetzt. Ein Beispiel für eine früher gängige SD-Studiokamera ist der Typ LDK 10/20 von Thomson/Philips. Hier wird ein FT-CCD-Wandler mit DPM eingesetzt, der eine einfache Umschaltung zwischen den Formaten 4:3 und 16:9 ermöglicht. Ein zweites Beispiel ist die SD-Studiokamera BVP 500/550 von Sony. Hier können wahlweise IT- oder FIT-CCD eingesetzt werden.

Heute werden für den Studiobereich fast nur noch Multiformat-Kameras entwickelt, die sich z. T. auch auf den SD-Betrieb umstellen lassen und in der Regel die beiden HD-Modi 1080i/25 und 720p/50 an (bzw. 29,97./59,94 für US-Einsatz) bieten.

HD-Kameras für den Studiobetrieb unterschieden sich äußerlich und bezüglich des Umgangs sehr wenig von SD-Kameras. Bekannte Typen sind die LDK 8000 von Grass Valley und die Kamera HDC 1500/2500 von Sony, die beide auch 1080p/50 unterstützten [99]. Die HDC 2500 arbeitet mit einem 2/3“-FIT-CCD und hat bei 50 Hz eine Empfindlichkeit von Blende 11 bei 2000 Lux. Die LDK 8000 arbeitet mit FT-CCD, die Grass Valley-Nachfolgeserie mit der Bezeichnung LDX soll mit CMOS-Sensoren ausgestattet werden, die als

Abb. 6.87. HD-Studiokameras Grass Valley LDK 8000 [219] und Sony HDC 2500 [208]

FT-CMOS so gebaut sind, dass der Rolling-Shutter-Effekt nicht auftritt. Abbildung 6.87 zeigt den kompakten Aufbau der Kameras und in Abbildung 6.88 wird deutlich, mit welchen Elementen sie zur Studiokamera umgebaut werden kann. Die Umschaltfunktion zwischen den HD-Modi wird herstellerabhängig verschieden realisiert. Meist ist nur eine Variante optimal, für die andere wird kameraintern interpoliert [48]. Die HDC 1500 gewinnt 720p aus einem H/V-downskalierten 1080i-Signal, die LDK 8000 arbeitet vertikal mit 720 Zeilen unskaliert [219]. Für einen Kamerazug mit Studioobjektiv (80 000 Euro) können Kosten von ca. 150 000 Euro entstehen.

Eine Besonderheit stellen Varianten der Kameras dar, die, wie die Kamera LDK 8300 von Thomson, speziell für den Sportbereich entwickelt wurden. Sie sind so modifiziert, dass sie keine 25, sondern 75 Bilder pro Sekunde liefern. Wenn diese hohe zeitliche Auflösung wieder auf ein Drittel reduziert wird, entsteht eine hervorragende Zeitlupenwiedergabe, bei der die Bewegungsabläufe weich und nicht ruckend erscheinen (Super Slow Motion).

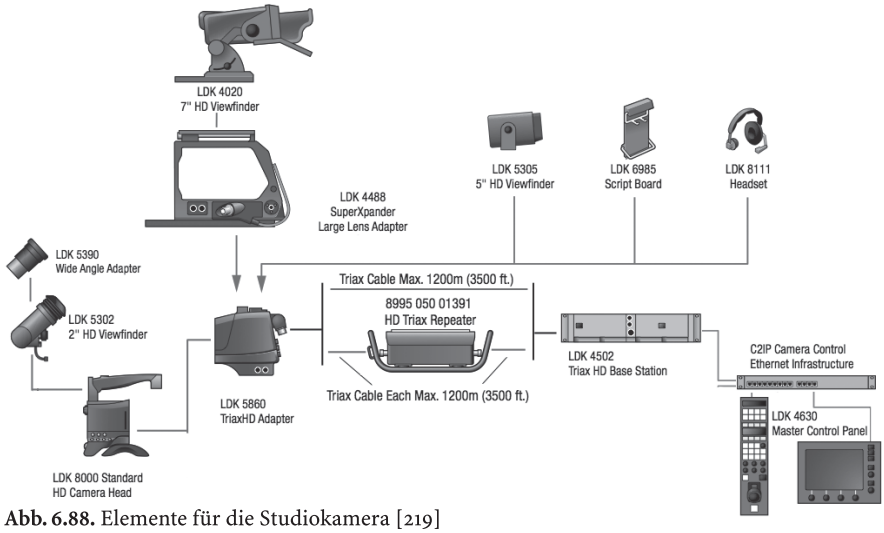


Abb. 6.88. Elemente für die Studiokamera [219]

6.3.8 Die EB/ EFP-Kamera

Während im Studio und bei großen Außenproduktionen meist mehrere Kameras verwendet werden, kommt bei der elektronischen Berichterstattung (EB) und bei kleineren elektronischen Außenproduktion (EAP, oder Electronic Field Production, EFP) meist nur eine Kamera zum Einsatz. Beispiele sind Industrie- und Werbefilme, Dokumentationen usw. Die Studiokameras bzw. ihre portablen Pendants können auch für die kleineren Produktionen verwendet werden, doch kommen aus Kostengründen oft einfachere Kameras zum Einsatz. Hier gibt es ein Überschneidungsgebiet mit den EB- und semiprofessionellen Kameras die immer leistungsfähiger werden. Im Vergleich zur elektronischen Berichterstattung werden bei der EFP meist höhere Anforderungen an Bildqualität und -gestaltung gestellt. Das Kamera-Set-Up weist eine große Spannweite auf. Die Kamera sollte leicht transportabel sein und viele Signalbeeinflussungsmöglichkeiten bieten. Um die Signalparameter gut kontrollieren zu können, ist es günstig, wenn die Kamera über eine Kamera-Kontrolleinheit (CCU) angeschlossen wird. Zur Farbbildkontrolle sollte ein Farbmonitor und ggf. ein Waveformmonitor/Vektorskop zur Verfügung stehen.

Die EFP-Kamera kann häufig als universelle Kamera eingesetzt werden. Sie ist für einfache Studioeinsätze tauglich und kann meist auch als EB-Camcorder verwendet werden. Für den Einsatz in der elektronischen Berichterstattung (EB) muss die Kamera mit einem Aufzeichnungssystem zum Camcorder, also mit direkt angeschlossenem Videorecorder, kombiniert werden. Früher kamen hier bandgestützte Recorder der Formate DigiBeta, IMX, DVCPro oder DVCAM (Abb. 6.89) zum Einsatz (s. Kap. 8). Heute wird in den meisten Fällen auf optische Platten und auf Festspeicher aufgezeichnet werden (XDCAM und P2 etc.). Die Qualität des Aufzeichnungsmediums sagt dabei nichts über die Qualität des Kamerateils aus. Wichtige Merkmale der Camcorder sind Handlichkeit und geringes Gewicht, Automatikfunktionen, hohe Lichtempfindlichkeit und geringe Leistungsaufnahme. Beim EB-Einsatz kann auf die Bildgestaltung weniger Rücksicht genommen werden als im Studio, denn in den meisten Fällen stehen nur zwei Personen als Bedienungspersonal zur Verfügung, eine für die Kamera und eine für den Ton. Bei der Kamerabedienung muss sich die Person also gleichermaßen um gestalterische wie technische Aspekte kümmern,

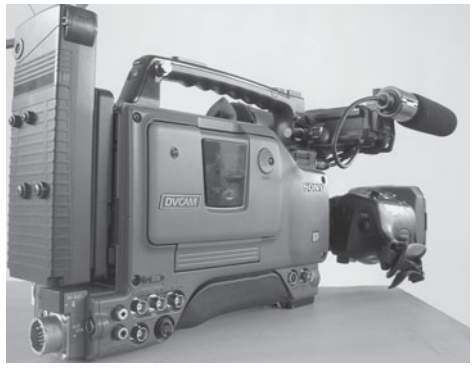


Abb. 6.89. EB-Digitalkamera [208]

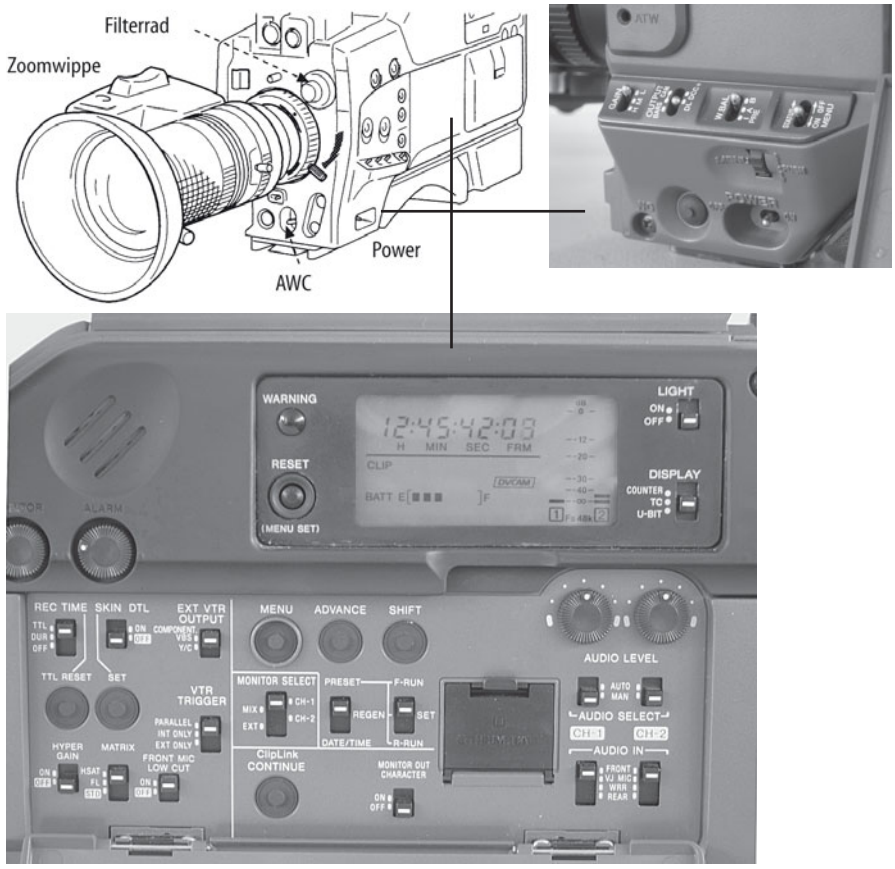


Abb. 6.90. Bedienelemente am EB-Camcorder DSR 570

d. h. neben der Bestimmung des Bildausschnitts, der Schärfe usw. die richtigen Voreinstellungen zu wählen und auch bei schnell wechselnden Aufnahmesituationen das Signal richtig auszusteuern.

Die Abbildungen 6.89 und 6.90 zeigen den Camcorder Sony DSR 570, ein nicht mehr aktuelles Modell, das sich aufgrund seiner großen und recht typischen Bedienelemente aber doch gut als Beispiel zur näheren Erläuterung eignet. Der Camcorder kann im EB-Einsatz ebenso arbeiten wie bei EFP mit abgesetztem Recorder oder als Studiokamera mit angeschlossener CCU. Die aktive Pixelanzahl beträgt 980/752 x 582, womit eine Horizontalauflösung von ca. 800 Linien erreicht wird. Die Empfindlichkeit wird mit F 11 bei 2000 Lux Beleuchtungsstärke angegeben. Die Kamera wird bei einer Leistungsaufnahme von insgesamt 24 W über NP1-Akkus mit einer Spannung von 12V versorgt. Das Gesamtgewicht incl. Akku, Objektiv und Cassette beträgt 6,4 kg.

Neben dem Objektivanschluss befinden sich im vorderen Teil des Kameragehäuses die Filterscheibe und unter dem Objektiv Taster für den automatischen Weißabgleich und zum Start des Recorders. Auf der linken Seite des

Kameragehäuses liegt unten der Einschalter. Die Schalterreihe darüber ermöglicht die Verstärkungszuschaltung, die Anwahl eines integrierten 100/75-Farbbalkengenerators (Bar), die Umschaltung der Weißabgleichs-Speicher und die Menü-Anwahl. Auf der rechten Seite des Kameragehäuses befinden sich das Cassettenfach und die Anschlüsse für den Videosignal-Ausgang, der Gen-Lock-Eingang sowie die Timecode-Buchsen.

Die Bedienelemente, die so wichtig sind, dass sie nicht im Software-Menü untergebracht werden, sind auf der linken Seite angeordnet. Hier befinden sich Einstellmöglichkeiten für den Timecode und die externe Signalausgabe, daneben Schalter für Bildparameter wie Skin Detail, Hyper Gain und Matrix.

Camcorder sind in jedem Falle auch mit einer Audioaufzeichnungsmöglichkeit ausgestattet. Für den EB-Einsatz sollte es hier möglich sein, mit professionellen Mikrofonen zu arbeiten. Das bedeutet, dass dreipolige XLR-Anschlüsse zur Verfügung stehen sollten sowie ein nachgeschalteter Verstärker, der für die geringen Mikrofonpegel ausgelegt ist. Für den Einsatz von Kondensatormikrofonen ist die Zuschaltmöglichkeit von 48 V-Phantomspeisung erforderlich. Meist stehen zwei Audiokanäle zur Verfügung, die mit zwei externen Buchsen an der Camcorder-Rückseite oder mit einer Buchse für das vorn aufgesteckte Mikrofon verbunden sind. Abbildung 6.90 unten zeigt rechts die Audiobedieneinheit mit zwei Pegelstellern. Die Pegelanzeige befindet sich im Display darüber. Unter den Stellern kann für beide Kanäle separat gewählt werden, ob das Signal manuell (über ein separates Mischpult) oder automatisch ausgesteuert werden soll. Darunter befinden sich die Zuordnungsschalter für die Audioquellen (Audio in) mit denen das Frontmikrofon z. B. auf beide Kanäle aufgezeichnet werden kann oder ein automatisch ausgesteuertes Mikrofonsignal auf einen Kanal und auf den zweiten ein über ein Mischpult gepegeltes Signal das dem Eingang mit Line-Pegel zugeführt wird.

Als Beispiel für einen HD-tauglichen EB-Camcorder lässt sich der Typ Sony PDW 700/800 anführen (Abb. 6.91). Er hat drei CCD-Sensoren mit je 1920 x 1080 Pixels. Die A/D-Umsetzung arbeitet mit 14 Bit und das Signal kann an zwei HD-SDI-Schnittstellen abgegriffen werden. Das integrierte Aufzeichnungsformat ist XDCAM HD422, das bedeutet, dass die Daten bei 50 Mbit/s mit 4:2:2-Farbsubsampling auf eine Professional Disc aufgezeichnet werden.

Abb. 6.91. HD-EB-Kamera [208]

6.3.9 Semiprofessionelle Kameras

Seit der Verfügbarkeit des DV-Aufzeichnungsverfahrens konnten sich hochwertige und doch recht preisgünstige 3-Chip-Camcorder als semiprofessionelle Geräte etablieren. Das einzige Manko war zunächst die mangelhafte Audiosignalzuführung. Mit der Nachrüstung von XLR-Buchsen und Signalzuordnungen wie bei EB-Camcordern konnten die Hersteller erfolgreiche semiprofessionelle SD-Geräte anbieten (z. B. Sony PD 170 oder Canon XL1).

Semiprofessionelle Camcorder haben eine fest integrierte Aufzeichnungseinheit, heute meist auf Basis von Festwertspeichern. Die Sensorgrößen sind relativ klein, i. d. R. werden 1/2"- oder 1/3"-Sensoren verwendet. Das ermöglicht den Einbau preisgünstiger Optiken, die meist fest mit der Kamera verbunden sind, nur wenige Modelle erlauben das Wechseln der Objektive.

Nach den SD-Geräten, die oft mit DVCAM oder DVCPro auf Band aufzeichneten, kam von Sony der Camcorder Z1 auf den Markt, der als Multiformatgerät auch HD-Signale, MPEG-2-codiert, auf HDV-Band aufzeichnete und anschließend der Sony Ex 1-Camcorder, der mit drei 1/2"-CMOS-Sensoren ausgestattet war. Dieser bot eine volle HD-Auflösung mit 2 Mio. Pixels, während die Kamera Z1 und viele Modelle anderer Hersteller Sensoren mit 1 Mio. Pixel verwendeten. Die Aufzeichnung erfolgt bei der EX 1 auf Festwertspeicherkarten mit MPEG-2 bei 35 Mbit/s und 4:2:0-Abtastung. Bei diesem Gerät war es bereits möglich, mit variabler Bildrate zu arbeiten.

Die Weiterentwicklung führte dann zu Geräten wie dem Camcorder Sony PMW 200 (Abb. 6.92 links), die auf der Kameraseite eine Ausstattung wie die EX1 aufweisen auf der Aufzeichnungsseite jedoch schon den Ansprüchen der Rundfunkanstalten genügen, da sie mit einem unskalierten HD-Bildformat bei einer Unterabtastung von 4:2:2 und 50 Mbit/s arbeiten. Diese Parameter sind von der EBU in ihrer Empfehlung R 132, trotz der Verwendung zeitabhängiger Bilder (Long GOP), als rundfunktauglich eingestuft [52, 112].

Die neueren Geräte haben SDI- und HDMI-Ausgänge meist schon als Standard, sodass neben einer externen Aufzeichnungseinheit, die ggf. mit geringerer Datenreduktion aufzeichnet und eine mögliche 10-Bit-Aufzeichnung nutzt, auch ein guter externer Suchermonitor einfach adaptiert werden kann. Dar-

Abb. 6.92. Semiprofessionelle Kamera und Amateurkamera

Tabelle 6.2. Parameter verschiedener semiprofessioneller HD- Kameras

Kamera	Sony EX1	Panas. HVX200	Sony PMW200	Canon XF105	JVC HM600
Bildwandler	3 x 1/2“ CMOS	3 x 1/3“ CCD	3 x 1/2“ CCD	1/3“ CMOS	3 x 1/3“ CM
Pixelzahl	1920 x 1080	960 x 540	1920 x 1080	2,07 MPix	2,2 MPix
Auflösungen	1080i/p, 720p	1080i/p, 720p	1080i /720p	1080i /720p	1080i /720p
HD-SDI out	ja	-	ja	ja	-
Genlock	-	-	ja	ja	-
Datenreduktion	MPEG-2/HDV	DVCProHD	XDCAM HD	MPPEG-2	MPEG/AVC
Datenrate	35/25 Mbit/s	100 Mbit/s	50 Mbit/s	50 Mbit/s	35/28 Mbit/s
Vari speed (fps)	1 ... 30/60 fps	1 ... 60 fps	1 ... 30/60 fps	1 ... 30/60 fps	-

über hinaus weisen einige Modelle auch eine Genlock-Funktion auf, sie können damit im Mehrkameraverbund betrieben werden. Technische Parameter bekannter Beispiele für weitere Camcorder sind in Tabelle 6.2 aufgeführt.

6.3.10 Amateurkameras

Amateurkameras sollen vor allem preiswert sein. Viele sind nur mit einem CMOS-Sensor ausgestattet, die Farbauszüge werden dann mit Hilfe von Mosaikfiltern gewonnen. Es kommen hier 1/3“- und 1/4“-Sensoren zum Einsatz, die kleine und damit preiswerte Objektive ermöglichen. Das Objektiv ist als Zoomobjektiv ausgeführt und kann nur bei teuren Typen gewechselt werden. Die Objektive sind gewöhnlich tief in das Gehäuse verlegt. Von den üblicherweise zur Bedienung vorhandenen Ringen steht oft nur noch einer zur Scharfstellung zur Verfügung, der eine elektronische Steuerung bewirkt. Die Blende kann manuell oft nur durch ein kleines Rädchen eingestellt werden und auch die Zoombedienung kann meist nicht manuell, sondern nur über die Zoomwippe erfolgen. Die Bedienung der Amateurgeräte ist für den professionellen Einsatz untauglich, oft wird auch beim optischen System gespart.

In den meisten Amateurkameras finden sich Shutter- und Automatikfunktionen. Der automatische Weißabgleich, die automatische Blende und das Autofokussystem können in vielen Fällen auf Knopfdruck abgeschaltet werden. Amateurkameras haben eine geringe Leistungsaufnahme von ca. 6 W und werden z. B. mit 6 V-Akkus versorgt. Sie sind leicht (weniger als 1 kg incl. Akku) und können auf einfachen Fotostativen befestigt werden. Ohne Stative sind sie nur schwer stabil zu halten (Abb. 6.92), daher enthalten hochwertige Typen Bildstabilisierungssysteme, die, wenn sie nicht als optisches Verfahren ausgeführt sind, mit einer Verringerung der effektiv nutzbaren Pixelzahl verbunden sind.

Ein Beispiel ist der Camcorder Sony HDR-CX 730 (Abb. 6.92 rechts). Er hat einen 1/2,88 CMOS-Sensor, ein Objektiv mit 10-fach optischem und 120-fach digitalem Zoom und einer Brennweite von 3,8 bis 38 mm. Aufgezeichnet wird auf Speicherkarten im AVCHD-Modus. Als Besonderheit erlaubt die Kamera auch Aufzeichnungen mit 50 HD-Vollbildern pro Sekunde. Weiterhin ist ein guter optischer Steady-Shot-Bildstabilisator inklusive Active- bzw. Anti-Roll-Modus integriert. Das Gewicht beträgt 625 g., die Maße 67,5 x 137,5 x 73,5 mm.

6.4 Digitale Cinematographie

Für dieses Kapitel wurde dieser Titel gewählt, weil hier die Verwendung digitaler Kameras als Ersatz für Filmkameras im Vordergrund steht. Angesichts der Tatsache, dass im Bereich der digitalen Filmbilder aus praktischen und ökonomischen Erwägungen bei Weitem die Verwendung der mit 2k aufgelösten Filmbilder dominiert, liegt es aufgrund der Nähe dieses Wertes zu den 1920 Bildpunkten für eine HDTV-Zeile nahe, daran zu denken, HDTV-Kameras für die direkte elektronische Akquisition von Filmbildern einzusetzen. Während aber eine HDTV-Kamera für den TV-Bereich im Wesentlichen nur hinsichtlich der höheren Auflösung modifiziert zu werden braucht, so gilt es beim Einsatz für den Filmbereich, mit den sehr spezialisierten Filmkameras zu konkurrieren. Bei der Geräteentwicklung sollten auch die Gegebenheiten, Einstellparameter und Handhabungsgewohnheiten berücksichtigt werden, die bei Filmkameras zu finden sind, da die Kameraleute allein durch technische Aspekte noch nicht von einem neuen Kameratyp zu überzeugen sind. Hier wird als Maßstab für die digitalen Kameras also zunächst die Filmkamera beschrieben. Es folgt die Darstellung der besonderen Anforderungen an digitale Kameras und anschließend werden real verfügbare Digital-Cinematography-Kameras vorgestellt.

6.4.1 Die Filmkamera

Filmkameras haben die Aufgabe, einen möglichst exakten intermittierenden Filmtransport zu gewährleisten und dabei den Film nur in seiner Ruhelage über eine definierte Zeit zu belichten. Bekannte Hersteller dieser mechanischen Präzisionsinstrumente sind Firmen wie Aaton, Arri (Arnold und Richter), Bolex, Moviecam, Panavision etc.

Der Filmantrieb erfolgt schrittweise mit einem Greiferwerk. Der Film wird von einer Vorratsrolle ab- und auf einem Wickelkern wieder aufgewickelt (Abb. 6.93). Die Filmvorratsrollen befinden sich gewöhnlich über- oder nebeneinander in einem Magazin (Kassette), das vom Kamerakörper getrennt werden kann, damit für einen Filmwechsel nicht die ganze Kamera in einen

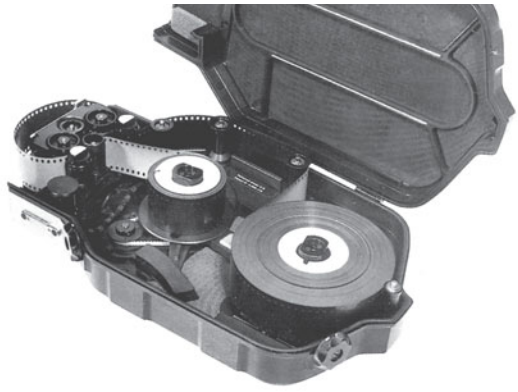


Abb. 6.93. Filmkassette

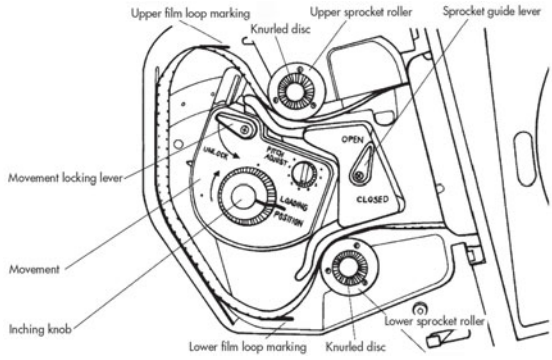


Abb. 6.94. Filmtransportmechanismus in der Kamera [9]

Dunkelsack gesteckt werden muss. Der Film wird beim Einlegevorgang in festgelegter Länge aus der Kassette heraus und in den Wickelkern geführt. Nach dem korrekten Einlegen und der Verbindung von Kamera und Kassette läuft er mit definierter Schlaufenlänge durch den Kamerakörper. In Abbildung 6.94 wird deutlich, wie der Film oben aus der Kassette austritt und innerhalb der Kamera geführt wird.

Im Betrieb wird der Film mit einer Zahnrolle kontinuierlich abgewickelt und dem Greiferwerk zugeführt. Ein darin enthaltener Doppel- bzw. Sperrgreifer hat die Aufgabe, den Film vor der Belichtung sehr genau in Position zu bringen. Zusätzlich wird in diesem Moment der Film durch eine Andruckplatte auch in der Tiefendimension fixiert. Nach der Belichtung werden die Fixierungselemente automatisch ausgekoppelt, und der Transportgreifer kann den Film um ein Feld weiterziehen. Um den Prozess möglichst geräuscharm realisieren zu können, sind sehr ausgeklügelte Transportsysteme erforderlich. Dabei muss eine Bildstandsgenauigkeit im Mikrometerbereich erreicht werden.

Das Greiferwerk ist direkt mit einer rotierenden Blende gekoppelt, die die Aufgabe hat, den Lichtweg während der Filmtransportphase abzudecken. Diese Umlaufblende hat einen Hellsektor, der nur bei sehr wenigen Geräten den Wert 180° übersteigt, jedoch aus unten genannten Gründen bei den meisten Kameras verringert werden kann.

Filmkameras können sehr einfach gebaut sein, wenn die Antriebsenergie durch ein Federwerk aufgebracht wird, kann die Kamera rein mechanisch

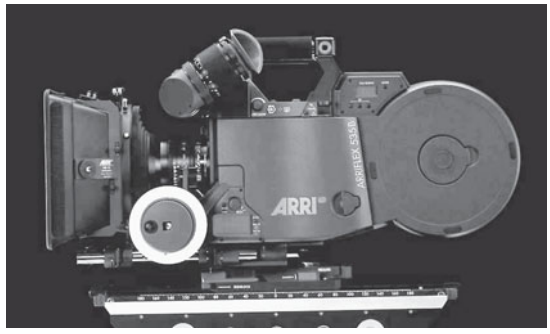


Abb. 6.95. Arri 35-mm-Kamera [9]

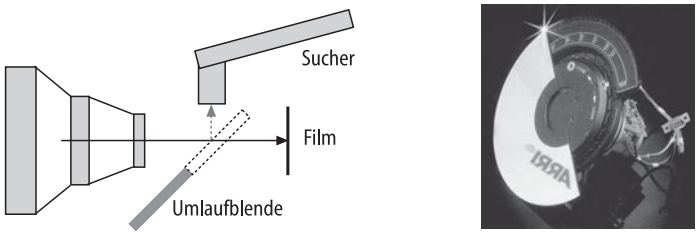


Abb. 6.96. Umlaufblende und Spiegelreflexsystem

betrieben werden. Auf der anderen Seite stehen moderne Kameras, die eine sehr aufwendige Kombination aus hochpräziser Mechanik und ausgefeilter Elektronik beinhalten. Ein Beispiel für die letztgenannte Art ist die hier etwas näher betrachtete Kamera Arriflex 535 der Firma Arri (Abb. 6.95). Dies ist eine 35-mm-Kamera, die als Studiokamera für den Stativbetrieb konzipiert ist [9]. Neben den dort üblicherweise verwendeten 300-m-Kassetten können auch eine 122-m-Kassette und ein Akku angebracht werden, sodass sie auch für den Schulterbetrieb geeignet ist. Diese Kamera wird elektronisch mit 24-V-Gleichspannungsmotoren betrieben, die quarzstabilisierte Bildwechselfrequenzen von 24; 25; 29,97 und 30 Bildern/s erzeugen. Darüber hinaus lässt sich die Geschwindigkeit auch stufenlos (3...50 B/s) und für Rückwärtslauf einstellen. Die variable Geschwindigkeit war einer der wesentlichen Vorteile von Filmkameras gegenüber vielen elektronischen HD-Videokameras.

Moderne Kameras sind so geräuscharm, dass sie kaum Störungen bei der Tonaufnahme verursachen. Die 16-mm-Kamera Arriflex SR3 erzeugt z. B. einen Geräuschpegel von nur 20 dB(A). Da die Geräuschisolation direkt in die Kamera integriert ist, spricht man von selbstgeblimpten Kameras, im Gegensatz zu „lauten“ Typen, die zur Minderung der Geräuschintensität mit einem schallabsorbierenden Gehäuse, dem sogenannten Blimp, umgeben werden müssen.

Die Arriflex-Kamera steht in der Tradition der Kameras, die durch Verwendung des bei Arri entwickelten Spiegelreflexsystems die optimale Kontrolle über das Bild im Sucher bieten (Abb. 6.96). Beim Spiegelreflexverfahren wird das durch das Objektiv fallende Licht in ein Suchersystem gelenkt, wenn der Lichtweg zum Film verschlossen ist, denn der Spiegel befindet sich auf der rotierenden Umlaufblende, die den Lichtweg während des Filmtransports unterbricht. Bei Verwendung eines 180°-Hellsektors wird somit das Licht in einer Zeitperiode je zur Hälfte auf den Film und auf eine Mattscheibe gelenkt, die wiederum durch eine Sucherlupe betrachtet wird. Vor oder auf der Mattscheibe befinden sich die Einzeichnungen für das jeweils verwendete Filmformat. Bei Kamerastillstand bleibt die Umlaufblende in einer Position stehen, über die das Bild in den Sucher gelangt. Bei älteren Kameras ist dies nicht automatisch der Fall, die Umlaufblende muss dann entsprechend nachgestellt werden.

Die Sucherlupe (engl.: Finder) zur Betrachtung der Mattscheibe kann vielfältige Formen haben, um einen flexiblen Einsatz zu ermöglichen. So kann z. B. bei Verwendung eines großen Magazins eine Lupenverlängerung angebracht werden, die bei teuren Kameras zudem in sehr viele Positionen auch quer zur

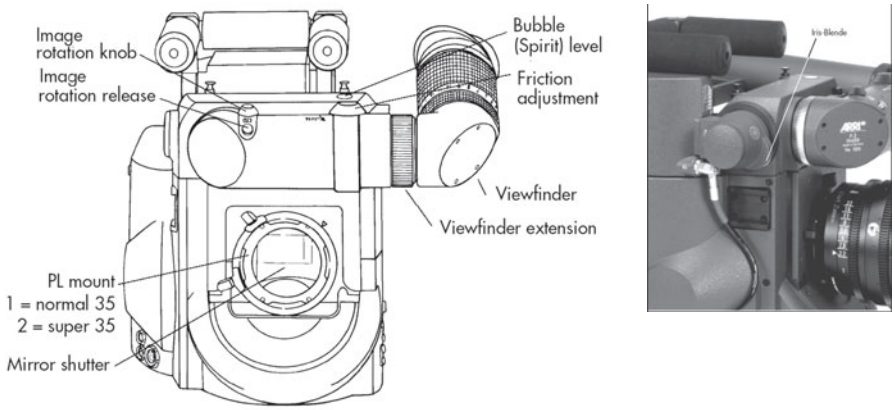


Abb. 6.97. Vorderansicht der Kamera und Seitenansicht mit Videoausspiegelung [9]

Kamera geschwenkt werden kann. Die Sucherlupen lassen sich z. T. auch auf ein horizontal entzerrtes Bild umschalten, zum Ausgleich der Verzerrung, die bei Verwendung eines anamorphotischen Objektivs entsteht. Es gibt auch beheizbare Lupen, die ein Beschlagen durch Feuchtigkeit verhindern.

Von dem zur Mattscheibe gespiegelten Licht wird oft ein Teil abgezweigt und zu einer angeschlossenen Videokamera gelenkt. Bei einer solchen Konstruktion spricht man von einer Videoausspiegelung, deren Bild also zusätzlich zum Suchersystem über einen angeschlossenen Videomonitor betrachtet werden kann (Abb. 6.97). Bei der Arriflex 535 kann festgelegt werden, ob 0%, 50% oder 90% der Lichtintensität die Videoausspiegelung erreichen. Letzterer Fall ist z. B. für eine fernbediente Kamera in Motion-Control-Systemen nützlich. Die Videobildqualität liegt allerdings nicht auf dem Niveau gewöhnlicher Videokameras, da Videosysteme in Europa immer mit einer Bildfrequenz von 25 Hz arbeiten, während sie bei der Filmkamera standardmäßig 24 Hz beträgt bzw. variiert werden kann, sodass eine Anpassung erforderlich wird, die die Bildqualität schmälert. Hinzu kommt das Problem, dass der Videobildwandler nur etwa während der halben Bildperiodendauer das Licht empfängt.

Wenn die Filmaufnahme nicht im Standardformat, sondern als Super 35 bzw. Super 16 gemacht werden soll, wird ein Bildbereich genutzt, der in Richtung der für die Tonspuren reservierten Fläche seitlich vergrößert ist. In diesem Fall muss das Zentrum des Objektivs ebenfalls ein kleines Stück seitlich verschoben werden. Moderne Kameras sind auf diesen Objektivversatz vorbereitet, ältere Kameras müssen umgebaut werden.

Die Belichtungszeit bei Filmkameras ist im Gegensatz zu Fotokameras nicht frei wählbar. Sie ergibt sich aus der Filmtransportgeschwindigkeit und der Größe des Hellsektors der Umlaufblende. Als Standardwerte gelten hier 24 Bilder/s und 180° Hellsektor, woraus eine Belichtungszeit von 1/48 s folgt. Die Filmgeschwindigkeit kann für Zeitraffer- und Zeitlupeneffekte variiert werden. Auch die Größe des Hellsektors kann verändert werden. Dieses geschieht einerseits, um mit kürzeren Belichtungsphasen entsprechend kürzere Bewegungsphasen

abzubilden, damit z. B. ein Ball, der sich in 1/48 s über die gesamte Bildbreite bewegt, nur an einer Stelle im Bild erscheint und damit deutlicher sichtbar wird (Shutter-Effekt).

Ein zweiter Grund für die Verringerung des Hellsektors besteht in der Notwendigkeit die Aufnahme an pulsierende Lichtquellen anzupassen, um störende Interferenzen und Helligkeitsschwankungen (Flicker-Effekt) zu vermeiden. Bestimmte Kunstlichtquellen, wie nicht flickerfreie HMI-Lampen, ändern ihre Intensität im Rhythmus der Netzfrequenz. Bei Nutzung der Standardbildfrequenz von 24 Frames per second (fps) ergibt sich damit sowohl bei der europäischen 50-Hz- als auch bei der amerikanischen 60-Hz-Netzfrequenz ein Interferenzproblem, da die Kamera diese quasi mit 24 Hz abtastet. Die entstehenden Interferenzerscheinungen erzeugen einen schwankenden Helligkeitseindruck und werden auch als Alias-Störungen bezeichnet. Das Problem wird nun gelöst, indem die Belichtungszeit mit Hilfe der Hellsektorgröße der Periodendauer der pulsierenden Quellen angepasst wird. Beim 50 Hz-Netz beträgt die Periodendauer 20 ms. Die Sektorenblende wird für die Anpassung so verkleinert, dass die Belichtungsdauer statt 1/48 s dann 1/50 s beträgt. Die dafür zu verwendenden Hellsektorgrößen, bezogen auf 360°, folgen aus der Multiplikation der Bildfrequenz mit der Belichtungsdauer. In diesem Fall gilt: 24 fps multipliziert mit 20 ms gleich 0,48, entsprechend 48 % von 360°, d. h. eine Hellsektorgröße von 172,8°. Bei einer Netzfrequenz von 60 Hz ist eine Hellsektorgröße von 144° zu verwenden. Hellsektoren für die wesentlichen Netzfrequenzen, also mit den Werten 180°, 172,8° und 144°, sind bei den meisten Kameras fest einstellbar, allerdings ist die verwendete Größe von außen oft nicht erkennbar. Sie sollte entsprechend deutlich an der Kamera vermerkt werden, da verringerte Hellsektorengrößen natürlich einen Einfluss auf die Lichtintensität haben, die den Film erreicht, und somit in die Berechnung der zu verwendenden Arbeitsblenden eingehen.

Im Zusammenhang mit der Änderung der Transportgeschwindigkeit ergibt sich ein besonderes Problem bei der Verwendung unterschiedlicher Bildfrequenzen, die natürlich ebenfalls einen Einfluss auf die Belichtung haben. Auch zum Ausgleich dieser Belichtungsänderungen kann der Hellsektor der Umlaufblende verwendet werden. Normalerweise wird die Hellsektorgröße vor der Aufnahme eingestellt, dann kann die Kamera nur mit einer festgelegten Geschwindigkeit arbeiten. Bei hoch entwickelten Kameras, wie der Arriflex 535, ist es aber auch möglich, die Hellsektorgröße während des Betriebs der Kamera zu ändern, sodass durch die kombinierte elektronische Steuerung von Filmgeschwindigkeit und Hellsektorgröße auch bei kontinuierlich veränderter Transportgeschwindigkeit eine definierte Belichtung beibehalten werden kann.

Moderne Kameras besitzen auch eine Einrichtung zur Aufbelichtung eines Zeitcodes (Timecode, s. Abschn. 9.3), der neben der Fußnummer eine wichtige Hilfe bei der Nachbearbeitung darstellt. Bei Arri-Kameras werden die zur Aufbelichtung des bei Arri eingesetzten sogenannten Arricodes verwendeten Leuchtdioden neben dem Bildfenster eingebaut. Die Timecodespur liegt daher dicht neben den Perforationslöchern, für die korrekte Platzierung in Längsrichtung sollte die vorgegebene Schlaufenlänge vor dem Bildfenster möglichst genau eingehalten werden. Um einen gut lesbaren Timecode zu gewährleisten,

muss die Empfindlichkeit des verwendeten Filmmaterials eingegeben werden. Eine alternative, besonders flexible TC-Form wird bei Kameras der Firma Aaton benutzt, die auch vom Kamerahersteller Panavision übernommen wurde. Beim Aaton-Zeitcode wird ein Informationsgehalt von 91 Bit in Form einer Matrix aus 7 x 13 Punkten geschrieben. Von Vorteil ist dabei die Tatsache, dass die Matrixform so gesteuert werden kann, dass sich Timecode-Daten oder direkt lesbare alphanumerische Zeichen ergeben.

6.4.2 Anforderungen an digitale Filmkameras

Die ideale elektronische HD-Kamera wird hier vom Standpunkt der Filmkameras aus betrachtet. Damit vor allem die Handhabung und Ergonomie der hoch entwickelten Filmkameras beibehalten werden kann, ist die naheliegendste Idee, eine Filmkamera mit einem Bildwandler anstelle des Filmmaterials auszustatten. Anstelle der Filmspule könnte die Kameraelektronik in das Gehäuse integriert werden, allerdings wäre dann kaum noch Platz für einen Recorder, der eine hoch qualitative Aufzeichnung, möglichst ohne Datenreduktion, gewährleisten kann. Bei dieser Konstruktion könnte also die Notwendigkeit einer externen Aufzeichnungseinheit zum Problem werden, die die Beweglichkeit der Kamera durch eine Kabelverbindung ggf. einschränkt.

Die Bildwandlung müsste bezüglich Zeit- und Ortsauflösung mindestens den Umständen der 2k-Abtastung von Filmbildern gerecht werden. Hinsichtlich der Zeitauflösung bedeutet dies, dass keine Videofrequenz, sondern die Bildfrequenz von 24 fps zugrunde liegen muss, die aber genau wie bei Filmkameras in möglichst weiten Bereichen veränderlich sein sollte, um Zeitlupen- und Zeitraffereffekte erzielen zu können, die nicht auf nachträglicher Bildinterpolation beruhen. In diesem Zusammenhang wäre weiterhin zu fordern, dass die Bildwandlung nicht auf dem Zeilensprungverfahren beruht, dass also eine progressive Abtastung verwendet wird.

Wenn die elektronische Kamera tatsächlich auf Basis der Filmkamera entstünde, würde über die dort enthaltene Umlaufblende bei einem 180°-Sektor sichergestellt, dass ein Einzelbild auch nur aus einer Belichtungsphase von 1/48 s entsteht, und zwar mit dem gleichen durch eine sin-Funktion geprägten Übergang zwischen Hell- und Dunkelphase. Unterschiede in der Zeitauflösung können die Wahrnehmung leicht irritieren. Und obwohl auch rein elektronische Kameras grundsätzlich in der Lage sind, die letzten Forderungen durch die progressive Abtastung mit 24 fps (24p) zu erfüllen, führt bereits die Verwendung des elektronischen Shutters anstelle der mechanischen Blende zu merklichen Differenzen. Die Umlaufblende bringt einen weiteren wesentlichen Vorteil, denn sie ermöglicht die Ausspiegelung eines optischen Sucherbildes, das an gewöhnlichen elektronischen Kameras oft vermisst wird, da dort die Abbildung immer über einen Suchermonitor betrachtet werden muss.

Bezüglich der Ortsauflösung wären mindestens 2048 x 1556 Bildpunkte zu fordern. Die 1920 Pixel pro Zeile, die bei HD-Video verwendet werden, kommen dieser Forderung bezüglich der Horizontalen weitgehend nahe, doch stehen bei HD-Video nur 1080 Zeilen zur Verfügung. Das ist bei Verwendung

von Standardbildformaten mit B/H = 1,66:1 oder 1,85:1 kaum ein Problem, doch wäre es für anamorphotische Aufnahmen wünschenswert, dass auch die Zeilenzahl für ein 4/3-Bildseitenverhältnis zur Verfügung stünde, also ca. 1500 Zeilen bei 2k. Im Hinblick auf die Vermeidung von Alias-Artefakten könnte zudem über eine unregelmäßige Anordnung der Fotorezeptoren nachgedacht werden.

Neben der Bildpunktanzahl auf dem Bildwandler ist auch dessen Größe von entscheidender Bedeutung, da die Bildgröße die Brennweite einer Normalabbildung bestimmt und die Brennweite neben der Blende wiederum der wichtigste Parameter für die Schärfentiefe ist (s. Abschn 6.2.1). Damit die Bilddiagonale mit der von 35-mm-Film vergleichbar wird, müsste die Kamera mit einem Bildsensor von ca. 24 mm Breite ausgestattet sein. Nur in diesem Falle ergibt sich, ohne die Verwendung weiterer Mittel, die gleiche Schärfentiefe wie bei der 35-mm-Filmkamera und die Möglichkeit, dieselben Objektive zu verwenden. Abbildung 6.47 zeigt die Auswirkung auf die Schärfentiefe bei Verwendung eines 2/3“-Wandlers anstelle eines Bildsensors mit 35-mm-Bilddiagonale.

Der Einsatz von hochwertigen Objektiven aus dem Filmbereich an der digitalen Filmkamera ist allerdings nicht nur mit der Bildwandlergröße verknüpft, sondern auch mit dem Abstand zwischen Objektiv und Bild. Bei einer typischen Videokamera ist dieser anders als bei Filmkameras, da vor dem eigentlichen Bildwandler der Prismenblock zur farbselektiven Strahlteilung angebracht werden muss, d. h., dass die Objektive für Film- und Videokameras unterschiedlich berechnet sind. Im Idealfall müsste auf den Strahlteiler verzichtet werden können, d. h. dass als Bildwandler ein einzelnes flaches Bauelement zum Einsatz käme. Dieses Ziel wäre erreichbar, wenn ein hoch auflösender Bildwandler mit 2000 Bildpunkten pro Zeile verwendet würde, die ihrerseits aus drei Subpixeln mit RGB-Farbfiltern bestehen, so wie es auch bei 1-Chip-Wandlern in Heimanwenderkameras realisiert wird. Bei der angestrebten Bildwandlergröße ist das Ziel erreichbar, allerdings ist sicherzustellen, dass die Filter vor den einzelnen Bildpunkten eine genügende Farbselektivität aufweisen.

Schließlich sollte der ideale Bildwandler auch eine Lichtempfindlichkeit und einen Kontrastumfang wie das Medium Film bieten. Moderne Videokameras erreichen heute bei der Belichtungszeit t = 1/25 s Vollaussteuerung bei Blende k = 11, ein Wert der zu einem filmäquivalenten Empfindlichkeitswert führt.

Ein viel größeres Problem als die Empfindlichkeit stellt der Kontrastumfang dar, da die Erzielung eines Belichtungsumfangs von über 10 Blenden auch beim Film nur dadurch möglich wird, dass eine Trennung in Negativmaterial mit flacher Gradationskurve und Positivfilm mit steiler Gradation vorgenommen wird. Bei elektronischen CCD- oder CMOS-Bildwandlern gibt es diese Trennung zunächst nicht, sie verhalten sich linear und bieten bei TV-Anwendungen mit einem Belichtungsumfang von sechs bis sieben Blenden einen ähnlich eingeschränkten Kontrastbereich wie Umkehrfilmmaterial.

Ausgehend von einer mittleren Belichtung bzw. Bezugsblende bei 18 % Remission ergibt sich zudem ein unsymmetrischer Belichtungsspielraum, der ohne den Einsatz einer Kniefunktion zu den Lichtern hin nur wenig mehr als zwei Blendenstufen aufweist. Bei gewöhnlichen Videosystemen ist daher besonders die Grauwertabstufung im Bereich der Lichter problematisch. Abbildung

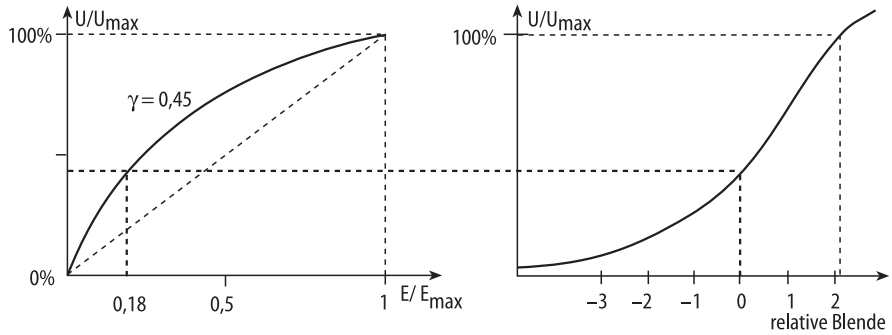


Abb. 6.98. Belichtungsspielräume bei Videokameras mit Gamma-Vorentzerrung

6.98 zeigt den Kennlinienverlauf der elektronischen Kamera mit Gamma-Vorentzerrung in linearer Skalierung und bezüglich der relativen Blende, der den eingeschränkten Spielraum für die Lichter deutlich macht.

In der idealen Kamera müsste ein speziell entwickelter Bildwandler zum Einsatz kommen, bei dem dafür gesorgt wird, dass die Ladungszunahme nicht proportional zur Belichtung abläuft, sondern in gleicher Weise, wie es für die Schwärzung bzw. Transparenz beim Filmnegativ gilt. Das elektronische Signal würde dann einem 10-Bit-log-Datensatz entsprechen und zur Betrachtung des elektronischen Bildes wäre ein spezieller Monitor erforderlich, der anstelle der Verwendung des Video-Gammas die Aufgabe hat das kontrastarme »elektronische Negativ« durch Kontrastanhebung zu entzerren. Durch den Einsatz von CMOS-Bildwandlern und der teillinearen Auslesung der Bildpunkte kann diese Funktion recht gut erreicht werden, des Weiteren auch mit sehr rauscharmen Kameras, die es erlauben, die Kennlinie in weitem Umfang elektronisch zu manipulieren [210]. In Analogie zum Filmnegativ sollten die Daten in möglichst unbearbeiteter Form, als Rohdaten (RAW) verfügbar gemacht werden.

6.4.3 Digitale Filmkameras

Die Entwicklung der digitalen Filmkamera ist eine der interessantesten Entwicklungen die in den letzten zehn Jahren im Videobereich stattgefunden hat. Sie wird hier in historischer Perspektive nachgezeichnet. Zu Beginn waren die meisten elektronischen HD-Kameras Weiterentwicklungen von Videokameras und daher in vielen Punkten weit von der idealen digitalen Filmkamera entfernt. Bei den ersten Versuchen im Rahmen der sogenannten elektronischen Cinematographie, also für Filmproduktionen auf Videokameras zurückzugreifen, wurde zunächst vor allem das Problem des Belichtungsbereichs bearbeitet, indem veränderliche Kniefunktionen und vielfältige Möglichkeiten zur Veränderung der Grauwertabstufungen geschaffen wurden. Damit wird eine erste Annäherung an die Gradationskurve von Film möglich, ohne allerdings den großen Belichtungsumfang von Film zu erreichen. Später kam die Anpassung der Bewegungsauflösung hinzu. Dazu gehört zum Ersten die Möglichkeit,

auch mit 24 Bildern pro Sekunde arbeiten zu können, und zum Zweiten die Notwendigkeit, dass jedes Bild aus nur einer Bewegungsphase stammt, die die gleiche Dauer hat wie das Filmbild. Das bedeutet, dass auf das Zeilensprungverfahren verzichtet und der Bildwandler progressiv ausgelesen werden muss, da das Zeilensprungverfahren gegenüber dem Film eine verdoppelte zeitliche Auflösung bewirkt. Die beiden Aspekte werden unter dem Schlagwort 24p zusammengefasst. Zur genaueren Erklärung muss außerdem hinzugefügt werden, dass die Belichtungsdauer bei 24 fps nicht 1/24 s betragen darf, wenn die Filmbedingungen erfüllt sein sollen, sondern i. d. R. nur die Hälfte, d. h. dass mit 1/48 s bzw. 1/50 s belichtet wird.

Das in diesem Zeitraum gewonnene Vollbild kann so aufbereitet werden, dass es mit 48 Hz bzw. 50 Hz ohne Großflächenflimmern wiedergegeben wird. Das Vollbild wird dazu wie beim Zeilensprungverfahren nachträglich in zwei Halbbilder aufgeteilt, die dann als segmented Frames (sf) bezeichnet werden, während zur Unterscheidung für Zeilensprungbilder, die bereits in diesem Modus aufgenommen wurden, die Bezeichnung interlaced (i) verwendet wird. Mit 24p-HD-Kameras und einer Belichtungsdauer von 1/48 s kann auf diese Weise also eine Bewegungsauflösung erreicht werden, die fast vollständig der von Filmkameras entspricht, die mit 24 fps und 180° Hellsektor arbeiten.

6.4.3.1 Digitale Filmkameras mit kleinem Sensor

Wie bereits dargestellt, entspricht die HD-Bildauflösung für die meisten Anforderungen weitgehend den Auflösungsverhältnissen bei 2k-Filmdaten, sodass insgesamt sowohl die Orts- als auch die Zeitauflösung von elektronischen HD-24p-Systemen als ausreichend angesehen werden kann. Entsprechend dieser Erkenntnis brachte die Firma Sony, verbunden mit dem Label CineAlta, als erste eine Kamera heraus, die dezidiert als digitale Filmkamera konzipiert war. Sie trägt die Bezeichnung HDW F900 (Abb. 6.99) und arbeitet mit drei 2/3"-FIT-CCD-Wandlern bei 24, 25 oder 30 fps im progressive oder interlaced-Mode mit einer festen Auflösung von 1920 x 1080 aktiven Bildpunkten. Eine Bildformatänderung geht hier auf Kosten der Auflösung. Das Gerät bietet vielfältige Einstellmöglichkeiten zur Veränderung der Bildparameter, dabei profitierte der Hersteller von den reichhaltigen Erfahrungen mit dem hoch entwickelten Digibeta-Camcorder der Reihe DVW 700.

Abb. 6.99. HD-Kamera Sony HDW F 900 [208]

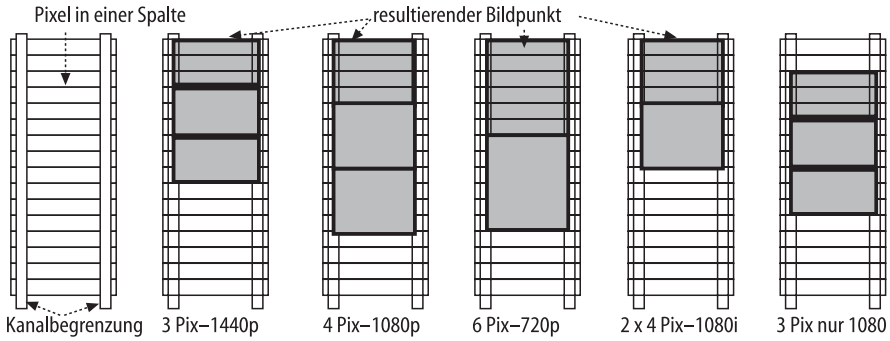


Abb. 6.100. Variable Zusammenfassung der Bildpunkte bei FT-CCD

Ein Vorteil der Kamera war der integrierte Recorder, mit dem das Gerät kabellos betrieben werden kann. Die Aufzeichnung erfolgt im HDCAM-Format auf eine 1/2“-Kassette aus der Betacam-Familie, die eine Laufzeit von ca. 50 Minuten bietet. Die Kompatibilität zu DigiBeta erlaubt jedoch nur eine Videodatenrate von 118 Mbit/s, die per Datenreduktion erreicht wird. Dazu wird zunächst durch Tiefpassfilterung die Horizontalauflösung von 1920 auf 1440 Bildpunkte herabgesetzt und anschließend eine DCT-basierte Reduktion um den Faktor 4,4 vorgenommen. Da das RGB-Signal der Kamera vor der Aufzeichnung noch in eine Komponentenform gewandelt und die Auflösung der Abtastwerte von 10 Bit auf 8 Bit reduziert wird, ist insbesondere die resultierende Farbauflösung für nachträgliche Bildmanipulationen, wie z. B. Stanzverfahren oder eine Farbkorrektur, als kritisch einzustufen. Die Kamera bietet jedoch neben analogen HD-Ausgängen optional auch direkte HD-SDI-Ausgänge, sodass der auflösungsbegrenzende Recorder umgangen und das Signal mittels einer Kabelverbindung auf ein abgesetztes Speichersystem aufgezeichnet werden kann.

Eine HD-Kamera, die trotz kleinem Sensor auch bereits früh einige Aspekte der idealen Kamera abdeckte, ist die Viper von Thomson (Abb. 6.101). Sie arbeitet mit drei 2/3-Zoll-CCD-Sensoren nach dem FT-Prinzip. Von Vorteil ist, dass bei FT-Chips das Dynamic Pixel Management genutzt werden kann, d. h. eine Umstellung der Bildformate durch veränderliche Bildpunktzuordnung. Dazu sind bei der Viper horizontal 1920 aktive Bildpunkte und vertikal 4320 (Sub)pixel verfügbar (Abb. 6.100), die in der Vertikalen auf verschiedene Arten

Abb. 6.101. Filmstyle-HD-Kamera Viper von Thomson [219]

Abb. 6.102. HD-Kameras AJ HDC 27V von Panasonic [164]

zusammengefasst werden können, sodass ohne Auflösungsverlust eine Vielzahl von Bildformaten bis hin zu Cinemascope bedient werden kann. So ergibt sich z. B. bei der Zusammenfassung von 4 Subpixeln zu einem Bildpunkt genau die HD-Auflösung 1920 x 1080, während bei 6 Subpixeln 720 Zeilen entstehen etc. Die Kamera lässt sich im 24p-Modus ebenso betreiben wie mit vielen Bildraten über 25 fps bis zu 30 fps mit und ohne Zeilensprungverfahren. Sie arbeitet mit 12-Bit-A/D-Umsetzung und intern mit 22 Bit. Damit sind vielfältige Möglichkeiten zur Anpassung der Kennlinie gegeben. Hier können die Daten im sogenannten Filmstream-Mode unbearbeitet (RAW-Modus) ausgegeben werden.

Die hinsichtlich der variablen Bildgeschwindigkeit als erste entwickelte Kamera ist der Camcorder AJ-HDC27V (Varicam) von Panasonic (Abb. 6.102). Er arbeitet stufenlos zwischen 4 fps und 33 fps und darüber hinaus mit 36, 40 und 60 fps. Mit der höchsten Bildfrequenz lässt sich gegenüber der Wiedergabe mit 24 fps eine 2,5-fache Zeitlupe erreichen. Bei der minimalen Bildrate ergibt sich ein 6-facher Zeitraffereffekt. Das Problem bei der variablen Bildrate bestand darin, dass die Aufzeichnung auf das Band eine feste Daten- bzw. Bildrate erfordert. Das Problem wird hier und auch bei anderen Kameras gelöst, indem die Aufzeichnung immer mit 60 fps erfolgt und bei Verwendung geringerer Bildfrequenzen neben den Originalbildern sogenannte Klone aufgezeichnet werden, die mittels Metadaten als solche gekennzeichnet werden. Trotz guter Farbreproduktion und filmähnlichen Gamma-Einstellungsmöglichkeiten ist dieser Camcorder nicht mit den vorher genannten Geräten vergleichbar, denn er arbeitet nur mit einer Bildauflösung von 1280 x 720 Bildpunkten und einer DV-Datenreduktion, die zu 100 Mbit/s führt. Er ist als elektronisches Pendant in der Klasse der 16-mm-Filmkameras konzipiert.

Für elektronische Kameras mit Strahlteilerprismen war zunächst die Verfügbarkeit von Prime Lenses etwas problematisch. Doch ist es, unter Lichtverlust, möglich, die gängigen Filmobjektive über einen optischen Adapter anzubringen und das Bild elektronisch wieder auf den Kopf zu stellen. Später wurden dann aber auch Prime Lenses für Strahlteilerkameras direkt entwickelt.

Das größte Problem im Bereich des Bildwandlers blieb der Kontrastbereich, der vor allem bei den Lichtern zu gering war. Bei einem mittleren Grauwert von 18%, der die Bezugsblende definiert, folgt, dass nach einer Vervielfachung, die 2 1/3 Blenden entspricht, bereits die Vollaussteuerung erreicht ist (Abb. 6.98). Dem Problem wurde mit der Kniefunktion begegnet, die für eine Abfla-

Abb. 6.103. Adaptionsoptik zur Erzielung einer filmähnlichen Schärfentiefe mit semiprofessionellen Kameras [163]

chung der Kennlinie bei Überschreitung von 80 % ... 100 % Signalpegel sorgt, sodass Gesamtpegel bis zu 600 % verarbeitet werden können (Abb. 6.72).

Für bildgestalterische Zwecke werden oft weitere Kennlinienänderungen benutzt. So kann z. B. neben dem Gesamt-Gamma mit dem Black Gamma eine Anpassung der Grauwerte allein im Schattenbereich vorgenommen werden. Damit die vielfältigen Grauwertänderungen und Kniefunktionen möglich werden, musste die elektronische Signalverarbeitung so weiter entwickelt werden, dass sie mit einer Quantisierung von mindestens 16 Bit arbeitet, auch wenn für die A/D-Wandlung z. B. nur 12 Bit verwendet werden.

Die bei Videokameras eingesetzte elektronische Anhebung der Konturen (Detailing), die eine visuelle Steigerung der Schärfewirkung hervorruft, wirkt bei der Umsetzung für das Kino oft künstlich. Das Detailing, das automatisch für Hauttonwerte herabgesetzt werden kann (Skin Tone Detailing), sollte daher ganz abschaltbar sein. Oft ist zusätzlich zu überlegen, ob die Annäherung an den Filmlook im Gegenteil nicht noch eine weitere Herabsetzung der Konturenschärfe durch außen angebrachte Black-Promist-Filter erforderlich macht.

In vielen Bereichen konnten sich diese Kameras so den Anforderungen an die ideale digitale Filmkamera nähern, doch unlösbar blieb das Problem der zu großen Schärfentiefe, solange mit direkter Abbildung auf 2/3"-Bildwandlern gearbeitet wurde. Die hierbei erzielte Schärfentiefe ist gerade mit der von 16-mm-Film vergleichbar. Um dem Problem der zu großen Schärfentiefe von Videosystemen zu begegnen, das umso gravierender wird, je kleiner der Bildwandler ist, wurde speziell für semiprofessionelle Camcorder, die mit sehr kleinen Bildwandlern arbeiteten (z. B. Canon XL1 und Sony DSR-PD 150), ein Kameraadapter entwickelt, der mit Mini 35 bezeichnet wurde (Abb. 6.103). Mit diesem System wird zunächst eine Zwischenabbildung auf eine Mattscheibe erzeugt, die mit den Abmessungen 24 mm x 18 mm den Dimensionen des 35-mm-Filmbildes entspricht. Das Zwischenbild wird danach wiederum auf den Bildwandler der Kamera abgebildet [163]. Auf diese Weise wird erreicht, dass die Objektive für 35-mm-Filmkameras an den Videocamcordern verwendet werden können und sich auch ein Schärfentiefeverhalten ergibt, das mit 35-mm-Film vergleichbar ist. Diese Technologie ist prinzipiell auch für 2/3"-HD-Kameras verwendbar, allerdings treten dann die Nachteile des Verfahrens noch

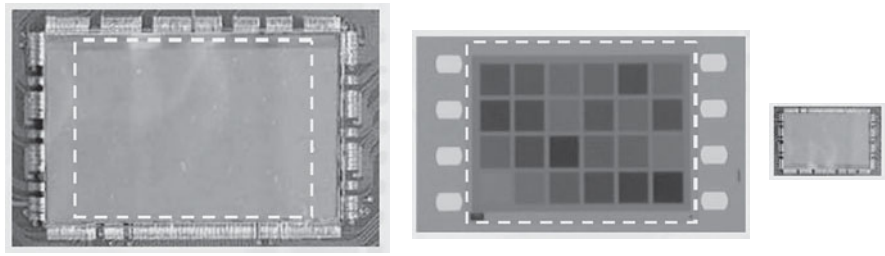


Abb. 6.104. Vergleich der Bildflächen von CMOS-Chip, 35 mm-Film und 2/3"-CCD-Wandler [9]

deutlicher hervor: Die Zwischenabbildung erzeugt einen Lichtverlust von etwa einer Blende bei ungleich verteilter Lichtintensität und macht vor allem das Bild unscharf, was zwar bei Standardvideoauflösung den Filmlook unterstützen kann, für den HD-Bereich aber ungeeignet ist. Bei kleinen Blendenwerten wird zudem die Struktur der Mattscheibe sichtbar, ein Problem, das durch die Bewegung der Scheibe gemindert wird.

6.4.3.2 Digitale Filmkameras mit großem Sensor

Zu Beginn des neuen Jahrhunderts wurde diese Enwicklung auf ein neues Niveau gehoben, da sich nun auch Hersteller von klassischen Filmkameras der digitalen Kamera zuwandten und dabei mit Sensorgrößen arbeiteten, die dem Super35-Format des Films entsprechen. Ein bekanntes Beispiel ist die Kamera Arri D21 (Abb. 6.105). Aufgrund der Tatsache, dass sich ein wichtiger Hersteller von Filmkameras der Konstruktion einer digitalen HD-Kamera widmete war gewährleistet, dass viele Akzeptanzprobleme, die im Bereich der Kameraumgebung und des Umgangs liegen, gelöst wurden. Die Basis der Konstruktion ist die Filmkamera Arriflex 435 [9]. Die Umlaufblende und der optische Sucher bleiben erhalten, anstelle der Filmspule und des Transportmechanismus befindet sich die Kameraelektronik. Der Kern des Gerätes ist ein Bildwandler der mit einer aktiven Fläche von 24 mm x 18 mm die gleiche Größe aufweist wie ein 35 mm-Filmbild (Abb. 6.104). Der Bildwandler arbeitet mit CMOS-Tech-

Abb. 6.105. Früh verfügbare Kamera für digitale Cinematographie [9]

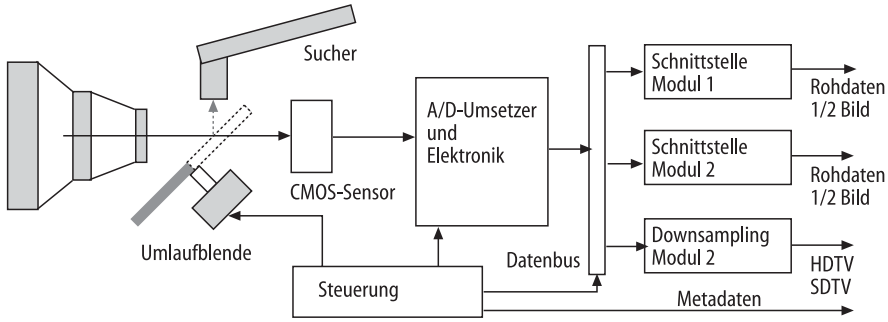


Abb. 6.106. Blockschaltbild der Kamera Arri D21

nologie und erlaubt einen Belichtungsumfang von ca. 10 Blenden. Es wird ein einziger, flacher CMOS-Chip verwendet und zur Gewinnung der drei Farbauszüge auf die Strahlteilung per Prisma verzichtet. Statt dessen wird jedem der ca. 6 Mio. Bildpunkten ein Farbfilter nach dem Bayer Pattern zugeordnet (s. Abschn. 6.1.4) und durch eine aufwendige elektronische Interpolation mittels einer sogenannten 3D-LUT (Look up table) wird für eine gute Farbwiedergabe gesorgt. Mit diesen Maßnahmen ergibt sich dasselbe Schärfentiefeverhalten wie bei 35 mm-Film. Das gesamte Zubehör der Arri-Kameras, inklusive der hochwertigen Objektive kann problemlos verwendet werden und bezüglich der Bedienung zeigt sich ein Gefühl wie bei Filmkameras.

Die Kamera ist auch für variable Bildraten ausgelegt. Die interne Verarbeitung erlaubt bis zu 150 fps, die nach außen darstellbare Bildrate hängt von den verwendeten Interface-Modulen ab. Intern folgt daraus eine Datenrate von bis zu 10 Gbit/s, die durch die Verwendung von 32 parallel arbeitenden 12-Bit-A/D-Umsetzern erreicht wird. Die Bilddaten des Sensors werden direkt digitalisiert, die gesamte weitere Verarbeitung findet auf der digitalen Ebene statt.

Aufgrund der Verwendung paralleler Ausgangsmodule ist es einerseits möglich, im Live Modus eine Betriebsart bereitzustellen, die der bei gewöhnlichen Videokameras ähnelt, nämlich, dass eine Echtzeitverarbeitung bezüglich der Farbe etc. durchgeführt und an einem der HD-SDI-Ausgänge ein Standard-HD-Signal ausgegeben wird (Abb. 6.106).

Als zweite Möglichkeit können die Daten auch unbeeinflusst, im sogenannten RAW-Modus, gespeichert bzw. über zwei parallele HD-SDI-Schnittstellen ausgegeben werden wobei eine höhere Bildauflösung als HD-CIF möglich ist. Da noch keine separierten RGB-Farbauszüge vorliegen ist in diesem Falle allerdings ein separater off-line Prozess erforderlich, um das korrekt kalibrierte und farbkorrigierte Bild zu erhalten. Diese Option ähnelt dem Vorgang bei der gewöhnlichen Filmproduktion wo das sichtbare Bild auch erst nach der Entwicklung und einer farbangepassten Positivkopie sichtbar wird, dafür aber ein viel größerer Spielraum für den Umgang mit den Bildparametern zur Verfügung steht. Im Live-Modus muss dagegen bereits am Drehort für die optimale Anpassung von Belichtung und Farbe gesorgt werden. Diese damals anspruchsvollen Parameter der Kamera waren dann das Orientierungsmuster für andere.

Im Laufe der weiteren Entwicklung gab es dann im Sektor der digitalen Filmkameras, nicht zuletzt durch die Einführung der Kamera Red One, einen erheblichen Preisdruck. Dieser Umstand veranlasste auch die Firma Arri dazu, preisgünstigere Geräte anzubieten. So wurde ab 2010 die Kamera Arri Alexa verfügbar (Abb. 6.107). Sie zeichnet sich durch einen sehr hochwertigen CMOS-Sensor im 35-mm-Format aus, der eine Lichtempfindlichkeit von 800 ASA aufweist. Wie bei der Arri D21 können also die bekannten 35mm-Objektive aus dem Filmbereich verwendet werden. Das Rauschen des Bildwandlers ist sehr gering, sodass als Belichtungsumfang mehr als 13 Blenden angegeben wird.

Die Kamera basiert auf der D21 und ist in drei Ausbaustufen erhältlich. In der unteren Stufe steht aus Kostengründen kein optischer Sucher zur Verfügung, allerdings ist die Ausnutzung der Sensorfläche so gestaltet, dass im Viewfinder, der eine Auflösung von 1280 x 720 Pixels bietet, ein größerer Bildausschnitt gezeigt wird als es dem ausgegebenen Signal entspricht. So können einige Vorteile eines optischen Suchers erhalten bleiben. Die Kamera arbeitet mit variabler Bildrate zwischen 0,75 und 60 fps. Die Aufzeichnung kann bei einer Auflösung von 3072 x 1728 Bildpunkten als Rohdaten im Format Arri-RAW mit einem Abtastratenverhältnis von 4:4:4 erfolgen. Mit Arri T-Link definiert der Hersteller für die damit verbundene hohe Datenrate ein Schnittstellenformat, das von Herstellern portabler Recorder, wie der Firma Codex, unterstützt wird. Unter Nutzung einer Auflösung von 2880 x 1620 Pixels kann hier das Signal via T-link aufgezeichnet werden. Daneben kann es auch als HD-SDI-Signal ausgegeben werden. Zusätzlich steht eine kamerainterne Aufzeichnung auf Flash-Cards zur Verfügung. Arri nutzt dafür den ProRes-Codec von Apple der eine 4:2:2- oder 4:4:4-Abtastung unterstützt, sodass auch auf diese Weise eine recht gute 2k-Bildqualität und ein einfacher Workflow in der Postproduktion, z. B. mit dem Schnittprogramm Apple Final Cut Pro sichergestellt werden kann. In der höchsten Ausbaustufe, als Alexa Studio, ist die Kamera mit rotierender Umlaufblende und optischem Sucher ausgestattet. Aufgrund der guten Bildqualität und Handhabung hat sich die Arri Alexa als eine der gefragtesten digitalen Filmkameras etabliert [131].

Abb. 6.107. Digitale Filmkamera Arri Alexa [9]

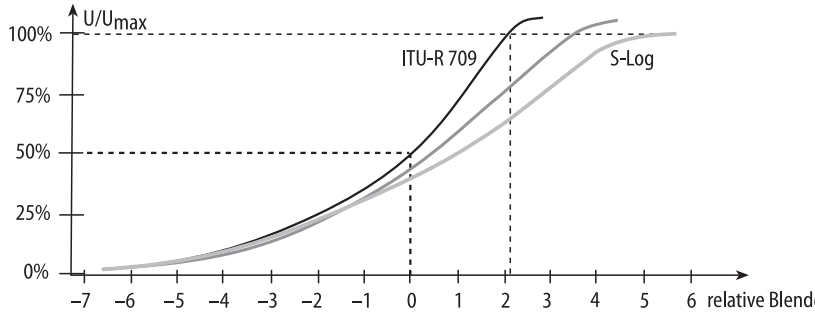


Abb. 6.108. Kennlinienveränderungen bei der HD-Kamera Sony F23 [41]

Eine bereits früh verfügbare Kamera, die für die Nutzung gewöhnlicher Filmobjektive und damit mit nur einem großem Bildwandler ausgestattet wurde, ist die digitale Filmkamera Panavision Genesis. Hier wird ein einzelner CCD-Sensor benutzt, der jedoch zur Gewinnung der Farbauszüge kein Bayer- sondern ein Stripe-Pattern aufweist. Er hat horizontal 3 x 1920 und vertikal 1080 Bildpunkte. Sie werden zu 1920 x 1080 Farbtripeln zusammengefasst, wobei die nebeneinander liegenden Pixel mit RGB-Farbfiltern versehen werden. Das Konzept hat den Vorteil, dass die RGB-Daten nicht mittels De-Bayering interpoliert werden müssen.

Ein ähnliches Konzept und eine Weiterentwicklung des Stripe-Pattern-Sensors wird bei der Kamera F35 von Sony eingesetzt (Abb. 6.109). Die Digitalisierung der Signale erfolgt bei der F35 mit 14 Bit und der Hersteller gibt an, dass ein Belichtungsumfang von 12 Blenden erreicht wird. Wie bei vielen anderen digitalen Kameras bestehen auch bei der F35 umfangreiche Möglichkeiten zur Beeinflussung der Kennlinie. Abbildung 6.108 zeigt die für alle digitalen Filmkameras typischen Kennlinien, die extrem flach gewählt werden können, um wie beim Negativfilm einen großen Belichtungsumfang zu erreichen (hier deshalb bei Sony S-Log genannt). Das Bild wird dann entsprechend auch sehr kontrastarm und muss beim Colour Grading korrigiert werden.

Abb. 6.109. Filmstyle-HD-Kamera Sony F23 / F35 [208]

Die Verwendung der vollen Bildauflösung für alle drei Farbkanäle soll zu einem Farbraum führen, der die ITU-Empfehlung 709 übersteigt. Die Kamera ist so konzipiert, dass neben HD-Komponentensignalen auch 4:4:4-RGB-HD-Signale aufgezeichnet werden können. Die Bildwandler arbeiten bis zu einer Bildrate von 60 fps, sodass, in z. B. Verbindung mit dem Recorder SRW 1, auch eine variable Bildrate zwischen 1 fps und 60 fps realisiert werden kann. Das Gehäuse ist kompakt und widerstandsfähig konstruiert. Der portable HDCAM-SR-Recorder Sony SRW 1 kann direkt am Gehäuse befestigt werden. Mit Zubehör kommen dann ca. 20 kg zusammen. Nachteilig ist die hohe Leistungsaufnahme von 56 W [208]. Als Pendant gibt es die Kamera Sony F23. Das Gehäuse ist das gleiche wie bei der Sony F35, die F23 sieht daher gleich aus (Abb. 6.109) und ist auch bei der Signalverarbeitung ähnlich. Der Unterschied zur F35 ist vor allem die Verwendung eines Strahlteilers und von drei 2/3"-CCD-Bildwandlern mit jeweils 1920 x 1080 Pixels.

6.4.3.3 Digitale Filmkameras mit 4k- Sensor

Die Entwicklung wurde bereits früh durch einen Filmkamera-fremden Hersteller namens Red Digital beeinflusst. Dieser hatte sich das Ziel gesetzt, eine sehr kompakte digitale Filmkamera zu einem sehr günstigen Preis anzubieten. Das erste Produkt ist die Kamera Red One (Abb. 6.110). Wie die Arri-Kameras hat auch sie einen Einzelbildwandler von der Größe des 35-mm-Filmbildes, allerdings kann sie nicht die Vorteile bieten, die durch einen optischen Sucher gewonnen werden. Dafür gehört sie zu den ersten Kameras die eine 4k-Auflösung boten.

Es wird ein CMOS-Sensor mit 4520 x 2540 aktiven Bildpunkten verwendet, der einen Signal-Rauschabstand von 66 dB aufweist. Die Bildfläche ist mit Farbfiltern nach dem Bayer-Pattern bedeckt und kann so eingerichtet werden, dass Bilder mit folgenden Auflösungen erzeugt werden: 4k+ (volle Auflösung), 4k, 2k, 1080p/i und 720p. Es wird eine variable Bildrate zwischen 1 fps und 60 fps erreicht, im sogenannten Windowed-Mode, bei 2 k und 720p, bis zu 120 fps. Die Verwendung der 2k-Auflösung ist mit der Bildung eines Bildausschnitts verbunden, sodass der Abbildungsmaßstab und ggf. auch die Schärfentiefe gegenüber dem 4k-Modus verändert ist.

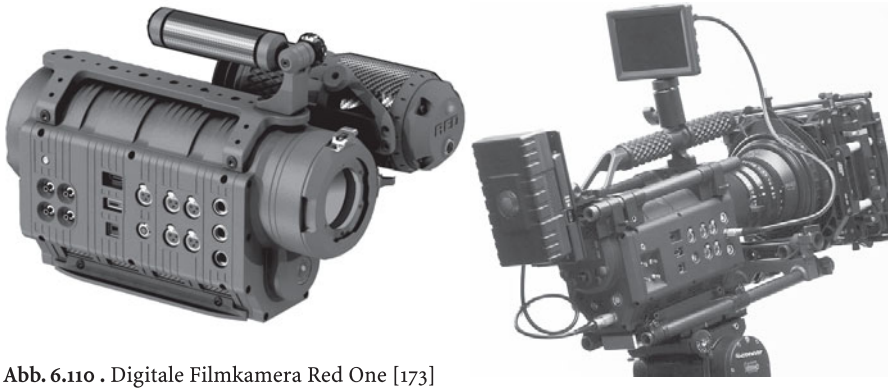


Abb. 6.110 . Digitale Filmkamera Red One [173]

Die Daten können im RAW-Modus auf Harddisk-Speicher oder Flashkarten-Speicher aufgezeichnet werden. Hier kommt dann unter dem Namen Redcode eine Datenreduktion mit variabler Bitrate zum Einsatz, die wie JPEG2000, auf der Wavelet-Transformation beruht. Damit werden im RedCode-RAW-Codec 4k-Daten bei 24 fps auf 27,5 MB/s bzw. 220 Mbit/s reduziert, wobei keine Artefakte sichtbar sein sollen (visually lossless) [173]. Redcode-RAW bedeutet auch hier, dass die Daten nicht direkt betrachtet werden können. Das Decoding, De-Bayering und die Umrechnung der RAW-Daten kann extern, z. B. mit Editing-Programmen vorgenommen werden.

Die Nachfolgeformate bei Red heißen Scarlet und Epic. Erstere ist als Hybridkamera für den Foto- und Filmmarkt ausgelegt. Die digitale Filmkamera Red Epic hat einen 5k-CMOS-Sensor mit einer Auflösung 5120 x 2700 Pixel. Die Sensorfläche entspricht mit 27,7 mm x 14,6 mm etwa dem S35-mm-Filmformat. Wie bei der Red One ergibt sich die gleiche Schärfentiefe aber nur beim 4k-Betrieb. Aufgrund der großen Vertikalen können Aufnahmen im Format 2:1 bzw. anamorphotische Aufnahmen realisiert werden. Der Rauschabstand wird mit 66 dB angegeben, damit soll ein Belichtungsumfang von 13 Blenden erreicht werden. Es werden hohe Frameraten möglich: Bei 4k beträgt die maximale Bildrate 150 fps, bei 2k-Auflösung 300 fps. Abbildung 6.111 zeigt den kompakten und modularen Aufbau, der es besonders leicht macht, den geringen Abstand der optischen Achsen der beiden Kameras zu erbringen, der für 3D-Aufnahmen erforderlich ist. Ein passendes 3D-Rig ist daneben sichtbar.

Auf den Erfolg der Red One reagierte die Firma Sony mit einer 4k-Kamera mit der Bezeichnung F65. Hier kommt ein 8K-CMOS-Sensor im S35-mm-Format zum Einsatz. Dieser erlaubt es, die für die Luminanz wichtige Grünkomponente für die 4K-Ausgabe ohne Interpolation zu gewinnen. Er ermöglicht es andererseits, zukünftig mit gewöhnlichem Debayering auch eine 8k-Ausgabe zu realisieren. Der Farbraum soll noch größer sein als bei der Sony F35, der Belichtungsumfang wird mit 14 Blenden angegeben und die Daten können als 16-Bit-RAW ausgegeben werden. Die Bildrate beträgt maximal 120 fps.

Optional kann die F65 mit einer rotierenden Blende ausgestattet werden, um den Rolling Shutter-Effekt des CMOS-Sensors zu eliminieren. Da die F65

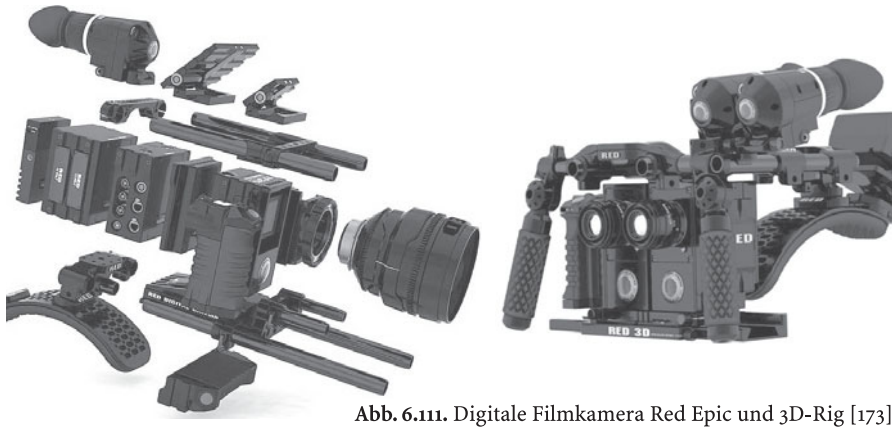


Abb. 6.111. Digitale Filmkamera Red Epic und 3D-Rig [173]

kleiner und leichter als die F35 ist, werden Anwendungen wie der Einsatz einer Steadycam erleichtert (Abb. 6.112). Aufzeichnungen können mit dem andockbaren SR-Master-Recorder SR-R4 ausgeführt werden. Die als Speichermedium dazu gehörigen SR-Memory-Karten garantieren einen Datendurchsatz von 5,5 GBit/s und erlauben die Aufzeichnung von 16Bit-RAW-Daten.

Auf die F65 ließ Sony schnell das 4k-Kameraduo F55/F5 folgen. Sie sind noch einmal wesentlich kompakter (Abb. 6.113 links). Der S35-mm-Bildsensor ermöglicht einer Auflösung von 4096 x 2160 Pixeln. Der Belichtungsumfang wird mit 14 Blenden angegeben. Beide Varianten der Kamera können 2k- und 4k-Daten ausgeben. Sie unterstützen eine interne Aufzeichnung auf Festwertspeicherkarten mit mehreren Codecs wie das XAVC-Format, den HDCAM-SR-Codec (MPEG4, SStP) sowie XDCAM422 mit 50 Mbit/s. Dabei können mit der F55, mit dem Codec XAVC, HD-Bildraten bis 180 Bilder/s erreicht werden, bei 4k beträgt die maximale Framerate 60 fps. XAVC bezeichnet bei Sony die professionelle Intraframe-Variante des H.264-Codecs, der verschiedene Auflösungen, auch 4k-Daten bei 4:4:4-Abtastung bis zu 12 Bit unterstützt. Zusätzlich zu der internen kann auch eine externe Aufnahme mit dem RAW-Recorder AXS-R5 realisiert werden. Dabei erreicht die F55 im 2k-Modus 240 Bilder/s. Für die Beurteilung des Bildes erlaubt die Kamera mittels mehrerer 3G-SDI-Anschlüsse die Ausgabe von 4k-Bildern bei 60fps [208].

Mit der FS700 bietet Sony auch eine 4k-Kamera im Preisbereich der semiprofessionellen Kameras (Abb. 6.113 rechts). Der CMOS-Chip dieser Kamera ist mit insgesamt 4352 x 2662 Bildpunkten ausgestattet. Diese befinden sich auf einer Sensorfläche von 23.6 mm x 13.3 mm, was in etwa den Abmessungen eines Super-35-Bildfensters im Filmbereich entspricht. Hier ergibt sich zwar die gewünschte, filmähnliche Schärfentiefe aber die interne Aufzeichnung arbeitet mit starker Kompression und liefert mit 8 Bit nur einen relativ geringen Dynamikumfang. Die Besonderheit der Kamera ist die Slow-Motion-Funktion, mit der Aufnahmen bis zu 240 fps in 1080p realisierbar sind. Bei reduzierter Auflösung kann sogar mit bis zu 960 fps aufgenommen werden

Mit einem Firmware-Upgrade soll die FS700 auch 4k-Bitstream-Daten über seine 3G-HD-SDI-Schnittstelle ausgeben können. Über einen Adapter können dieses Daten dann mit dem Recorder AXS-R5 aufgezeichnet werden.

Abb. 6.112. Digitale Filmkamera Sony F65 mit optionaler Umlaufblende (rechts) [208]

Durch die Entwicklung der Digitalelektronik, die immer kleinere Geräte ermöglichte, gelang es im Laufe der Zeit, eine Videofunktionalität auch in Fotokameras einzubauen. Bei den hochwertigen Typen, die in Anlehnung an die Spiegelreflextechnik als DSLR (Digital Single Lens Reflex) bezeichnet werden, ist die Signalverarbeitung auf HD-Niveau, sodass diese Geräte auch für den Bereich der digitalen Cinematographie verwendet werden. Ursache für diesen Trend ist der günstige Preis sowie die Tatsache, dass die Qualität der Abbildung im Fotobereich traditionellerweise sehr hoch ist. Wenn die Bildwandler in den digitalen Kameras die gleiche Größe aufweisen wie das Kleinbildformat, können auch die dafür vorhandenen Objektive verwendet werden, die in großer Zahl und guter Qualität zur Verfügung stehen. Die Bilder haben dann entsprechend auch eine große selektive Schärfe bzw. geringe Schärfentiefe. Diese ist beim Kleinbildformat, bei 36 mm Breite, noch extremer als bei 35 mm-Film mit 24 mm und damit nicht der idealen Filmkamera entsprechend.

Zunächst erscheint der DSLR-Einsatz sehr günstig, er ist aber doch mit einigen Nachteilen verbunden. So ist zu bedenken, dass Adaptionsmechaniken zu beschaffen sind, wenn professionelles Filmequipment und ggf. Follow-Focus etc. benutzt werden sollen. Die Videotauglichkeit wird durch die Verfügbarkeit eines Live-Bildes im Sucher unterstützt, die Bildfrequenz ist jedoch oft auf 30 fps fixiert. Es werden meist 1080p- oder 720p-Signale auf Speicherkarten aufgezeichnet und zwar mit Datenreduktion. Auch beim Einsatz von H.264 ist diesbezüglich jedoch z. B. eine 8-Bit-Datenrate von 40 Mbit/s (Canon EOS 5D) als kritisch einzustufen. Abhängig vom Kameratyp und der Implementierung wäre es aber denkbar, ein ausgegebenes HDMI-Signal unkomprimiert aufzuzeichnen. Zu bemängeln ist die Tatsache, dass die Tonaufzeichnung bei Fotokameras meist in keiner Weise professionellen Ansprüchen genügt. Kritisch ist weiterhin, dass im Bild räumliches Aliasing auftreten kann, da die Tiefpassfilter, die dieses verhindern, für die höhere Fotoauflösung von z. B. 20 Mpixels ausgelegt sind und nicht für HD-Videobilder mit 2 Mpixels.

Bezüglich der Häufigkeit der Verwendung von DSLR-Kameras sticht besonders die Kamera EOS 5D der Firma Canon hervor, die aufgrund ihres guten Bildprozessors trotz der Wortbreite von nur 8 Bit Bilder von guter Qualität ermöglicht. Diese Kamera wurde oft für Filmaufnahmen verwendet.

Abb. 6.113. Digitale Filmkameras Sony F55 und FS700 (rechts) [208]

Angesichts dieses Erfolgs entschied sich Canon, Kameras zu entwickeln, die direkt für den Bereich der digitalen Cinematographie ausgelegt sind und damit die oben genannten Nachteile vermeiden.

So entstand die Kamerareihe EOS C100, C300, C500. Die C500 ist als 4k-Kamera ausgelegt (Abb. 6.114). Der Foto-Objektivanschluss, der Canon EF-Mount, wird beibehalten, das Gerät kann aber auch mit PL-Mount erworben werden. Als Bildwandler kommt ein S35-mm-CMOS-Sensor mit 4206 x 2340 Pixeln zum Einsatz. Für die Gewinnung der Farbauszüge wird das Bayer-Pattern benutzt. Der Debayering-Prozess ist jedoch beim Einsatz als 2k-Kamera nicht erforderlich, da für jeden Bildpunkt die volle RGB-Auflösung, im Grünkanal sogar verdoppelt, zur Verfügung steht. Die Canon C500 bietet eine Slow- und Quick-Motion Funktion von bis zu 60 fps im 4k-RAW-Modus und bis 120 fps im 2k-Modus. Die Aufzeichnung erfolgt auf Compact-Flash-Speicherkarten als RGB-RAW oder RGB-4:4:4 mit 10 oder 12 Bit über einen 3G-SDI-Ausgang.

Die Entwicklung der Digitaltechnik ermöglichte schließlich noch weitere Verkleinerungen und extrem kompakte Kameras. Als ein Beispiel sei hier die GoPro-Kamera erwähnt (Abb. 6.114). Sie ist als Hero3 bei 15 fps in der Lage, eine 4k-Auflösung zu liefern, zählt aber eigentlich nicht zu den Digital Cinematographie-Kameras. Sie ist als Action-Kamera konzipiert. Entsprechend ist es eine Fix-Focus-Kamera mit großer Schärfentiefe, die den gesamten Bereich von 60 cm bis Unendlich vor der Kamera scharf abbildet. Alle weiteren Bildparameter sind vorgegeben bzw. automatisiert und die Aufnahme erfolgt vergleichsweise weitwinklig, sodass die Kamera eigenständig arbeiten kann. Sie ist sehr robust und befindet sich in einem staub- und wassergeschützten Gehäuse, daher kann sie an Fahrzeugen, Helmen etc. in ungewöhnlichen Postionen angebracht werden, um spektakuläre Bilder aus besonderen Perspektiven zu liefern. Die Variante GoPro Hero3 zeichnet im H.264-Codec mit bis zu 45 Mbit/s auf Speicherkarten auf. Ein 1/2,3“ HD-CMOS-Sensor erlaubt als Videoauflösungen und Framerates 4k bei 15 fps, 2,7k bei 25fps, 1080p bei 60 fps und 720p bei 120 fps. Daneben ist sie in der Lage, auch Audiosignale aufzuzeichnen.

6.4.3.4 Kameradatenverarbeitung

Bei der Betrachtung der Entwicklung digitaler Filmkameras ist insgesamt festzustellen, dass sich diese im Laufe der Zeit stark beschleunigt hat, vor allem durch die Verfügbarkeit sehr leistungsfähiger CMOS-Sensoren, die hohe Sig-

Abb. 6.114. Digitale Filmkameras Canon C500 (links) und Action-Kamera GoPro Hero [37]

nal/Rauschabstände aufweisen und durch die stark gestiegene Leistungsfähigkeit der Digitalprozessoren. Damit verbunden ist der Trend zu immer höherer Auflösung, der Trend zu 4k-Kameras. Hier kann man kritisch fragen, wo diese Auflösung gesehen werden kann, doch unterliegen Trends natürlich auch Marktstrategien und so kann der 4k-Trend auch bei Displays beobachtet werden. Die gestiegene Leistungsfähigkeit führt weiterhin, durch die Verwendung von Codecs wie Apple Pro Res und Sony XAVC, die hohe Auflösungen auch bei 12 Bit unterstützen, zu immer besserer Signalaufzeichnung, sowohl auf kompakten externen Geräten wie auch in den Kameras selbst. Daneben wurde der verarbeitbare Dynamikumfang und auch die maximal aufzeichenbare Bildrate immer weiter gesteigert, sodass weiche Bewegungen auch bei Zeitlupeneffekten gewährleistet werden können und dem Trend zur Verwendung von mehr als 48 fps im Kino, der unter dem Schlagwort High Frame Rate (HFR) bekannt ist, entgegengekommen wird.

Die Entwicklung geht zudem hin zu einer Verkleinerung der Kameras, was oft mit einem modularen Aufbau verbunden ist und auch dazu geführt hat, dass die digitalen Filmkameras zu immer günstigeren Preisen erworben werden können. Die geringe Baugröße ermöglicht den Einsatz der Filmkameras in sehr beengten Situationen, für die Fixierung oder die ruhige Führung per Hand sind aber umfangreiche Befestigungssysteme zusätzlich zu erwerben, was besonders bei DSLR-Kameras deutlich wird, die ja eigentlich nicht für den Filmbereich konzipiert sind.

Trotz guter Datenreduktionscodecs sind die aufgezeichneten Datenraten gestiegen, das wird auch erforderlich durch den Trend, bei den Kameras neben gewöhnlichen HD-Signalen auch RAW-Formate auszugeben, also Daten, die möglichst unbeeinflusst direkt vom Sensor kommen. Bilder im RAW-Format sind nicht für die unmittelbare Betrachtung geeignet sondern müssen einen Bildverarbeitungsprozess durchlaufen. Der Vorteil ist, dass dieser Prozess eine optimierte Qualität der Bilder hervorrufen kann, die mit gewöhnlicher Codierung in der Kamera nicht erreicht wird. Nun wird zwar die Bezeichnung RAW von allen Herstellern verwendet, aber die ausgegebenen Daten unterliegen verschiedenen Vorverarbeitungsprozessen und sind nicht einheitlich. Um diesem Mangel Abhilfe zu schaffen, schlug die Firma Adobe einen offenen RAW-Format-Standard mit der Bezeichnung Digital Negative (DNG) vor, die an das analoge digitale Negativ erinnert, welches auch erst nach der Farbkorrekturbearbeitung Bilder der endgültigen Qualität hervorbringt. DNG ist ein Fileformat für unkomprimierte Daten und beruht auf dem bekannten digitalen Einzelbildformat TIFF (Tagged Image File Format). Es kann RAW-Daten verschiedener Art und verschiedener Anordnung der Farbfilter, aber auch non-RAW-Daten (linearDNG) enthalten, weiterhin Metadaten und ein JPEG-codiertes Vorschaubild. Der DNG-Datenstrom kann videoorientiert in einem MXF-Container oder als Einzelbildfolge in einem Ordner gespeichert werden.

Ein zweiter Ansatz, zu konsistenten Daten und einem einheitlichen Workflow zu kommen, wurde unter dem Namen ACES (Academy Color Encoding System) erarbeitet und soll auch von der SMPTE standardisiert werden. Hier geht es vor allem darum, auch bei der Verwendung verschiedener Farbräume Endergebnisse mit konsistenter Qualität zu erzielen [172]. Das Format soll

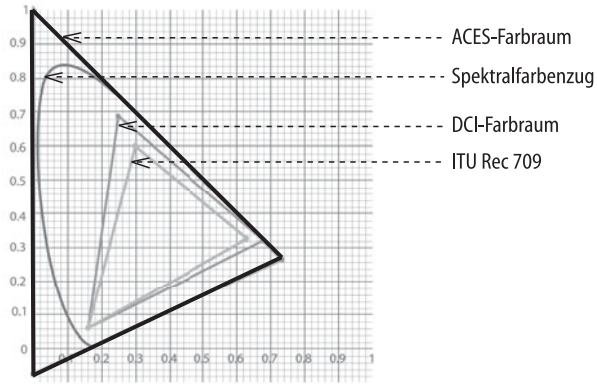


Abb. 6.115. ACES-Farbraum

möglichst auch für in Zukunft zu entwickelnde Systeme geeignet sein. Dazu definiert ACES einen RGB-Farbraum, der den gesamten hufeisenförmigen Farbraum im CIE-XY-Modell umfasst (Abb. 6.115). Die Farbwerte werden hier intensitätsskaliert verwendet, ohne Gamma- oder log-Verzerrung, dabei werden die Daten in jedem Kanal als 16-Bit-Werte dargestellt und erlauben damit einen Dynamikumfang, der den der menschlichen Wahrnehmung übersteigt.

Beim ACES-Verfahren wird das Eingangssignal mittels Input-Device-Transformatio (IDT) in ACES-Daten überführt, dieser Vorgang ist kameraspezifisch und muss vom Kamerahersteller bestimmt werden. Danach folgt die Reference-Rendering-Transformation (RRT). Sie hat die Aufgabe, die Daten für ein idealisiertes Displaymodell bereitzustellen, als Vorbereitung für den dritten Schritt, der Output-Device-Transformation (ODT), die der Anpassung an das tatsächliche Display dient. Für heutige Displays sind ODT verfügbar, die der Kalibrierung nach Rec 709, P3, P3 D60 und DCDM (XYZ) entsprechen.

Bezüglich der Farbräume war die Situation zu Beginn des Jahrhunderts erheblich einfacher. Für den Video- und TV-Bereich gab es mit ITU Rec 601 für SD und Rec 709 für HD wenige eindeutige Farbraumformate. Sie werden als verlässliche Basis auch weiterhin Bestand haben, während bezüglich der hier vorgestellten Bemühungen um eine bessere Bildqualität noch abzuwarten bleibt, welche sich in der Praxis auch durchsetzen werden und wo die neuen Qualitäten überhaupt sichtbar werden.

Die neuen Verfahren und Kameras machen aufgrund ihrer Vielfalt die Arbeit manchmal auch komplizierter, doch sollen zum Schluss dieses Abschnitts noch einmal die Vorteile dargestellt werden, die elektronische Kameras gegenüber Filmkameras bieten. Zu nennen sind hier die sofortige Verfügbarkeit eines hervorragenden HD-Bildes am Drehort und die sehr hohe Bildstandsgenauigkeit, die nicht wie bei Film ggf. die Schärfewirkung herabsetzt oder Encodierungsprobleme für nachgeschaltete Datenreduktionsstufen hervorruft, die z. B. bei der Ausstrahlung im digitalen Fernsehen (DVB) oder für Digital Cinema eingesetzt werden. Von Vorteil sind weiterhin die langen Aufzeichnungsdauern, die Möglichkeit, hochqualitativen Ton parallel mit aufzuzeichnen und bei elektronischer Wiedergabe schließlich die Umgehung des Kopierwerkes und damit die Vermeidung von Entwicklungs- und Abtasterkosten.

6.5 Kamerasupport

Der Begriff Kamerasupport steht für die Stütz- und Haltesysteme der Kameras. In den meisten Fällen sollen die Bilder nicht verwackeln, sodass die Kameras während der Aufnahme stabil gehalten werden müssen. Dazu dienen im Fernsehbereich für den Studioeinsatz vor allem Pumpstative und im Außeneinsatz Dreibeinstative. Fernsehstudios sind so gebaut, dass sie sehr glatte Böden aufweisen. Hier können mit schweren Stativen (Pedestals) ruhige Fahrten durchgeführt werden. Die Kameras, die hier meist mit schweren Objektiven ausgestattet sind, können auf ihnen leicht beweglich und gleichzeitig stabil gelagert sein. Die Stative sind einfach verfahrbar, die Kamera kann geneigt (tilt) und geschwenkt (pan) werden. Auch die Position in der Höhe kann mit geringem Kraftaufwand verändert werden. Aufgrund dieser Funktion werden sie als Pumpstative (Pumpen) bezeichnet, denn sie werden durch Aufpumpen so austariert, dass ein Gasgegendruck genau dem Schweredruck der Kamera entgegenwirkt. Diese Systeme können auch automatisiert und motorisch getrieben werden, sodass kein Personal erforderlich ist, um Fahrten, z. B. bei der Einführungsszene einer täglich ausgestrahlten Nachrichtensendung, immer in gleicher Weise zu wiederholen (Abb. 6.116) [181]. Bekannte Hersteller sind die Firmen Vinten und Radamec.

Im Bereich der elektronischen Berichterstattung kommen vor allem leicht transportable Dreibein-Stative (Tripods), z. B. der Firma Sachtler, zum Einsatz. Außer in dem Fall, dass Rollen montiert werden, kann hier die Kamera nicht gefahren werden. Für eine neue Einstellung muss der Standort gewechselt werden. Damit die Kameras geneigt und geschwenkt werden können, müssen auf dieses und die anderen Supportsysteme noch die Stativköpfe montiert werden. Bei guter Qualität ermöglichen sie sehr ruckfreie Schwenk- und Neigebewegungen, z. B. durch Fluid-Dämpfung (Abb. 6.117).

Bei anspruchsvolleren Fernsehproduktionen, aber vor allem im Filmbereich, soll die Kamera oft spektakuläre Bilder liefern und daher möglichst beweglich

Abb. 6.116. Studiopedestals für manuellen und automatischen Betrieb [146]

Abb. 6.117. Steadycam, Dreibeinstativ und Stativkopf [187]

sein. Zunächst gelingt dies, wenn die Kamera einfach auf der Schulter getragen wird, dabei ist sie allerdings insbesondere bei Teleaufnahmen kaum ruhig zu halten. Für besonders große personengebundene Beweglichkeit bei gleichzeitig hoher Bildstabilität wird die sogenannte Steadycam eingesetzt. Die Kamera wird dabei auf einem Stützkorsett befestigt, welches auf den Schultern und der Hüfte der bedienenden Person ruht. Die Kamerabewegungen können dann präzise, schnell und vor allem sehr weich ausgeführt werden. Die Bildkontrolle geschieht über einen separaten Monitor, der ebenfalls am Korsett angebracht wird (Abb. 6.117).

Im Filmbereich sind sehr oft Fahrten auch außerhalb von Studios erforderlich. Der klassische Support dafür ist der Dolly (Fahrwagen), der auf Rädern fährt oder bei unebenem Gelände auf Schienen (Abb. 6.118). Die Kamera kann geneigt und geschwenkt und auch in der Höhe verfahren werden. Ein bekannter Hersteller ist die Firma Panther, deren Dolly es ermöglicht, dass der Kameraoperator direkt an der Kamera mitfährt. Um die Funktion eines ein-

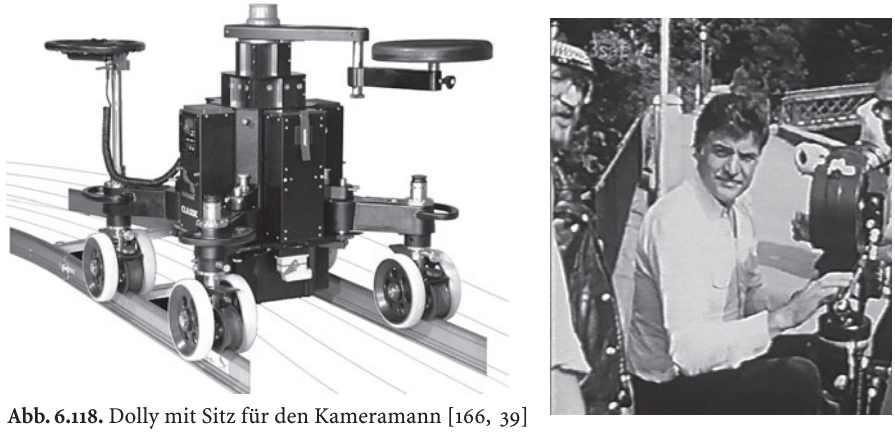


Abb. 6.118. Dolly mit Sitz für den Kameramann [166, 39]

Abb. 6.119. Kamerakran [146]

fachen kleinen Kamerakrans zu erreichen, kann ein kurzer beweglicher Arm (Jib-Arm) montiert werden. Bei schienenlosem Betrieb können die Räder auf verschiedene Weise verkoppelt werden (crab and steer). Damit ist der Dolly einfach zu manövrieren und Kurvenfahrten sind exakt festlegbar [166]. Die für die Bewegung zuständigen Personen sind dem Bereich Dolly Grip zugeordnet, der auf Deutsch auch als Kamerabühne bezeichnet wird.

Im Filmbereich treten oft auch Fahraufnahmen mit Fahrzeugen auf. Hier gibt es eine Vielzahl von speziellen Befestigungen von Klemmstativen bis zu Montagen auf Autokühlerhauben oder vorausfahrenden Fahrzeugen. Für besonders große Höhenänderungen und hohe Beweglichkeit kommen Kamerakräne zum Einsatz, z. B., um in der Einführungssequenz eines Spielfilms die ganze Szene von oben zu überblicken und dann ohne Schnitt hinabzufahren und auf den Protagonisten zu verdichten. Hier kann die Kamera nicht mehr direkt bedient werden, Position, Zoom und Fokus werden mittels Remotehead ferngesteuert (Abb. 6.119).

Speziell für den Sportbereich sind die Systeme GoCam und SpiderCam entwickelt worden (Abb. 6.120). Ersteres arbeitet mit einer langen Schiene, an der

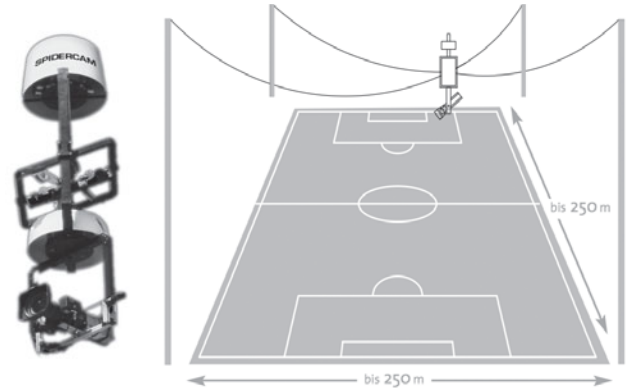


Abb. 6.120. Spidercam [146]

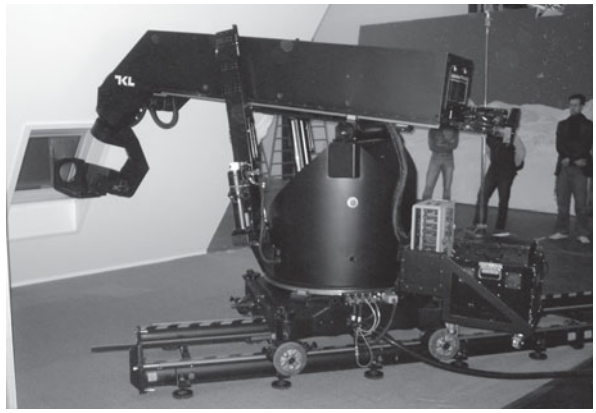


Abb. 6.121. Motion Control-System

sie sehr schnelle Fahrten, z. B. parallel zu einem Rennrad ausführen kann. Spidercam liefert spektakuläre Aufnahmen von oben [146]. Es ist ein Drahtseilsystem, das z. B. ein ganzes Fußballstadion überspannt und alle Positionen über dem Rasen mit hoher Geschwindigkeit erreichen kann (Abb. 6.120).

Besonders aufwendig sind Supportsysteme, mit denen die Fahrten exakt reproduziert werden können. Sie werden als Motion-Control-Systeme bezeichnet. Abbildung 6.121 zeigt als Beispiel das Milo-System. Die Reproduzierbarkeit erfordert eine große Stabilität, die über eine hohe Masse von ca. einer Tonne erreicht wird. Dagegen ist der motorisch gesteuerte Arm leicht und schnell beweglich. Das System wird vor allem für Effekt-Aufnahmen verwendet, z. B. um dieselbe Person mehrfach in einer Szene agieren zu lassen, wobei die Kamera immer in derselben Weise bewegt wird.

Oben wurde beschrieben, dass die Hersteller immer kompaktere und modulare Kameras bauen. Um diese bei gewöhnlichen Filmproduktionen auf einem Stativ oder handgeführt einsetzen zu können, sind umfangreiche Stütz- und Haltesysteme erforderlich für die sich ein eigener Angebotssektor entwickelt hat, der sehr schnell auf neue und aktuelle Kameramodelle reagiert. Abbildung 6.122 zeigt ein solches System für die sehr kompakten DSLR-Kameras.

Abb. 6.122. Rigging- und Follow-Focus-Systeme für leichte Kameras

6.5.1 Stereoskopische Aufnahmen

Stereoskopische Aufnahmen sind vom Grundprinzip recht einfach: Es müssen gleichzeitig zwei Aufnahmen gemacht werden, die wie bei der menschlichen Wahrnehmung die Szene aus leicht unterschiedlichen Blickwinkeln abbilden. Dafür können zwei gewöhnliche Kameras zum Einsatz kommen, die zusammen mit ihren Objektiven und Aufzeichnungssystemen exakt gleich eingestellt und möglichst synchronisiert sein müssen. Stereoskopische Aufnahmen sind daher vor allem eine Frage des Supportsystems, das oft auch 3D-Rig genannt wird. Die Rigs müssen eine sehr exakte Positionierung der Kameras ermöglichen, die Positionen sollten pixelgenau konstant bleiben. Auch eine Differenz bezüglich der Bildhöhen (Höhenparallaxe) muss vermieden werden.

Der unterschiedliche Blickwinkel wird, wie bei den menschlichen Augen, durch einen seitlichen Versatz der beiden optischen Achsen erzeugt. Der Versatz wird Stereobasis, Basis oder interaxiale Distanz (engl. interaxial, IA) genannt. Gebräuchlich ist auch der Begriff Interokularabstand. Dieser kann jedoch irreführend sein, da er nahelegt, dass die Stereobasis dem Augenabstand entsprechen muss. Tatsächlich genügt bei vielen Nahaufnahmen für eine Großbildprojektion eine Basis von wenigen Zentimetern, während eine ausreichende Raumtiefenwirkung einer Totalen erst mit einer Basis von einigen Dezimetern erreicht wird. Trotzdem ist der Abstand der menschlichen Augen von ca. 6 cm eine erste Orientierung für die Stereobasis. Mehr als die Hälfte der gewöhnlichen Produktionen wird mit einer Stereobasis zwischen 1 cm und 6 cm durchgeführt.

Im einfachsten Fall werden die Kameras nebeneinander auf kleine Gestelle montiert, die als Side by Side-Rigs bezeichnet werden (Abb. 6.123 oben). Allerdings können nur sehr kleine Kameraköpfe, bzw. Objektive, so eng zusammengebracht werden, dass ihre Distanz dieser Größenordnung entspricht.

Stereoskopische Aufnahmen mit größeren Kameras werden mit Mirror-Rigs realisiert. Die beiden Perspektiven werden dabei über halbdurchlässige Spiegel zu den Kameras geführt, wobei eine Kamera wie gewöhnlich und die andere senkrecht dazu steht, also von oben oder unten auf den Spiegel gerichtet ist (Abb. 6.123 unten). Das Mirror-Rig wird in der Mehrzahl der Fälle eingesetzt,

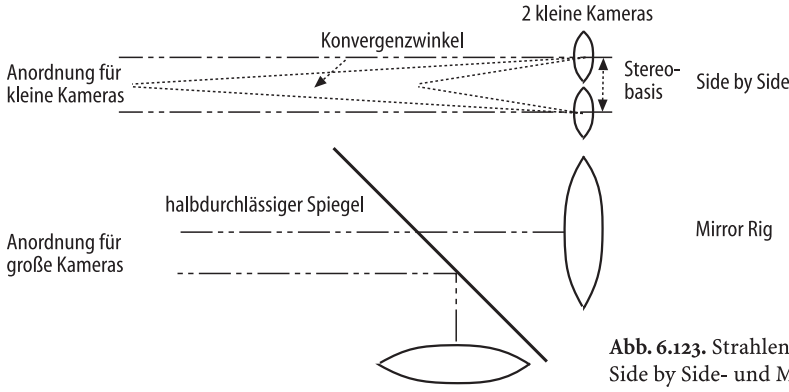


Abb. 6.123. Strahlengänge von Side by Side- und Mirror-Rigs

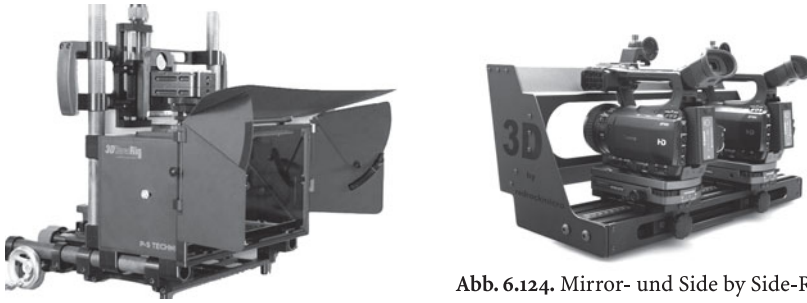


Abb. 6.124. Mirror- und Side by Side-Rigs [163]

da es eine beliebige Wahl der Stereobasis erlaubt. Nachteilig ist, dass ein Lichtverlust von mindestens einer Blende am Spiegel entsteht. Ein zweites Problem ist die mögliche Verschmutzung des Spiegels (Abb. 6.124).

Die Größe des Kameraabstandes, ist direkt proportional zur empfundenen Raumtiefe, denn große Abstände bewirken auf der Bildwand stark voneinander abweichende Konturen. Eine Übertreibung der Separation und damit der Tiefenwirkung wird Hyperstereoskopie genannt. Dagegen führt ein Kameraabstand, der kleiner ist als der Augenabstand, zur Hypostereoskopie. Große Stereobasisbreiten von bis zu einem halben Meter werden für Landschaftsaufnahmen verwendet, die aus großen Entfernungen geschehen.

Durch die Stereobasisbreite wird also das Tiefenbudget festgelegt (Abb. 6.125). Daneben wird die Produktion erheblich durch den Konvergenzwinkel zwischen den Kameras bestimmt. Wenn die Kameraachsen nicht konvergieren, also parallel ausgerichtet sind, ergibt sich für alle Objekte in der Raumtiefe eine negative Parallaxe. Die gewünschte Tiefenanordnung eines Objekts muss dann in der Postproduktion bestimmt werden, wo mittels der horizontalen Verschiebung (Horizontal Image Translation, HIT) der beiden Bilder festgelegt wird, in welcher Entfernung das Scheinfenster liegt und welche Objekte sich davor oder dahinter befinden sollen. Wenn dagegen mit konvergierenden Kameraachsen gearbeitet wird, wird diese Anordnung direkt am Set bestimmt, was z. B. für Live-Übertragungen günstig bzw. erforderlich ist (Abb. 6.125). Die konvergierenden Kameraachsen bieten den Vorteil, dass beide Kameras den gleichen Bildausschnitt erfassen. Der Nachteil ist, dass Bildverzerrungen durch trapezförmige Verzerrungen auftreten (Key-Stoning). Dieser Effekt tritt bei der Parallelausrichtung nicht auf. Nachteilig ist hier nur der unterschiedliche Bildausschnitt, der in der Postproduktion korrigiert werden muss.

Um das Wahrnehmungsvermögen der Betrachter nicht zu überfordern, sind bei der Aufnahme einige Regeln einzuhalten, denn stereoskopische Systeme zwingen die menschlichen Augenachsen, zu konvergieren. Zu große Konvergenzwinkel, d. h. zu große Parallaxen, können zu Unwohlsein führen. Als Näherung kann davon ausgegangen werden, dass die Parallaxe 70 Winkelminuten bzw. 1/30 der Bildbreite nicht überschreiten sollte. Noch weniger dürfen die Augen zu einer Divergenz gezwungen werden.

Besonders zu bedenken ist zunächst, für welche Betrachtungssituation produziert wird, denn auch der Betrachtungsabstand ist für die Tiefenwirkung entscheidend. Die Tiefenwirkung steigt mit dem Abstand zur Bildwand, die

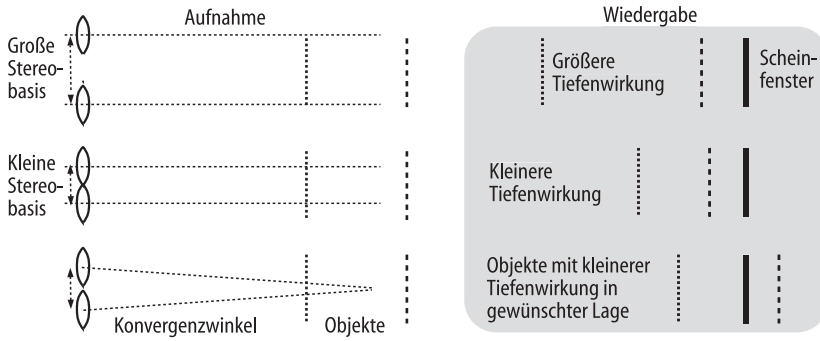


Abb. 6.125. Auswirkungen der Veränderung von Stereobasis und Kovergenzwinkel

Wahl der Kameraparameter hängt vom Präsenzfaktor, ab dem Verhältnis von Bildbreite zum Betrachtungsabstand. Das ist der Grund dafür, dass in den meisten Fällen Kinoproduktionen im Fernsehen zu wenig Tiefe aufweisen.

Bei der Aufnahme sollte mit großer Schärfentiefe gearbeitet werden, denn im Gegensatz zur natürlichen Wahrnehmung wird bei der Konzentration auf eine andere Bildtiefe der Vorder- oder Hintergrund bei der Stereoskopie nicht unscharf. Hier sind Kameras mit kleinem Bildwandler vorteilhaft. Weiterhin ist es erforderlich, darauf zu achten, dass sich die binokularen und monokularen Tiefenhinweise verstärken, statt sich zu behindern.

Zoomfahrten sollten sehr sparsam verwendet werden. Es wird vorzugsweise mit Festbrennweiten gearbeitet. Wenn ein Objekt sich in der Tiefe der Szene bewegt, muss neben der Schärfe ggf. auch die Stereobasis nachgeführt werden. Bei negativer Parallaxe ist es wichtig, dass die Objekte, die sich vor der Bildwand befinden, nicht vom Rahmen beschnitten werden, dass also die Scheinfensterverletzung vermieden wird. Damit das alles gelingt, sollte sich die Bildgestaltung auf ein im Vorfeld erarbeitetes Tiefenscript stützen.

Generell muss während der Produktion darauf geachtet werden, dass möglichst wenige Unterschiede zwischen den Bildern bezüglich Farbe, Form, vertikaler Ausrichtung etc. auftreten. Schon kleine Abweichungen behindern die Bildfusion im Auge. Derartige Fehler müssen bei der Postproduktion aufwendig korrigiert werden. Vollständig sind die Fehler nicht vermeidbar, daher werden bei 3D-Liveübertragungen spezielle Bildanalyse- und Korrektursysteme verwendet, die die Korrektur in Echtzeit vornehmen. Hier ist natürlich auch besonders auf die Synchronisation der Kameras zu achten.

Als Hilfsmittel zur Bestimmung der richtigen Aufnahmeparameter stehen Programme und Werkzeuge wie STAN (Stereoscopic Analyser) zur Verfügung. STAN beruht auf einer Bildanalyse, aus der die Visualisierung der Tiefe im Bild und die Festlegung der Konvergenzebene gewonnen werden, auch Regelverletzungen werden angezeigt. Trotz dieser Hilfsmittel sind am Set aber erfahrene Stereoskopie-Spezialisten, die Stereographer, erforderlich. Ihre Aufgabe besteht darin, den 3D-Effekt herauszuarbeiten, die Aufnahme und Postproduktion aber auch so zu gestalten, dass sich für die meisten Menschen ein möglichst komfortables Seherlebnis einstellt.

7 Bildwiedergabesysteme

Bildwiedergabegeräte dienen zur Rekonstruktion des auf der Aufnahmeseite erzeugten Bildes, zur Umsetzung des Videosignals in sichtbares Licht. Auch wenn die Bildpunkte, je nach Verfahren, in verschiedener Form nacheinander auf dem Bildschirm erscheinen, entsteht aufgrund der Trägheit des Wiedergabewandlers und des menschlichen Auges der Eindruck eines ganzen Bildes.

Wiedergabewandler lassen sich in aktive und passive Typen unterscheiden. Aktive Typen emittieren selbst Licht, während passive mit elektronisch gesteuerter Lichttransmission oder -reflexion arbeiten. Abbildung 7.1 zeigt eine Übersicht [147]. Im gesamten 20. Jahrhundert dominierte die zu den Aktivsystemen gehörige Kathodenstrahlröhre, ihr Einsatz war dafür verantwortlich, dass das analoge Videosignale eine recht komplizierte Form hat. Im 21. Jahrhundert wurden die Flachdisplaytechnologien immer ausgereifter, mit denen die Nachteile der Röhrentechnik, also hohes Volumen, hohes Gewicht und großer Energiebedarf, überwunden wurden. Heute dominieren flache LC-, Plasma- und OLED-Displays.

Als allgemeine Forderungen an Wiedergabewandler gelten: große Helligkeit, kleine Trägheit, großer Betrachtungswinkel, hohes Auflösungsvermögen, großes Bild und geringer Leistungsbedarf. Die Bezeichnung »großes Bild« meint bei Heimanwendung eine Bilddiagonale von bis zu 1 m oder mehr, die besonders für die HDTV-Wiedergabe gewünscht ist. Die Einführung von HDTV wurde nicht zuletzt durch die Schwierigkeit der Entwicklung eines passenden Displays gehemmt. Heute stehen kostengünstige Systeme für diese Anwendung zur Verfügung. Noch größere Bilder werden mit speziellen Großbildprojektionen erreicht. Die dazu eingesetzten Verfahren werden zum Schluss dargestellt.

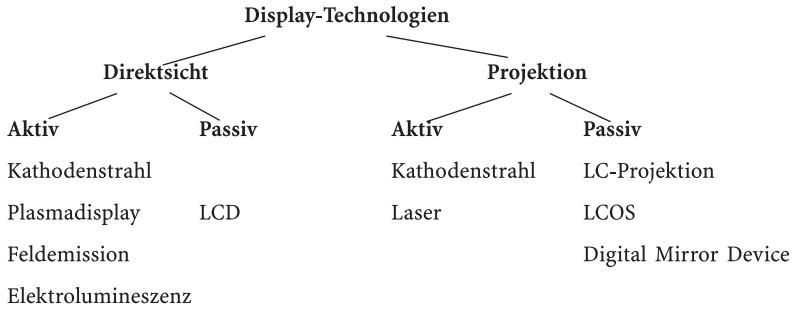


Abb. 7.1. Übersicht über die Bildwiedergabewandler

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4_7, © Springer-Verlag Berlin Heidelberg 2013

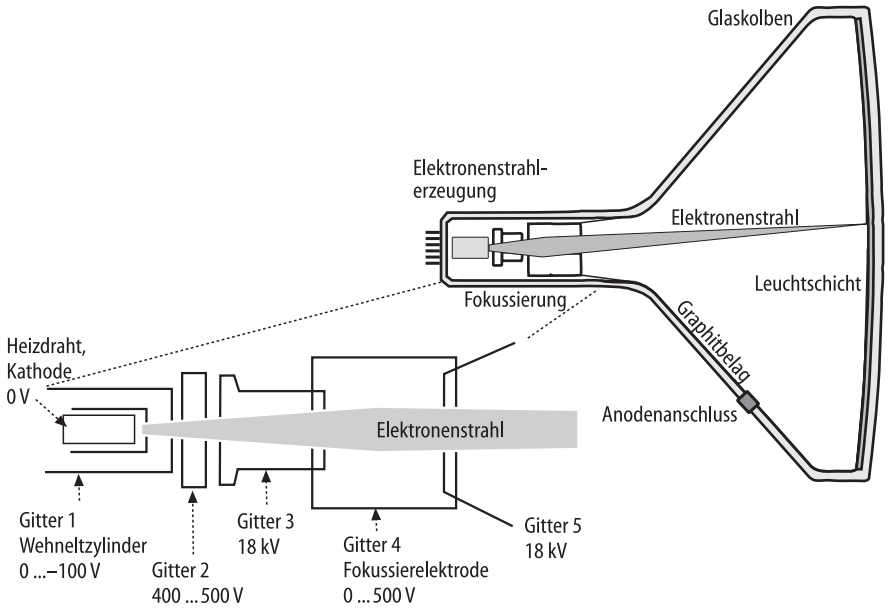


Abb. 7.2. Kathodenstrahlröhre und Elektrodensystem

7.1 Bildwiedergabe mit Kathodenstrahlröhren

7.1.1 Das Funktionsprinzip

Die Bildwiedergabe mit einer Kathodenstrahlröhre (Cathode Ray Tube, CRT) (Abb. 7.2) beruht darauf, dass beim Auftreffen von hochbeschleunigten Elektronen Leuchtstoffe, die sogenannten Phosphore, zur Lichtemission angeregt werden. Die Farbe des emittierten Lichtes ist dabei von der Art des Leuchtstoffes bestimmt. Die Lichtemission erfolgt ungerichtet, damit ergibt sich der Vorteil einer weitgehenden Richtungsunabhängigkeit bei der Betrachtung. Damit der Elektronenstrom nicht von Luftmolekülen behindert wird, muss die Röhre luftleer sein. Kathodenstrahlröhren werden als evakuierte Glaskolben hergestellt. Sie müssen dem hohen Außendruck standhalten, die Glaswände sind entsprechend dick und schwer.

Zur Elektronenerzeugung wird eine Metallfläche (Kathode) mit einem Glühdraht erhitzt, aufgrund der thermischen Energie treten die Elektronen aus der Kathode aus. Die Kathode ist vom Wehneltzylinder umgeben, der gegenüber der Kathode negativ geladen ist (Abb. 7.2). Je nach Größe des negativen Potenzials werden mehr oder weniger Elektronen zurückgehalten, d. h. über die Spannungsdifferenz zwischen Wehneltzylinder und Kathode kann die Strahlintensität gesteuert werden. Die Elektronen werden bei S/W-Röhren mit ca. 18 KV auf die Anode hin beschleunigt und gelangen dabei in den Bereich der

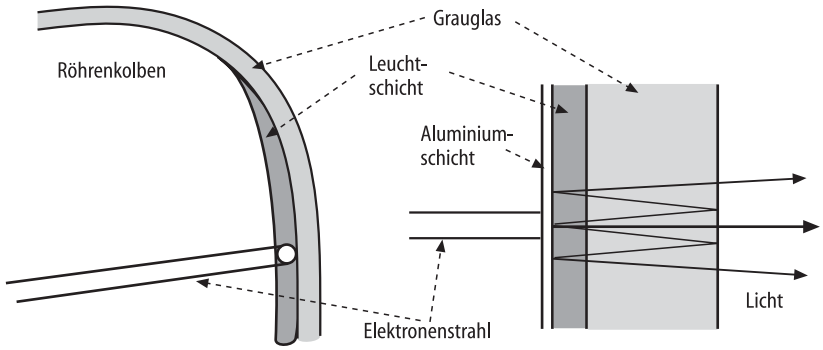


Abb. 7.3. Glaskolben der Kathodenstrahlröhre und Beschichtungen

Fokussierelektroden, die ein elektrisches Feld erzeugen, welches zur Bündelung des Elektronenstrahls führt. Nachdem sie die Anode passiert haben, bleibt ihre Geschwindigkeit konstant. Die Elektronen fliegen weiter, bis sie auf die Leuchtschicht treffen und diese zur Lichtemission anregen. Die resultierende Leuchtdichte steigt mit der Beschleunigungsspannung und der Strahlstromstärke. Es werden Strahlstromstärken unter 1 mA verwandt, denn bei großem Strahlstrom wird auch der Elektronenstrahldurchmesser groß, was wiederum die Bildauflösung verringert. Farbbildröhren erreichen damit Leuchtdichten bis zu 100 cd/m². Die Beschleunigungsspannungen am Anodenanschluss (Abb. 7.7) liegen unterhalb 30 kV, denn beim Auftreffen auf die Schicht und den Glaskolben werden die Elektronen so stark abgebremst, dass Röntgenstrahlung entsteht. Die Röntgenstrahlung muss abgeschirmt werden, was bei Spannungen oberhalb 30 kV nur mit großem Aufwand möglich ist.

Die Leuchtschicht leitet selbst oder ist zusammen mit einer leitfähigen Schicht auf dem Glaskolben aufgebracht. Über diese Schicht fließen die Elektronen zur Anode ab und damit wird der Stromkreis geschlossen. Zusätzlich ist die Leuchtschicht mit einer dünnen Aluminiumschicht hinterlegt, die von den Elektronen durchschlagen wird. Diese dient zur Ableitung von Sekundärelektronen, die beim Auftreffen des Elektronenstrahls entstehen. Außerdem wirft sie das Licht zurück, das die Leuchtstoffe in den Röhreninnenraum abstrahlen, sodass die Bildhelligkeit erhöht wird (Abb. 7.3). Um den Kontrast zu steigern, wird für die Herstellung der meisten Röhren Grauglas verwendet. Damit wird

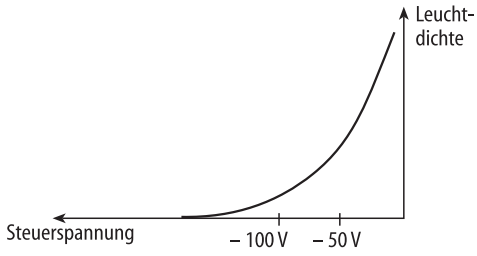


Abb. 7.4. Steuerkennlinie der Kathodenstrahlröhre

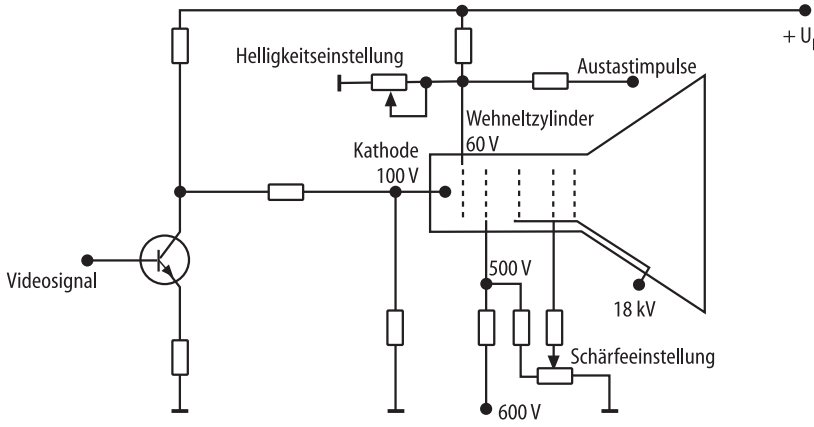


Abb. 7.5. Beschaltung der Bildröhre

das emittierte Licht zwar geschwächt, für von außen auftreffendes Fremdlicht wirkt die Schwächung jedoch doppelt, sodass insgesamt eine Kontrastverbesserung eintritt.

Die Intensität des Elektronenstrahls, der Strom I und damit die Bildpunkthelligkeit, wird gesteuert, indem das Videosignal zum Wehneltzylinder oder zur Kathode geführt wird. Die Steuerkennlinie (Abb. 7.4) bewirkt dabei einen nichtlinearen Zusammenhang zwischen Videosignal und Leuchtdichte, der mit Hilfe einer Potenzfunktion mit dem Exponenten γ angenähert beschrieben werden kann. Als Standardwert gilt γ_W = 2,2, ein Wert, der in der Kamera durch Vorentzerrung mit γ_A = 0,45 ausgeglichen wird (s. Abschn. 2.2).

Abbildung 7.5 zeigt eine Möglichkeit zur Beschaltung einer Schwarz/Weiß-Bildröhre [108]. Das von den Synchronsignalen befreite Videosignal steuert über die Kathode die Elektronenstrahlintensität. Die Gesamthelligkeit und der Schwarzwert können über das Steuergitter beeinflusst werden. Die Spannung an den Fokussierelektroden bestimmt die Bildschärfe.

Damit der Elektronenstrahl nicht nur einen Punkt, sondern eine zweidimensionale Fläche zum Leuchten bringt, wird er gemäß dem genormten Fernsehzeilenraster horizontal und vertikal abgelenkt. Zur Ablenkung dient eine Spulenanordnung, die auf dem Röhrenhals angebracht ist (Abb. 7.6 und 7.7). Durch

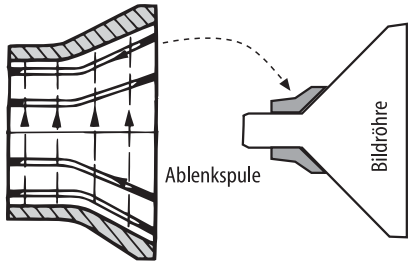


Abb. 7.6. Ablenkspulen auf dem Röhrenhals

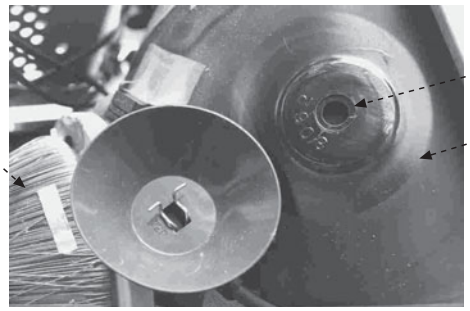


Abb. 7.7. Ablenkspule und Anodenanschluss

die Ströme in den Spulen entstehen Magnetfelder, mit denen der Elektronenstrahl abgelenkt wird. Der Stromverlauf für die Ablenkung ist annähernd sägezahnförmig. Der Strahl wird während der aktiven Zeilendauer von links nach rechts über den Schirm geführt und springt während der Austastlücke unsichtbar zurück, analog funktioniert die Vertikalablenkung.

Zur Konstruktion möglichst flacher Wiedergabegeräte ist ein großer Ablenkwinkel erforderlich. Mit den modernsten Bildröhren wurden diagonal Werte von bis zu 110° erreicht. Hier besteht ein wesentlicher Unterschied zu Bildaufnahmeröhren, bei denen mit sehr viel kleineren Ablenkwinkeln gearbeitet werden kann, da die Wandlerfläche wesentlich kleiner ist. Bei modernen, flach gewölbten Bildschirmen ist der Strahlweg zum Rand hin erheblich länger als zur Mitte. Daraus resultieren Bildverzerrungen, die aufgrund ihrer Gestalt als Kissenverzerrungen bezeichnet werden. Diese Verzerrungen werden über seitlich an der Ablenkeinheit angebrachte Permanentmagnete und über Korrekturen des sägezahnförmigen Ablenkstroms ausgeglichen.

7.1.2 Farbbildröhren

In Farbbildröhren werden genormte rot, grün und blau leuchtende Leuchtstoffe verwendet, die in kleinen Strukturen sehr eng beieinander angeordnet werden. Jeder Bildpunkt besteht aus drei Pixeln (Farbtripel), die Licht der entsprechenden Farben emittieren. Bei genügend großem Betrachtungsabstand erscheinen die drei Pixel als ein einziger Punkt, dessen Farbe durch die additive Mischung der Farbwerte bestimmt ist.

Wird, wie bei der S/W-Bildröhre, auch zur Farbbildwiedergabe nur ein Elektronenstrahl verwendet, so muss dieser nicht nur von Bildpunkt zu Bildpunkt seine Intensität verändern können, sondern bei vergleichbarer Auflösung dreimal so schnell wie bei der S/W-Röhre steuerbar sein, da er jeden Tripelpunkt ansprechen muss. Außerdem muss immer feststehen, welche Art Phosphor der Strahl gerade trifft, damit eine richtige Zuordnung zur Strahlsteuerung vorgenommen werden kann. Eine Einstrahlröhre ist aus diesen Gründen sehr aufwendig und wird praktisch nicht verwendet.

In der Praxis wird mit Dreistrahlröhren gearbeitet. Im Röhrenkolben gibt es drei separate Strahlerzeugungssysteme für die drei Grundfarben. Die drei eng

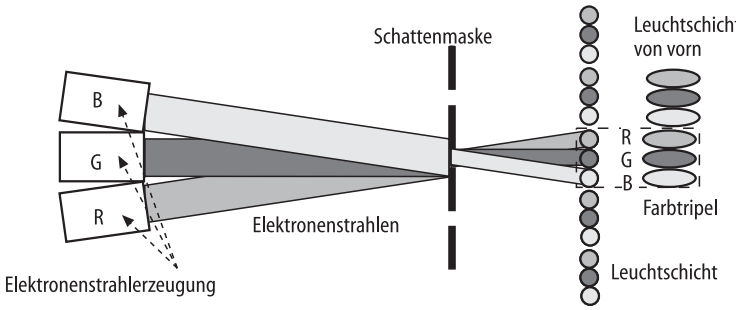


Abb. 7.8. Farbbildröhre mit Schattenmaske

benachbarten Elektronenstrahlen werden gemeinsam abgelenkt. Das wesentliche Problem besteht nun darin, dafür zu sorgen, dass die jeweiligen Elektronenstrahlen nur die zugehörigen Leuchtstoffe treffen. Die Lösung besteht in der Verwendung einer Schattenmaske, die in einem fest definierten Abstand ca. 15 mm vor der Leuchtschicht angebracht wird. Jedem Farbtripel ist genau eine Öffnung in der Maske zugeordnet, die Elektronenstrahlen werden gemeinsam durch diese Maskenöffnung geführt. Da die Strahlen wegen der separaten Strahlerzeugung vor der Maske nicht in einer Linie laufen, sondern um ca. 1° gegen die Röhrenachse geneigt sind, ergibt sich hinter Maske auch für jeden Strahl eine andere, fest definierte Richtung. Die Anordnung kann so gewählt werden, dass sich die Strahlen genau am Ort der Maske kreuzen und jeder Strahl hinter der Maske nur den ihm zugeordneten Leuchtstoff trifft (Abb. 7.8). Bezüglich der Anordnung der Strahlerzeugungssysteme bzw. der Farbtripel unterscheidet man Farbbildröhren in Delta-, In Line- und Trinitron-Typen.

Ein Nachteil der Maske ist das Einbringen eines starren Musters, das bei der Wiedergabe zu Alias-Störungen führen kann, die bei S/W-Röhren nicht auftreten. Ein zweiter Nachteil ist die Verdeckung großer Bereiche durch die Maske, je nach Typ wird nur noch eine Transparenz von ca. 20% erreicht. Der Helligkeitsverlust der Farbbildröhre gegenüber der S/W-Röhre wird durch eine hohe Beschleunigungsspannung (ca. 30 kV) teilweise kompensiert. Durch Aufbringen einer, von vorn gesehen, schwarzen Schicht zwischen den Löchern der Maske wird bewirkt, dass von außen einfallendes Licht nur geringfügig von der Maske reflektiert wird. Diese sogenannte Black Matrix erhöht damit den Wiedergabekontrast.

Die Maske muss überall in definierter Entfernung von der Leuchtschicht angebracht sein und ihre Position sehr genau beibehalten. Verformt sie sich, z. B. aufgrund thermischer Ausdehnung, so führt dies zu Farbreinheitsfehlern. Die mechanische Befestigung der Maske wird daher aufwendig temperaturkompensiert, als Material wird z. T. Invar verwendet. Auch eine Magnetisierung der Maske führt zu Strahlabweichungen und damit zu Farbreinheitsfehlern. Unvermeidlicher Restmagnetismus wird automatisch mit Hilfe einer Entmagnetisierungsspule beseitigt, die beim Einschalten kurzzeitig ein schnell abklingendes Magnetfeld erzeugt. Die Entmagnetisierung kann bei einigen Geräten auch über eine Taste (Degauss) ausgelöst werden.

7.1.2.1 Delta-Röhre

Der Name dieses ältesten praktisch eingesetzten Farbbildröhrentyps resultiert aus der Strahlsystemanordnung als gleichseitiges Dreieck, ähnlich dem griechischen Buchstaben Delta (Δ). Die Elektronenstrahlen werden gemeinsam durch kreisrunde Löcher geführt, daher auch der Name Lochmaskenröhre. Der Strahlerzeugungsanordnung entsprechend müssen die Farbtripel auch in Delta-Anordnung liegen (Abb. 7.9). Der Abstand zwischen zwei Leuchtpunkten einer Reihe wird Pitch genannt. Bei einer typischen Farbbildröhre mit 60 cm Bildschirmdiagonale und ca. 400 000 Löchern in der Maske beträgt der Pitch etwa 0,7 mm.

Damit über die sogenannte Landungsreserve eine gewisse Fehlertoleranz gegeben ist, ist der Lochdurchmesser mit ca. 0,2 mm kleiner als der Strahldurchmesser und der Leuchtpunktdurchmesser (beide ca. 0,3 mm). Die Delta-Röhre ermöglicht aufgrund der sehr dicht liegenden Tripelpunkte eine hohe Auflösung und wurde vorwiegend in Computermonitoren eingesetzt. Nachteil ist die geringe Maskentransparenz von ca. 17% sowie der hohe Aufwand, der getrieben werden muss, um mit Zusatzspulen auf der ganzen Bildschirmfläche eine Deckungsgleichheit der Farbauszüge (Konvergenz) zu erzielen, d. h. zu erreichen, dass sich zu jedem Zeitpunkt die drei Strahlen in nur einem Loch der Maske kreuzen.

7.1.2.2 In-Line-Röhre

Dieser auch als Schlitzmaskenröhre bezeichnete Typ wurde entwickelt, um die Konvergenzeinstellung vereinfachen zu können. Wie der Name andeutet, sind die Strahlerzeugungssysteme in einer Reihe angeordnet. Die Elektronen treten durch eine Schlitzmaske und fallen auf die Leuchtstoffe, die nicht mehr im Dreieck, sondern nebeneinander angeordnet sind (Abb. 7.9). Wegen der In-Line-Anordnung sind die Konvergenzfehler von drei auf zwei Dimensionen reduziert und die Ablenkeinheiten können selbstkonvergierend gebaut werden. Aufgrund dieser Vereinfachung war die In Line-Röhre der Standardtyp in Fernsehempfängern. Die Maskentransparenz hat etwa den gleichen Wert wie bei der Delta-Röhre.

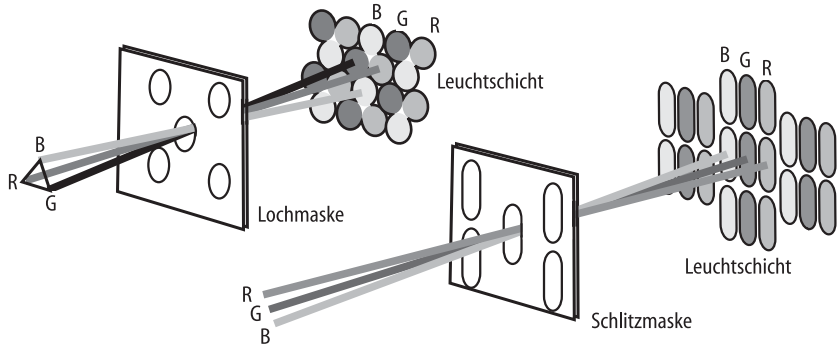


Abb. 7.9. Prinzip von Loch- und Schlitzmaskenröhre

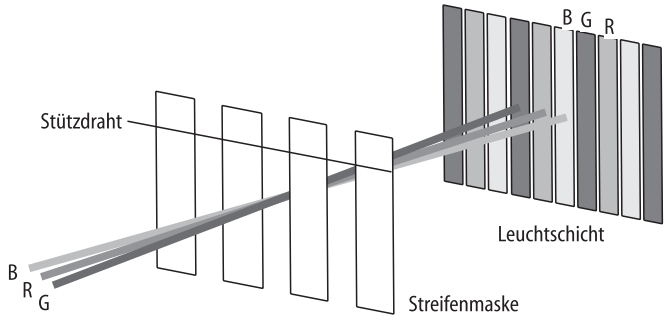


Abb. 7.10. Prinzip der Trinitron-Röhre

7.1.2.3 Trinitron-Röhre

Noch vor der In-Line-Röhre wurde von der Firma Sony die Trinitron-Röhre entwickelt, die prinzipiell ähnlich arbeitet. Die Elektronenquellen sind in einer Reihe angeordnet, die Fokussierung geschieht für alle Strahlen gemeinsam. Wie beim In-Line-System liegen die Leuchtstoffe auf dem Bildschirm nebeneinander, die Phosphore sind aber in vertikaler Richtung nicht unterbrochen, sondern laufen als durchgehende Streifen über den gesamten Schirm. Die Schattenmaske besteht dementsprechend auch nicht aus kurzen Schlitzen, sondern ist ein Blendengitter mit durchgehenden, vertikalen Streifen (Abb. 7.10). Die Schattenmaske ist insgesamt aufgebaut wie eine Harfe. Zur Stabilisierung werden die Metallstreifen unter Spannung gesetzt. Um Schwingungen zu verhindern, sind sie mit ein oder zwei dünnen Querdrähten verbunden.

Die Trinitron-Röhre bietet mit einer Maskentransparenz von ca. 22% eine relativ hohe Lichtausbeute. Kontrast und Vertikalauflösung sind sehr gut, und es sind selbstkonvergierende Wiedergabesysteme realisierbar. Problematisch ist die mechanische Empfindlichkeit der Maskenkonstruktion, insbesondere die Neigung zu Schwingungen bei mechanischer Erschütterung.

7.1.3 Signalverarbeitung

Neben der Bildwandlerröhre befinden sich im Videomonitor oder einem Fernsehgerät Signalverarbeitungsstufen zur Wandlung des Videosignals in ein RGB-Signal, das wiederum der Ansteuerung der Röhre dient, sowie die Stufen zur Erzeugung der Ablenkströme.

Je nach Art muss das Videosignal mehr oder weniger decodiert werden, um schließlich das zur Ansteuerung der Röhrenkathoden erforderliche RGB-Signal zu erhalten. Am aufwendigsten ist die Decodierung des Composite-Signals. In Kapitel 2 wurde gezeigt, wie das von den Synchronsignalen befreite Videosignal mit Hilfe von Filtern in Luminanz- und Chrominanzanteile aufgespalten wird. Die Filterqualität bestimmt wesentlich die Bildqualität.

Das Luminanzsignal wird verzögert, um die Laufzeiten in den Farbverarbeitungsstufen auszugleichen und gelangt dann direkt zur RGB-Matrix. Das

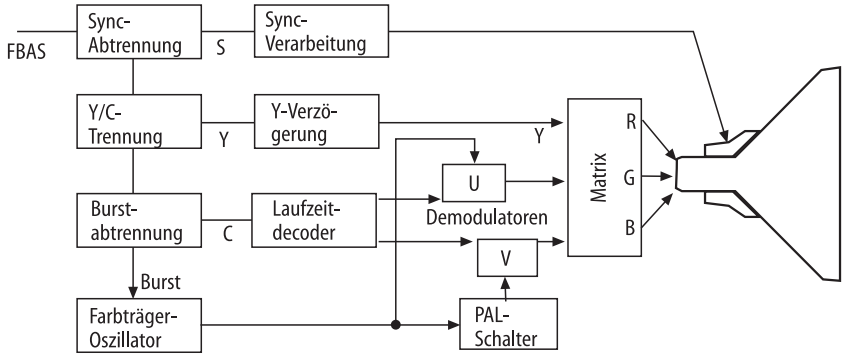


Abb. 7.11. Blockschaltbild eines PAL-Videomonitors

Chrominanzsignal wird zunächst verstärkt, dann wird der Burst abgetrennt. Der Burst synchronisiert den im Monitor enthaltenen Farbhilfsträgeroszillator, sodass dieser, bezogen auf den Sender, mit richtiger Phase und Frequenz schwingt. Weiterhin steuert der Burst die PAL-Kippstufe derart, dass insgesamt immer ein phasenrichtiger Hilfsträger zur Synchrondemodulation zur Verfügung steht. Der Farbhilfsträger wird schließlich dem Demodulator zugesetzt, wobei die V-Phase automatisch in jeder 2. Zeile zurückgeschaltet wird (Abb. 7.11).

Das vom Burst befreite Chrominanzsignal wird mit Hilfe des PAL-Laufzeitdecoders in zwei Signale F_U und F_V zerlegt, d. h. in die Farbdifferenzkomponenten U und ± V, die noch auf den Farbhilfsträger moduliert sind. Am Eingang des Decoders wird das Signal auf zwei direkte Wege aufgeteilt wobei bei einem eine 180°-Phasendrehung vorgenommen wird (Abb. 7.12). Der dritte Weg für das Signal geht über ein Ultraschall-Laufzeitglied, das das Signal um die Dauer einer Zeile, bzw. um genau 64,056 µs, verzögert. Bei der Zusammenführung der Signale der verzögerten und unverzögerten Zeile fällt jeweils eine der beiden Komponenten weg, während die andere übrig bleibt. Im eigentlichen Synchrondemodulator wird zu F_U und F_V der Farbhilfsträger hinzugefügt und mit Hilfe zweier Gleichrichter ergeben sich schließlich die Farbdifferenzsignale B – Y und R – Y. Aus diesen und dem Luminanzsignal wird in einer Matrix dann das RGB-Signal gewonnen (Abb. 7.11). Über Verstärkerstufen werden diese Signalkomponenten zum Schluss zu den Kathoden der Bildröhre geführt.

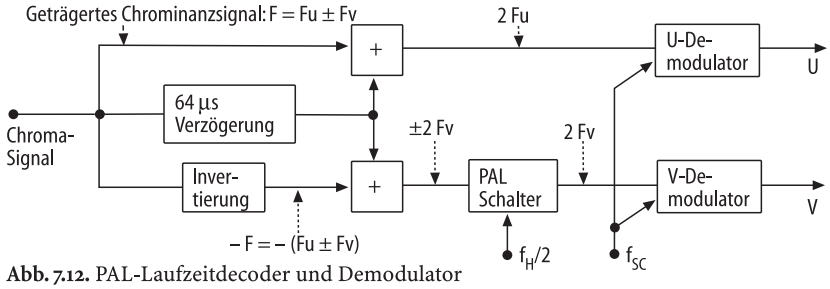


Abb. 7.12. PAL-Laufzeitdecoder und Demodulator

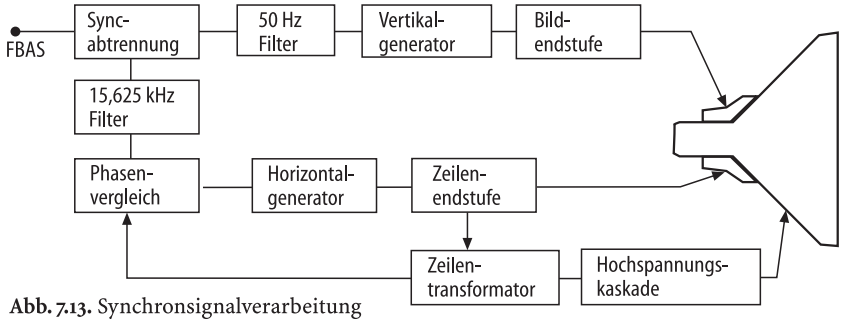


Abb. 7.13. Synchronsignalverarbeitung

7.1.3.1 Synchronsignalverarbeitung

Im sogenannten Amplitudensieb (Sync-Abtrennung) werden die Synchronimpulse vom Videosignal getrennt und das H- und V-Synchronsignal gewonnen (Abb. 7.13). Die Vertikalsynchronimpulse synchronisieren einen Bildgenerator, der einen sägezahnförmigen Ablenkstrom mit der Frequenz 50 Hz erzeugt, der wiederum über die Bildendstufe den Bildablenkspulen zugeführt wird. In ähnlicher Weise synchronisieren die H-Synchronimpulse einen Zeilenoszillator, der die Zeilenendstufe ansteuert, in der dann ein sägezahnförmiger Ablenkstrom mit der Frequenz 15625 Hz entsteht. Dieser Strom wird zusammen mit Korrekturströmen durch die H-Ablenkspulen geschickt.

Das Videosignal ist mit einem Zeitfehler behaftet, dem Jitter, der besonders stark auftritt, wenn Signale aus einfachen Videorecordern wiedergegeben werden. Videomonitore und TV-Geräte sind in der Lage, auch in diesem Fall ein ruhiges Bild zu zeigen, da sie über einen eigenen Ablenkoszillator verfügen. Der Oszillator wird über eine Regelschaltung mit einer gewissen Zeitkonstante quasi an den Mittelwert der schwankenden Synchronimpulse angepasst.

Das Signal aus der Zeilenendstufe wird auch zur Ansteuerung des Zeilentransformators verwendet. Im Zeilentransformator wird die hohe Anodenspannung zur Elektronenstrahlbeschleunigung erzeugt, indem zunächst die Spannung auf ca. 10 kV_{ss} hochtransformiert und dann über eine Hochspannungskaskade auf den gewünschten Wert gebracht wird (Abb. 7.14). Wegen der hohen Spannungen, die auch im ausgeschalteten Zustand einige Zeit bestehen bleiben, bergen Videomonitore ein Gefahrenpotenzial. Außerdem haben die Monitorchassis manchmal keine galvanische Netztrennung, d. h. es besteht die Gefahr, dass Netzspannung an den Metallrahmen liegt.

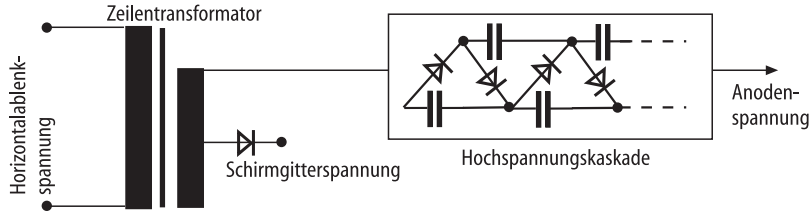


Abb. 7.14. Zeilentransformator und Hochspannungskaskade

7.2 Flüssigkristall-Displays

Flüssigkristallbildschirme (Liquid Crystal Display, LCD) sind heute die dominanten Videodisplays. Ab etwa dem Jahre 2000 verdrängten sie die Röhrenmonitore im Bereich der Computerdisplays. Einige Jahre später konnte die Bewegungsauflösung verbessert werden, sodass sie videotauglich wurden. Bereits ab ca. 2005 wurden im Heimanwenderbereich mehr LC- als CRT-Displays verkauft. Zu dieser Zeit begannen die LCD, die Röhre auch im Bereich der Video- und TV- Produktion zu verdrängen. Sie wurden zunächst für minderqualitative Zwecke eingesetzt, bis schließlich ab 2009 auch LCD für höchstqualitative Anwendungen, also als Referenzmonitore, verfügbar waren.

7.2.1 LCD-Funktionsprinzip

Flüssigkristallbildschirme oder -projektoren sind passiver Art. Sie strahlen nicht selbst Licht ab, sondern wirken als Lichtventile, dunkeln also eine Hintergrundbeleuchtung mehr oder weniger ab oder bewirken eine veränderliche Auflichtreflexion, indem sie die Intensität des transmittierenten Licht verändern. Zur Schwächung des Lichts wird der Umstand ausgenutzt, dass polarisiertes Licht, d. h. Licht dessen elektrischer Feldvektor stets in einer einzigen Ebene schwingt, nur dann eine Polarisationsfolie (Analysator) durchdringt, wenn diese parallel zur Schwingungsrichtung der Lichtwelle liegt.

Wenn nun das Hintergrundlicht wiederum mittels einer Polarisationsfolie (Polarisator) gewonnen wird, muss diese parallel zur Polarisationsrichtung des Analysators liegen, damit Licht transmittiert wird. Stehen die Polarisatoren senkrecht zueinander, wird das Licht gesperrt (Abb. 7.15).

Die Flüssigkristalle werden dazu benutzt, die Polarisationsebene des durch die Kristalle laufenden Lichts zu drehen. Die Flüssigkristallschicht wird zwischen zwei Glasplatten gebracht, auf denen außen auch die Polarisationsfolien angebracht werden. Zunächst wird dafür gesorgt, dass die Polarisationsebene des Lichts durch die Kristalle um genau 90° gedreht wird, sodass bei parallel stehenden äußeren Polarisationsfolien das Lichtventil im Idealfall völlig ge-

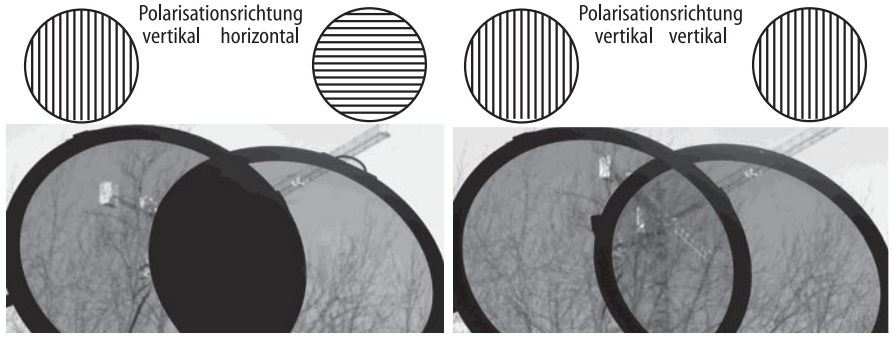


Abb. 7.15. Absorption und Transmission bei senkrecht und parallel stehenden Polarisatoren

sperrt wird (Normally Black Mode). Bei senkrecht zueinander stehenden Folien kann das Licht gerade passieren (Normally White Mode).

Flüssige Kristalle sind Materialien, die sich in einer mesomorphen Phase befinden, d. h. sie zeigen in einem Temperaturbereich um 300 K, d. h. bei Raumtemperatur, sowohl Eigenschaften von Kristallen als auch von Flüssigkeiten. Die Moleküle dieser Flüssigkeiten können sich in verschiedenen Zuständen (Phasen) befinden. In der hier wichtigen nematischen Phase weisen sie eine langgestreckte Form auf und haben die Eigenschaft, einerseits in der Flüssigkeit beweglich zu sein und andererseits eine regelmäßige Struktur zu bilden, indem sie sich aneinander ausrichten.

Durch das Aufbringen von sehr feinen Rillenmustern (Ausrichtungsschichten, Alignment Layer) auf die Innenseiten der beiden Glasplatten werden die Moleküle dazu gezwungen, sich parallel zu den Rillen auszurichten und die Richtung in der Flüssigkeit bis zur gegenüberliegenden Scheibe beizubehalten. Dadurch, dass das Rillenmuster einer Scheibe dann um 90° gegenüber dem anderen verdreht wird und sich die Moleküle weiterhin aneinander ausrichten, wird bewirkt, dass sich der Molekülfaden innerhalb der Flüssigkeit schraubenförmig um 90° dreht. Daher rührt die Bezeichnung Twisted Nematic (TN). Wenn nun zusätzlich dafür gesorgt wird, dass das einfallende Licht in Richtung des Rillenmusters polarisiert ist, der Polarisator also parallel zum Muster liegt, wird die Lichtpolarisation über alle Moleküle weitergegeben, sodass sich insgesamt die Polarisationsrichtung auch um 90° dreht. Bei parallel stehenden äußeren Polarisationsfolien bleibt so die Zelle dunkel (Abb. 7.16 links).

Die Steuerung des Lichtventils wird nun durch den Umstand ermöglicht, dass sich die Moleküle nicht nur mechanisch an den Rillen, sondern auch durch ein äußeres elektrisches Feld ausrichten lassen (Abb 7.16. rechts). Bei einer Spannung von ca. 5 V werden die Moleküle in Richtung des Feldes gedreht. Die Schraubenform wird aufgebrochen und eine Normally-Black-Zelle

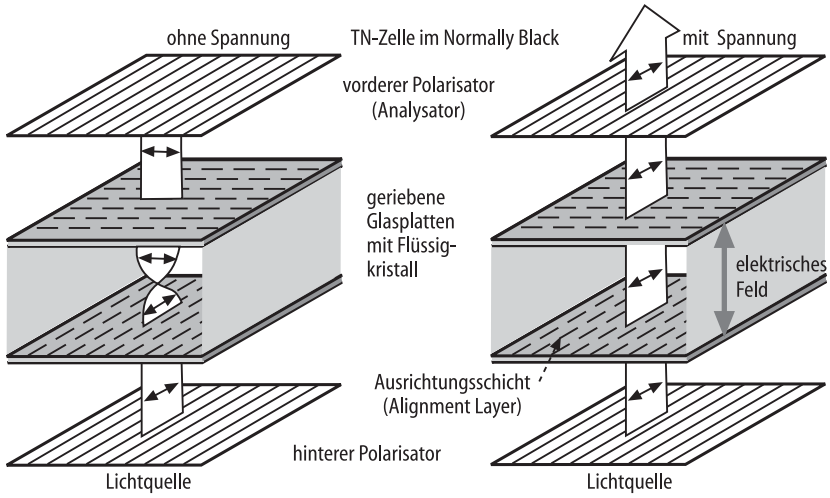


Abb. 7.16. Twisted Nematic LCD-Funktionsprinzip

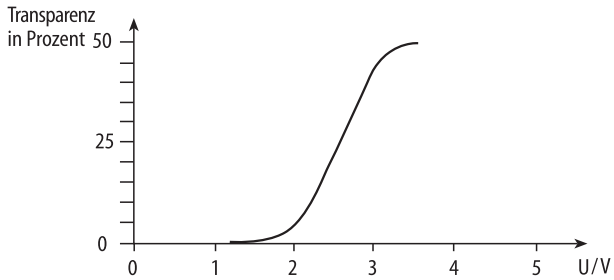


Abb. 7.17. Beispiel einer Steuerkennlinie beim LCD

transmittiert Licht. Mit Hilfe der Spannung kann also die Lichttransmission ein- und ausgeschaltet werden, eine Funktion, die z. B. bei elektronisch gesteuerten Shutterbrillen zur Betrachtung von 3D-Projektionen genutzt wird.

Die Ausrichtung der Moleküle im elektrischen Feld resultiert daher, dass das Feld die Verteilung der Ladungen innerhalb des Moleküls verändert, die Ladungen werden getrennt und es entsteht ein Dipol. Falls ein Molekül nicht in Richtung des äußeren Feldes liegt, wird es durch die Wechselwirkung zwischen Dipol und Feld wieder in die Feldrichtung gedreht. Allerdings gelingt das nicht vollständig, denn zusätzlich zur Spannung wirkt die Kraft an der Ausrichtungsschicht, sodass die Ausrichtung am Feld vor allem in der Mitte der Flüssigkristallschicht stattfindet und zu den Rändern hin abnimmt. Sie ist von der Höhe der angelegten Spannung abhängig, jedoch nicht von der Polarität. Daher kann über die Spannungsvariation auch die Lichttransmission und damit der Grauwert gesteuert werden. Der Betrieb erfolgt mit Wechselspannung, da sich die Flüssigkristallschicht bei Gleichspannung zersetzen kann. Als Wechselfrequenz kann z. B. die Halbbildfrequenz von 50 Hz gewählt werden.

Die Steuerkennlinie des LCD ist völlig verschieden von der Kennlinie einer Bildröhre (Abb. 7.17). Da das Videosignal über die Gamma-Vorentzerrung der Röhre angepasst ist, liegt bei LCD eine Fehlanpassung vor, die elektronisch für die LCD-Kennlinie ausgeglichen werden muss. Um dabei eine sehr gute Farbdarstellung zu erreichen, die mit der von CRT-Monitoren konkurrieren kann, wird bei modernen Displays die Kennlinie mit Hilfe von speziellen Look up Tables (LUT) angepasst, die ggf. eine nichtlineare Verteilung der RGB-Werte enthält.

Aufgrund der Tatsache, dass die Moleküle in der Nähe der Ausrichtungsschichten in deren Richtung verharren, ist eine optimale Zellwirkung nur bei verdrehtem Flüssigkristallfaden, also im spannungslosen Zustand gegeben. Mit einer Steuerspannung ergibt sich ein Gemisch aus den Wirkungen der Ausrichtungsschicht und der Spannung. Daraus folgen unterschiedliche Vorteile für den Normally White und den Normally Black Mode. Im ersten Fall wird das Optimum der Lichttransmission erreicht, die Zelle wird durch die Spannung abgeschaltet, was aber nur unvollkommen gelingt. Umgekehrt ist bei Normally Black die Transmission nicht optimal, dafür gelingt das Abschalten der Zelle besser. Daher erzielt man bei diesem Modus den besseren Schwarzwert und damit den höheren Kontrast.

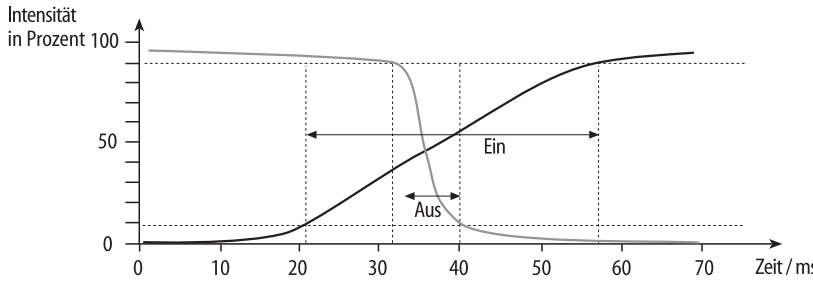


Abb. 7.18. Beispiel für die Reaktionszeit beim Ein- und Auschalten der Spannung am LCD

Für den Einsatz von LCD im Videobereich sind die Reaktionszeiten der Bildpunkte von großer Bedeutung, Ein Aspekt, der zu Beginn der LCD-Entwicklung noch recht problematisch war. Die Schaltzeiten eines LCD sind proportional zu g und d^2 und sinken quadratisch mit der Steuerspannung. Das bedeutet, dass eine hohe Schaltgeschwindigkeit durch eine geringe Viskosität g und eine möglichst kleine Schichtdicke d erreicht werden kann. Abbildung 7.18 zeigt, dass das Schaltverhalten unsymmetrisch ist. Die Drehung der Moleküle durch das Anlegen der Steuerspannung dauert länger als die Rückkehr in den Grundzustand. Die Verkürzung der Einschaltzeit kann hier durch eine kurzzeitige Überhöhung der Einschaltspannung erzielt werden.

Die Farbdarstellung bei LC-Displays wird mit Hilfe von Farbfiltern erreicht, die bei weißem Hintergrundlicht die Grundfarben RGB transmittieren. Ein Bildpunkt wird aus drei nebeneinander liegenden Pixeln gebildet (Abb. 7.19). Dazu wird das Glassubstrat bei der Herstellung mit einem undurchsichtigen Gitter aus Streifen von einigen Mikrometer Dicke versehen. Die Zwischenräume in dieser sogenannten Black Matrix bestimmen die Größe der Bildpunkte und dort werden die Farbfilter platziert (Abb. 7.19). Eine Schutzschicht darüber gleicht Unebenheiten aus, anschließend folgt die Elektrode für die Zuführung der Steuerspannung. Diese wird aus Indium-Zinn-Oxid (ITO) hergestellt, da sie transparent sein muss.

Das hier beschriebene LCD-Grundprinzip wird als TN bezeichnet. Aufgrund der strengen Ordnung der LC-Moleküle und der damit verbundenen Lichtrichtung ist ein optimales Bild nur bei senkrechter Betrachtung eines LCD gegeben. Bei abweichendem Blickwinkel sinkt der Kontrast. Da die Sperrwirkung des Lichtventils dabei vermindert ist, erscheinen dunkle Bildbereiche aufgehellt und zwar für jeden Farbauszug unterschiedlich, sodass das Schwarz, abhängig vom Blickwinkel, auch verschieden gefärbt erscheinen kann. Im Videobereich ist allerdings ein relativ großer Betrachtungswinkel erforderlich. An hochwertigen, sogenannten Grade 1-Monitoren müssen farbverbindliche Bildbeurteilungen auch von Personen möglich sein, die nebeneinander sitzen. Dies ist der Grund, warum die LCD erst relativ spät im Produktionsbereich Verwendung fanden. Erst im Laufe der Zeit gelang es, u. a. durch Technologien wie MVA (Multi-Domain Vertical Alignment), PVA (Patterned Vertical Alignment), IPS (In Plane Switching) etc. den Schwarzwert farbunabhängig niedrig bzw. den Kontrast ausreichend über einen großen Betrachtungswinkel hoch zu halten.

7.2.2 Pixelansteuerung und -beleuchtung

Das beschriebene LCD-Funktionsprinzip gilt für jeden Bildpunkt, der sich automatisch an der Stelle bildet an dem die Spannung liegt. Die Bildpunkte müssen nicht als separate Zellen konstruiert werden. Jedem Bildpunkt auf der Bildschirmfläche muss einfach eine Steuerspannung zugeführt werden. Dazu kann eine aktive oder passive Matrix benutzt werden. Die passive Matrix wird aus Leiterbahnen gebildet, die horizontal auf der unteren und vertikal auf der oberen Glasplatte aufgedampft sind. Die Bildpunkte einer Zeile werden einerseits durch die Auswahl der richtigen Zeile angesprochen. Wenn andererseits auch die zugehörige Spalte adressiert wird, bildet sich am Kreuzungspunkt das elektrische Feld. Dieses Verfahren wird auch bei anderen Displays eingesetzt und führt dazu, dass im Gegensatz zur Bildpunktadressierung des CRT-Displays hier eine zeilenweise Adressierung vorliegt.

Mit einer einfachen, passiven Matrix ist nur eine geringe Bewegungsauflösung möglich. Ein weiterer Nachteil ist, dass das elektrische Feld nicht eng genug auf den gewählten Kreuzungspunkt beschränkt bleibt, sondern sich auch entlang der Leiterbahn ausbildet, sodass Störlinien sichtbar werden können. Diese Nachteile können durch die Verwendung von Activ-Matrix-Displays (AM-LCD) vermieden werden. Eine aktive Matrix enthält an jedem Matrix-Kreuzungspunkt einen eigenen Feldeffekttransistor, der dort das Feld gezielt ein- und ausschaltet. Die Transistoren werden direkt auf der Glasschicht aufgebracht. Sie werden daher in Dünnschicht-Technik hergestellt, woraus die Bezeichnung TFT-Display (Thin Film Transistor) resultiert. Anstelle der Transistoren kann die LC-Ansteuerung auch mit Hilfe der im nächsten Abschnitt näher erläuterten Plasma-Entladung realisiert werden. Es entsteht dann das sogenannte plasmaadressierte LCD (PALC, auch Plasmatron) das heute aber keine Bedeutung mehr hat. Für den LCD-Einsatz im Videobereich ist die TFT-Ansteuerung unverzichtbar.

Abbildung 7.19 zeigt den Aufbau einer LCD-Zelle mit TFT-Ansteuerung. Damit der LCD-Zustand bis zur nächsten Ansteuerung erhalten bleibt, wird die von den Zellen gebildete Eigenkapazität ggf. durch Hinzufügen eines Dünnfilmkondensators erhöht. Alle TFT einer Zeile sind mit dem Gate an eine gemeinsame Zeilenleitung angeschlossen. Welches Pixel in welcher Form ange-

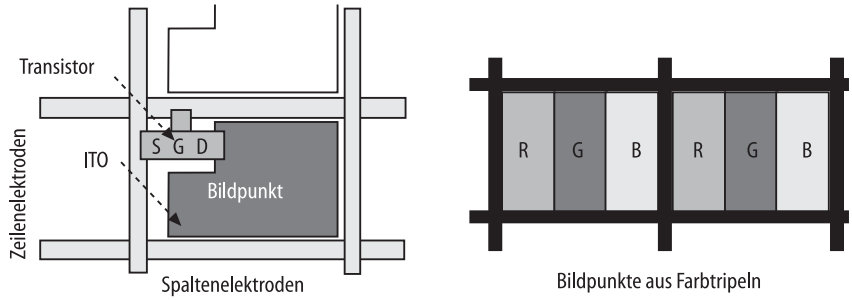


Abb. 7.19. Aktive Matrixansteuerung und Farbtripel

sprochen wird, hängt dann von den Informationen an den Spaltenelektroden ab, die mit den Source-Anschlüssen der TFT verbunden sind. Wenn der Transistor über die Zeilen/Spaltenkombination angesteuert wird, wird die Source-Drain-Strecke niederohmig. Am Drain liegt die Elektrode aus lichtdurchlässigem Indium Zinnoxid (ITO), über die der Flüssigkristall die Steuerspannung bekommt. Die Größe dieser Bildpunktelektrode bestimmt die Größe des Bildpunktes [142].

Die Adressierung geschieht auch für die Farbwerte zeilenweise. Alle RGB-Werte einer Zeile werden über die Gates mit Spannungen um + 10 V eingeschaltet, wenn die entsprechende Zeile selektiert ist. Die Gates der anderen Zeilen werden dabei auf negatives Potenzial von z. B. – 20 V gelegt, damit sie nicht angesprochen werden. Die Informationen jeder Zeile werden über die Bilddauer von z. B. 1/60 Hz gehalten. Aufgrund dieser Tatsache gehört das LCD zu den sogenannten Hold Type Displays, die sich in dieser Beziehung erheblich von den CRT-Displays unterscheiden, die ja bildpunktweise reagieren [68].

Das unterschiedliche Verhalten führt dazu, dass LCD besondere Bewegungsartefakte aufweisen. Bei horizontal bewegten Objekten befinden sich z. B. die zuerst adressierten Konturen im oberen Bildbereich und sind weiterhin sichtbar, auch wenn die unten adressierten Konturen seitlich verschoben sind. Das Auge versucht, solchen Strukturen zu folgen und wird irritiert. Meist erscheint ihm eine unscharfe Kontur (Motion Blur). Dieser Effekt ist schwer zu minimieren. Ein Ansatz ist, mit hohen Bildwiederholfrequenzen von z. B. 120 Hz zu arbeiten. Dabei kann jedes zweite Bild schwarz geschaltet werden und es ergibt sich eine ähnliche Wirkung wie bei einer Flügelblende. Weiterhin besteht die Möglichkeit, das Hintergrundlicht periodisch ein- und auszuschalten und so eine höhere Bewegungsauflösung vorzutäuschen. Für diese Funktion ist eine LED-Hintergrundbeleuchtung erforderlich, da gewöhnliche Beleuchtungen zu langsam reagieren (s. u.) [140].

Die LCD-Adressierung Zeile für Zeile bewirkt eine progressive Bilddarstellung. Interlaced-Bilder mit Zeilensprung müssen daher gewöhnlich vor der Darstellung de-interlaced werden, was nur bei hohem Aufwand mit guter Qualität gelingt. Dies ist u. a. ein Argument dafür, dass HD-Bilder mit progressivem Bildaufbau favorisiert werden (z. B. 720p statt 1080i). Eine echte Darstellung im Zeilensprungverfahren gelingt aber auch bei LCD. In diesem Falle muss halbbildweise wechselseitig jede zweite Zeile schwarz gesteuert werden. Diese Funktion wird bei Referenzmonitoren erforderlich, von denen erwartet wird, dass sie jede Signalform weitestgehend unverändert anzeigen.

Die Hintergrundbeleuchtung muss mit hoher Lichtintensität erfolgen, denn eine LC-Zelle transmittiert insgesamt nur ca. 6% des Lichts. Wenn ein LCD eine Leuchtdichte von 300 cd/m² erzeugen soll, muss die Hintergrundbeleuchtung also ca. 5000 cd/m² erreichen. Die hohen Transmissionsverluste setzen sich multiplikativ aus folgenden Faktoren zusammen: Zunächst halbiert der Polarisator auf der Leuchtfläche die Lichtmenge, da er eine Schwingungsrichtung der Welle sperrt. Zusätzlich dämpft er, ebenso wie der Analysator das Licht auf ca. 80%. Die Flüssigkristallschicht lässt ca. 95% des Lichts durch, aber die lichtwirksame Öffnung beträgt im Verhältnis zur Gesamtpixelfläche nur ca. 80% und schließlich wirken die Filter mit einer Transmission von ca. 25%.

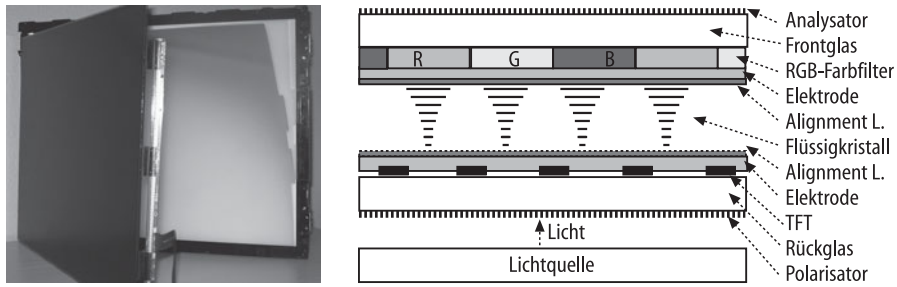


Abb. 7.20. LCD geöffnet und in Seitenansicht

Für die Hintergrundbeleuchtung (Backlight) in LCD reicht es in den meisten Fällen aus, eine gleichmäßige weiße Ausleuchtung zu erzielen, die mit geringer Wärmebelastung verbunden ist. Es kommen dafür Cold Cathode Fluorescent Lamps (CCFL) zum Einsatz. Sie werden z. B. an den Seiten des Displays platziert und das Licht wird über diffuse Kunststoffflächen verteilt. Diese ist in Abbildung 7.20 links zusammen mit den Ansteuerelektroden sichtbar. Rechts sind alle Schichten des LCD in Seitenansicht dargestellt.

Aufgrund der in den letzten Jahren stark gestiegenen Lichtausbeute von Leuchtdioden (LED), ist es auch möglich, eine Vielzahl von LED-Kombinationen in den Farben Rot Grün und Blau als Hintergrundbeleuchtung einzusetzen. Das bietet gegenüber den CCFL den Vorteil, dass das Backlight abgeschaltet oder geschwächt und so die Bilddarstellung beeinflusst werden kann. Durch Steuerung der RGB-Werte können auch Farbwertverschiebungen und Anpassungen an verschieden Farbräume erzielt werden [138].

LCD zeichnen sich durch geringe Betriebsspannungen von 3 V bis 15 V und geringen Leitungsbedarf von ca. 0,1 mW pro cm² Bildschirmfläche aus. Die Bilder sind über die gesamte Fläche gleichmäßig scharf, Konvergenzprobleme treten nicht auf. Die Pixeldichte kann sehr groß sein, d. h. es sind kleine, hoch auflösende Displays herstellbar, es ist aber auch möglich, große Bilddiagonalen von mehr als 60“ (1,5 m) und sogenannte Full-HD-Auflösung mit 1920 x 1080 Pixeln, bei Leuchtdichten über 500 cd/m² zu erreichen. Die Herstellung von TFT-Farbdisplays ist allerdings aufwendig, denn die Transistoren müssen sehr ähnliche Kennlinien aufweisen, und kein einzelner Transistor darf ausfallen.

Als LCD-Weiterentwicklung wird mit ferroelektrischen Flüssigkeiten (FLC) gearbeitet, bei denen sich die Moleküle statt in der nematischen in der smektischen Phase befinden und geneigt zum Bildschirm stehen. FLC haben die Eigenschaft der Bistabilität, d. h. ein Bild bleibt auch ohne äußere Spannung solange erhalten, bis erneut eine Steuerspannung angelegt wird. Das Verhalten ähnelt dem von magnetisiertem Eisen, woher auch die Bezeichnung rührt. Ein weiterer Vorteil ist dass die Schaltzeiten bei FLC sehr kurz sind. Weitere LCD-Entwicklungen beziehen sich darauf, das Glas als Träger des Flüssigkristalls durch andere, flexiblere Materialien zu ersetzen. Hier wird insbesondere an polymerdispersierten Flüssigkeiten (PDLC) geforscht, bei denen die Lichttransmission nicht über die Polarisationsrichtung, sondern über einen veränderten Brechungsindex beeinflusst wird.

7.3 Selbstleuchtende Flachbildschirme

Nicht selbstleuchtende Flüssigkristallanzeigen sind prädestiniert für die Einzelbetrachtung, dort wo eine hohe Pixeldichte gefragt ist, wie bei Computer-Displays. Aufgrund der im Laufe der Zeit verbesserten Bewegtbilddarstellung eignen sie sich inzwischen auch für den Einsatz im Videobereich. Zur Fernsehbildwiedergabe mit flachen Bildschirmen werden aber auch aktive Anzeigen eingesetzt. Weit verbreitet ist hier das Plasma-Display, verstärkt kommt aber auch das OLED-Display zum Einsatz. Es sind auch andere Verfahren in der Diskussion, ihre Entwicklung ist aber noch nicht so weit fortgeschritten.

7.3.1 Plasmadisplays

Das Plasmadisplay (PDP) war das erste Videoflachdisplay das mit Bilddiagonalen von mehr als 1 m realisierbar war. PDP sind aktiv leuchtend, ein elektrisch leitendes Gas strahlt im elektrischen Feld bei hohen Feldstärken Licht ab. Für Farb-Displays wird Xenon verwendet, das im UV-Bereich strahlt. Wie auch bei anderen Lichtquellen werden Elektronen energetisch angeregt, die die aufgenommene Energie bei der Rückkehr in ihren Grundzustand in Form einer elektromagnetischen Welle abstrahlen. Die Anregungsenergie erfolgt wie bei einer Leuchtstofflampe durch Stoßionisation. Das bedeutet, dass bei elektrischen Spannungen von ca. 300 V Elektronen von ihren zugehörigen Atomen getrennt werden. Die Gasionen werden stark beschleunigt, treffen auf andere Gasatome und trennen durch die Stoßenergie dort weitere Elektronen ab. Insgesamt entsteht ein Lawineneffekt, man sagt, das Gas sei gezündet. Zu Erhaltung des Zustands ist anschließend nur noch eine kleinere Spannung erforderlich. Das Gemisch aus Gasionen und abgetrennten Elektronen wird als Plasma bezeichnet. Die abgetrennten Elektronen kehren immer wieder spontan in ihren Grundzustand zurück und können dann wieder angeregt werden. Bei der Rückkehr in den Grundzustand wird die Anregungsenergie als UV-Licht abgestrahlt, das anschließend in sichtbares Licht umgesetzt wird.

Jedem Bildpunkt sind drei Subpixel und diesen wiederum je eine einzelne Glaszelle zugeordnet (Abb. 7.21) [132]. Die Innenwände der Zellen sind mit

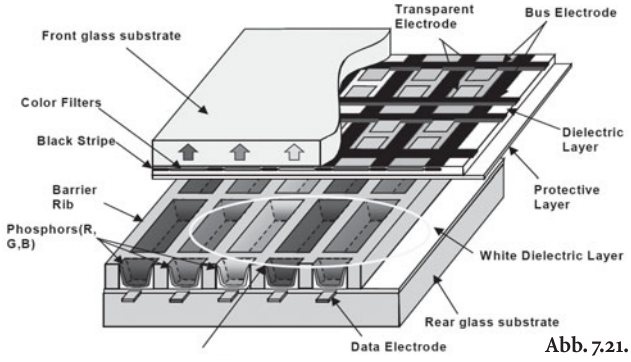


Abb. 7.21. Aufbau des Plasmadisplays

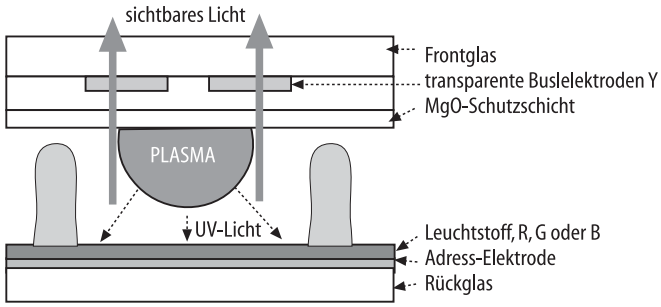


Abb. 7.22. Plasmazelle mit Entladung zwischen nebeneinanderliegenden Elektroden

rot, grün oder blau leuchtenden Phosphoren beschichtet, die vom unsichtbaren UV-Licht zur Abstrahlung von Licht der genannten Farben angeregt wird, wobei ein Bildpunkt wieder aus einem RGB-Farbtripel besteht.

Das Gasgemisch eines einzelnen Bildpunktes kann gezündet werden, indem einfach eine Gleichspannung an die entsprechenden Spalten und Zeilen der Bildpunkt am Kreuzungspunkt angelegt wird. Diese Art der unipolaren Ansteuerung ist wenig aufwendig und wurde beim DC-PDP eingesetzt. Der Nachteil ist, dass die Elektroden einem hohen Verschleiß unterliegen, da sie direkt dem heißen Plasma ausgesetzt sind. Daher wird heute das bipolare AC-Verfahren angewandt. Dabei werden die Elektroden durch ein Isolationsmaterial geschützt und bilden einen Kondensator, der eine aufwendigere Wechselspannungs-Ansteuerung erfordert.

Nicht nur die Elektroden, sondern auch die Leuchtstoffe müssen vor dem Plasma geschützt werden Dazu wurde die Elektrodenanordnung so modifiziert, dass die Entladung nicht mehr zwischen den gegenüberliegenden Elektroden sondern zwischen nebeneinander liegenden stattfindet und das Plasma die Leuchtstoffe auf der anderen Seite nicht direkt erreicht.

Die Ansteuerung geschieht in drei Phasen mit Hilfe zweier paralleler Busleitungen Y1 und Y2. In der Initialisierungsphase werden zunächst alle Zellen über Y1 vorgeladen, die im nächsten Vollbild aktiv sein sollen (Abb. 7.22). Diese Adressierung ist mit einem Memory-Effekt verbunden, der eine zeilenweise Adressierung ermöglicht. In der Displayphase wird eine so hohe Wechselspannung zwischen die Buselektroden gelegt, dass die adressierten Pixel zum Leuchten gebracht werden. Durch die Spannungsinversion pendeln die Elektronen zwischen den Buselektroden und es kommt zur Plasmabildung mit der Emission von Lichtimpulsen. Nicht adressierte Pixel bleiben dunkel, da hier die äußere Spannung als Zündspannung nicht ausreicht. Trotz zeilenweiser Adressierung erfolgt die Darstellung damit vollbildweise. Die dritte Phase ist die Löschphase, in der für das gesamte Bild schlagartig wieder ein neutraler Zustand hergestellt wird.

Da ein Pixel im PDP entweder gezündet ist oder nicht, kann durch eine Variation der Steuerspannung keine Graustufendarstellung erzielt werden. Daher wird die Dichte der Impulse in der Displayphase (bis zu 400 Impulse in 20 ms) variiert, sodass dem Auge die Zelle heller oder dunkler erscheint.

Plasmadisplays sind für großflächige Darstellungen prädestiniert, die Zellen lassen sich aber schlecht verkleinern. Sie erreichen hohe Leuchtdichten. Gegenüber LCD ist von Vorteil, dass der Blickwinkel sowohl horizontal als auch vertikal bei über 160° liegt. Das schnelle Schalten der Pixel verhindert Großflächenflimmern. Als Nachteile sind die relativ hohen Spannungen und Leistungen zu nennen, die zur Zündung verfügbar sein müssen, sowie die mechanische Empfindlichkeit. Teilweise werden die Geräte mit Lüftern ausgestattet, die eine Geräuschbelästigung mit sich bringen. Bei lang gehaltenen statischen Bildinhalten kann es zur Einbrenneffekten kommen, d. h., dass die Struktur auch nach Abschalten des Signals noch sichtbar ist. Als Gegenmaßnahme wird der Bildinhalt vom Monitor bei einigen Ausführungen langsam hin- und her bewegt. Typisch für PDP ist auch die Regelung der Leuchtdichte in Abhängigkeit von der Größe der hellen Fläche. Wenn diese nur 1% der sonst schwarzen Bildfläche beträgt, kann sie in einem Messbeispiel 360 cd/m² erreichen. Dieser Wert sinkt auf z. B. 140 cd/m² wenn 50% der Fläche hell ist. Die Helligkeitssteuerung mittels der Variation der Einschaltdauer führt außerdem zu Bewegungsartefakten. Die genannten Aspekte lassen das PDP als Monitore im Produktionsumfeld für die meisten Zwecke als ungeeignet erscheinen, dagegen können sie ihre Vorteile als Präsentationsdisplays auf Messen etc. ausspielen.

7.3.2 OLED-Displays

Auch durch eine flächige Anordnung rot-, grün- und blau-leuchtender Leuchtdioden (LED) lässt sich ein Flachbildschirm aufbauen. Bei Verwendung gewöhnlicher Leuchtdioden haben diese als diskrete Bauelemente jedoch sehr große Abstände und die Auflösung ist sehr schlecht, bzw. es werden sehr große Flächen und sehr große Betrachtungsabstände benötigt, um die Pixelstruktur unsichtbar werden zu lassen.

Die Weiterentwicklung erlaubt es heute, LED mit organisches Stoffen (Organic Light Emitting Diode, OLED) herzustellen, sodass sie sehr kleine Strukturen aufweisen und nicht aus separaten Einzeldioden bestehen. Man unterscheidet hier Small-Molecule-OLED (SM-OLED) und Polymer-OLED (P-OLED oder PLED). Die Entwicklung der OLED verspricht viele Vorteile. Insbesondere durch die Tatsache, dass es sich, wie bei der Kathodenstrahlröhre, um ein aktiv Licht abstrahlendes Displayprinzip handelt, ist es möglich, sehr hohe Kontraste und auch große Betrachtungswinkel zu realisieren.

Der Aufbau ist einfach: Zwischen zwei Elektroden befindet sich eine sehr dünne Schicht aus organischen Leuchtstoffen. Das ist der Emitter Layer (EL), der das Licht abstrahlt. Die Schichtdicke liegt im Nanometerbereich. Eine der beiden Elektroden muss lichtdurchlässig sein und besteht meist aus Indium-Zinn-Oxid (ITO). Die Gegenelektrode besteht aus einem Metall mit niedriger Austrittsarbeit, wie z. B. Aluminium. Diese, und auch der organische Leuchtstoff selbst, sind sehr empfindlich gegen Umwelteinflüsse, wie z. B. Feuchtigkeit, daher müssen die OLED gut verkapselt werden. Die Leuchtschicht und die Elektroden werden meist auf Glas als Träge aufgebracht, es können aber auch andere, sogar biegsame Trägermaterialien verwendet werden [121].

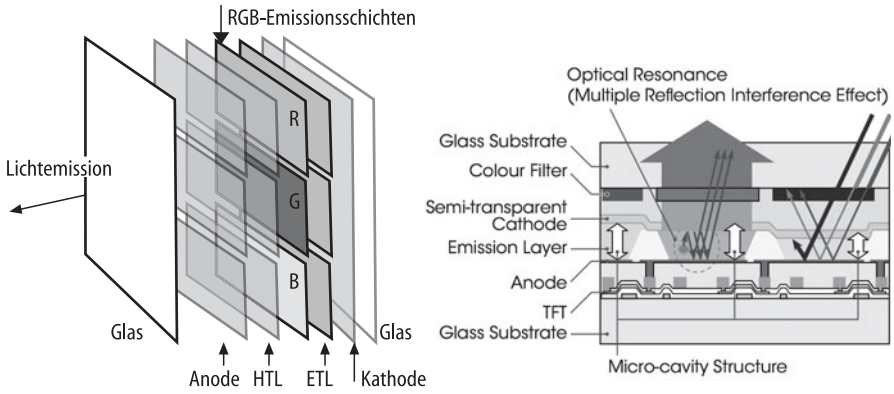


Abb. 7.23. Aufbau von OLED-Bildpunkten, links als BES- und rechts als TES-OLED [208]

Die eigentliche Lichtemission des OLED wird als Elektrolumineszenz bezeichnet. Wenn eine Spannung an den Elektroden liegt werden positive und negative Ladungsträger (Löcher und Elektronen) gebildet, wofür die einfache Schichtanordnung durch Lochleiter- bzw. Elektronenleiter-Schichten (Hole und Electron Transport Layer, HTL bzw. ETL) erweitert wird (Abb. 7.23). Die Ladungsträger bewegen sich unter dem Einfluss der Spannung in Richtung der jeweiligen Gegenelektrode und wenn sie aufeinandertreffen, rekombinieren die negativ geladenen Elektronen mit den positiven Löchern. Sie bilden dabei ein Exciton in einem energetisch angeregten Zustand. Dieser ist instabil und beim Übergang in den Grundzustand wird die Energie in Form einer elektromagnetischen Welle abgestrahlt. Durch die Wahl des OLED-Materials werden dabei die Energiebandabstände und damit die emittierte Wellenlänge bzw. die Farbe des Lichts bestimmt. Die gewünschten Farben können durch Dotierung der organischen Stoffe fein abgestimmt werden.

Bezüglich des Aufbaus wird zwischen BES-OLEDs (Bottom Emission Structure OLED) und TES-OLED (Top Emission Structure OLED) unterschieden. Bei TES-OLED ist die Anode opak und die transparente Kathode ermöglicht den Lichtaustritt (Abb. 7.23). Dieses Verfahren erlaubt eine relativ hohe Lichtausbeute. Bei BES-OLED ist die Anode transparent, das in der Nähe der Kathode erzeugt Licht muss die Steuerelektroden passieren, wofür relativ wenig Platz bleibt. Der Vorteil ist hier, dass diese Anordnung mit druckerbasierten Techniken relativ leicht herzustellen ist, während der Herstellungsprozess bei TES-OLED aufwendiger ist. Der Herstellungsaufwand bestimmt auch die Wahl zwischen SM-OLED mit den kleinen Molekülen und den P-OLED mit den langen Molekülen. Für letztere ist der Herstellungsprozess einfacher, da die Polymere wasserlöslich sind.

Die Pixelansteuerung erfolgt bei OLED-Displays im Prinzip ähnlich wie bei LCD, sie werden zeilenweise adressiert und gehören damit zu den Hold-Type-Displays mit ihren typischen Bewegungsartefakten. Ein wesentlicher Unterschied besteht aber darin, dass es für LCD ausreicht, mittels TF-Transistoren eine Spannung zu den Bildpunkten zu führen, während OLED stromgesteuerte

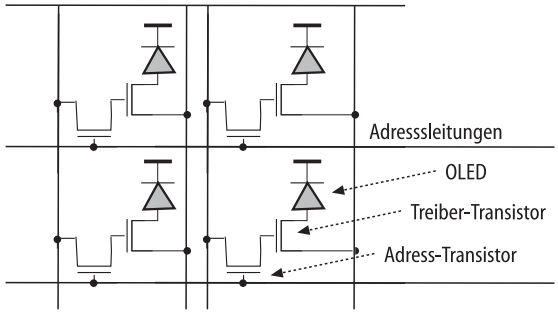


Abb. 7.24. Ansteuerung von vier OLED-Bildpunkten

Bauelemente sind, sodass für jeden Bildpunkt ein zweiter Transistor erforderlich wird, der die für die jeweilige Helligkeit erforderliche Stromstärke an die einzelne Diode liefert. Ein derartig angesteuertes Display wird AMOLED (Active Matrix-OLED) bezeichnet. Einen einfachen Aufbau zeigt Abbildung 7.24. Einer der Transistoren dient als Zugriffs- oder Adressierungstransistor. Dieser steuert den zweiten, den Treiber-Transistor, der den Stromfluss im LED regelt.

Farbig wird das Display, wie bei LCD, durch die Verwendung von drei Subpixeln in den Farben RGB pro Bildpunkt. Die Farben selbst sind durch die Wahl der emittierenden Materialien bestimmt (Abb. 7.23), es ist aber auch möglich, alle Bildpunkt weiß leuchten zu lassen und die Farbigkeit mit Hilfe von Farbfilterschichten vor den Subpixeln zu gewinnen.

OLED wurden über eine sehr lange Zeit erforscht, die Entwicklung hat sich in den letzten Jahren extrem beschleunigt. Die entsprechende Festkörper-Lichtemission wurde zum ersten Mal 1907 beobachtet, 1935 wurde sie als Elektrolumineszenz bezeichnet. In den 80er Jahren gelang es dann, die für die Emission erforderliche Spannung auf unter 10V zu senken, und 1990 entstand die erste P-LED. Im Jahre 1996 war das erste Display mit sehr kleiner Auflösung verfügbar, ab 2002 wurden gering aufgelöste Displays z. B. in Suchermonitoren eingesetzt. Ab 2011 war die Technologie so ausgereift, dass sie sogar in den professionellen Sektor Eingang fand. Die rasante Entwicklung wurde dabei erheblich durch die Weiterentwicklung der TF-Transistoren beeinflusst. Als stromgesteuerte Bauelemente müssen sie bei OLED höheren Anforderungen genügen, als bei LCD. Auch die Beseitigung der letzten Problembereiche hat viel mit der Optimierung der LED-Steuerung zu tun [121].

Tabelle 7.1. Veränderung des Kontrasts bei Variation des V- und H-Betrachtungswinkels

		CRT-Monitor	LCD-Monitor	OLED-Monitor
vertikal / horizontal				
+20°	+50°	1 : 1.660	1 : 22	1 : 51.050
	0°	1 : 1.950	1 : 567	1 : 92.680
	-50°	1 : 2.050	1 : 20	1 : 68.100
+15°	+30°	1 : 1.960	1 : 88	1 : 70.660
	0°	1 : 1.970	1 : 613	1 : 70.280
	-30°	1 : 2.040	1 : 86	1 : 76.040

OLED-Displays werden vor allem im professionellen Bereich immer mehr favorisiert. Sie sind prinzipiell gut für professionelle Anwendungen geeignet, da sie im Gegensatz zu LCD aktiv Licht abstrahlen und somit, wie die Bildröhre, das Licht auf einen breiten Raumbereich verteilen. Zudem kann die Lichtemission auf Null gesenkt werden, was einen extrem tiefen Schwarzwert und damit einen sehr hohen Kontrast ermöglicht. Das lässt sich auch messtechnisch zeigen. Tabelle 7.1 stellt dazu Kontrastwerte und ihre Betrachtungswinkelabhängigkeit an CRT-, LCD- und OLED Monitoren dar [121].

7.3.3 FED und SED

Um den hoher Kontrast und den großen Betrachtungswinkel zu erreichen, den CRT-Monitore bieten, befinden sich auch das Field Emission Display (FED) und das neuere Surface Conduction Electron Emitter Display (SED) in der Entwicklung. Sie arbeiten sehr ähnlich und basieren auf dem Funktionsprinzip gewöhnlicher Bildröhren, d. h. schnelle Elektronen regen bestimmte Phosphore in den Farben R, G, B zum Leuchten an. Im Unterschied zur Röhre ist hier aber jedem Pixel eine eigene Elektronenquelle zugeordnet (Abb. 7.25). Die Elektronen werden durch den quantenmechanischen Tunneleffekt aus einer sehr feinen Spitze emittiert. Das bedeutet, dass sie einen Potenzialwall, der sie eigentlich am Austritt hindert, durchdringen (tunneln) und austreten können. Eine Heizung zur Erzielung der Austrittsarbeit, wie sie bei CRT-Displays verwendet wird, ist nicht erforderlich, und damit wird zum Betrieb nur wenig Energie gebraucht. Jedem Pixel sind beim FED mehrere Tausend Spitzen zugeordnet. Die daraus austretenden Elektronen werden auf sehr geringer Distanz und somit durch ein starkes elektrisches Feld beschleunigt. Aufgrund der feinen Spitzen (Microtip) wird auch die Bezeichnung Microtip Fluorescent Display (MFD) verwendet. Die Spitzen haben Krümmungsradien unter 50 nm, sodass mit Spannungen von bis zu 400 V elektrische Feldstärken von 10 MV/cm erreicht werden, die für diese sogenannte kalte Elektronenemission ausreichen. Die Graustufendarstellung wird über die Pulsbreitenmodulation oder, wie bei CRT, über die Variation der Gitter-Kathodenspannung realisiert. Im Gegensatz zu OLED haben FED und SED in der Praxis gegenwärtig keine Bedeutung.

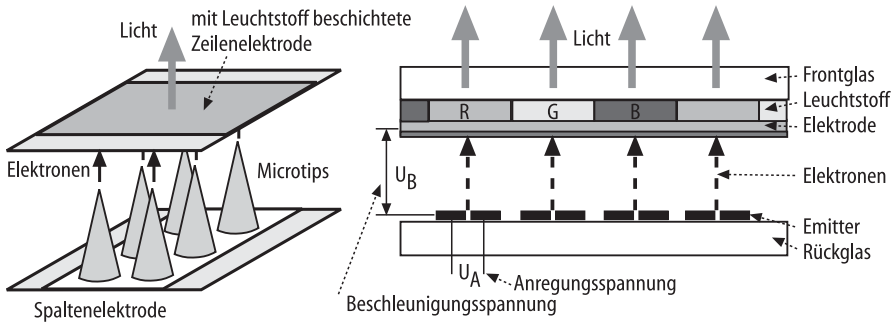


Abb. 7.25. Funktionsprinzip bei FED und SED

Tabelle 7.2. Betrachtungsabstände und Blickwinkel bei SD- und HD-Wiedergabe

Format	aktive Pixel	B/H	Betrachtungsabstand		Winkel vertikal	Winkel horizontal
SD	720 x 576	4/3	a = 6 H	a = 4,5 B	9,5°	12,7°
HD	1280 x 720	16/9	a = 4,5 H	a = 2,5 B	12,7°	22,4°
HD	1920 x 1080	16/9	a = 3 H	a = 1,7 B	18,9°	32,8°

7.4 Videomonitore

Monitore dienen allgemein der Bildwiedergabe, d. h. der Darstellung des Bildes, sowohl im Studio als auch als Fernsehgerät beim Zuschauer zuhause. Im Produktionsumfeld erfüllt der Monitor neben der Sicht- auch noch eine Kontrollfunktion bzw. ist die Referenz für die endgültige Bildqualität. In diesem Zusammenhang hat der Monitor die Funktion eines Messgerätes.

Bis etwa zum Jahre 2000 wurden die genannten Aufgaben ausschließlich von Röhrenmonitoren übernommen, die entsprechend mehr oder weniger aufwendig konstruiert waren Heute können verschiedene Wiedergabeprinzipien verwendet werden. Als Entscheidungskriterium für ein bestimmtes Display dienen folgende Qualitätsparameter: Helligkeit und Kontrast, bzw. Schwarz- und Weißwert, Grauwertübertragung (Gamma), Farbraum und Farbtemperatur am Weißpunkt sowie Farbreinheit, Betrachtungswinkelabhängigkeit, Bewegungsartefakte und die Verzögerungszeit bei der Darstellung (Latenz). Im professionellen Bereich ist zudem die Langzeitstabilität, die Art der Schnittstellen und die exakte Darstellbarkeit von i- und p-Signalformen von Bedeutung.

Abhängig von der Bildauflösung (SD oder HD) des jeweiligen Gerätes ist der Betrachtungsabstand zu wählen. SD-Videomonitore sind für große Betrachtungsabstände konzipiert, der Betrachtungsabstand a beträgt etwa das fünf- bis siebenfache der Bildhöhe H (s. Abschn. 2.2). Bei diesem Abstand ist das menschliche Gesichtsfeld nur zu einem kleinen Teil vom Bild erfüllt. Das ändert sich bei der Betrachtung von HD-Displays, die, insbesondere beim „großen“ HD-Format mit 1080 Zeilen, bei einem Betrachtungsabstand a von etwa dem dreifachen der Bildhöhe H ein kinoähnliches Sehgefühl hervorrufen.

Tabelle 7.2 zeigt einen Vergleich der Betrachtungsabstände und Blickwinkel für SD- und HD-Displays. Die Abstände in Relation zur Bildbreite B ergeben sich aus dem Bildseitenverhältnis. Obwohl im professionellen Umfeld, bei Displayvergleichen etc., auch kleinere Abstände, bis hinab zu a = 4 H, für SD-Signale verwendet werden, wird der Vergleich in der Tabelle in Bezug auf a = 6 H für Heimanwendungen vorgenommen, nicht zuletzt auch deshalb, weil im professionellen Umfeld oft die Sehbedingung im Heim vorweggenommen werden müssen. Die Berechnungen in der Tabelle erfolgten weiterhin unter der Annahme, dass auf üblichen SD-(Röhren-) Monitoren von den 576 aktiven Zeilen wegen des Overscan nur 540 wirklich sichtbar sind.

Seit der Einführung von HDTV müssen die Monitore heute Signale mit zwei Bildseitenverhältnissen verarbeiten, nämlich 4/3 und 16/9. Während im Heimbereich von den Geräteherstellern zur Anpassung alle möglichen Skalierungen bis hin zu Bildverzerrungen ermöglicht werden, müssen im professio-

nellen Umfeld klare Vorgaben und Sprachregelungen getroffen werden, wobei Bildverzerrungen natürlich nicht zulässig sind.

Die HD-Produktionen im 16/9-Format sind inzwischen Standard. Sie können auf HD-Displays unverzerrt dargestellt werden. Zur Anpassung an eine 4/3-SD-Darstellung für das analoge Fernsehen wird das Bild so verkleinert, dass die Bildbreite genau ausgefüllt ist. Zwangsläufig erscheinen dann beim 16/9-Programm auf 4/3-Displays schwarze Ränder am oberen und unteren Bildrand. Sie sind jeweils 72 Zeilen breit, denn nur 432, d. h. $576 \cdot 4/3 : 16/9$ sind im derart skalierten 16/9-Programm aktiv. Diese Darstellung von 16/9-Programmen für das analoge Fernsehen wird als 4/3-Letterbox bezeichnet [157].

Abbildung 7.26 zeigt, wie auch SD-Produktionen an die neuen 16/9-Dispalys angeglichen werden. Um das Problem von 4/3-Letterbox, nämlich die verminderte vertikale Auflösung, zu vermieden, wird das 4/3-Format sowohl vertikal als auch horizontal mit 16/9-Programmen gefüllt. Dabei muss das Bild horizontal gestaucht werden und es entstehen „Eierköpfe“. Das geschieht bereits bei der Produktion. Diese Form wird als 16/9-Vollformat bezeichnet, oder in Anlehnung an das Kino (Cinemascope) auch als anamorphotisch komprimiert. Die Verzerrung ist auf das Signal beschränkt und wird nicht dargestellt. Ihr Sinn ist nur, die volle vertikale Auflösung zu erhalten. Für die Darstellung wird das 16/9-Vollformat horizontal gedehnt und korrekt auf 16/9-Displays angezeigt. Dies gilt im Studio und auch bei digitalen Fernsehempfängern. Digitale Empfänger sind in der Lage, das Bild richtig zu skalieren. Damit dies automatisch geschieht, wird den Empfängern mittels eines Wide-Screen-Signalling-Signals (WSS) mitgeteilt, wann sie das Bild horizontal strecken sollen.

Das Begriffspaar 16/9-Vollformat Letterbox bezeichnet schließlich ein 16/9-Vollformat das ein Quellbild, z. B. aus Spielfilmen (Cinemascope), wiedergibt das im Original oben und unten bereits schwarze Streifen enthält.

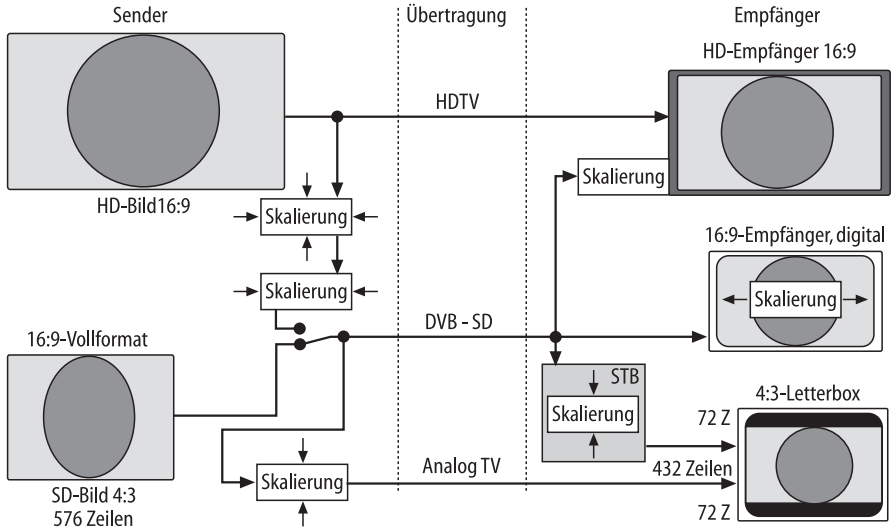


Abb. 7.26. Produktion und Empfang bei verschiedenen Bildformaten

Tabelle 7.3. Anforderungen an Studiomonitore nach EBU-Tech 3320

	Grade 1	Grade 2	Grade 3
Luminanzbereich	70 bis 100 cd/m²	70 bis 200 cd/m²	70 bis 250 (400) cd/m²
Schwarzpegel	< 0,1 cd/m²	< 0,4 cd/m²	< 0,7 cd/m²
Vollkontrast	> 1000 : 1	> 500 : 1	> 300 : 1
Simultankontrast	> 200 : 1	> 100 : 1	> 100 : 1
Betr.-Winkel Vertikal	±20°	±20°	
Betr.-Winkel Horiz.	±45°	±45°	
Farbtemperatur	D65	D65 opt. 3200K	D65 opt. 3200K
Unbuntwiedergabe	0.0010 Δu‘, 0.0015 Δv	0.003 Δu‘, 0.004 Δv	0.004 Δu‘, 0.006 Δv‘

7.4.1 Produktionsmonitore

Bei Monitoren für den Produktionsbereich gibt es Anforderungen, die in Empfehlungen der EBU (Tech 3320) beschrieben sind [52]. Hier wird in drei Klassen, nämlich nach Referenz-, Kontroll- und Beobachtungsmonitoren unterschieden, eingeteilt in Grade 1, 2 und 3.

Referenzmonitore (Grade 1) sind Geräte bestmöglicher Bildqualität. Sie stellen ein „Messgerät“ zur Bildbeurteilung dar, d. h. dass diese Monitore mindestens den Qualitätsstand der zu kontrollierenden Signale und Systeme entsprechen müssen. Im Gegensatz zu Heimgeräten dürfen sie Signal- oder Gerätemängel nicht verschleiern. Alle im Signal enthaltenen Fehler müssen dargestellt werden. Dem entsprechend muss auch der Abtastmodus eines Signals nativ wiedergeben können, d. h., dass bei Interlaced-Signalen ein bei LCD- oder OLED typischerweise vorhandenen De-Interlacer umgangen und das Signal interlaced dargestellt wird. Referenzmonitore werden überall dort eingesetzt, wo Videobilder beurteilt, kontrolliert, verglichen oder verändert werden, wie z. B. bei der Farbkorrektur und der technischen Abnahme von Programmen.

Die Anforderungen an die technischen Parameter sind in Tabelle 7.3 dargestellt. Die dort angegebene Unterscheidung zwischen Vollkontrast (Full screen contrast) und Simultankontrast bezieht sich darauf, dass im ersten Fall in einem völlig dunklen Raum der Unterschied zwischen 100%-Weiß bei 100 cd/m² und 0%-Schwarz ermittelt wird, während im zweiten Fall Weiß und Schwarz in einem Testmuster gleichzeitig angezeigt werden.

Referenzmonitore müssen mindestens eine Auflösung, bzw. Bildpunktanzahl haben, die dem zu kontrollierenden Signal entspricht, zudem ist erforderlich, dass sie ein 1:1-Pixelmapping unterstützen, d. h, dass z. B. auf HD-Monitoren SD-Signale unskaliert dargestellt werden, wobei das Bild nur 720 x 576 Pixel füllt. Der Rest ist typischerweise schwarz [68].

Die Grauwertübertragung des Monitors (Electro Optical Transfer Function, EOTF) ist über das Display-Gamma bestimmt, das bei Referenzmonitoren 2,35 betragen soll. Mit einem Kamera-Gamma (OETF) von 0,45 ergibt sich mit der Gleichung System-Gamma = OETF · EOTF insgesamt ein System-Gamma von 1,06. Bei Heimgeräten beträgt der Wert meist 1,2. Die Farbwiedergabe muss mit den Grundfarben und Weißpunkten erfolgen, die für das dargestellte Signal den jeweiligen Signalstandards, wie ITU-R BT.601 oder 709, entsprechen.

Tabelle 7.4. Anforderungen an Schnittstellen nach EBU-Tech 3320

Interface	Grade 1	Grade 2	Grade 3
SDI	mind. 2	mind. 2	1
HD-SDI	mind. 2	mind. 2	1
Dual HD-SDI / 3Gb/s	optional	optional	optional
HDMI 1.3	optional	optional	optional
DVI 1.0	optional	optional	optional
Component, RGB	optional	optional	optional
FBAS	optional	optional	optional

Kontrollmonitore (Grade 2) sollten ähnlichen Ansprüchen genügen wie Referenzmonitore. Um einen günstigeren Preis zu ermöglichen, dürfen sie aber höhere Toleranzen aufweisen und sind ggf. nicht so aufwendig ausgestattet. Der geringere Aufwand bezieht sich auch auf die Signalverarbeitung. Da diese einfacher ist, kann bei Grade 2- und Grade 3-Monitoren besonderer Augenmerk auf schnelle Reaktion gelegt werden. Die EBU legt diesbezüglich fest, dass die Latenzzeit zwischen Signal und Anzeige des Bildes nicht mehr als 10 ms betragen darf. Sicht- und Kontrollmonitore erlauben daher eine gute Beurteilung der Lippensynchronität. Als Einsatzgebiete für Kontrollmonitore werden Vorschauanwendungen, Regiewände, Schnittplätze und Schalträume angegeben.

Beobachtungsmonitore (Grade 3) sind Geräte mit einer Qualität, die mindestens der von guten Heimempfängern entspricht. Im Unterschied zu diesen sollten sie aber stabiler und robuster konstruiert sein und studiorelevante Schnittstellen aufweisen. Tabelle 7.4 zeigt die erforderlichen Interfaces. Beobachtungsmonitore werden in der Tonregie, in Lichtsteueranlagen und Kommentatorplätzen eingesetzt. Zudem als Mitschaumonitore im Studio und für das Publikum. Da die Monitore dabei ins Bild kommen können, ist eine Umschaltmöglichkeit auf die Farbtemperatur von 3200 K vorgesehen. Für schlechte Sichtbedingungen in heller Umgebung können sie auch mit bis zu 400 cd/m² betrieben werden (Tabelle 7.3).

Als weitere Ausstattungsmerkmale kommt für alle Grades die Umschaltbarkeit des Bildformats (4/3 oder 16/9) hinzu, für Grade 1 und 2 dazu noch 1:1-Pixelmapping, Safe Title Markierungen, Tally-Licht, Blue Only-Modus und externe Synchronisierbarkeit. Im Blue Only-Modus wird nur der Blau-Farbauszug dargestellt (Abb. 7.27). Mit dieser Betriebsart wird eine bessere Beurteilbarkeit des Farbrauschens möglich, sowie eine einfache Einstellung der richtigen Farbintensität und der Farbphase bei NTSC. Eine Besonderheit bei Röhrenmonitoren ist das H/V-Delay (Impulskreuz-Darstellung). In dieser Betriebsart wird der Bildschirm heller gesteuert und die Ablenksteuerung der Bildröhre wird vertikal um die halbe Bildhöhe und horizontal um etwa ein Drittel der Bildbreite verzögert (Abb. 7.27). Damit werden die Austastlücken und die Synchronimpulse sichtbar und können auch ohne Oszilloskop überprüft werden. Schließlich können die meisten professionellen Monitore im Underscan-Modus betrieben werden, d. h. das Bild erscheint verkleinert und damit unbeschnitten auf dem Bildschirm. Die unbeschnittene Darstellung ist auch Standardbetriebsart, die von der EBU empfohlen wird.

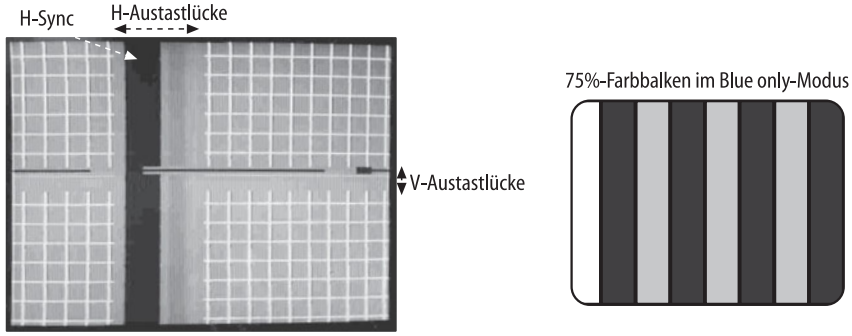


Abb. 7.27. H-V-Delay- und Blue Only-Darstellung

7.4.1.1 Ausführung von Monitoren

Während früher fast ausschließlich Röhrenmonitore verwendet wurden, gilt dies heute für die LCD und zwar im professionellen Umfeld und im Heim. Die letzten CRT-Monitore wurden noch als Referenzmonitore gebaut (Abb. 7.28 links), doch auch in diesem Bereich waren dann einige LCD verfügbar, die die Grade 1-Ansprüche erfüllten. In Abbildung 7.28 ist rechts ein LCD mit Grade 2-Qualität dargestellt. Vor allem die Grade 1-Monitore werden aber heute von OLED-Monitoren abgelöst, die die hohen Anforderungen als aktive Displays besser erfüllen.

Als Beispiel für Grade 1-Monitore auf OLED-Basis sei hier die BVM E/F-250 -Serie von Sony angegeben. Die Displays können 10-Bit-Signale direkt darstellen, während gewöhnliche LCD auf 256 Werte pro Farbe beschränkt sind. Die Signalverarbeitung erfolgt auf 12-Bit-Basis und erlaubt eine exakte Anpassung an die erforderlichen Gamma-Kurven und auch mehrere Farbräume. Dies Geräte akzeptieren Signale via 3G-SDI und über Dual link und auch 4:4:4-RGB-Daten. Sie erlauben eine native Interlaced-Darstellung, sowie auch die native Darstellung von 24p-Signalen bei einer Bilddarstellung mit 72 Hz. Die Lichterzeugung erfolgt über TES-OLED und hat damit eine hohe Effizienz. Hinzu kommt eine micro-cavity-Stuktur, die die Lichtfilterwirkung für die Farbwiedergabe über einen Resonanzeffekt optimiert.

Abb. 7.28. Grade 1-Röhren- und Grade 2-LC-Monitor [208]

Abbildung 7.29 zeigt oben das Anschlussfeld eines typischen analogen SD-Monitors. Man erkennt vier BNC-Buchsen zum Anschluss von Composite-Signalen, eine Hosidenbuchse für ein Y/C-Signal und acht Buchsen für ein Komponentensignal. Die Doppelbuchsen dienen zum Durchschleifen des Signals an andere Geräte. In den Monitor sind oft ein Audioverstärker und ein kleiner Abhörlautsprecher eingebaut, sodass das Gerät auch eine einfache Audiokontrollmöglichkeit bietet.

Das Anschlussfeld des HD-LCD-Monitors aus Abbildung 7.28 ist in Abbildung 7.29 unten dargestellt. Hier gibt es zwei SDI-Inputs, die wegen automatischer interner Umschaltung auch als HD-SDI-Eingänge genutzt werden können. Des Weiteren stehen ein digitaler DVI-Eingang und diverse analoge Eingänge zur Verfügung. Die Gerätevorderseiten zeigen bei beiden Geräten die Einsteller für Helligkeit, Kontrast und Farbsättigung sowie die Eingangswahlschalter und Tasten für Blue Only, H/V-Delay, Overscan etc. (Abb. 7.30). Der Helligkeitssteller dient der Erhöhung (Verringerung) von Schwarz- und Weißwert bei festgehaltener Differenz der Werte, er ist also für den Schwarzwert bestimmend. Der Kontraststeller bewirkt die Veränderung des Weißwertes bei festem Schwarzwert und ändert damit die Differenz.

7.4.1.2 Monitorabgleich

Der Produktionsmonitor ist das wichtigste Gerät zur Bildbeurteilung, daher muss er gut positioniert und eingestellt sein. Videomonitore werden im Studio auf eine Spitzenleuchtdichte von 80 cd/m² abgeglichen. Das menschliche Auge passt sich der Lichtintensität an, wobei die Kontrastwahrnehmung aber eingeschränkt ist, wenn der Bildschirm wesentlich heller oder dunkler ist als die Umgebung. Das Umfeld sollte daher so aufgehellt sein, dass die Umfeldleuchtdichte ca. 10–20 % der Spitzenleuchtdichte des Monitors beträgt. Das gilt insbesondere für Monitore, die im Produktionsumfeld eingesetzt werden.

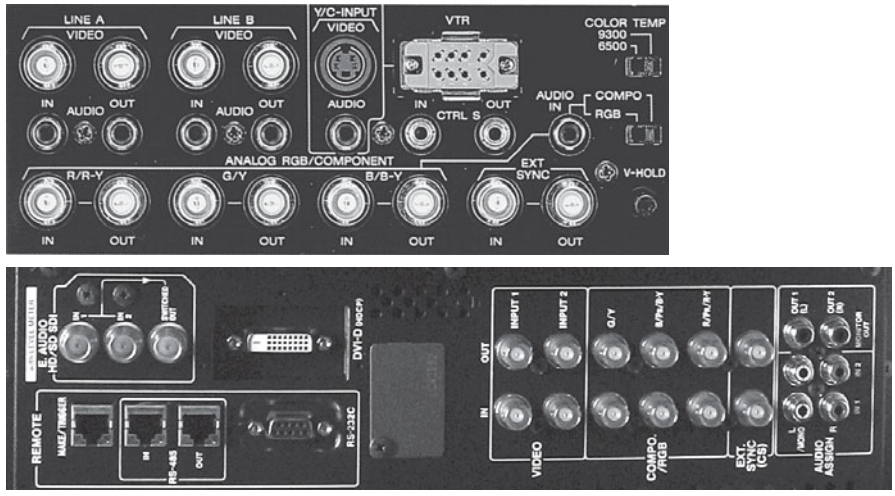


Abb. 7.29. Anschlussfelder eines SD Röhren- und eines HD-LCD-Videomonitors [208]

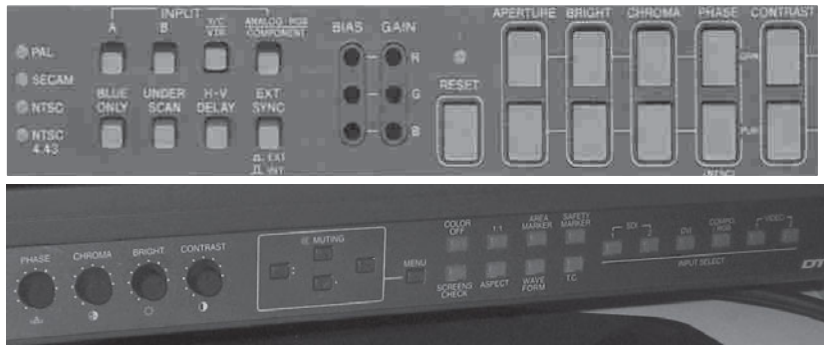


Abb. 7.30. Bedienfelder eines SD-Röhren- und eines HD-LCD-Videomonitors [208]

Die Einstellung kann mit einem 75%-Farbbalken folgendermaßen vorgenommen werden: Nach Möglichkeit wird der Underscan-Modus gewählt und der Kontraststeller auf Mittelstellung gebracht. Helligkeits- und Farbsättigungssteller werden auf Linksanschlag gedreht, sodass der Farbbalken Schwarz/Weiß erscheint. Mit dem Steller für die Helligkeit wird der Schwarzwert nun so gewählt, dass der schwarze Balken gerade vom Umfeld getrennt erscheint und sich vom Blau-Balken abhebt. Dann wird der Kontrast so justiert, dass sich Gelb- und Weißbalken voneinander abheben. Schließlich wird noch die Farbsättigung eingestellt, sodass im Blue Only-Modus die Balken für Gelb, Grün, Rot und Schwarz gleichermaßen schwarz und die Balken für Cyan, Magenta und Blau gleich intensiv (75%) blau oder grau sind. Der Weißbalken muss heller erscheinen (100%, 80 cd/m²). Sehr hilfreich ist dabei auch das in Kap. 2.5 beschriebene PLUGE-Signal (Picture Line-up Generating Equipment).

Zum Schluss des Abschnitts über professionelle Monitore sei noch einmal allgemein auf die Latenzproblematik hingewiesen, die im Digitalzeitalter wegen anspruchsvoller Rechenprozesse immer häufiger auftritt. Sowohl Audio- als auch Videosysteme können zu einem Bild-Tonversatz führen. Die EBU hat festgelegt, dass das Audiosignal im Produktionsbereich höchstens 5 ms früher (sound before picture) und 15 ms später (sound after picture) als das zugehörige Bild erscheinen darf. Für den Heimbereich bzw. direkt vor der Ausstrahlung gelten Grenzen von ≤ 40 ms für sound before picture und ≤ 60 ms für sound after picture [139].

7.4.2 Heimmonitore und Fernsehempfänger

In diesem Bereich kann eine Unterscheidung in SD- und HD-Empfänger vorgenommen werden. Damit korrespondiert zum großen Teil auch die eingesetzte Displaytechnologie: Röhren bei alten SD- und LCD oder auch PDP bei neuen HD-Geräten. Während früher ein Fernsehgerät immer ein Fernsehempfangsteil enthielt, werden moderne HD-Geräte als universelle Heimdisplays verkauft, an die Module für den Fernsehempfang wie Satellitenreceiver aber auch PC, USB-Sticks, und weitere Speicher mit Videodaten angeschlossen werden können.

7.4.2.1 SD-Fernsehgeräte

Die klassischen SD-Fernsehempfänger arbeiteten in den meisten Fällen mit In Line- oder Trinitronröhren. Sie unterscheiden sich von SD-Videomonitoren dadurch, dass sie zusätzlich eine analoge HF-Empfangseinheit (Tuner) für Bild und Ton sowie eine Tonsignalverarbeitungsstufe enthalten.

Der Tuner dient zur Selektion des gewünschten Kanals aus dem VHF- oder UHF-Bereich (Abb. 7.31). Im Tuner wird die Empfangsfrequenz auf eine Bild-Zwischenfrequenz von 38,9 MHz umgesetzt. Die Zwischenfrequenz (ZF) kann hochselektiv ausgefiltert werden, da sie ihren Wert unabhängig vom eingestellten Kanal immer beibehält (Superhet-Prinzip). Bei der ZF-Umsetzung wird die Frequenzlage umgekehrt. Der Tonträger liegt damit um 5,5 MHz unterhalb des Bildträgers, d. h. bei 33,4 MHz bzw. 33,158 MHz für den 2. Tonträger. Die ZF-Durchlasskurve hat eine schräg abfallende, sogenannte Nyquist-Flanke, die eine Amplitudenverdopplung aufgrund des Restseitenbandverfahrens vermeidet. Das Bild-ZF-Signal wird verstärkt und demoduliert. Nach der AM-Demodulation entsteht das FBAS-Signal, welches den im Abschnitt über Röhrenmonitore beschriebenen Signalverarbeitungsstufen zugeführt wird. Für die Wiedergabe des Signals aus einem Videorecorder werden die Geräte häufig über die (Euro)-AV-Buchse verbunden.

Bei der Tonsignalverarbeitung unterscheidet man das Intercarrier-, das Parallelton- und das Quasiparalleltonverfahren. Beim Intercarrier-Verfahren wird ein gemeinsamer Verstärker für Bild- und Ton-ZF verwendet. Es besteht hier die Möglichkeit, dass sich Bild und Ton gegenseitig stören. So treten z. B. Tonstörungen (Intercarrier-Brummen) auf, wenn im Bildsignal hohe Amplituden und Frequenzen vorliegen (Schrifteinblendungen). Das Paralleltonverfahren vermeidet die Nachteile des Bild-Ton-Übersprechens durch die Verwendung von zwei separaten ZF-Stufen. Das Quasi-Paralleltonverfahren kombiniert die Vorteile von Intercarrier- und Paralleltonverfahren, d. h. der Bild- Tonträgerabstand bleibt konstant, und die Signale beeinflussen sich nicht gegenseitig.

Modernere Farbfernsehempfänger arbeiteten trotz analoger Signalübertragung intern digital. Die A/D-Wandlung wird dabei nach der Demodulation vorgenommen. Die PAL-Decodierung und das Signal-Processing finden dann auf der digitalen Ebene statt, sodass schließlich das D/A-umgesetzte Komponentensignal nur noch der analogen Matrix und das resultierende RGB-Signal der Röhre zugeführt wird. Die Bildqualitätsverbesserung durch diese sogenannten digitalen Fernsehempfänger resultiert vor allem daraus, dass sich in der digitalen Ebene hochwertige Filter realisieren lassen.

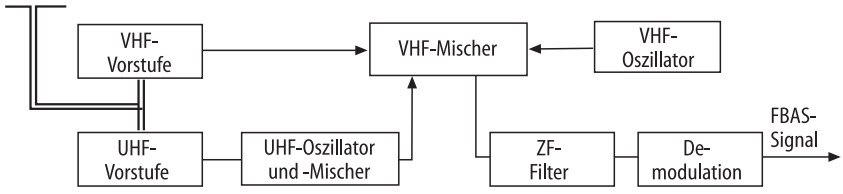


Abb. 7.31. Blockschaltbild des analogen Fernsehtuners

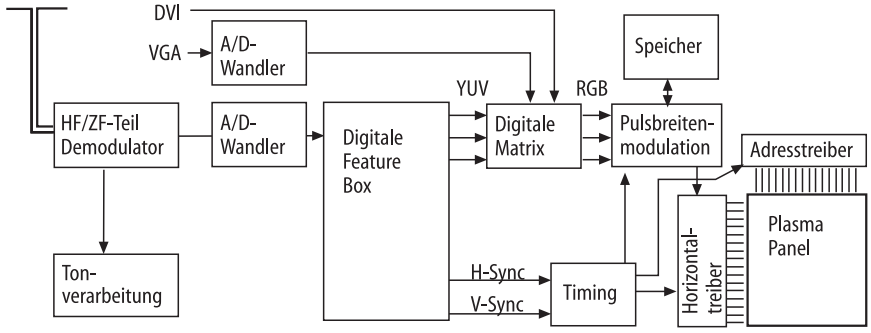


Abb. 7.32. Blockschaltbild eines digitalen Fernsehempfängers mit Plasma Display

Die digitale Signalverarbeitung erlaubt auch die einfache Speicherung eines Videobildes. Hier kann das Bild mit doppelter Frequenz ausgelesen und damit das Großflächenflimmern quasi beseitigt werden (100 Hz-Technik). Weiterhin erlaubt die digitale Stufe die Bildformatumschaltung zwischen 3:4 und 16:9 mit Hilfe der Zoomfunktion. Diese Funktionen, zu denen auch der Picture in Picture-Modus gehört, werden in der sogenannten Feature Box ausgeführt, die auch im digitalen HD-Geräten enthalten ist (Abb. 7.32).

7.4.2.2 HD-Fernsehgeräte

HD-Fernsehdisplays arbeiten wie die Videomonitore für den Produktionsbereich, sie werden aber preisgünstiger hergestellt und beinhalten einen analogen oder digitalen Fernsehempfangsteil (Tuner). In den meisten Fällen kommen LCD zum Einsatz, seltener auch PDP (Abb. 7.32) und OLED.

Angesichts der zwei HD-Standards gibt es eine Verunsicherung der Konsumenten bezüglich der Leistungsfähigkeit der Geräte. Um diese zu beseitigen, hat die Herstellervereinigung namens Digitaleurope Mindestgrenzen für die technische Qualität entwickelt, verbunden mit Logos, die nur bei Einhaltung der Grenzen vergeben werden. So entstand zunächst das HD-Ready-Logo (Abb. 7.33), das besagt, dass das Display unskaliert mindestens 1280 x 720 Pixel darstellen können muss sowie 1080i-HD-Daten mit Skalierung. Des Weiteren werden Digitalschnittstellen wie DVI und HDMI vorgeschrieben (s. Kap. 3). Der Nachfolger heißt HD Ready 1080p, was anzeigt, dass ein solches Gerät ohne Skalierung die 1080-Zeilen-HD-Variante beherrscht und das sogar bei progressiver Abtastung. Für diese Funktion ist auch Bezeichnung „Full HD“ gebräuchlich, aber von Digitaleurope nicht spezifiziert.

Abb. 7.33. HD Ready-Logos

7.4.3 3D-Displays

Die Grundlagen der räumlichen visuellen Wahrnehmung sind in Kap. 2 dargestellt. Das dreidimensionale Sehen gelingt den Menschen u. a. durch die Tatsache, dass die beiden Augen einen Gegenstand aus leicht unterschiedlicher Perspektive sehen, was einen Effekt erzeugt, der mit dem Begriff Disparität beschrieben wird. Bei der technischen Umsetzung können, den beiden Augen entsprechend, zwei zweidimensionale Abbildungen mit unterschiedlicher Perspektive gezeigt werden, sodass sie vom Wahrnehmungssystem zu einem Bild mit Tiefendimension verschmolzen werden.

Ideal wäre es, wenn das Bild wirklich dreidimensional wiedergegeben werden könnte, also mit Bildpunkten, die nicht nur in einer Ebene, sondern auch in der Tiefe des Raumes liegen (Volumendisplay). Zur Realisierung einer solchen Objektsubstitution wäre es denkbar, ein Gas in drei Dimensionen zur Fluoreszenz zu bringen, indem alle Raumpunkte durch ein System aus sich kreuzenden Laserstrahlen angeregt werden. Eine zweite Realisierungsidee nutzt die Holographie. Sie beruht darauf, dass von vielen Streuzentren ausgehende Lichtwellen interferieren. Die Streuzentren sollten dabei einen Abstand haben, der kleiner ist als die Lichtwellenlänge, also 0,5 µm. Hier wird das Grundproblem der echten 3D-Wiedergabe deutlich: allein zur Bereitstellung der Bildpunkte ist ein Rechenaufwand erforderlich, der bisher realisierbare Größenordnungen weit übersteigt [199]. Daher finden in der Praxis nur stereoskopische Wiedergabeverfahren Verwendung, bei denen zunächst in Verfahren mit und ohne Sehhilfen unterschieden werden kann.

7.4.3.1 Autostereoskopische Displays

Autostereoskopische Verfahren verdanken ihre Bezeichnung dem Umstand, dass sie keine Sehhilfen erfordern. Die Halbbildtrennung erfolgt direkt bei der Betrachtung. Es wird ein einzelnes Display verwendet, das im einfachen Fall mit vertikal verlaufenden Streifen oder Linsen versehen ist, die eine Parallaxenbarriere oder ein vertikal verlaufendes Linsenraster darstellen. Die Konstruktion ist so gewählt, dass bei geeignetem Betrachtungsabstand jeder zweite Streifen, bzw. Spalte jeweils für ein Auge ausgeblendet ist (Abb. 7.34). Das Bild muss also aus zwei Teilbildern bestehen, die horizontal spaltenweise verkämmt werden. Die Verkämmung halbiert die Horizontalauflösung.

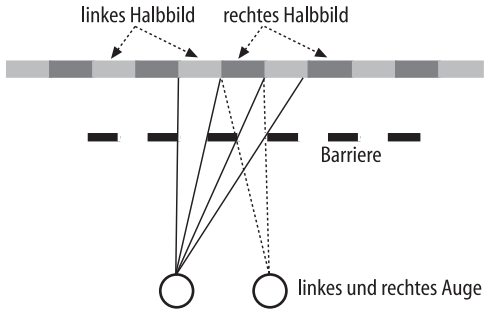


Abb. 7.34. Grundprinzip der autostereoskopischen Displays

Wie Abbildung 7.34 deutlich macht, ergibt sich die richtige Zuordnung der Augen zu den Teilbildern nur an einer fest definierten Betrachterposition (Sweet Spot). Der Sehbereich kann erweitert werden, indem mehrere Bilder verschachtelt werden, die die 3D-Szene aus mehreren Blickwinkeln zeigen [122]. Anstatt einem einzelnen Betrachter mehrere Single Views anzubieten, kann das Verfahren auch verwendet werden, um mehrere Sweet Spots für mehrere Betrachter zu erzeugen. Solche Multiview-Ansichten sind flexibel an verschiedene Displays anpassbar, werden als zukunftsträchtig angesehen und sind auch bei MPEG standardisiert. Es ist aber zu bedenken, dass mit der Anzahl der zu zeigenden Bilder die Horizontalauflösung nicht nur halbiert, sondern noch stärker eingeschränkt wird oder ein Display mit sehr hoher Auflösung erforderlich wird.

Die Einschränkung der Betrachterposition beim autostereoskopischen Display kann durch Detektion der Kopfposition gemindert werden. Ein Trackingsystem registriert dabei fortwährend die Lage des Kopfes und steuert entsprechend die Position der Linsenrasterplatte (Barrier). Allerdings kann ein solches System i. d. R. wieder nur auf einzelne Betrachter reagieren.

Moderne autostereoskopische Displays lassen sich zwischen 3D- und gewöhnlicher 2D-Wiedergabe umschalten. Das gelingt dadurch, dass die Parallaxenbarrieren als Flüssigkristallstreifen ausgeführt werden. Mittels einer Steuerspannung werden sie zwischen opak und transparent, also zwischen 3D- und 2D-Darstellung umgeschaltet (Switchable Barrier).

7.4.3.2 Polarisationsverfahren

Wie der Name sagt, wird hier die Trennung der Teilbilder mit Hilfe von Polarisationsfolien vorgenommen. Sie werden in den Brillen angeordnet, die die Betrachter tragen. Wichtig ist, dass sie dabei bezüglich ihrer Transmissionsrichtung um 90° gegeneinander verdreht sind. Bei der Bilddarstellung am Display wird mit Hilfe von Zeilenfiltern dafür gesorgt, dass zeilenweise abwechselnd jeweils Licht mit unterschiedlicher Polarisation abgestrahlt wird. Bei Betrachtung mit Polarisationsbrillen erhält jedes Auge nur Licht jeder zweiten Zeile. Die stereoskopischen Teilbilder werden dafür wie beim Zeilensprungverfahren verschachtelt. Die Vertikalauflösung ist halbiert. Die Brillen sind preiswert und brauchen keine Energieversorgung (passives System). Nachteilig ist der starke Lichtverlust durch das Durchlaufen zweier Polarisationsfilter. Das Polarisationsverfahren ist im Kino verbreitet, bei den Displays für den Heimbereich dominiert aber das Shutterverfahren, das keinen Zeilenfilter erforderlich macht.

7.4.3.3 Shutterverfahren

Bei diesem Verfahren werden Brillen verwendet, die über ein elektrisches Signal, meist mit LC-Elementen, verschließbar sind. Die stereoskopische Darstellung wird erreicht, indem die zwei Halbbilder mit doppelter Bildrate nacheinander gezeigt werden und der Verschluss in der Brille immer nur für das richtige Teilbild geöffnet wird. Die dafür erforderliche Synchronisation mit dem Bildgeber kann mittels Infrarotimpulsen erreicht werden, die zu den Brillen geschickt werden. Dieses Verfahren bietet eine gute Trennung der Halbbilder, hat allerdings den Nachteil, dass es auf teuren Brillen beruht, die eine eigene Energieversorgung benötigen. Der Lichtverlust ist auch hier erheblich.

7.5 Großbildprojektion

Zur Darstellung von Großbildern mit Diagonalen über 1,5 m werden vorwiegend Projektionsverfahren eingesetzt. Die Großbildprojektion ist als Auf- oder Rückprojektion möglich. Bei der Aufprojektion befinden sich Projektor und Betrachter vor der Projektionswand (Abb. 7.35). Die Wand sollte einen möglichst hohen Reflexionsgrad aufweisen. Häufig werden Projektionswände mit ausgeprägter Richtcharakteristik eingesetzt, die aber nur nahe der Projektionsachse ein gutes Bild bieten. Die Rückprojektion erfordert dagegen einen hohen Transmissionsgrad der Projektionswand, da sich der Projektor hinter der Wand und die Zuschauer davor befinden. Meist ist die Lichtausbeute kleiner als beim Aufprojektions-Verfahren. Die Rückprojektion erfordert eine spiegelverkehrte Abbildung. Geräte, die für beide Projektionsarten tauglich sind, müssen auf spiegelverkehrte Bilddarstellung umschaltbar sein (Abb. 7.35).

Wie bei den Flachbildschirmen können auch bei den Projektionsverfahren aktive und passive Systeme unterschieden werden, wobei die passiven Systeme (Lichtventilprojektoren) aufgrund der begrenzten Leuchtdichte aktiver Systeme stark überwiegen. Qualitätskriterien sind auch hier wieder die mögliche Bildauflösung, der Kontrast und die erzielbar Lichtstärke. Letztere ist besonders dann von Interesse, wenn gegen starkes Raumlicht angekämpft werden muss, wie es in typischen Präsentationsfällen oder bei Shows vorkommt.

Als Beispiel kann angenommen werden, dass im Raum eine Beleuchtungsstärke von 100 lx vorliegt. Wenn dann die Darstellung des Bildes fünfmal heller erscheinen soll, muss die Beleuchtungsstärke auf der Projektionsfläche (Leinwand) E = 500 lx betragen. Die erforderliche Lichtstärke des Projektors ergibt sich nun über die beleuchtete Fläche A_2 mit der Beziehung $E = \Phi/A_2$. Bei einer Fläche von 2 m x 1,5 m ist also ein Lichtstrom von 1500 lm erforderlich, wenn der Verstärkungsfaktor (Gain) durch eine richtungsabhängige Reflexion mit eins angenommen wird. Bei ideal diffuser Reflexion der Leinwand lässt sich mit $L = E \cdot R/\pi$ weiter die Leuchtdichte L = 127 cd/m² ausrechnen, wobei der Reflexionsfaktor R = 0,8 angenommen wurde.

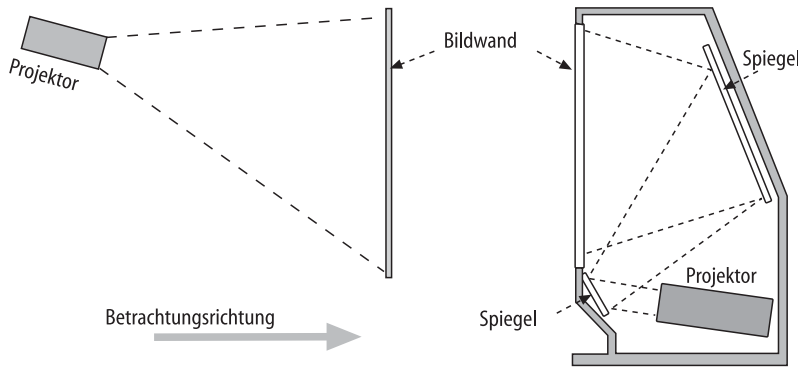


Abb. 7.35. Schematische Darstellung von Beispielen für Auf- und Rückprojektion

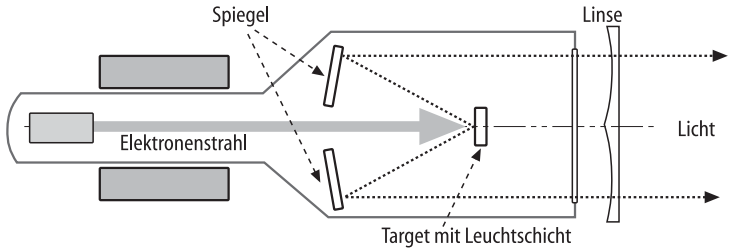


Abb. 7.36. Kathodenstrahlröhre für die Projektion

7.5.1 Aktive Großbilderzeugung

7.5.1.1 Projektion mit Kathodenstrahlröhren

Bei CRT-Projektoren besteht das Arbeitsprinzip darin, dass auf der Leuchtschicht einer Kathodenstrahlröhre ein möglichst helles Bild erzeugt und über eine optische Anordnung auf die Projektionswand abgebildet wird, wobei für jeden Farbauszug eine eigene Röhre verwendet wird. Der Elektronenstrahl trifft das Target, das mit einem dem Farbauszug entsprechenden Leuchtstoff beschichtet ist. Über einen Hohlspiegel wird das Licht reflektiert, gefiltert und auf die Projektionswand geleitet, wobei eine Korrekturlinse die Abbildungsfehler ausgleicht. Die Spiegel können in die Röhre integriert werden (Abb. 7.36). Die drei Farbauszüge werden deckungsgleich übereinander projiziert (Abb. 7.37). Mit Beschleunigungsspannungen von mehr als 30 kV und Strahlströmen bis zu 2 mA werden Maximallichtströme bis 500 lm erreicht. Eine Steigerung dieser Werte ist kaum möglich, da die Gefahr besteht, dass die Leuchtschicht zerstört wird. Dieser Lichtstrom erscheint sehr klein. CRT-Projektoren werden aber für Videoprojektion eingesetzt, bei der es, im Gegensatz zum o. g. Beispiel weniger auf die Helligkeit als den Kontrast ankommt. Dieser ist sehr gut, sodass bei verdunkeltem Raum ein gutes Bild erzeugt werden kann.

Die verwendeten Kathodenstrahlröhren haben Diagonalen von 7“ oder 9“. Röhrenprojektoren. Sie arbeiten ohne Schattenmaske und zeigen daher keine Pixelstruktur wie LCD-Projektoren. Sie weisen aber alle Nachteile der Röhrentechnologie auf, sind aufwendig einzustellen und kaum portabel. In der Praxis sind sie heute kaum noch zu finden.

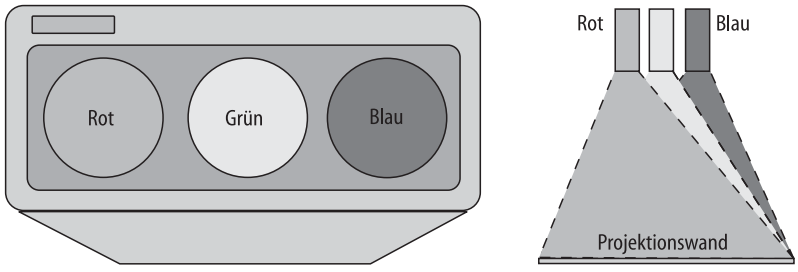


Abb. 7.37. Röhrenprojektor

7.5.1.2 Großbilddarstellung mit dem Laser

Manche Hersteller bieten Laserprojektoren an. Damit ist aber nicht immer ein wirkliches Bildumsetzungsprinzip mit dem Laser gemeint, sondern oft nur die Tatsache, dass für passiv arbeitende Projektoren auf Basis von LCD oder DMD Laser, meist Laserdioden, als Lichtquelle verwendet werden (s. u.).

Bei der direkten Bilderzeugung mittels Lasern werden stark gebündelte Lichtstrahlen aus drei Lasern verwendet, die rotes, grünes und blaues Licht für die additive Mischung der drei Farbauszüge erzeugen. Die Intensitäten der Strahlen können vom Videosignal mit Hilfe von Lichtmodulatoren gesteuert werden. Ein Lichtmodulator lässt sich mit einem Kristall realisieren, der in der Lage ist, die Polarisationsrichtung des Lichtes in Abhängigkeit von einer Steuerspannung zu drehen. Damit ist es unter Nutzung eines zweiten, feststehenden Polarisators möglich, die Intensität in Abhängigkeit vom Polarisations-Drehwinkel zu steuern. Die farbigen Teilstrahlen werden über teildurchlässige Spiegel zu einem Strahl zusammengefasst und mit Hilfe von rotierenden Spiegeln horizontal und vertikal über die Bildfläche geführt (Abb. 7.38). Die Abbildung mit Lasern ermöglicht hohe Bildfrequenzen und eine hohe Auflösung, dabei ist wegen der scharfen Bündelung des Strahls praktisch keine Begrenzung der Schärfentiefe vorhanden. Aufgrund der Kohärenz des Laserlichts entstehen aber durch Interferenz Flecken (Speckles), die mit einigem Aufwand unterdrückt werden müssen. Die technische Weiterentwicklung des Systems hängt eng mit der Entwicklung preiswerter Laser für Grün und Blau zusammen.

7.5.1.3 Video-Wand-Systeme

Große, aktiv leuchtende Flächen lassen sich auch mit einer Vielzahl konventioneller Monitore erzeugen. Die Monitore werden dafür mit besonders flachen Bildschirmen und sehr dünnem Rand gebaut, sodass beispielsweise 9 Monitore eine nur geringfügig gestörte Bildfläche mit einer Diagonalen von über 3 m ergeben (Abb. 7.39). Das Videobild wird natürlich nicht vervielfacht dargestellt, sondern das Videosignal wird digitalisiert und mit einem sogenannten Splitrechner wird eine der Monitoranzahl entsprechende Zahl von Teilbildern

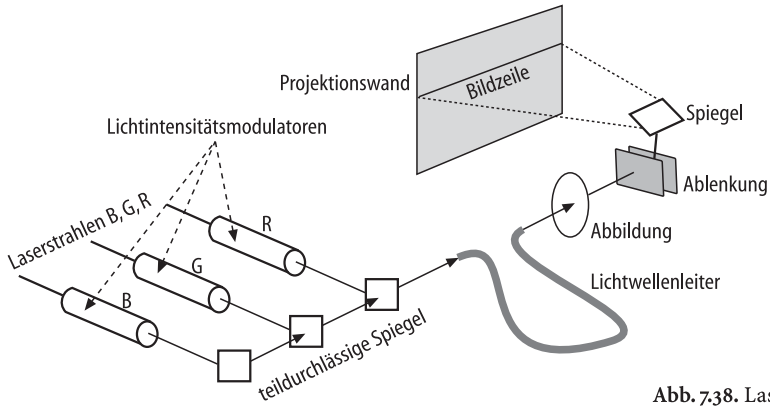


Abb. 7.38. Laserprojektor

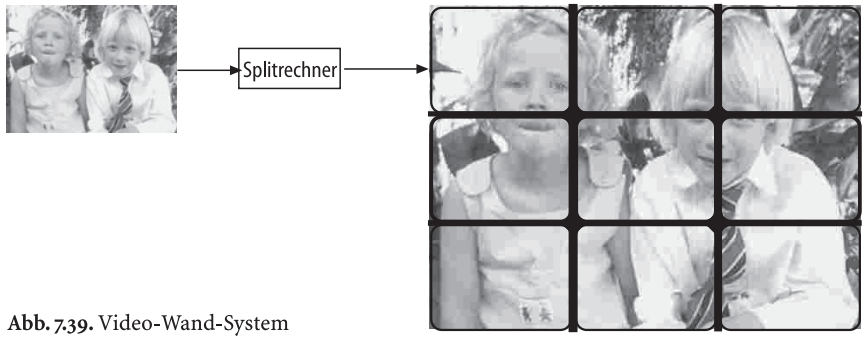


Abb. 7.39. Video-Wand-System

erzeugt, die den einzelnen Wiedergabegeräten zugeführt werden. Video-Wand-Systeme bieten ein helles Bild das auch bei Tageslicht gut sichtbar ist. Der Einsatz ist aber teuer und erfordert einen hohen Aufwand um gute Konvergenz und für alle Monitore den gleichen Helligkeits- und Farbeindruck zu erzielen.

Zu dieser Klasse aktiver Großbildsysteme, die keine Projektionssysteme sind, können auch die LED-Walls gezählt werden, die pro Bildpunkt ein LED-Tripel mit den Grundfarben enthalten. Bei genügend großem Abstand erscheint ein homogenes Bild während aus der Nähe die LED deutlich erkennbar sind.

7.5.2 Passive Großbilderzeugung

Systeme zur passiven Bilderzeugung arbeiten als Lichtventile, die das Licht aus einer externen Lampe steuern. Die Lampen, ihr Wirkungsgrad und ihre Lebensdauer bestimmen wesentlich die Kosten des Projektors. Sie beeinflussen auch die Farbwiedergabe. Normale Halogenlampen haben eine Farbtemperatur um 3000 K, Metalldampflampen liegen bei 6000 K und UHP-, UHE-Lampen bei 6500 bis 8000 K. Xenonlampen kommen mit 6500 K dem Farbspektrum des Sonnenlichts am nächsten.

Halogenlampen arbeiten, wie gewöhnliche Glühlampen, mit einem Heizdraht. Im Gegensatz dazu hat die Metalldampflampe zwei spitze, eng benachbarte Wolframelektroden, die sich in einer Atmosphäre aus Füllgas, Edelgas und einer Gaszumischung befinden. Mittels einer hohen Zündspannung wird das Gas wie beim Plasmadisplay zum Leuchten gebracht.

Für besonders anspruchsvolle Projektionen kommen Xenon-Hochdrucklampen zum Einsatz. Wegen des entstehenden extrem hohen Drucks ist die Lampe für den Heimbereich nicht geeignet. Sehr verbreitete, kleinere Hochdruck-Gasentladungslampen, mit einer Aufnahmeleistung von 50W bis 10 kW, sind die UHP-Lampen (Ultra high Peformance). UHE-, VIP- und SHP-Lampen sind ihnen ähnlich, aber von verschiedenen Herstellern entwickelt.

Alle haben eine bessere Lichtausbeute als Halogenlampen. Das bedeutet, dass sie auch weniger gekühlt werden müssen. Der Elektrodenabstand wird möglichst klein gehalten, um eine besonders punktförmige Lichtquelle zu erhalten. Durch den Reflektor kann dann das Licht sehr gut gebündelt werden.

UHP- und UHE-Lampen können Betriebsdauern bis zu 4000 Stunden erreichen.

Durch die Entwicklung neuer und leistungsstarker Leuchtdioden ist es inzwischen möglich, auch diese als Projektionslichtquellen zu verwenden. Damit können sehr kompakte Projektoren gebaut werden. Sie haben zwar zunächst eine beschränkte Helligkeit, bieten aber den Vorteil einer sehr geringen Wärmeentwicklung und einer Lampenlebensdauer von über 30000 Stunden.

7.5.2.1 LCD-Projektoren

Flüssigkristall-Lichtventile werden oft mit der Abkürzung LCLV (Liquid Crystal Light Valve) bezeichnet. In Abbildung 7.40 ist der mögliche Aufbau eines LC-Projektors dargestellt. In den meisten Fällen arbeiten diese Geräte nach dem Prinzip des Diaprojektors: Flüssigkristallanzeigen, die genauso aufgebaut sind, wie die in Abschn. 7.2 beschriebenen Flüssigkristall-Flachbildschirme, werden anstelle des Dias durchleuchtet und das transmittierte Licht wird mit Hilfe eines Linsensystems auf die Projektionswand abgebildet.

Das Licht wird meist mit einer Metalldampflampe hoher Leistung erzeugt und zunächst mit Filtern von Ultraviolett- und Infrarotanteilen befreit. Das damit entstehende sogenannte Kaltlicht belastet die Lichtventile viel weniger als ungefiltertes Licht. Die Aufteilung der von der Lichtquelle kommenden Strahlung in die Rot-, Grün- und Blauanteile wird mit dichroitischen Spiegeln erreicht. Dichroitische Spiegel sind mit dünnen Schichten bedampft, um ein frequenzselektives Reflexionsverhalten zu erreichen. Mit den Filtern wird ein darstellbarer Farbumfang erreicht, der mit dem Umfang gewöhnlicher TV-Bildröhren vergleichbar ist, aber eine etwas andere Lage aufweist [74]. Nachdem das Licht der drei Farbauszüge die zugehörigen Lichtventile durchdrungen hat, wird es wieder zu einem Lichtbündel zusammengefasst und mit dem Projektionsobjektiv auf die Bildwand abgebildet. Abbildung 7.41 zeigt einen LCD-Projektor mit XGA-Auflösung, 2000 ANSI-Lumen und Kontrast 400:1.

Der Vorteil von LCD-Projektoren gegenüber Kathodenstrahl-Projektoren ist der kompakte, leichte Aufbau sowie die Abbildung der drei Farbauszüge über eine gemeinsame Optik. Hochwertige LC-Projektoren, z. B. des Herstellers Barco, erreichen bis zu 4000 ANSI-Lumen mit einem Kontrast von 400:1 bei XGA-Auflösung, also 1024 x 768 Bildpunkten. Der Hersteller Christie bietet ein Gerät an, das 7700 ANSI-Lumen mit einem Kontrast von 800:1 bei 1600 x

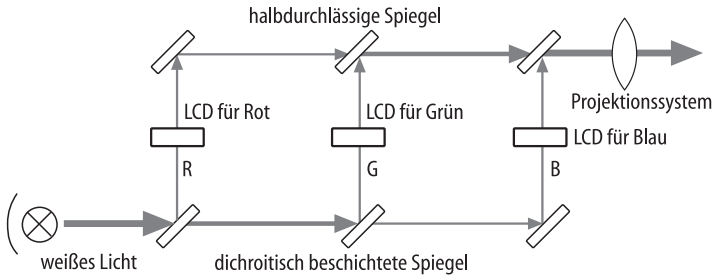


Abb. 7.40. LCD-Projektor

Abb. 7.41. LCD-Projektor

1200 Bildpunkten erzielt. Weitere Steigerungen der Lichtströme lassen sich nur schwer erreichen. Das Problem bei dieser Art von LCD-Projektoren ist, dass die Lichtenergie im Flüssigkristallelement stark absorbiert wird. Die Lichtleistung kann daher nicht beliebig gesteigert werden, denn die Erwärmung kann zur Zerstörung des Lichtventils führen. Die aktiven Flächen der Pixel und damit auch die Panels dürfen nicht zu klein sein, denn die zur Ansteuerung benutzten TF-Transistoren beanspruchen Platz auf dem Panel. Je kleiner das Pixel, desto ungünstiger ist das Verhältnis von aktiver zu lichtundurchlässiger Fläche. Ein weiterer Nachteil ist, dass die zur Bildpunktansteuerung benutzte aktive Matrix als Punktmuster im Bild sichtbar ist.

LCD-Projektoren mittlerer Leistungsklasse bieten Lichtstärken von 1500 bis 3000 lm. Die Geräte bieten meist eine komfortable Ausstattung. Sie stellen sich i. d. R. auf verschiedene Zeilenfrequenzen ein, verarbeiten sowohl progressive als auch interlaced Signale und sind somit als Video- und Datenprojektor geeignet. Für das Videosignal stehen FBAS- und Y/C- bzw. digitale DVI- oder HDMI-Eingänge zur Verfügung, Computersignale werden über einem VGA- oder DVI-Anschluss zugeführt. Die Entwicklung führt zu immer kompakteren und leichteren Geräten, die mobil eingesetzt werden können.

7.5.2.2 Image Light Amplifier

Ein großer Teil der genannten Probleme üblicher LCD-Projektoren kann beseitigt werden, wenn das LCD-Element nicht durchstrahlt wird, sondern mit Reflexion arbeitet. Das diesbezüglich zuerst entwickelte System nennt sich Image Light Amplifier (ILA).

Beim Reflexions-Betrieb kann die Ansteuerung von hinten erfolgen. So steht die gesamte Display-Fläche für die Ansteuerung, aber auf der anderen Seite auch für die Lichtreflexion zur Verfügung. Das verwendete Flüssigkristall-Element hat beim ILA-Verfahren eine homogene Schicht, einen Festkörper-Film, der nicht in Bildpunkte aufgeteilt ist. Damit gibt es keine Fehlstellen und keine Auflösungsbegrenzung durch das Bildpunkt-Raster. Die Ansteuerung der Flüssigkristallschicht geschieht dadurch, dass zunächst mit einer gewöhnlichen Kathodenstrahlröhre ein Bild erzeugt wird, das auf die Rückseite des LCD-Elementes abgebildet wird. Hier befindet sich eine fotoelektrische Schicht, in der ein dem optischen Bild entsprechendes Ladungsbild entsteht, das die Spannung innerhalb der Flüssigkristallschicht und damit die Transmissions- und Reflexionseigenschaften von polarisiertem Licht beeinflusst.

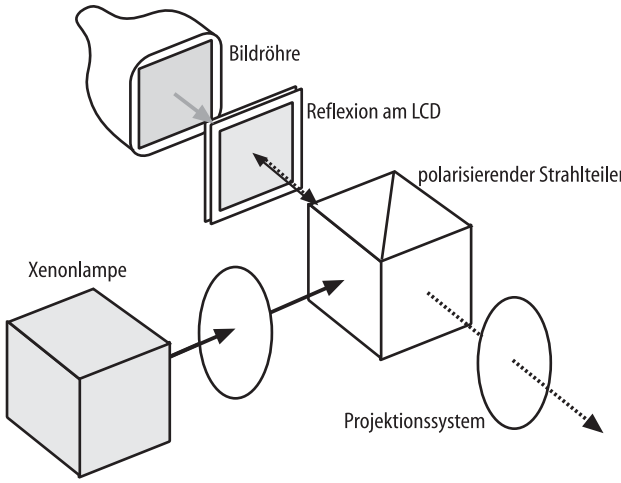


Abb. 7.42. ILA-Projektor

Das Projektionslicht wird oft durch eine Xenonlampe (z. B. 1500 W) erzeugt und zunächst von UV- und IR-Anteilen befreit. Es fällt seitlich zum Projektionsstrahlengang auf einen polarisierenden Strahlteiler und wird auf das Lichtventil gelenkt (Abb. 7.42). Hier muss es nicht noch einmal einen Polarisator durchlaufen. Das Licht dringt in die sehr dünne LC-Schicht ein, wird auf der Rückseite gespiegelt und durchläuft die LC-Schicht ein zweites Mal. Wenn die Flüssigkristalle so angeordnet werden, dass sich die Polarisationsebene des Lichts bei jedem Durchlauf um 45° dreht, so ist das austretende Licht um 90° verdreht und kann den Polarisationsstrahlteiler passieren. Der Grad der LC-Verdrehung und damit die Helligkeit des Bildpunktes kann durch die Steuerspannung beeinflusst werden. Die Bildhelligkeit ist damit unabhängig von der Ansteuerung und wird im Wesentlichen durch die Projektionslichtstärke bestimmt. Das modulierte, reflektierte Licht gelangt schließlich durch den Strahlteiler auf das Projektionsobjektiv. Zur Farbbilddarstellung werden drei Systeme für Rot, Grün und Blau parallel eingesetzt. Das entsprechend farbige Projektionslicht wird mit Hilfe dichroitischer Spiegel gewonnen. Die Erwärmung des ILA-Panels ist erheblich geringer als bei LCD-Projektoren, da das Licht nur von der LC-Schicht absorbiert wird.

Als Alternative wurde mit Direct Drive-ILA (D-ILA) ein Verfahren auf Halbleiterbasis entwickelt, das preiswerter ist. Es wird das gleiche Arbeitsprinzip wie bei ILA verwendet, nur die Röhren werden durch Halbleiter-Steuerelemente ersetzt und direkt mit dem LC-Element verbunden, wobei allerdings wieder eine feine Pixelstrukur sichtbar wird. D-ILA-Projektoren werden gegenwärtig oft dort eingesetzt, wo es wichtig ist, möglichst wenig von der Pixelstruktur zu sehen und wo man nicht von einem schnellen Schalten der Bildpunkte irritiert werden darf (s. DLP), z. B. beim Colour Grading. Für diesen Zweck werden die Geräte sehr aufwendig gebaut, sodass sie kalibrierbar sind und auf Farbprofile, die über Look Up Tables (LUT) eingegeben werden, reagieren können. Abbildung 7.44 zeigt links einen derartigen D-ILA-Projektor von JVC.

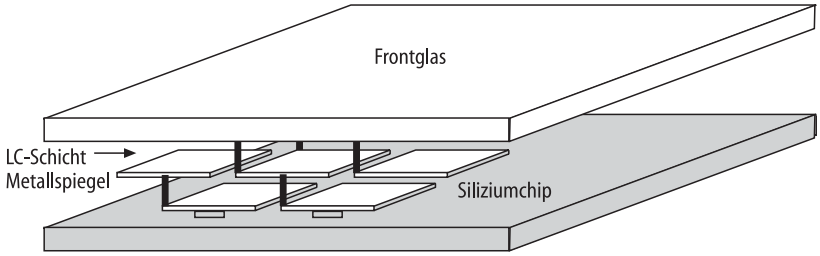


Abb. 7.43. Aufbau des LCOS-Chips

7.5.2.3 LCOS

Diese Abkürzung steht für Liquid Crystal on Silicon, d. h. dass hier, ähnlich wie bei D-ILA, mit Flüssigkristallen gearbeitet wird die direkt mit der Ansteuerelektronik verbunden sind. Das Licht kann auch hier die Elektronik nicht durchdringen, daher ist ein LCOS-Chip nur für den Reflexionsbetrieb geeignet. Das Licht gelangt zunächst zu einer Glasplatte die innen mit einer durchsichtigen und leitfähigen Schicht aus Indium-Zinn-Oxid belegt ist, die als Elektrode dient. Dahinter befindet sich eine sehr dünne Flüssigkristallschicht. Das Licht trifft dann auf eine Vielzahl von spiegelnden Aluminiumgegenelektroden über die die Spannungen an die Bildpunkte gelangen, da sie direkt mit dem Ansteuerchip verbunden sind (Abb. 7.43). Es liegt hier also eine aktive Ansteuerung wie bei einem TFT-Display vor, nur dass die Transistoren nicht in der Bildfläche liegen, was wiederum der Vorteil hat, dass der Füllfaktor hoch ist und wenig Pixelstruktur sichtbar wird. Das Funktionsprinzip ist ähnlich wie bei gewöhnlichen LCD: Die Polarisationsebene des Lichts wird auf seinem Weg durch die LC-Schicht verdreht. Oft wird, auch bei D-ILA, das Ventil so konstruiert, dass ohne Steuerspannung kein Licht transmittiert wird. Das gelingt z. B. dadurch, dass die Flüssigkristalle die Polariationsebene nicht drehen, weil sie senkrecht zu den Elektroden ausgerichtet sind (Vertical Alignment, VA). Ohne Steuerspannung, wird das Licht so in die Lampe zurückgeworfen. Über die Spannung lässt sich die Helligkeit variieren, weil darüber die Moleküle mehr oder weniger in die für die Transmission erforderliche Polarisationsebene gedreht werden. Das Funktionsprinzip wird bei Sony mit SXRD (Silicon X-tal Reflective Display) bezeichnet (Abb. 7.44 rechts) [208].

Abb. 7.44. D-ILA und LCOS-Projektoren mit jeweils 4K-Auflösung [208]

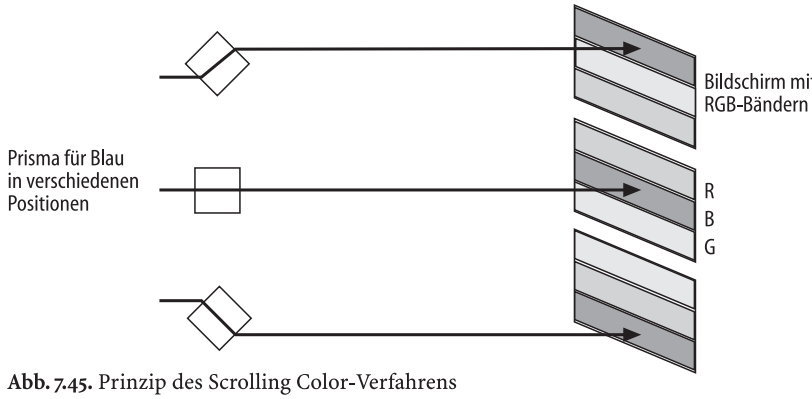


Abb. 7.45. Prinzip des Scrolling Color-Verfahrens

Ein großer Vorteil der direkten Verbindung mit einem Chip aus einkristallinem Silizium ist die hohe Elektronenbeweglichkeit und damit eine entsprechend große Schaltgeschwindigkeit.

Zur Farbdarstellung kann man wie bei gewöhnlichen LCD mit einzelnen Farbfiltern vor den Pixel arbeiten oder mit drei Chips und nach RGB aufgeteiltem Licht wie bei LC-Projektoren. Die hohe Schaltgeschwindigkeit mit einer Ansprechzeit von weniger als 0,7 ms ermöglicht bei LCOS jedoch auch ein besonderes Verfahren zur Farbdarstellung, bei dem nur ein Chip und keine aufwendige Strahlteilung erforderlich ist. In diesem sogenannten „Scrolling Color Mode“ wird das Bild vertikal in drei Streifen aufgeteilt und jeweils ein horizontaler Streifen mit einer der drei Farben Rot, Grün und Blau beleuchtet. In den entsprechenden Bildspeicherteil auf dem Chip wird dabei an drei Stellen simultan die zugehörige Teilbildinformation eingegeben. Die Farbstreifen wandern sequenziell von oben nach unten über das Bild indem sie von synchron rotierenden Prismen gesteuert werden (Abb. 7.45).

7.5.2.4 Spiegelprojektion

Diese Art der Projektion ist durch die Entwicklung digital gesteuerter, mikromechanisch arbeitender Kleinspiegel (Digital Micromirror Device, DMD) möglich geworden. Jedem Bildpunkt ist ein eigener Spiegel zugeordnet. Es ist gelungen, die Spiegel soweit zu verkleinern, dass auf einem Silizium-Chip mit 22 mm-Bilddiagonale über zwei Millionen Spiegel untergebracht werden können. Die Spiegel werden aus Aluminium gefertigt und jeder einzelne hat eine Fläche von ca. 13 μm^2.

Die Spiegel ruhen auf zwei Stützpfosten und können durch elektrostatische Anziehung, gesteuert durch ein CMOS-RAM, um ca. $\pm$ 12° verkippt werden. Sie lenken das Projektionslicht entweder auf die Bildwand oder daran vorbei in ein Absorptionsgebiet. Ein elektrostatisches Feld erzeugt das erforderliche Rückstell-Drehmoment. Die Darstellung von Graustufen wird durch einen schnellen Schaltrhythmus und die Variation der Einschaltdauer realisiert, was, wie bei Plasma-Display, mit eingeschränkter Bewegungsauflösung bei der Darstellung verbunden sein kann.

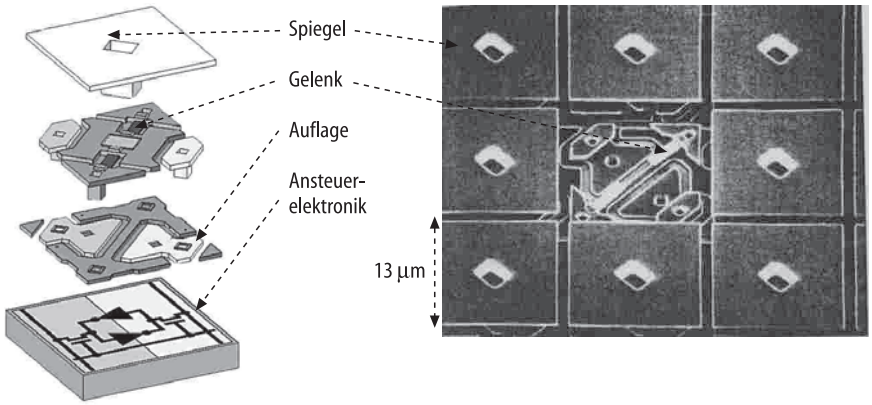


Abb. 7.46. DMD-Element und Ausschnitt aus der Spiegelfläche [106]

Die Ansteuerelektronik ist auf dem Silizium-Chip untergebracht. Jedem Spiegel sind zwei Adressierelektroden zugeordnet, sodass eine XY-Matrix-Ansteuerung möglich wird. Die Adressierung erfolgt wie bei anderen Prinzipien wieder zeilenweise. Es werden Kippzyklen bis hinab zu 15 µs erreicht, sodass auch 10-Bit-Signale verarbeitet werden können.

Abbildung 7.46 zeigt einen Ausschnitt aus der Spiegelfläche und die schematische Darstellung der beweglichen Spiegelauflage mit den zugehörigen Gelenken. Aufgrund der erforderlichen Abstände kann hier nicht so eine hohe aktive Flächendichte wie bei D-ILA erreicht werden.

DMD werden mit einer Ansteuerelektronik und dem Beleuchtungssystem zu einer Digital Light Processing-Einheit (DLP) zusammengeführt. DLP-Systeme unterscheiden sich durch die Anzahl der verwendeten DMD-Elemente für die Farbdarstellung. Die optimale Lösung ist die Trennung der drei Farbauszüge, die Reflexion an drei separaten DMD und anschließende Zusammenführung der Anteile. Dazu kann ein System aus drei verbundenen Prismen verwendet werden, die, wie das Strahlteilerprisma in der Kamera, mit dichroitischen Schichten versehen sind, die als Farbfilter dienen.

Das Verfahren wird bei professionellen Geräten eingesetzt: Im Mittelklassebereich z. B. beim Projektor PT-DW 7000 von Panasonic (Abb. 7.47), der eine Auflösung von 1366 X 768 Pixels bei 6000 ANSI-Lumen und einem Kontrast von 4000:1 erzielt und im Oberklassebereich, vor allem für Digital Cinema.

Abb. 7.47. DLP-Projektoren von Panasonic und Barco [164]

Hier haben sich die D-Cine-Projektoren von Barco, wie der Typ DP 100 etabliert, der auf DLP-Basis 2048 x 1080 Pixels bei 18000 Lumen und einem Kontrastverhältnis von 2000:1 darstellen kann (s. Abschn. 7.6.2).

DLP-Projektoren werden auch im Heimbereich oft eingesetzt. Aus Kostengründen wird dabei mit 2- und 1-Chip-Lösungen gearbeitet. Hier werden Farbräder in den Strahlengang gebracht, die als Filter dienen. Beim 1-Chip-System besteht dieses aus drei Sektoren, die die RGB-Anteile transmittieren. Der Chip wird zeitsequentiell mit den entsprechenden Farbauszügen angesteuert und das Filterrad muss synchron dazu den richtigen Sektor aufweisen. Jede Primärfarbe wird nur für ein Drittel der Bilddauer projiziert, was gegenüber dem Scrolling Color-Mode bei LCOS einen Nachteil darstellt. Beim 2-Chip-Verfahren hat das Farbfilter zwei Sektoren und transmittiert abwechselnd Rot und Blau und Rot und Grün, bzw. ein Chip ist allein für Rot zuständig, um den geringen Rotanteil der eingesetzten Lampe auszugleichen [106]. Bei 1-Chip DLP-Geräten kann durch die zeitsequentielle Farbseparation bei der Bildbetrachtung eine Irritation auftreten. Denn wenn man den Blick über die Bildwand gleiten lässt, können aufgrund der Interferenz dieser Bewegung mit der Bewegung des Farbrads die Farben kurzzeitig separat wahrgenommen werden.

7.5.2.5 Dunkelfeldprojektion

Die Dunkelfeldprojektion beruht auf der Lichtbeugung, die bewirkt, dass sich das Licht auf einen größeren Raumbereich verteilt (Abb. 7.48). Der bekannteste Projektor nach diesem Prinzip ist der bereits in den 30er Jahren entwickelte heute aber veraltete Eidophor-Projektor. Das seitlich zum Projektionsstrahlengang einfallende Licht einer Xenon-Lampe gelangt hier über ein Spiegelbarrensystem auf einen Hohlspiegel, der mit einem dünnen Ölfilm überzogen ist. Bei gleichmäßig glatt verteiltem Öl wird das Licht über die Barren zur Lampe zurückgeworfen, der Bildschirm bleibt dunkel. Wird nun die Dicke des Ölfilms z. B. mit einem Elektronenstrahl verändert, so wirkt die räumlich verschieden dicke Ölschicht wie ein Beugungsgitter. Das auf den Hohlspiegel fallende Licht wird damit gebeugt und gelangt durch die Barren hindurch zur Bildwand.

Der Elektronenstrahl wird im Fernsehzeilenraster über den Spiegel geführt und verändert in Abhängigkeit von seiner Intensität die Ölfilmdicke und damit das Maß der Beugung, bzw. die Helligkeit des Bildes (Abb. 7.49). Nach der Abtastung eines Teilbildes zerläuft der Ölfilm, er wird zusätzlich durch ein Rakel und die langsame Drehung des Spiegels geglättet.

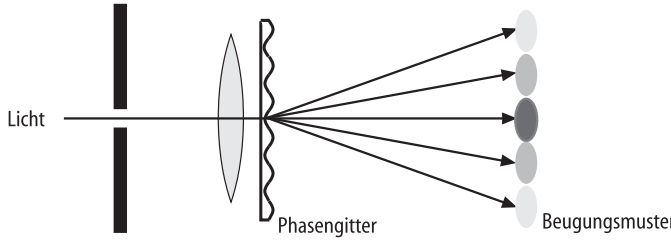


Abb. 7.48. Beugung am Phasengitter

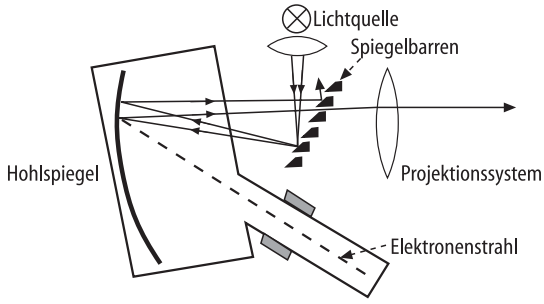


Abb. 7.49. Eidophor-Projektor

Neuere Varianten zielen darauf ab, die Ölschicht durch verformbare Festkörper zu ersetzen. Hier wird, u. a. unter der Bezeichnung Grating Light Valve (GLV), vor allem an Spiegeln geforscht, die durch Anlegen eines Bildsignals verformt werden, sodass ein Beugungsgitter entsteht [66].

7.5.2.6 Praxis der Projektion

Beim Einsatz der Projektoren in der Praxis spielt vor allem die Umgebung und ihre Lichtverhältnisse ein Rolle. Grob kann man den Einsatz für Präsentationszwecke von der kinoähnlichen Videoprojektion unterscheiden, die auch bei Heimanwendern vorzufinden ist. Präsentationen finden oft bei großer Umgebungslichtstärke statt. Der Projektor muss hier vor allem eine große Helligkeit aufweisen, um heller als die Umgebung erscheinen zu können. Dagegen ist es dann nicht so wichtig, dass das Gerät sehr kontrastreich arbeitet, da die dunklen Bildpartien vom Umgebungslicht aufgehellt werden.

Bei Heimprojektionen ist es umgekehrt: hier ist vor allem ein sehr geringer Schwarzwert erwünscht, der einen guten Kontrast gewährleistet. Um kinoähnliche Sehbedingungen zu erreichen, sind dabei geringe Leuchtdichten von ca. 50 cd/m² ausreichend. Durch sogenannte Silberleinwände kann erreicht werden, dass das Licht eine Vorzugsrichtung bekommt und stärker nach vorn als zu den Seiten reflektiert wird. Dafür wird ein Gain-Faktor angegeben. Bei hohem Gain steigert sich der Helligkeitseindruck in einem engen Raumwinkel.

Damit ein Projektor seinen geringen Schwarzwert zeigen kann, muss auf eine entsprechende Raumverdunkelung geachtet werden. Bereits das Streulicht, das auf die Wände fällt, hellt das Bild wieder auf. Der Projektor sollte so aufgestellt sein, dass sich die Unterkante des Bildes für sitzende Betrachter bei etwa 1,2 m und für stehendes Publikum bei 1,6 m Höhe befindet.

In vielen Fällen wird der Projektor an der Decke des Raums befestigt. Dafür benötigt er die horizontale und die vertikale Bildumkehr. In den meisten Fällen wird zudem eine Schrägprojektion stattfinden, die das Bild trapezförmig verzerrt. Zum Ausgleich der Verzerrung bieten hochwertige Projektoren Shift-Objektive, die das ganze Objektiv nach oben bzw. nach unten schieben, ohne dass der Projektor bewegt wird. Die elektronische Variante ist die Trapezkorrektur (Keystone) die auf einer Bildberechnung beruht, was qualitativ schlechter ist als der optische Ausgleich. Beim Erwerb eines Projektors sollte auf den möglichen Projektionsabstand (Zoom) und die Lüftergeräusche geachtet werden.

7.6 Kinoprojektion

Wie die Kinoton- so hat auch die Kinobildwiedergabe den Vorteil, dass sie recht optimalen und standardisierbaren Umgebungsbedingungen unterliegt. Der Raum ist weitgehend abgedunkelt und es kommt eine Aufprojektion mit reflektierenden Bildwänden zum Einsatz. Nach DIN soll die Leuchtdichte auf der Leinwand etwa 45 cd/m², mindestens jedoch 30 cd/m² betragen und zum Rand hin kaum abnehmen. Abhängig vom Reflexionsgrad der Leinwand sind dafür Beleuchtungsstärken erforderlich, die bei Verwendung von Xenon-Lampen und Reflektoren mit elektrischen Leistungen erreicht werden, die bei 30 W pro ausgeleuchtetem Quadratmeter Bildwand liegen. Die Lichtquelle und der Spiegel befinden sich im Lampenhaus des Projektors. Es können Lampenleistungen zwischen 1000 W und 10000 W verwendet werden, ab 2500 W sind Hitzefilter zum Schutz des Bildfensters und Kühleinrichtungen erforderlich.

7.6.1 Filmprojektion

Ebenso wie die Filmkamera hat auch ein Filmprojektor die Aufgabe, den Film schrittweise zu transportieren. Während des Bildstillstandes durchdringt das Licht einer hellen Quelle den Film und das Filmbild kann auf die Bildwand abgebildet werden. Die Geschwindigkeit ist auf 24 oder 25 Bilder/s festgelegt. Die verwendeten Antriebe sind wie bei der Kamera mit einer Umlaufblende gekoppelt. Die Umlaufblende hat einen Hellsektor, der in zwei Teile geteilt ist (Abb. 7.50). Alternativ wird eine Blende verwendet, die mit doppelter Geschwindigkeit rotiert. Innerhalb der Bildwechselperiodendauer treten damit zwei Hell- und zwei Dunkelphasen auf. Während der Hellphasen durchdringt das Licht einer starken Xenonlampe den Film und über das Projektionsobjektiv entsteht auf der Bildwand das für die Zuschauer sichtbare Bild. Die Dunkelphasen sind zeitlich so gewählt, dass die eine genau im Zeitraum des Filmtransportes liegt und diesen damit unsichtbar macht. Die zweite Phase dient nur der symmetrischen wiederholten Unterbrechung des Lichtstroms, womit auf der Leinwand die Hell/Dunkel-Phasen mit einer Frequenz von 48 Hz entstehen, wenn die Filmgeschwindigkeit 24 Bilder/s beträgt. Jedes Bild wird also zweimal gezeigt. Die Verdopplung der Bildfrequenz ist deshalb erforderlich, weil eine Frequenz von 24 Bildern/s zwar ausreicht, um bezüglich der abgebildeten Bewegungen einen Verschmelzungseindruck zu erzeugen, aber nicht bezüglich der Helligkeitswahrnehmung des gesamten Bildes. Ein mit 24 Hz projiziertes Bild er-

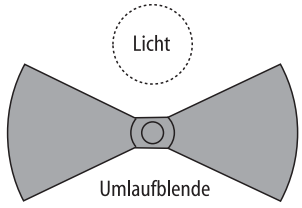


Abb. 7.50. Umlaufblende im Filmprojektor

scheint flackernd, das Auge empfindet das störende Großflächenflimmern. Die Stärke dieser Erscheinung hängt vor allem von der Lichtintensität und der Bildfrequenz ab. Bei den im Kino verwendeten geringen Leuchtdichten von ca. 50 cd/m² reichen 48 Hz aus, um die Flimmerverschmelzungsgrenze zu erreichen, sodass also ein doppelt projiziertes Bild weitgehend flimmerfrei ist und das Großflächenflimmern nur noch bei geringem Betrachtungsabstand in sehr hellen Bildteilen auffällt. Teilweise wird aber auch mit einer Verdreifachung der Bilddarstellung, also mit 72 Hz gearbeitet.

Das Verfahren der Unterbrechung des Lichtstroms ist simpel, kann aber bei Videosystemen nicht angewendet werden bzw. nur dann, wenn Videobilder gespeichert vorliegen. Hier wird stattdessen das Zeilensprungverfahren angewandt. Auch auf diese Weise lässt sich die Flimmerverschmelzungsgrenze erreichen, doch erzeugt das Verfahren eine gegenüber der Filmtechnik erheblich veränderte Bildwirkung, da durch das Zeilensprungverfahren aufnahmeseitig eine verdoppelte zeitliche Auflösung der Bewegungen entsteht.

Die technische Ausführung eines Kinoprojektors zeigt Abbildung 7.51 am Beispiel des verbreiteten 35-mm-Projektors FP 30 der Firma Kinoton. Oben und unten sind jeweils die Filmspulen zu erkennen, die mit Filmlängen von 600 m, 2000 m oder 4000 m verwendet werden können. Die Spulenachsen können entfernt werden, sodass ein Betrieb mit Filmtellereinrichtungen oder Spulentürmen möglich wird. Die Verwendung von Spulen hat den Nachteil, dass der Film nach dem Programmende wieder zurückgespult werden muss, damit der Filmbeginn wieder außen liegt. Bei Filmtellereinrichtungen wird der Beginn nach innen gewickelt und von dort aus dem Projektor zugeführt. Auf dem Aufwickelteller kommt der Beginn automatisch wieder nach innen, sodass der Film direkt für die nächste Projektion bereit ist. Die horizontal liegenden Filmteller (Abb. 7.51) fassen bis zu 7000 m Film, was 2,5 Stunden Programm entspricht [126].

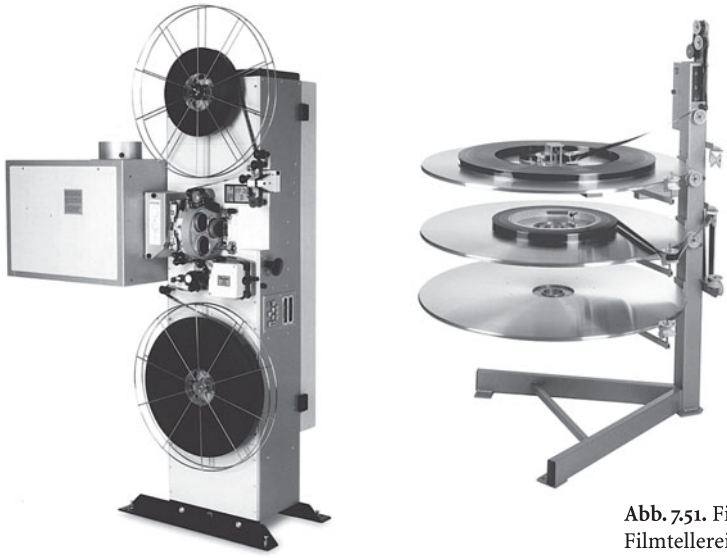


Abb. 7.51. Filmprojektor und Filmtellereinrichtung [111]

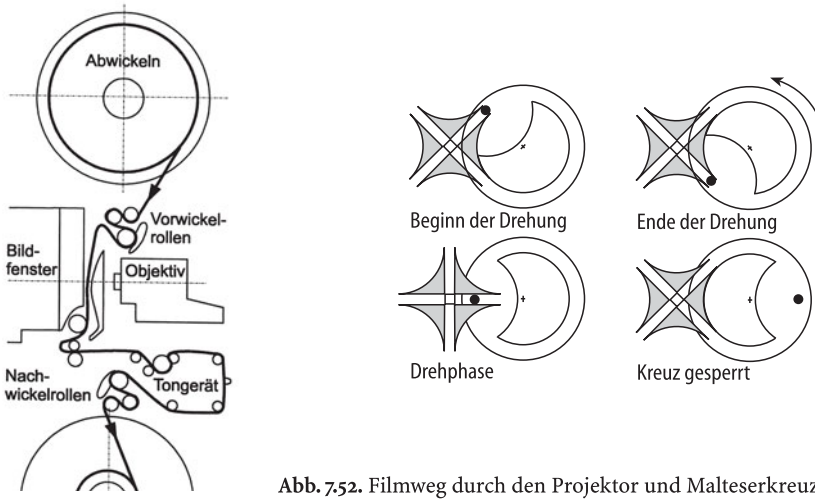
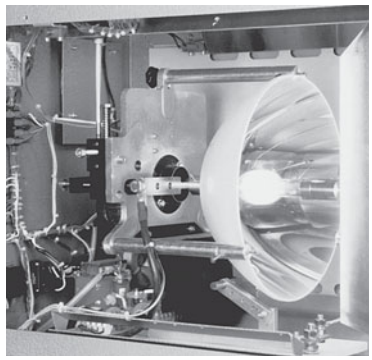


Abb. 7.52. Filmweg durch den Projektor und Malteserkreuzantrieb

Abbildung 7.52 zeigt ein Schema in dem der Filmweg im Projektor deutlich wird. Der Film läuft von oben nach unten und wird über eine Schaltrolle bewegt, die über ein Malteserkreuz oder über einen elektronischen Antrieb bewegt wird.

Das Malteserkreuzgetriebe ist ein recht ungewöhnlicher Antrieb, der sich aber seit dem Beginn der Filmtechnik bewährt hat. Es hat die Funktion, eine gleichmäßige Rotation in eine intermittierende Bewegung umzusetzen. Das Arbeitsprinzip beruht darauf, dass ein Stift in das Malteserkreuz greift und über eine Welle die Schaltrolle um eine Viertelumdrehung weiterdreht, wobei der Film um genau ein Bild bewegt wird. Für den restlichen Teil der vollen Umdrehung wird der Bildstillstand gewährleistet, indem der Sperrbogen verhindert, dass sich das Kreuz weiter dreht (Abb. 7.52). Statt des Malteserkreuzes kann ein elektronischer Antrieb verwendet werden, wobei sich ein verbesserter Bildstand und eine höhere Lichtleistung durch längere Standzeiten ergeben. Der Antrieb ist justierbar, damit das Filmbild auch bei unkorrekten Klebestellen oder ähnlichen Fehlern wieder in das Bildfenster gebracht werden kann und der Bildstrich verschwindet.

In Abbildung 7.53 wird das Innere des Lampenhauses, ein Revolverkopf für drei Projektionsobjektive und das Bildfenster sichtbar. Der Film wird hier über einen Andruckmechanismus an eine gekrümmte Filmbahn gedrückt, um eine optimale Bildschärfe zu erreichen. Mit Hilfe des Revolverkopfes wird die Auswechslung der Objektive erleichtert, zur Anpassung an verschiedene Bildformate. Typische Brennweiten der Objektive lassen sich mit der Beziehung $a/a' = y/y'$ leicht berechnen, wenn die zulässige Näherung gemacht wird, dass die Brennweite f der Gegenstandsweite a des abzubildenden Filmbildes entspricht. Für die Abbildung eines 35-mm-Films mit y = 13,25 mm Bildhöhe auf eine Projektionswand der Höhe $y' = 4$ m im Abstand $a' = 24$ m ist danach eine Brennweite von f = a = 80 mm erforderlich. Eine anamorphotische Bildentzerrung wird durch die Verwendung von Vorsatzlinsen erreicht.



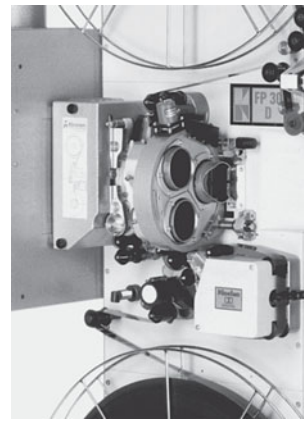


Abb. 7.53. Lampenhaus, innen, Bildfenster und Projektionsobjektiv [111]

Das Einlegen des Films erfolgt von oben nach unten. Bei korrekt eingelegtem Film stehen die Bilder im Bildfenster auf dem Kopf und die Lichttonspur liegt in Projektionsrichtung gesehen an der rechten Seite. Die beschichtete Seite des Films ist der Lichtquelle zugewandt [100]. Zur Anpassung des schrittweisen Filmtransports im Bildfenster an den kontinuierlichen Lauf, der für die Tonwiedergabe erforderlich ist, wird eine Schlaufe gebildet, mit einer Länge, die gewährleistet, dass die Lichttonwiedergabe um 20 Bilder gegenüber dem Bildfenster versetzt ist. Die Wiedergabeeinheit mit zwei Abtastsystemen für Analogton und Dolby Digitalton (s. Kap. 5) wird in Abbildung 7.53 unterhalb des Revolverkopfes sichtbar. Zur Wiedergabe von Audiosignalen nach DTS oder SDDS muss eine weitere Abtasteinheit montiert werden.

7.6.2 Digitalprojektion im Kino

Hinsichtlich der mit dem Begriff Electronic Cinema bzw. Digital Cinema verknüpften Vision einer volldigitalen Kette zwischen Filmakquisition und -präsentation seien hier Bildwiedergabeprinzipien für die Projektion lichtstarker hoch aufgelöster Bilder dargestellt. Die wesentlichen Beurteilungsparameter sind neben der erzielbaren Beleuchtungsstärke auch hier wieder die Bildauflösung, der darstellbare Kontrastumfang und der Farbbereich.

Während auf der Produktionsseite auf möglichst hohe Bildauflösung Wert gelegt werden sollte, damit genügend Bearbeitungsspielraum bei der Postproduktion erhalten bleibt, steht die Ortsauflösung bei der Projektion nicht so stark im Mittelpunkt. Für eine gute Ortsauflösung des Bildes genügt es, dass 2048 x 1556 bzw. nach den DCI-Vorgaben (Digital Cinema Initiative, s. Abschn. 7.6.3) 2048 x 1080 Bildpunkte zur Verfügung stehen, sodass unter Einbeziehung möglicher Projektionsverluste mindestens eine mit 1,5 k vergleichbare Auflösung auf der Leinwand erreicht wird, so wie es auch für gängige Filmprojektionen angenommen werden kann. Es ist wünschenswert, dass die Bildauflösung der elektronischen Quelle direkt der möglichen Projektorauflö-

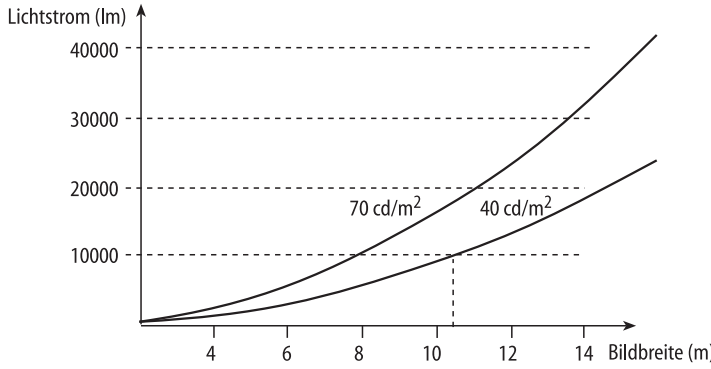


Abb. 7.54. Für Kinoprojektionen erforderliche Lichtstärken in Abhängigkeit von der Bildbreite

sung entspricht, da eine Anpassung und Interpolation die Bildqualität immer verschlechtert. Generell ist es ideal, an der Quelle mit der Auflösung 2k oder besser noch 4k zu arbeiten, die während die Bildbearbeitung beibehalten und für die Projektion direkt übernommen bzw. ohne Interpolation einfach halbiert wird. Für Auflösungsanforderungen, die über den normalen Kinobereich hinausgehen, lassen sich Parallelprojektionen verwenden.

Von großer Bedeutung bei der Bildwiedergabe ist die erreichbare Bildhelligkeit und auch die Dunkelheit, d. h. der Kontrastumfang. Die Lichtstärke des Projektors und die damit erzielbare Beleuchtungsstärke E oder Leuchtdichte L hängt von der Größe der bestrahlten Fläche ab. Abbildung 7.54 zeigt die erforderlichen Lichtströme zur Erzielung der Leuchtdichten 40 cd/m² und 70 cd/m² in Abhängigkeit von der Bildwandbreite [74]. Um beispielsweise auf einer Kinoleinwand von ca. 10 m Breite die Kino-Soll-Leuchtdichte von mindestens der von der DCI geforderten 48 cd/m² zu erreichen, ist danach ein Projektorlichtstrom von mehr als 10 000 ANSI-Lumen erforderlich. Die Be-

Abb. 7.55. Digitale Kinoprojektoren von Barco, Christie und Sony [186]

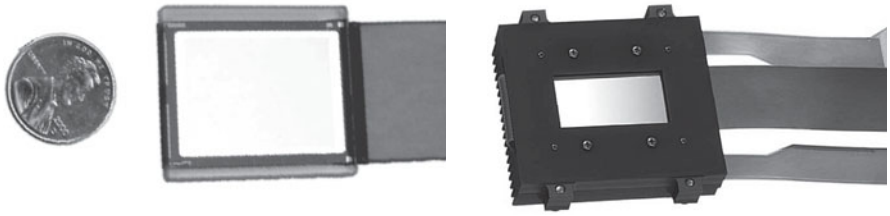


Abb. 7.56. D-ILA-Panel im Vergleich zu einer Münze und SXRD-Panel [105]

zeichnung ANSI-Lumen bezieht sich auf eine Standardisierung, bei der die ggf. ungleichmäßige Leuchtdichteverteilung über der Fläche durch Mittelung berücksichtigt wird. Auch der Kontrast- und Farbumfang der Projektion sollte mit der von Filmbildern vergleichbar sein.

In der Praxis dominieren für die digitale Kinoprojektion Systeme nach dem DLP-Prinzip mit drei separaten DMD für RGB. Bekannte Hersteller sind Christie und Barco deren Typen CP 2000 und DP 100 in Abbildung 7.55 dargestellt sind. Der Christie-Projektor erzeugt 12000 lm, das Barco-Gerät 22000 lm.

Die Konkurrenz im Kino ist das D-ILA- bzw. LCOS-Prinzip. Abbildung 7.56 zeigt ein D-ILA-Panel. Aufgrund der Ansteuerung von hinten werden hier 93 % der Gesamtfläche als Bildwandlerfläche genutzt, d. h. dass die Struktur der 2048 x 1536 Pixels bei D-ILA nur schwer wahrnehmbar ist. Die Bildwandlergröße beträgt 26 mm x 20 mm. Als Kontrastumfang wird 2000 : 1 angegeben.

Oft sind die erzielbaren Leuchtdichten der D-ILA und LCOS-Systeme für große Kinoleinwände nicht ausreichend. Eine Ausnahme macht da das SXRD-System der Firma Sony, der es gelang, auf LCOS-Basis den ersten 4k-Projektor zu bauen (Abb. 7.55 rechts). Hier werden bis zu 18000 lm erzielt. Das LC-Element arbeitet im Vertically Aligned Nematic (VAN)-Modus. Dabei ist der Strahlengang ohne Steuerspannung gesperrt, wodurch ein guter Schwarzwert erreicht wird. Die LC-Schicht ist bei SXRD mit 2 µm sehr dünn, sodass ein sehr gutes Schaltverhalten gegeben ist (Abb. 7.57). Das Bild besteht beim SXRD-Projektor aus vier Teilbildern mit je 2048 x 1080 Pixels, die für Messeanwendungen etc. separat und für hochaufgelöste Videodarstellung als Einheit betrieben werden können (Abb. 7.56). Das ermöglicht auch die Zweiteilung der Fläche und damit eine besondere 3D-Darstellung mit Polariationsverfahren, das die zwei Teilbilder mit einer Doppeloptik gleichzeitig zeigen kann (s. Abschn. 7.6.4).

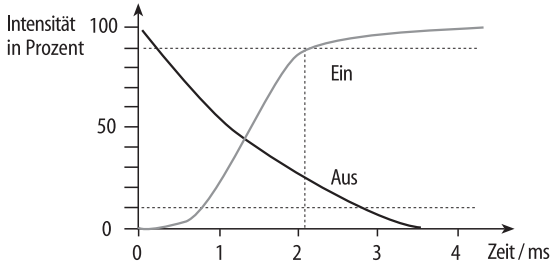


Abb. 7.57. SXRD-Schaltzeiten

7.6.3 Digital Cinema

Dieser Begriff bzw. der Begriff Electronic Cinema beschreibt den Einsatz der Digitaltechnik am Ende der Filmproduktions- und Übertragungskette, also im Kino. Während die Verwendung der Digitaldaten im Bereich Postproduktion bereits früh etabliert und auch bei der Bildaufnahme in den letzten Jahre Einzug gehalten hat, ist das Kino das letzte Glied einer möglichen volldigitalen Übertragungskette, das noch nicht ganz digitalisiert ist.

Als generelle Anforderung an Digital Cinema gilt, dass die Bild- und Tonqualität nicht schlechter sein darf als im konventionellen Kino. Aufgrund der Tatsache, dass hochwertige und lichtstarke elektronische Projektoren verfügbar sind (s. Abschn. 7.6.2), gilt die Entwicklungsarbeit vor allem der Frage nach den digitalen Übertragungsverfahren und der Verhinderung von Piraterie. Um zu kostengünstigen Lösungen zu gelangen, wird dabei von technischen Systemen ausgegangen, die sich bereits etabliert haben.

Für die Übertragung gilt es zunächst die Frage zu klären, ob die Filmdigitaldaten auf großen Datenträgern ins Kino gebracht oder in datenreduzierter Form direkt übertragen werden. Da ein wesentlicher Vorteil des elektronischen Kinos darin gesehen wird, dass es den aufwendigen Versand von Kinokopien nicht mehr erforderlich macht, sollte eine Datenübertragung über Netzwerke oder drahtlos, z. B. via ATM oder per Satellit, erfolgen können. Dies impliziert aus ökonomischen Gründen die Verwendung von Datenreduktion. Zunächst geht es also um die Signaldefinition. Hier liegen mit den Parametern für HDTV Bilddatenbeschreibungsformen vor, die für das elektronische Kino geeignet erscheinen. Man geht davon aus, dass die zugehörige Bildauflösung von 1920 x 1080 oder 2k Bildpunkten auf jeden Fall für Leinwandbreiten über 10 m geeignet ist. Für diese Dimensionen sind Projektoren mit Lichtstärken von mindestens 10 000 ANSI-Lumen erforderlich.

Nachdem die Daten in das digitale Videoformat umgesetzt sind, wird die Datenreduktion eingesetzt (Abb. 7.58). Für einen Bereich wie Digital Cinema, in dem es nur darum geht, ein fertiges Endprodukt zu den Endkunden zu bringen, kann mit datenreduzierten Signalen gearbeitet werden. Wegen der besseren Qualität gegenüber MPEG-2 werden modernere Codecs wie H.264/

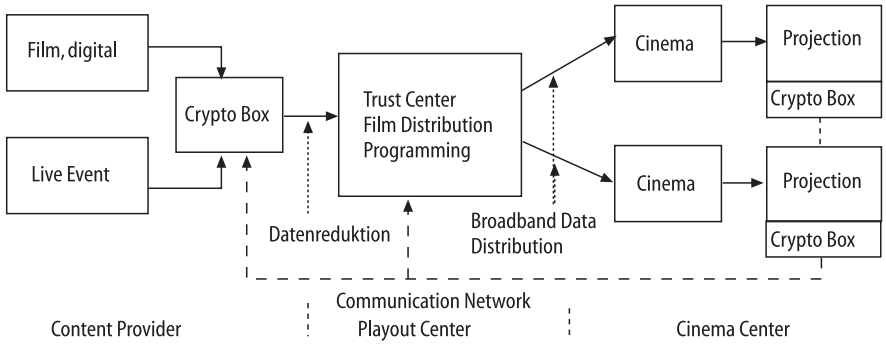


Abb. 7.58. Konzept zu Digital Cinema

Tabelle 7.5. DCDM-Parameter nach DCI-Specification 1.2

Auflösung	2k	1998 x 1080 Pixel oder 2048 x 858 Pixel (Scope)
	4k	3996 x 2160 Pixel oder 4096 x 1706 Pixel (Scope)
Aspect Ratio	Bild: 1,85:1 oder 2,39:1	Pixel: 1:1
Bildfrequenz	24 Hz, 48 Hz	
Farbraum	XYZ	
Weißpunkt	6500 K	
Gamma	2,6	
Codierung	12 Bit	

AVC und JPEG2000 bevorzugt; letzterer insbesondere deshalb, weil er sehr gut für hoch aufgelöste Bilder geeignet und zudem leicht skalierbar ist.

Neben der Klärung der technischen Daten muss ein Konzept für Digital Cinema auch berücksichtigen, dass hier kein gewöhnlicher Broadcastfall vorliegt, sondern nur Kunden das Datenmaterial erhalten dürfen, die dafür bezahlt haben. Daher ist ein Management der Aufführungslizenzen erforderlich, das der Generierung und Verwaltung der Verschlüsselungsdaten dient (Abb. 7.58). Weiterhin ist es erforderlich, auch die Fragen des Programmabrufs zu bedenken, d. h., ob die Kinoprogrammierung nach einem Push- oder Pull-Modell ablaufen soll, ob also der Distributor das übertragene Programm zentral bestimmt oder ob der Programmabruf vom Kinobetreiber aus erfolgt.

Damit diese Überlegungen zu einem internationalen Standard führen können, hat eine Vereinigung bekannter Filmstudios die Digital Cinema Initiatives (DCI) gegründet [45]. In deren Digital Cinema System Specification werden entsprechende Parameter festgelegt und zwar unter der Prämisse, dass die Wiedergabe der digitalen Bilder im Kino einen Qualitätsstandard erreicht, der mindestens der Wiedergabe von 35-mm-Film entspricht. Hier ist das für die Verteilung benutzte Digitalformat definiert, die verwendete Datenreduktion, die Verpackung der Daten sowie die Anforderungen an das Theater System.

Bei der DCI wird das Quellmaterial als Digital Source Master (DSM) bezeichnet. Es enthält alle Bild-, Ton- und Zusatzdaten ist aber wenig spezifiziert. Es ist davon auszugehen, dass es unkomprimiert mit einer Qualität vorliegt, die nicht schlechter ist, als die der nächsten Schritte der DCI-Kette. Sehr viel mehr Raum nimmt die Beschreibung des Digital Cinema Distribution Masters (DCDM) ein. Hier sind wesentliche Signalparameter festgelegt. Für die Ortsauflösung wurden zwei Stufen definiert, nämlich 2k und 4k, mit 2048 x 1080 und 4096 x 2160 Bildpunkten. Die Bildpunkte sind immer quadratisch. Verschiedene Bildseitenverhältnisse werden durch Beschneidung erreicht, zugelassen sind erstens das Bildformat 1,85:1 [92]. Es wird beim 2k-Level durch Nutzung der vollen Bildhöhe mit 1080 Zeilen, aber Beschneidung der Horizontalen auf 1998 Pixel erreicht. Zweitens ein Scope-Format mit 2,39:1 unter Nutzung aller 2048 Bildpunkte in der Horizontalen, während die Vertikale auf 858 aktive Zeilen eingeschränkt wird. Jeder Bildpunkt wird mit 12 Bit codiert. (Tabelle 7.5)

Der Bildaufbau ist in zeitlicher Hinsicht auf 24,000 fps festgelegt, bei 2k-Daten werden auch 48 fps unterstützt. An dieser engen Beschränkung gibt es Kritik, denn die Umsetzung von Produktionen mit 25 Hz Bildwechselfrequenz für das digitale Kino ist mit einer zeitlichen Konversion und damit mit Qua-

Tabelle 7.6. DCI-Parameter für die Projektion

Auflösung	2k mit 2048 x 1080 Pixels oder 4 k mit 4096 x 2160 Pixels
Leuchtdichte	48 cd/m²
Weißpunkt	x = 3,14 y = 3.51
Gamut minimal:	Rot: x = 0.680, y = 0.320 Grün: x = 0.265 y = 0.690 Blau: x = 0.150, y = 0.060
Kontrast	sequenziell: min. 2000:1; intraframe: min. 150:1
Gamma	2,6

litätsverlusten verbunden. Insbesondere aus Europa wird daher gefordert, die DCI-Spezifikationen um 25 fps und auch Bildraten zwischen 8 und 22 fps zu erweitern, um historisch wertvollen alten Filmen auch originalgetreu Rechnung tragen zu können.

Als Farbraum (Gamut) wurde für das DCDM ein XYZ-Koordinatensystem zugrunde gelegt. Der Vorteil ist, dass geräteunabhängig Farbbereiche genutzt werden können, die z. B. ein erheblich größeres Gamut als HD-Videosysteme aufweisen [204]. Nachteilig ist, dass sowohl auf der Postproduktionsseite eine RGB-XYZ- als auch bei der Wiedergabe eine XYZ-RGB-Konversion erforderlich ist. Zudem könnte theoretisch ein sehr kleines Gamut verwendet werden. Um diesen Fall zu verhindern, legt die DCI einen Minimalfarbraum fest, dessen Koordinaten in Tabelle 7.6 neben weiteren Parametern für die Projektion aufgeführt sind.

Die DCDM-Audiospezifikationen legen Abtastraten von 48 kHz oder 96 kHz für bis zu 16 Audiokanäle fest. Tabelle 7.7 zeigt, dass bisher sechs oder acht in ihrer Zuordnung spezifiziert sind, die Nummerierung links bezieht sich auf digitale Audiopaare nach AES/EBU.

Als Kompressionsalgorithmus für das Bild ist nur JEG2000 spezifiziert, der so eingesetzt werden soll, dass die damit verbundene Qualitätsminderung visuell verlustlos, also keinesfalls sichtbar ist. Die maximale Datenrate ist auf 250 Mbit/s festgelegt. Die Skalierbarkeit von JPEG2000 erlaubt die Ableitung von 2k- aus 4k-Daten. Die komprimierten Videodaten werden zusammen mit den Audio- und Subtiteldaten in ein DCP (Digital Cinema Package) verpackt. Als Fileformat ist MXF (s. Kap. 10) festgelegt [14].

Für den Transport der Daten werden mehrere Möglichkeiten offengelassen. Die DCI fordert vor allem, dass dabei Sicherheitsaspekte beachtet werden. Bezüglich der Anlieferung des Materials bestehen folgende Möglichkeiten: via

Tabelle 7.7. DCI-Tonkanalzuordnung

1/1	1	L/Left	Far left screen loudspeaker
1/2	2	R/Right	Far right screen loudspeaker
2/1	3	C/Center	Center screen loudspeaker
2/2	4	LFE/Screen	Screen Low Frequency Effects subwoofer lsp
3/1	5	Ls/Left	Surround left wall surround loudspeakers
3/2	6	Rs/Right	Surround right wall surround loudspeakers
4/1	7	Lc/Left	Center mid left to center screen loudspeaker
4/2	8	Rc/Right	Center mid right to center screen loudspeaker
5/1	9	Unused	
	10 ... 16	Unused/User	Defined

Satellit, über terrestrische Glasfasernetzwerke oder via Speichermedien. Die Daten werden im Theater gespeichert, die Speicher werden in zentral und lokal unterschieden. Erstere speichern die Filmsequenzen für einen ganzen Theaterkomplex, Letztere für eine einzelne Leinwand. Die Kapazität muss hier für einen dreistündigen Film plus 20 Minuten Zusatzfilmmaterial und 16 Audiokanäle bei 24 bit/48 kHz ausreichen. Für einen Film, der mit einer typischen Bitrate von 200 Mbit/s codiert ist, muss also insgesamt ein drittel Terabyte verfügbar sein.

Die zentrale Rolle im Theatersystem spielt der Media Block. Hier werden die Daten ausgepackt, die dann mit dem Kompressionsdecoder in die Ausgangsdaten zurückgeführt und entschlüsselt werden. Da diese Stufen innerhalb des Media Blocks liegen, der auch das Projektionssystem enthält, ist sichergestellt, dass die entschlüsselten Daten an keiner Stelle zugänglich sind, was wiederum eine sehr wichtige Komponente für die Sicherheit darstellt. Bekannte Kinoserver bzw. Kino-Showplayer stammen von der Firma XDC (Abb. 7.59).

Die Projektion selbst soll eine native, d. h. unskalierte Auflösung von 4096 x 2160 oder 2048 x 1080 Bildpunkten aufweisen. Die Geräte müssen in der Lage sein, eine Anpassung der Daten mit dem Faktor 2 durchzuführen. Für die Leuchtdichte, die schließlich auf der Leinwand entsteht, wurde ein Wert von 48 cd/m² festgelegt. Der sequenzielle Kontrast, folgend aus abwechselnd schwarzen und weißen Bildern, sollte dabei mindestens bei 2000:1 liegen, der Kontrast im Bild (Schachbrettmuster) bei 150:1 (Tabelle 7.6). Die Projektionsleinwand muss nach den DCI-Spezifikationen das ganze sichtbare Lichtspektrum gleichmäßig reflektieren und eine variable schwarze Maskierung zur Anpassung an verschiedene Bildformate ermöglichen.

Trotz hoher Kosten für die Projektoren sollen sich mit Digital Cinema langfristig auch Kostenvorteile ergeben – einerseits durch Vermeidung der Herstellung und des Vertriebs vieler teurer Kinokopien und andererseits durch neue Nutzungsformen, da mit der neuen Technik auch Live-Events wie z. B. Opernkonzerte in die Kinos übertragen werden können und sich eine individuellere Anpassung der Werbung an das Kino und seine Zuschauer erreichen lässt.

Die DCI-Spezifikation unterstützen aufgrund der Bildrate von 48 Hz bei 2k-Daten auch die 3D-Bildwiedergabe. Da die Projektoren sehr hohe Bildraten verarbeiten, können die zwei 3D-Halbbilder abwechselnd mit jeweils 24 Hz bzw. Mehrfachem davon dargestellt werden, dabei soll in jeder Sequenz das linke Bild zuerst kommen und sie soll mit dem rechten abschließen. Im nächsten Abschnitt ist dargestellt, wie die Teilbilder separat den beiden Augen der Betrachter zugeführt werden und dass dies im Zuge von HFR (High Frame Rate) ggf. auch mit weiter erhöhter Bildrate geschehen kann.

Abb. 7.59. Kinoshowplayer für das digitale Kino

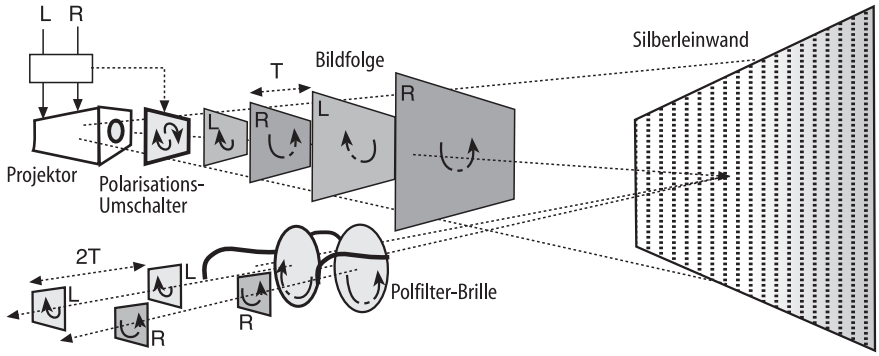


Abb. 7.60. 3D-Wiedergabe im Kino mit dem Polarisationsverfahren (RealD)

7.6.4 3D-Kino

Mit der Verfügbarkeit schnell reagierender digitaler Projektoren ist verknüpft, dass die stereoskopische Darstellung bezüglich des Wiedergabegerätes keine besonderen Kosten verursacht. Bei der Umrüstung eines Kinos auf Digitaltechnik ergibt sich die 3D-Tauglichkeit, abgesehen von einer ggf. erforderlichen Umrüstung der Bildwand, mit geringem Aufwand. Nicht zuletzt aus diesem Grunde wurde S3D zu einem wichtigen Thema der Filmbranche, wieder, könnte man sagen, aber diesmal mit mehr Aussicht auf Dauerhaftigkeit.

Der letzte Satz deutet an, dass S3D nicht völlig neu ist. Die frühesten bekannten und dokumentierten Versuche zur stereoskopischen Darstellung stammen von Wheatstone, der bereits im Jahre 1838 ein Stereoskop vorstellte. Stereoskope sind Geräte, in die man hineinschaut, sie dienen, wie heute die Brillen, der Separation der Teilbilder. Ein Stereoskop ist direkt vor dem Kopf zu verwenden und ähnelt den Head Mounted Displays (HMD), einer auf dem Kopf zu tragenden Helmkonstruktion, die einen separaten Bildschirm für jedes Auge enthält. Stereoskope und Head Mounted Displays nutzen also den einfachsten Weg der Bildtrennung, indem sie den beiden Augen jeweils direkt das zugehörige Halbbild zuordnen und alles andere ausblenden. Von Vorteil ist die optimale Trennung der Halbbilder. Nachteilig ist, dass das Betrachten ermüdend und wenig kommunikativ ist, was selbst viele Nutzer von 3D-Spielen stört.

In den 50er-Jahren des letzten Jahrhunderts gab es eine kurze Blüte des 3D-Kinos durch die Einführung des Anaglyphen-Verfahrens. Dabei beruht die Bildtrennung auf einer Farbseparation. Die beiden Halbbilder werden hier unterschiedlich eingefärbt und dann überlagert. Die doppelten Konturen in den Bildern, die sich aus den Differenzen zwischen den Teilbildansichten ergeben, erscheinen dabei farbig. Damit die Betrachter nun pro Auge nur ein Teilbild sehen, benutzen sie Brillen, die auf jeder Seite jeweils eine der beiden Farben ausfiltern. Die meistens verwendeten Farben sind Rot und Cyan, wobei sich durchgesetzt hat, Rot für das linke Auge zu verwenden. Anaglyphen haben den Vorteil, unabhängig von der Verbreitungsart, im Printbereich ebenso einsetzbar zu sein wie im Kino. Allerdings können Farben kaum wiedergegeben werden.

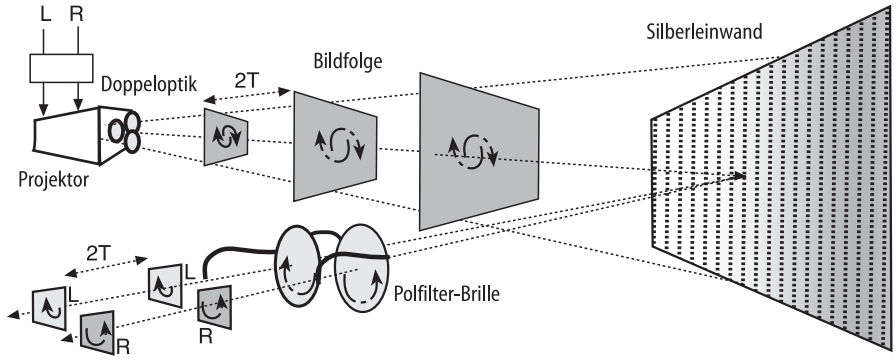


Abb. 7.61. 3D-Wiedergabe im Kino nach dem Polarisationsverfahren mit dem 4k-Projektor

7.6.4.1 Polarisationsverfahren

Das Polarisationsverfahren ist im Gegensatz zum Anaglyphen-Verfahren voll farbtauglich. Wie der Name sagt, wird hier die Trennung der Teilbilder mit Hilfe von Polarisationsfolien vorgenommen. Sie werden in den Brillen angeordnet, die die Betrachter tragen. Wichtig ist, dass sie dabei bezüglich ihrer Transmissionsrichtung um 90° gegeneinander verdreht sind. Bei der Bilddarstellung muss dafür gesorgt werden, dass die beiden Halbbilder so auf der Bildwand erscheinen, das das Licht der beiden Quellen wiederum um 90° gegeneinander verdreht polarisiert ist. Daraus resultiert dann, dass jeweils die Schwingungsrichtung der Lichtwelle der einen Teilabbildung parallel zu einem der Polarisatoren in der Brille liegt, was dann bedeutet, dass sie für das andere Auge gesperrt wird.

Damit eine Kopfneigung die Trennung der Teilbilder und damit den 3D-Effekt nicht mindert, wird in der Praxis mit rechts/links-zirkularer statt mit linearer Polarisation gearbeitet (Abb. 7.60 und 7.61). Unter Bezeichnung RealD hat dieses Verfahren eine sehr große Verbreitung in den Kinos gefunden. Von Vorteil ist hier, dass die Polarisationsbrillen, die an die Zuschauer verteilt werden, preisgünstig sind. Nachteilig ist der starke Lichtverlust durch das Durchlaufen zweier Polarisationsfilter. Bereits eine einzelne Polarisationsfolie absorbiert mehr als die Hälfte des Lichts. Wiedergabeseitig arbeitet das System mit einem, vor dem Projektor angebrachten Z-Screen, einer Projektorvorsatzeinheit auf Basis von Flüssigkristallen, die das Projektorlicht zirkular polarisiert und die Polarisationsrichtung synchron zu den ankommenden Halbbildern nach jedem Teilbild zwischen rechts- und links-zirkular umschaltet.

RealD kann mit üblichen Digitalprojektoren problemlos verwendet werden, solange sie schnell genug arbeiten. Die DCI-Spezifikationen sehen für den 3D-Betrieb vor, dass die Bildfrequenz für beide Halbbilder jeweils 24 fps beträgt, d. h. bei zeitsequentieller Darstellung (Frame Sequential) auf insgesamt 48 Hz verdoppelt wird. Wie im Kino üblich, wird dann zusätzlich die Darstellungsfrequenz durch Bildwiederholung erhöht, um ein Großflächenflimmern zu verhindern. Bei RealD wird mit dem Triple-Flash-Verfahren, also der dreifachen Wiedergabe eines jeden Halbbildes gearbeitet, was insgesamt eine Frequenz

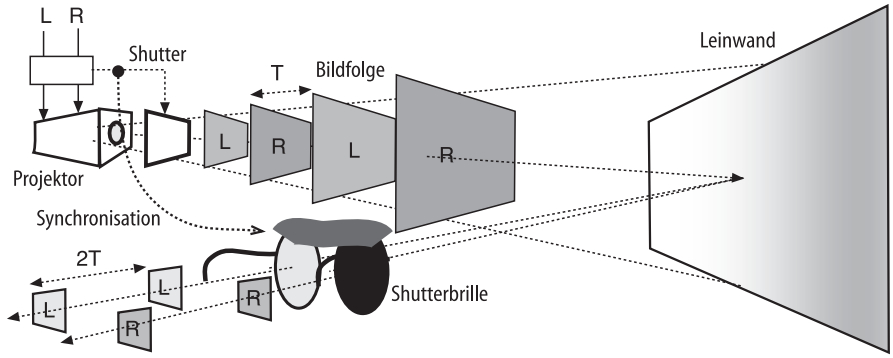


Abb. 7.62. 3D-Wiedergabe im Kino mit dem Shutter-Verfahren

von 144 Hz bewirkt. Während hier die Stereo-Halbbilder, üblicherweise mit DLP-Projektoren, zeitlich nacheinander gezeigt werden, kann bei Verwendung eines Sony-SXRD-Projektors dagegen die 4k-Auflösung den Projektors dazu genutzt werden, mit einer Doppeloptik gleichzeitig, und nicht nacheinander, zwei polarisierte Bilder mit jeweils 2k-Auflösung zu zeigen (Abb. 7.61).

Die Voraussetzung für die Bildtrennung mittels des Polarisationsverfahrens ist, dass die Polarisationsrichtung auf dem Lichtweg erhalten bleibt. Das bedeutet insbesondere, dass spezielle Bildwände (Silberleinwände) zum Einsatz kommen müssen, die das Licht polarisationserhaltend zurückwerfen. Eine übliche, mattweiße Projektionswand streut das Licht. Beim Einsatz der Silberleinwand ist nachteilig, dass die Außenbereiche eines Bildes dunkler wirken und es somit zu sogenannten Hot Spots, d. h. übermäßig hellen Bildbereichen, kommen kann, die abhängig von der Betrachtungsposition, mit einem verkleinerten Betrachtungswinkel verbunden sind. Diesem Umstand kann durch eine Krümmung der Leinwand entgegengewirkt werden. Silberleinwände sind teurer als einfache Leinwände und lassen sich i. d. R. auch nicht aufrollen.

7.6.4.2 Shutterverfahren

Das Shutterverfahren bietet den Vorteil, dass keine spezielle Bildwand erforderlich ist und die Trennung der Halbbilder besser gelingt als beim Polarisationsverfahren. Bei diesem Verfahren werden Brillen verwendet, die über ein elektrisches Signal schnell verschließbar sind. Die Verschlüsse werden mit LC-Elementen realisiert. Die stereoskopische Darstellung wird erreicht, indem nacheinander die zwei Halbbilder gezeigt werden und der Verschluss in der Brille immer nur für das richtige Teilbild geöffnet wird. Die dafür erforderliche Synchronisation mit dem Projektor kann im Kino z. B. mittels Infrarotimpulsen erreicht werden, die zu den Brillen geschickt werden (Abb. 7.62). Auch dieses Verfahren wird mit nur einem digitalen Projektor im Kino realisiert.

Aufgrund der Shutterbrille ist die Nutzung aber wieder mit Lichtverlust verbunden. Nachteilig ist auch, dass die Brillen eine eigene Energieversorgung benötigen und relativ schwer sind. Kinobetreiber müssen die teuren Brillen wiederverwenden und vor jeden Einsatz reinigen. Das Shutterverfahren wird

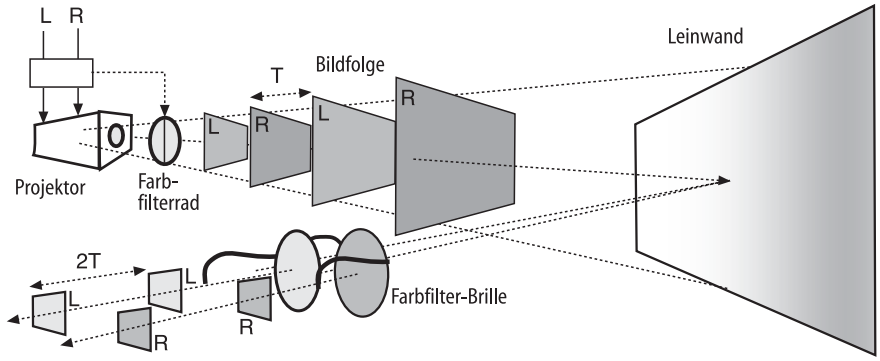


Abb. 7.63. 3D-Wiedergabe im Kino nach dem Dolby-3D-Verfahren

unter der Bezeichnung Xpand in den Kinos verwendet, ist aber nicht sehr weit verbreitet. Es arbeitet wie RealD mit zeitsequentieller Darstellung bei 144 Hz.

7.6.4.3 Dolby 3D

Das Projektionsverfahren der Firma Dolby knüpft an das Anaglyphenverfahren an. Auch hier wird mit Farbfiltern gearbeitet. Diese werden auf einem Rad im Lichtweg vor dem Projektor installiert, das sich synchron mit den Halbbildern dreht (Abb. 7.63). Auf diese Weise wird jedem Auge ein Farbsignal zugeführt, das aufgrund einer sehr speziellen Filtertechnik aus zwei sehr schmalbandigen RGB-Triplets gebildet wird, die in ihren Frequenzen geringfügig gegeneinander verschoben sind. Die Zuschauer bekommen passive Brillen, die keine Elektronik oder Stromversorgung benötigen. Allerdings sind sie doch vergleichsweise teuer, da sie, entsprechend der projizierten Lichtwellen, wieder sehr hochwertige Filter enthalten. Die Brillen werden wiederverwendet, was für das Kino mit einem Aufwand verbunden ist, der dem beim Shutterverfahren ähnelt.

Dolby 3D kann mit gewöhnlichen Bildwänden verwendet werden, zudem ermöglicht es eine einfache Umschaltung auf einen 2D-Betrieb, da nur das Filterrad entfernt werden muss. Nachteilig ist der Umstand, dass vor der Projektion eine Farbkorrektur erforderlich ist, da die beiden RGB-Triplets aufgrund der Frequenzverschiebung relativ zur Intensitätskurve der Lichtquelle zu unterschiedlichen Intensitäten bei der schmalbandigen Ausfilterung führen.

Bei den meisten der genannten Verfahren werden für den 3D-Betrieb also jedem Auge 24 fps dargeboten, wobei die Darstellung jedes einzelnen Bildes verdoppelt oder verdreifacht wird. Dieses Verfahren ändert sich im Zuge der Filmproduktion mit High Frame Rate (HFR). Durch die Erhöhung der Aufnahmefrequenz auf z. B. 48 fps soll hier die Bildwirkung bezüglich der Bewegungsphasen, insbesondere bei der 3D-Wiedergabe verbessert sein. Die Projektionssysteme müssen ebenso so schnell bzw. schneller arbeiten. Ob sich HFR auf der Aufnahmeseite mit 48 fps durchsetzen wird, ist ungewiss. Die DCI definiert in 2013 mehrere Optionen: bei 2k-Aulösung und 2D-Darstellung neben 48 fps die Bildfrequenzen 60, 72, 96 und 120 fps und für den 3D-Betrieb bei 2k-Auflösung 48 und 60 fps ebenso wie für 2D bei 4k-Darstellung.

8 Bildaufzeichnungsgeräte

8.1 Entwicklungsgeschichte

Für den Fortschritt der Videotechnik ist die Verfügbarkeit eines Speichersystems von größter Bedeutung. Die Entwicklung begann nach dem 2. Weltkrieg in den USA, wo zunächst versucht wurde, einen Video Tape Recorder (VTR) auf Basis eines Audiobandgerätes zu bauen, dessen Arbeitsprinzip in Abbildung 8.1 dargestellt ist. Die größte Schwierigkeit war dabei die hohe Frequenzbandbreite des Videosignals, die eine Erhöhung der Bandlaufgeschwindigkeit um den Faktor 5 MHz/20 kHz = 250 erfordert, wenn ein Tonbandgerät unmodifiziert zur Videoaufnahme eingesetzt werden soll. Das Problem konnte zu Beginn der 50er Jahre mit der Erkenntnis gelöst werden, dass nicht nur das Band, sondern auch die Videoköpfe in Bewegung gesetzt werden können, um die erforderliche hohe Relativgeschwindigkeit zwischen Band und Kopf zu erzielen. So konnte bei Ampex das erste studiotaugliche Magnetaufzeichnungsgerät (MAZ) entwikkelt und 1956 als Quadruplex-Format vorgestellt werden. Das Format wurde in verbesserter Form bis in die 80er Jahre hinein benutzt.

Die nächste Generation studiotauglicher Geräte arbeitete mit Schrägspuraufzeichnung. Die entsprechenden Formate B und C wurden Ende der 70er Jahre eingeführt und waren bis in die 90er Jahre hinein aktuell. In den 70er Jahren entwickelte sich auch ein Massenmarkt für Videorecorder [241]. Es wurden preiswerte Cassettengeräte gebaut (Betamax, VHS), die die Luminanz- und Chrominanzanteile des Videosignals getrennt verarbeiten (Colour Under) und eine typische Luminanzsignal-Bandbreite von 3 MHz aufweisen.

Im professionellen Bereich war der nächste Schritt die Einführung der Aufzeichnung von Luminanz- und Chrominanzkomponenten auf getrennten Spu-

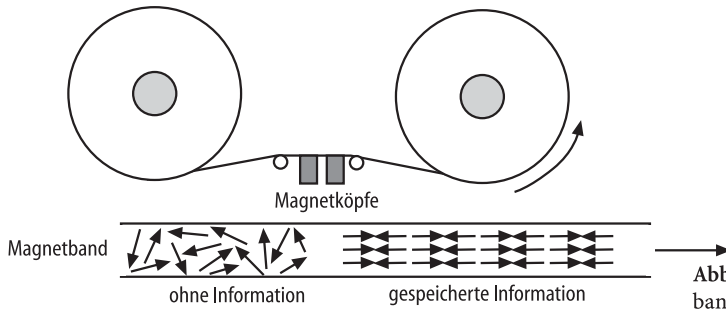


Abb. 8.1. Magnetbandspeicherprinzip

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4_8, © Springer-Verlag Berlin Heidelberg 2013

Tabelle 8.1. Historische Entwicklung der Magnetbandaufzeichnungsformate

	1950	1960	1970	1980	1990	2000
FM-Direkt	Quadruplex			1“ B, 1“ C		
Colour Under			U-Matic VCR	Betamax VHS	Video8	Hi8 S-VHS
Komponenten					Betacam (SP) MI	MII
Digital Composite					D2	D3
Digitale Komponenten ohne Datenreduktion					D1	D5 (HD)
Digitale Komponenten mit schwacher Datenreduktion					DCT	D-Beta
Digitale Komponenten mit stärkerer Datenreduktion						DV D9 D8 BetaSX D10
High-Definition-Aufzeichnung					D6	HDCam

ren. Auf dieser Basis entwickelte Sony Anfang der 80er Jahre das 1/2“-Betacam-Verfahren, das erste, professionellen Ansprüchen genügende Cassettensystem, welches als weiterentwickeltes Betacam-SP das letzte professionelle analoge Format war. Die 90er Jahre waren bestimmt von der Entwicklung digitaler Videoaufzeichnungsverfahren, die entweder auf der Basis von Magnetbändern (D1-3, D5, DigiBeta, D7, D9, D10) oder bandlos arbeiten, wobei in zunehmendem Maße auch Geräte mit Datenreduktion eingesetzt werden (Digital Betacam, DVCPro, Betacam SX, D10). Für die HD-Aufzeichnung stehen die Formate DVCProHD, HD-D5 und HDCam zur Verfügung (Tabelle 8.1).

8.2 Grundlagen der Magnetaufzeichnung

Die Signalspeicherung geschieht überwiegend mit Hilfe ferromagnetischer Materialien, die entweder auf Bänder oder Platten aufgetragen sind. In beiden Fällen erzeugt das zu speichernde Signal ein Magnetfeld, das in die bewegte Magnetschicht übertritt und dort die Magnetpartikel kollektiv ausrichtet.

8.2.1 Das magnetische Feld

Unter bestimmten Umständen können Kraftwirkungen beobachtet werden, die nicht mechanisch zu erklären sind. Wird z. B. ein Elektron in die Nähe eines zweiten gebracht, so entsteht eine abstoßende Kraftwirkung, die bei Abwesenheit des zweiten Elektrons nicht bemerkt wird; ähnliches gilt für magnetisierte Eisenteilchen. Die Anwesenheit des Elektrons oder des Eisenstückes ist also

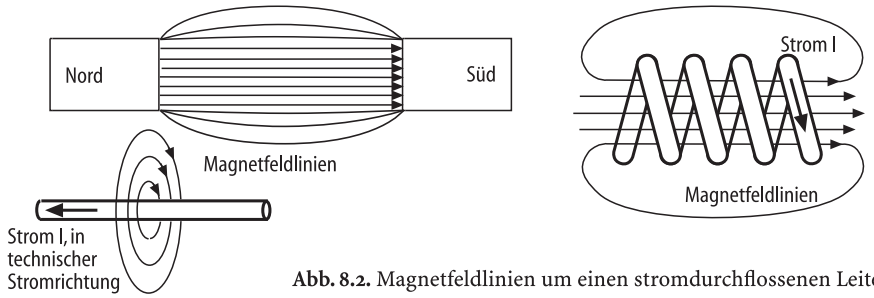


Abb. 8.2. Magnetfeldlinien um einen stromdurchflossenen Leiter

mit einer potenziellen Kraftwirkung verbunden, die mit Hilfe elektrischer bzw. magnetischer Felder beschrieben wird (Abb. 8.2). Wenn im Feld zeitliche Änderungen auftreten (z. B. bei der Wellenausbreitung), so sind beide Phänomene als elektromagnetisches Feld verbunden, welches theoretisch durch die Maxwellschen Gleichungen beschrieben wird. Ein elektrischer Strom besteht aus bewegter elektrischer Ladung, daher ist jeder Strom von einem Magnetfeld umgeben, das sich verstärkt, wenn der Leiter zu einer Spule aufgewickelt wird (Abb. 8.2). Im Vakuum ist die magnetische Feldstärke H dem Strom I und der Spulenwindungszahl N proportional. Wird andererseits ein Magnetfeld relativ zum Leiter bewegt, so wird in diesem eine Spannung induziert, die einen Stromfluss bewirken kann.

Die unsichtbaren elektromagnetischen Felder werden durch Feldlinien veranschaulicht, deren Richtung beim elektrischen Feld vom Plus- zum Minuspol und beim magnetischen Feld vom Nord- zum Südpol verlaufen. Die Dichte der Linien repräsentiert die Feldstärke. Im Magnetfeld werden immer zwei gegensätzliche Pole (Nord und Süd) wirksam, zwischen denen sich die Feldlinien ausbreiten (Abb. 8.2). Bei Anwesenheit von Materie im magnetischen Feld fällt das Verhalten der sogenannten ferromagnetischen Stoffe Eisen, Nickel und Kobalt auf. Diese Materialien bewirken eine Konzentration des Magnetfeldes, sie haben einen geringen magnetischen Widerstand (Abb. 8.3) und können auch ohne äußeres Magnetfeld eine makroskopische Magnetisierung aufweisen.

Die Magnetisierung entsteht aus einer Kopplung der Eigendrehimpulse der Elektronen, den Spins, die sich dadurch über weite Bereiche parallel ausrichten. Diese Bereiche (Domänen, Weißsche Bezirke) sind wiederum begrenzt, angrenzende Bereiche weisen andere Magnetisierungsrichtungen auf, sodass in unmagnetisiertem Zustand makroskopisch keine Magnetisierung festzu-

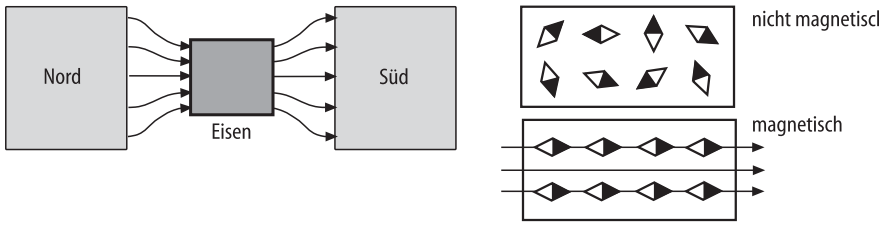


Abb. 8.3. Magnetfeldlinien in Eisen, magnetisches und nichtmagnetisches Material

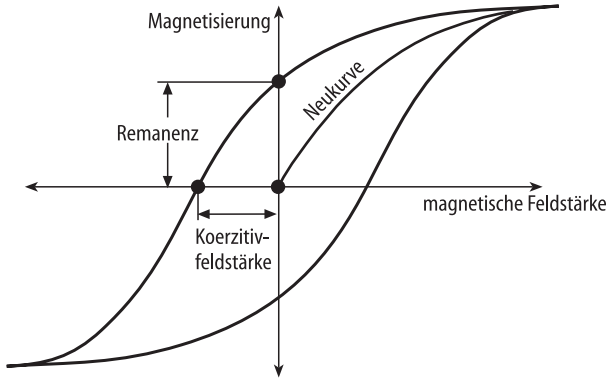


Abb. 8.4. Hysteresekurve

stellen ist. Durch äußere Magnetfelder können die Domänen kollektiv parallel ausgerichtet werden (Abb. 8.3). Eine solche Magnetisierung bleibt erhalten, bis eine regellose Umordnung der Domänen durch Überschreitung der sogenannten Curie-Temperatur (bei Eisen 1042 K), durch mechanische Erschütterung oder durch möglichst unregelmäßige Ummagnetisierung (Löschen) erzeugt wird. Ferromagnetische Materialien für Speichersysteme (Eisenoxyd, Reineisen, Chromdioxyd oder kobaltdotiertes Eisenoxyd) werden zu Teilchen mit einer Größenordnung von weniger als 1 μm zermahlen und zusammen mit einem Bindemittel auf Bänder oder Platten aufgetragen.

Für die Langzeitspeicherung von Informationen ist die verbleibende Magnetisierung nach Abschalten des magnetisierenden Feldes (Remanenz) bzw. die nötige Feldstärke zur Beseitigung der Magnetisierung von Bedeutung, zwei Parameter, deren Zusammenhang aus der Magnetisierungskennlinie deutlich wird. Die Magnetisierungskennlinie beschreibt die Abhängigkeit der Magnetisierung M von der magnetischen Feldstärke H (Abb. 8.4 und 8.5). Mit steigender Feldstärke steigt zunächst die Magnetisierung stark an, dann flacht die Kurve ab, weil die meisten der Weißschen Bezirke gleichgerichtet sind. Nach Abschalten der Feldstärke geht die Magnetisierung etwas zurück, der Großteil, die Remanenz, bleibt aber erhalten. Um diese zu beseitigen, muss eine entgegengesetzte Feldstärke aufgewandt werden (Koerzitivfeldstärke). Wird die

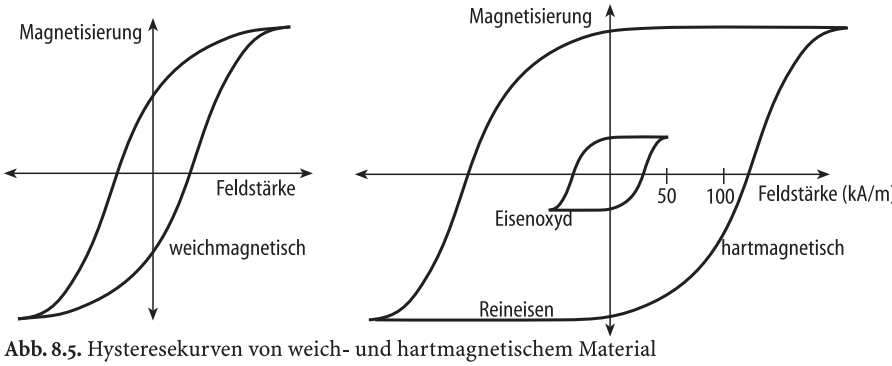


Abb. 8.5. Hysteresekurven von weich- und hartmagnetischem Material

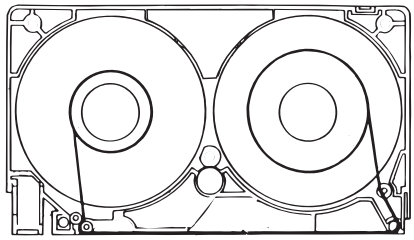


Abb. 8.6. Magnetbandcassette

magnetische Feldstärke in entgegengesetzter Richtung schließlich noch weiter gesteigert, so ergibt sich eine negative Magnetisierung, symmetrisch zur positiven. Der gesamte Kennlinienverlauf wird als Hysteresekurve bezeichnet.

Bei einer breiten Kurve ist eine hohe Koerzitivfeldstärke erforderlich, das Material ist magnetisch hart und bietet eine hohe Speichersicherheit. Magnetisch weiche Materialien mit einer schmalen Hysteresekurve werden dagegen dort gebraucht, wo eine Ummagnetisierung mit möglichst wenig Energie gefordert ist, wie z. B. bei Magnetkopfmaterial, das dazu dient, das Magnetfeld zum Band zu führen (Abb. 8.5).

Der Magnetismus bietet bis heute einen der wichtigsten physikalischen Effekte zur Speicherung von Signalen und Daten. Die Magnetspeicherung ist sehr gut an digitale Daten angepasst, deren zwei mögliche Zustände direkt den beiden magnetischen Ausrichtungen Nord und Süd zugeordnet werden können.

8.2.2 Das MAZ-Grundprinzip

Die Entwicklung der Videorecorder (Magnet-Aufzeichnungsgeräte, MAZ) ist von der Verkleinerung der Geräte, der Steigerung der Signalqualität bestimmt. Beide Aspekte sind eng mit der Verfügbarkeit hochwertigerer Magnetbänder verknüpft. Hinzu kommt die Vereinfachung der Handhabung des Bandes. Anstelle der früher verwendeten offenen Spulen als Träger werden Cassetten eingesetzt (Abb. 8.6), in denen das Band geschützt ist. Es wird im Video-Cassettenrecorder (VCR) automatisch aus dem Gehäuse gezogen.

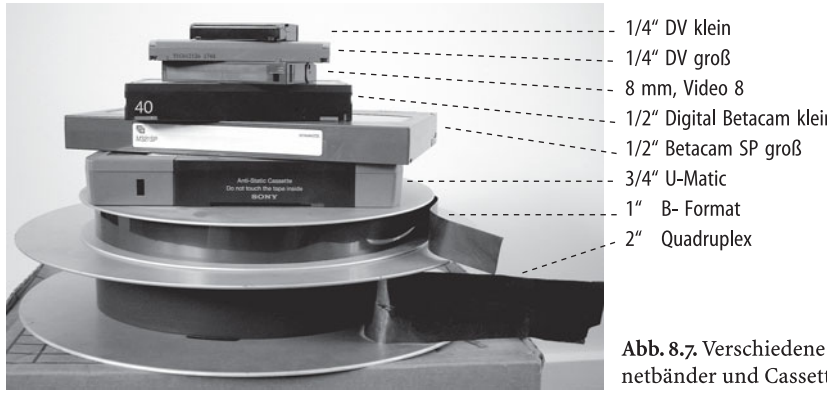


Abb. 8.7. Verschiedene Magnetbänder und Cassetten

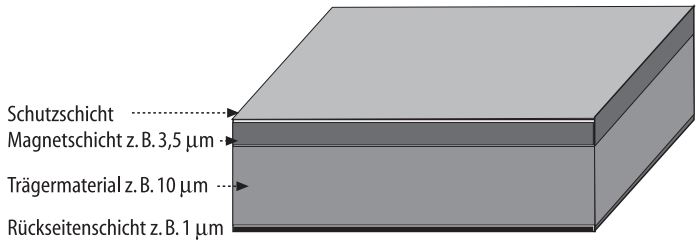


Abb. 8.8. Magnetbandbeschichtung

Die Breite des Bandes ist ein wichtiges Charakteristikum und gibt vielen Formaten den Namen, denn die Breite bestimmt wesentlich die Ausmaße des Laufwerkes und die Einsatzmöglichkeiten außerhalb des Studios. Die erste studiotaugliche MAZ-Maschine arbeitete mit einem Band der Breite 2“ (50,8 mm), das heute nicht mehr verwendet wird. MAZ-Maschinen nach Standard B und C arbeiten mit 1“-Bändern (25,4 mm), die ebenso wie die 2“-Bänder auf offenen Spulen untergebracht sind. Das nächst kleinere Maß ist das 3/4“-Band (19 mm), das für U-Matic und die Digitalformate D1, D2, D6 und DCT eingesetzt wird. Die am häufigsten benutzte Bandbreite ist 1/2“ (12,7 mm), die bei den Formaten VHS und Betacam SP sowie bei den Digitalformaten D3, D5, Digital Betacam und HDCam verwendet wird. Für den Amateur-Bereich gibt es das 8 mm-Format und die Bandbreite 1/4“ (6,35 mm), die für das Digitalformat DV eingesetzt wird. Abbildung 8.7 zeigt Beispiele für verschiedene Magnetbänder und Cassetten, und Abbildung 8.8 zeigt die Beschichtung der Bänder.

8.2.2.1 Der Magnetkopf

Zur Erzeugung des Magnetfeldes dient der Aufnahmekopf der auch als Ton- oder Videokopf bezeichnet wird. Abbildung 8.9 zeigt einen montierten Videokopf im Vergleich zu einem Tonkopf aus einem Audio-Cassettengerät. Das zu speichernde Signal erzeugt im Kopf ein entsprechend wechselndes Magnetfeld das aus dem Kopfspalt austritt. Das Feld tritt in das Magnetband ein und die Weißschen Bezirke in der Magnetschicht werden kollektiv ausgerichtet. Die entstehende Magnetisierung bleibt erhalten, die Information ist gespeichert.

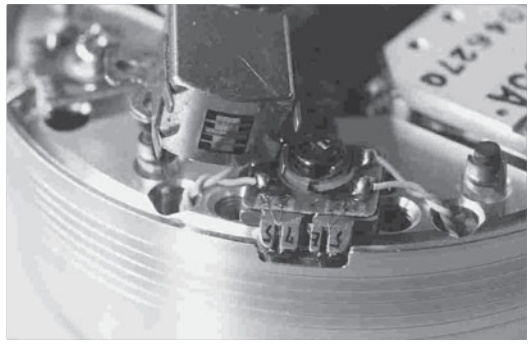


Abb. 8.9 Video- und Tonmagnetkopf im Vergleich

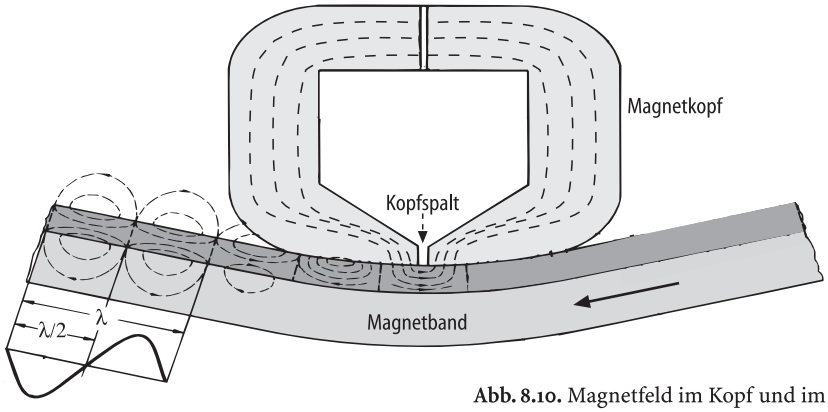


Abb. 8.10. Magnetfeld im Kopf und im Band

Das Band wird über die als Kopfspiegel bezeichnete Kontaktfläche hinweg bewegt, dabei wird die zeitliche Änderung des Magnetfeldes in eine örtlich veränderliche Magnetisierung umgesetzt [229]. Auf dem Band ergibt sich mit der räumlichen Periodizität die Bandwellenlänge λ (Abb. 8.10). Sie ist der Kopf-Band-Relativgeschwindigkeit v proportional und umso größer, je kleiner die Signalfrequenz f ist. Es gilt: $\lambda = v/f$. Die Bewegung des Bandes muss äußerst gleichmäßig erfolgen, das Band sollte möglichst eben sein.

Damit die Magnetfeldlinien aus dem Kopf austreten und sich im Magnetband ausbreiten können, besteht der Kopf aus einem weichmagnetischen Werkstoff (Ferrit) mit einem glasgefüllten Spalt. Das ferromagnetische Kernmaterial des Kopfes dient zur möglichst verlustlosen Übertragung des Feldes zum Spalt, an dem dann das Feld in das Magnetband übertritt. Das weichmagnetische Material weist geringe Ummagnetisierungsverluste auf und stellt einen geringen magnetischen Widerstand dar, es hat eine hohe Permeabilität. Auch die hartmagnetische Schicht des Bandes hat eine hohe Permeabilität, während sie dagegen für den glasgefüllten Spalt sehr gering ist. Im Bereich des Spaltes treten die Magnetfeldlinien daher aus dem Kopfspalt aus und befinden sich vorzugsweise im aufliegenden Magnetband (Abb. 8.10).

Die Spaltlänge bestimmt die Breite der Videospur. Die Geräte arbeiten mit Spurbreiten zwischen 10 µm und 80 µm. Aus fertigungstechnischen Gründen ist die Kopfspiegelbreite oft größer als die Spaltlänge, die dann durch die in Abbildung 8.11 gezeigte Verjüngung erreicht wird.

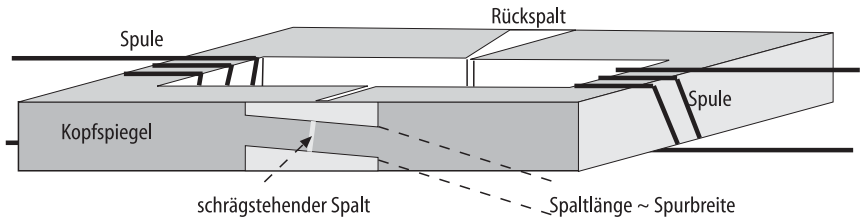


Abb. 8.11. Aufbau des Videokopfes

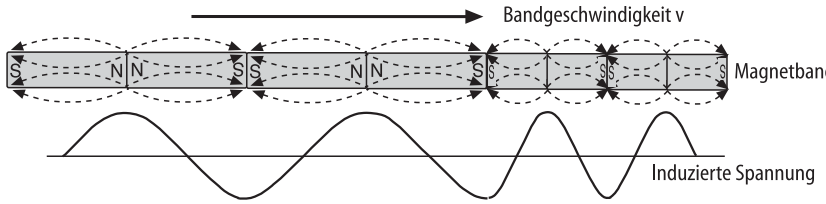


Abb. 8.12. Bandwellenlängen und Wiedergabefrequenzen

Die Spaltbreite d bestimmt die kleinsten magnetisierbaren Bereiche auf dem Band, daraus resultiert mit $d \approx \lambda/2$ die kleinste räumliche Periodizität λ_{min}. Je schneller das Band bewegt wird, desto schneller wechseln die Perioden, und desto höher liegt die aufzeichenbare Grenzfrequenz $f_{gr} = v/\lambda_{min}$ (Abb. 8.12). Videoköpfe sollten für eine Verarbeitung eines FM-Frequenzspektrums von bis zu 20 MHz geeignet sein, Audioköpfe für Frequenzen bis zu 20 kHz. Tonköpfe haben Spaltbreiten zwischen 3 µm und 8 µm, Videoköpfe zwischen 0,3 µm und 0,5 µm. Kleinere Spaltbreiten sind schwer realisierbar, da die Feldstärke im Magnetband zu klein wird.

8.2.2.2 Der Aufnahmevorgang

Das Magnetfeld wird im Ton- oder Videokopf mit Hilfe einer um die Kopfschenkel gewickelten stromdurchflossenen Spule erzeugt. Das Signal und damit die Magnetisierung ändern sich zeitlich in schneller Folge, sodass auch das Material des Tonkopfes ständig ummagnetisiert wird. Der Zusammenhang zwischen Signal und Magnetisierung ist nichtlinear und so weist auch der remanente Bandfluss eine nichtlineare Magnetisierungskennlinie auf. Ein Signal wird hier verzerrt. Abbildung 8.13 zeigt als Beispiel die Verformung eines Sinus-Signals an der Magnetisierungskennlinie. Die Signalverzerrungen stellen, abhängig von der aufzuzeichnenden Signalart, ein mehr oder weniger großes Problem dar. Bei der analogen Videosignalaufzeichnung wird das Kennlinien-Problem umgangen, indem die Information nicht in der Signalamplitude, sondern in der Frequenz verschlüsselt wird (FM, s. Abschn. 8.4.1). Bei der üblichen analogen Audioaufzeichnung auf Längsspuren werden die durch die Kennlinie bewirkten Signalverzerrungen durch ein HF-Vormagnetisierungsverfahren weitgehend vermieden (s. Abschn. 8.4.4).

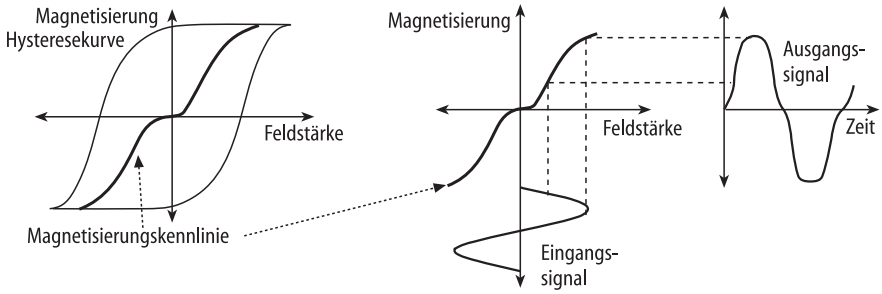


Abb. 8.13. Signalverzerrung an der Magnetisierungskennlinie

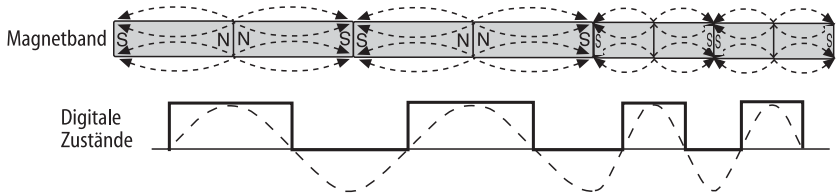


Abb. 8.14. Digitale und Magnetisierungszustände

Die geringsten Schwierigkeiten ergeben sich bei der Aufzeichnung digitaler Signale. Hier sind nur zwei Zustände erlaubt, die ohne Rücksicht auf die Kennlinie den zwei Magnetisierungsrichtungen (Nord und Süd) zugeordnet werden können (Abb. 8.14).

8.2.2.3 Der Wiedergabevorgang

Zur Wiedergabe wird das Band in gleicher Weise wie bei der Aufnahme am Magnetkopf vorbei geführt. Die auf dem Band gespeicherte, veränderliche Magnetisierung erzeugt bei diesem Bewegungsvorgang im Kopf ein wechselndes Magnetfeld, das wiederum zur Induktion einer Spannung U führt. Das Induktionsgesetz lautet: $U = - N \cdot d\Phi/dt$. Darin steht N für die Windungszahl der Spule und $d\Phi/dt$ für die Veränderung des magnetischen Flusses mit der Zeit. Daraus folgt, dass die induzierte Spannung umso größer ist, je schneller sich der Fluss, bzw. die Magnetisierung ändert. Schnelle Änderungen entstehen bei hohen aufgezeichneten Signalfrequenzen. Die Spannung hängt damit linear von der Frequenz ab, sie verdoppelt sich bei Frequenzverdopplung (pro Oktave). Bei der Frequenz o Hz beträgt die induzierte Spannung o V. Der zugehörige Spannungsverlauf über der Frequenz wird als Omega-Gang bezeichnet.

Der Gesamt-Frequenzgang wird von Verlusten bestimmt, die vor allem bei hohen Frequenzen wirksam werden. Abbildung 8.15 zeigt die Veränderung des Omega-Gangs und den resultierenden Gesamt-Frequenzgang.

Ein wichtiger Verlustfaktor ist sowohl bei der Aufnahme als auch bei der Wiedergabe der aufgrund des mechanischen Antriebs nicht optimale Band-Kopf-Abstand. Daneben tritt als Verlustfaktor der Spalteffekt auf, d. h., dass bei hohen Frequenzen die Bandwellenlängen so klein werden können, dass unterschiedlich ausgerichtete magnetische Bereiche im Kopfspalt wirksam wer-

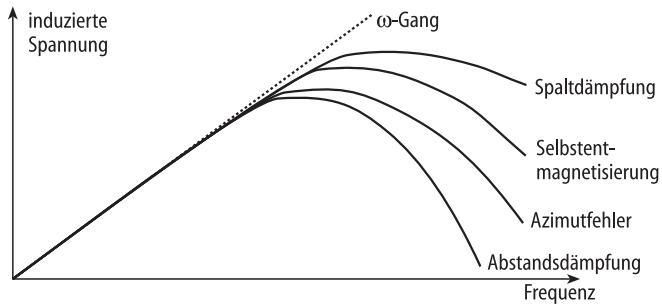


Abb. 8.15. Omega-Gang und Verluste

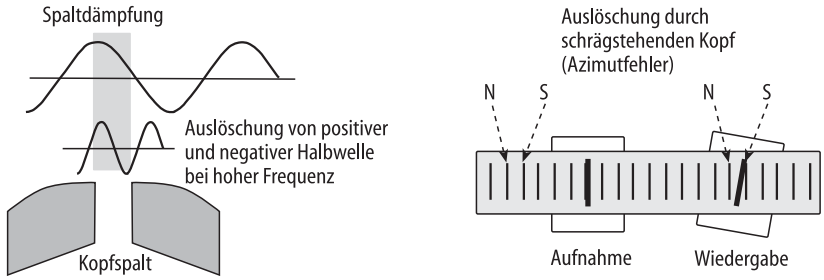


Abb. 8.16. Verluste durch Spaltdämpfung und Azimutfehler

den und sich in ihrer Wirkung gegenseitig kompensieren. Im Extremfall wird dadurch die induzierte Spannung zu Null (Abb. 8.16). Ein weiterer wichtiger Verlustfaktor ist der sogenannte Azimut-Fehler. Der Azimut bezeichnet den Winkel zwischen Kopfspalt und der magnetisierten Spur auf dem Band. Hier kommt es vor allem darauf an, dass die Köpfe zur Aufnahme und Wiedergabe denselben Azimut aufweisen. Ein im Bezug auf die Aufnahme gekippter Wiedergabekopf erfasst in Spurlängsrichtung zu große Bereiche und wirkt so wie ein verbreiterter Kopfspalt mit entsprechender Dämpfung hoher Frequenzen.

8.2.2.4 Der Löschvorgang

Ein gelöschtes Band trägt keine gespeicherten Signale. Dieser Zustand wird erreicht, wenn alle Magnetpartikel so ungeordnet sind, dass sich die Magnetisierung im Mittel aufhebt. Die Signale auf dem Band werden mit Hilfe eines Löschkopfes gelöscht, der entweder als stationärer Kopf die gesamte Breite des Bandes umfasst und damit ggf. mehrere Spuren zugleich löscht, oder wie der Aufnahmekopf, die Spuren einzeln erfasst. Dazu muss er als sog. fliegender Löschkopf auf der Kopftrommel montiert sein. Ein Löschkopf hat einen breiten Spalt oder Doppelspalt. Gelöscht wird mit einem Hochfrequenzfeld (f > 100 kHz). Beim Vorbeiziehen des Bandes am Spalt wird es bis zur Sättigung durchmagnetisiert, dann entfernt sich das Band stetig aus dem Feld. Dabei nimmt die Feldstärke ab, sodass immer kleinere Hystereseschleifen durchlaufen und damit alle Bereiche in der Magnetschicht erfasst werden (Abb. 8.17).

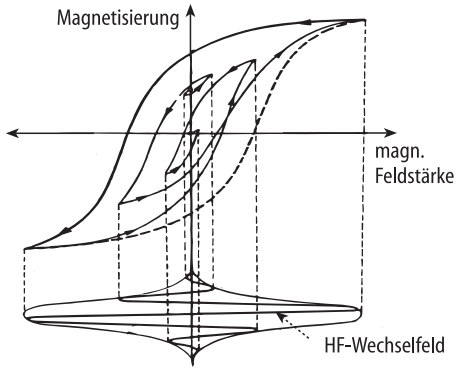


Abb. 8.17. Magnetbandlöschung mit HF-Wechselfeld

8.2.3 Magnetband- und Kopfführung

Bei der Magnetbandaufzeichnung ist vor allem eine konstante Bandgeschwindigkeit erforderlich. Meistens wird diese über einen kraftschlüssigen Bandantrieb mit Hilfe eines Capstans erreicht. Eine aus Hartgummi bestehende Andruckrolle presst dabei das Band an eine über einen geregelten Motor angetriebene Welle (Abb. 8.18). Das Band bewegt sich mit konstanter Geschwindigkeit, allerdings ist ein Schlupf nicht auszuschließen. Das Band wird von separat angetriebenen Wickeltellern aufgenommen. Insgesamt erfordert der Bandantrieb hohe mechanische Präzision, alle Wellen und Führungsrollen müssen exakt justiert und vor allem sauber sein.

8.2.3.1 Längsspuraufzeichnung

Die Längsspuraufzeichnung ist das Aufzeichnungsverfahren, das beim konventionellen Tonbandgerät zum Einsatz kommt. Es wird mit feststehenden Magnetköpfen gearbeitet. Nur das Band wird bewegt, sodass sich eine ununterbrochene Längsspur parallel zum Band ergibt (Abb. 8.19). Ausgehend von einer Spaltbreite $d = 5\ \mu m$ und $d \approx \lambda/2$ kann beim Tonbandgerät nach der Beziehung $f_{gr} = v/\lambda$ die Audio-Grenzfrequenz von 20 kHz mit einer Bandgeschwindigkeit $v = 20$ cm/s erreicht werden.

In Videorecordern wird das Längsspurverfahren zur Aufzeichnung von Audiosignalen, von Timecode (s. Kap. 9.3) und von Kontrollsignalen (CTL) eingesetzt. Für Videosignale ist die Längsspuraufzeichnung nicht geeignet, denn da die Kopfspaltbreite relativ zum Tonkopf nur etwa um den Faktor 10 reduziert werden kann, ist zur Erzielung einer Grenzfrequenz von 5 MHz für die Videoaufzeichnung eine Erhöhung der Relativgeschwindigkeit zwischen Kopf und Band um etwa Faktor 25 erforderlich. Eine entsprechende Erhöhung der Bandgeschwindigkeit bei stehendem Magnetkopf ist unpraktikabel.

8.2.3.2 Quer- und Schrägspuraufzeichnung

Der Gedanke, dass nicht die Bandgeschwindigkeit, sondern nur die Relativgeschwindigkeit zwischen Kopf und Band entscheidend ist, und dass auch der Magnetkopf in Bewegung gesetzt werden kann, führte dann zum ersten MAZ-Format von praktischer Bedeutung.

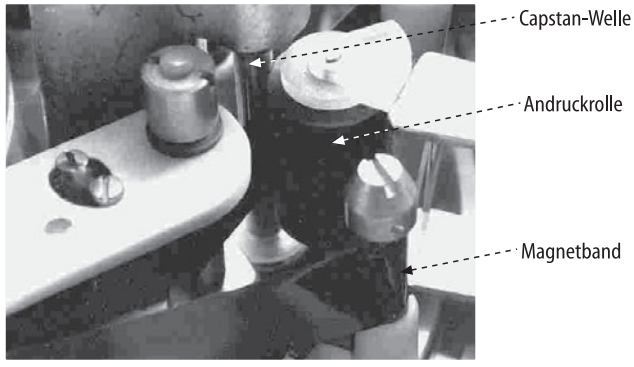


Abb. 8.18. Capstan-Antrieb

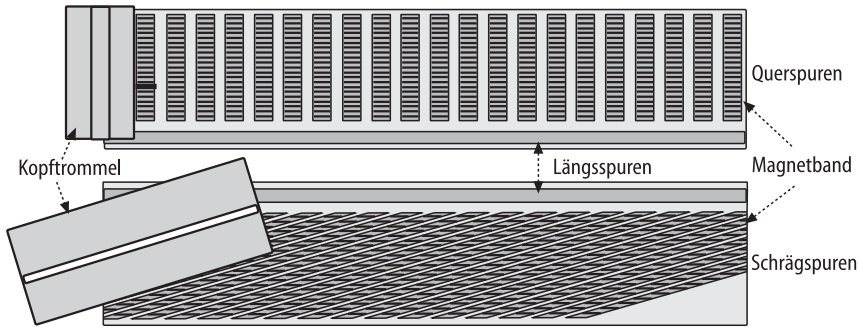


Abb. 8.19. Quer- Längs- und Schrägspuren

Die Videoköpfe werden auf einer Kopftrommel (Drum) befestigt und in schnelle Rotation versetzt, während die Bandvorschubgeschwindigkeit auf dem Niveau der Audioaufzeichnungsgeräte bleibt. Beim zunächst entwickelten Querspurverfahren (Quadruplex-Format) rotierte die Kopftrommel senkrecht zum Band (Abb. 8.19). Alle anderen relevanten Formate arbeiten mit einer relativ zum Band schräg stehenden Kopftrommel. Bei diesem Schrägspurverfahren werden Spur-Band-Winkel zwischen 2° und 15° (meist 5°) verwendet. Der Winkel ergibt sich aus der Lage der Kopftrommel und hängt außerdem von der Bandvorschubgeschwindigkeit ab.

Die Schrägspuren unterbrechen das kontinuierliche Signal. Damit sich daraus möglichst wenige Probleme ergeben, wird die Schrägspurlänge bei den meisten Formaten so gewählt, dass eine Spur ein vollständiges Halbbild enthält. Falls dies nicht der Fall ist (z. B. beim 1“ B-Format), spricht man von segmentierter Aufzeichnung. Spuren werden von der unteren zur oberen Bandkante geschrieben. Am Spuranfang liegt der vertikale Synchronimpuls.

Um eine hohe Aufzeichnungsdichte zu erzielen, sollen die Schrägspuren möglichst schmal sein und dicht nebeneinander liegen. Es werden Spurbreiten zwischen 160 µm (1" C) und 10 µm (DV) verwendet. Der Abstand zwischen den Spuren beträgt etwa ein Viertel der Spurbreite (Abb. 8.20).

Die Aufzeichnungsdichte kann erhöht werden, indem auf den Spurzwischenraum (Rasen) verzichtet wird. Allerdings muss dann verhindert werden, dass bei einer Fehlabtastung das Signal der Nachbarspur im Videokopf wirksam wird. Dieses Übersprechproblem wird durch die Slanted Azimut-Technik minimiert, bei der die Videokopfspalte nicht wie üblich senkrecht zur Spur stehen, sondern bewusst um Winkel zwischen ± 6° und ± 20° gegen die Spur verkippt werden. Die Vorzeichen zeigen an, dass die Verkippung nach rechts

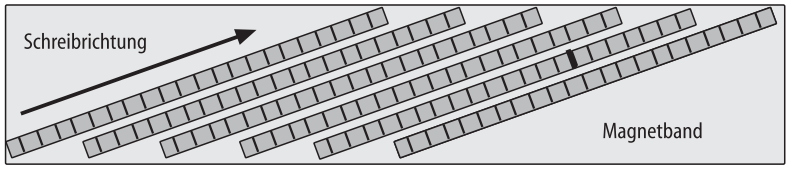


Abb. 8.20. Schrägspuren mit Sicherheitsabstand

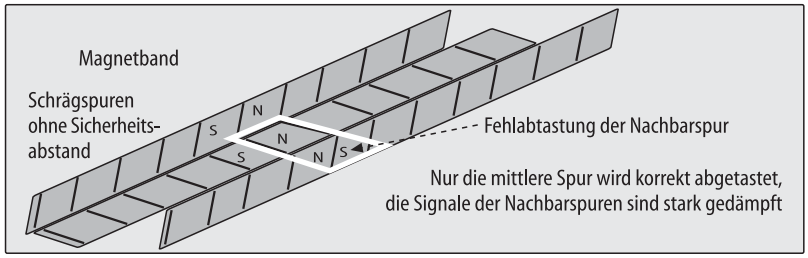


Abb. 8.21. Dämpfung des Spur-Übersprechens durch Slanted Azimut

und nach links erfolgt, sodass sich die Spuren mit den zugehörigen Winkeln jeweils abwechseln (Abb. 8.21). Wenn ein Kopf durch eine Fehlabtastung auf die Nachbarspur gerät, so wird diese mit einem Azimutfehler von mindestens 12° abgetastet, was zu einer Übersprechdämpfung führt, die besonders bei hohen Frequenzen wirksam ist.

8.2.3.3 Kopftrommelumschlingung

Insgesamt befinden sich auf Videobändern Schrägspuren für das Videosignal und Längsspuren für Audio- und Zusatzsignale. Die Schrägspuren liegen in der Mitte, die Längsspuren an den Rändern des Bandes. Im Gegensatz zum Längsspurverfahren ergeben sich bei der Quer- und Schrägspuraufzeichnung unterbrochene Videospuren. Wenn nur ein Videokopf benutzt wird, muss das Band möglichst weit um die Trommel geschlungen werden, wie es beim 1“ C-Format geschieht (Umschlingung 350°), damit das Videosignal bei der Unterbrechung nur für möglichst kurze Zeit ausfällt. Bei den meisten Formaten wird mit mehreren Köpfen aufgezeichnet, die abwechselnd eingesetzt werden. Die einfachste, z. B. bei VHS und Betacam SP eingesetzte Variante ist die 180°-Umschlingung, bei der sich zwei Köpfe abwechseln (Abb. 8.22).

Der Umschlingungsgrad hängt mit dem fest definierten Kopftrommeldurchmesser zusammen. Wird bei einem Format eine verkleinerte Kopftrommel eingesetzt, um z. B. ein kompakteres Gerät entwickeln zu können, muss der Umschlingungsgrad entsprechend verändert werden. Bei Geräten, die mit offenen Spulen arbeiten, muss das Band von Hand eingelegt werden, bei Cassettensystemen wird die Kopftrommelumschlingung automatisch durchgeführt. Führungsstifte greifen hinter das Band und ziehen es um die Kopftrommel herum, wobei die Bandführung bei der 180°-Umschlingung nach U- und M-Loading unterschieden wird (Abb. 8.23).

Abb. 8.22. 350°- und 180°- Kopftrommelumschlingung

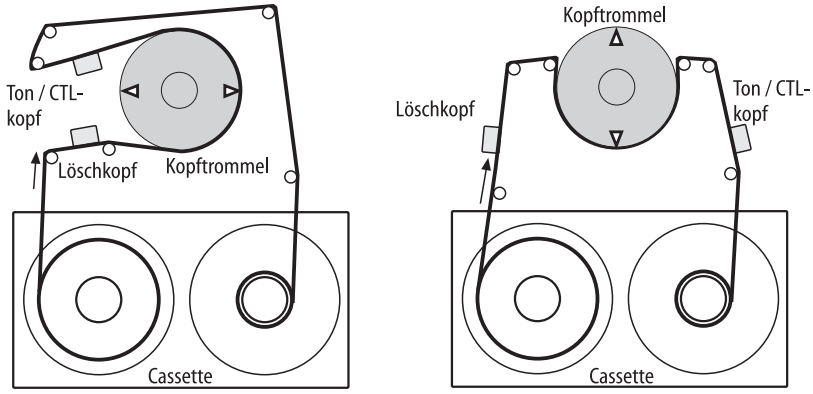


Abb. 8.23. U- und M-Loading-Prinzip

8.2.3.4 Die Kopftrommel

Die Kopftrommel besteht aus einem rotierenden und einem feststehenden Teil. Häufig rotiert die gesamte obere Kopftrommelhälfte, bei einigen Formaten aber nur eine dünne Kopfscheibe im Inneren (Scanner). Die Videoköpfe ragen dann aus einem sehr dünnen Spalt heraus. Sie haben gegenüber der Trommel einen Kopfüberstand von ca. 30 µm bis 60 µm. Die Signalübertragung vom stationären zum rotierenden Teil geschieht mit Hilfe von Übertragern. Auf den beiden Kopftrommelteilen sind jeweils Spulen montiert. Wenn sich die rotierende Spule über die feststehende hinweg dreht, wird eine kontaktlose Signalübertragung bewirkt. Häufig befinden sich neben den sog. Kombiköpfen für Aufnahme und Wiedergabe weitere Köpfe auf der Trommel, z. B. separate (evtl. bewegliche) Wiedergabeköpfe, HiFi-Tonköpfe und fliegende Löschköpfe, die eine selektive Löschung einzelner Spuren ermöglichen. Abbildung 8.24 zeigt eine entsprechend bestückte Kopftrommel (Betacam SP).

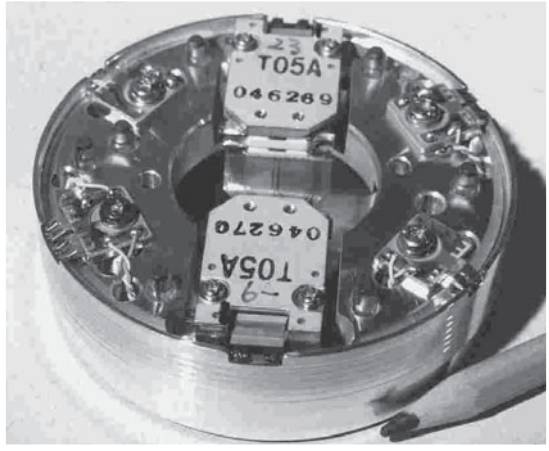


Abb. 8.24. Kopftrommel mit drei Kopfpaaren (Betacam SP)

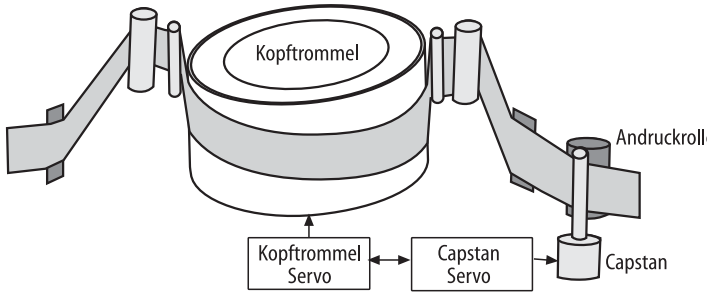


Abb. 8.25. Magnetbandführung an der Kopftrommel

8.2.3.5 Servosysteme

Für die Qualität eines Videorecorders ist die Erzeugung und Abtastung einer gleichmäßigen Spuranordnung von großer Bedeutung. Bei der Aufnahme soll ein exaktes und vor allem normgerechtes Spurbild entstehen, damit Bänder zwischen Geräten desselben Formates ausgetauscht werden können. Um diese Ziele zu erreichen, ist es erforderlich, dass die Bandführung sehr präzise erfolgt. Die Führungselemente müssen exakt justiert und sauber sein, insbesondere gilt dies für die Führungselemente, die die Schräglage des Bandes relativ zur Kopftrommel bestimmen. Abbildung 8.25 zeigt eine typische Bandführung. Im Vergleich zum Tonbandgerät ergibt sich eine Besonderheit durch das Vorhandensein zweier Antriebssysteme für Kopf und Band. Bezüglich beider Systeme wird dafür gesorgt, dass die Rotationsgeschwindigkeiten über Regelkreise konstant gehalten werden. Die Regelkreise (Servosysteme) beeinflussen aber nicht nur die Rotationsfrequenz, sondern auch die Rotationsphase. Sie sind auch miteinander verkoppelt damit die Kopftrommel gerade soweit rotiert, dass der nächste aktive Videokopf genau die nächste Spur abtastet.

Professionelle Geräte bieten außerdem die Möglichkeit, die Spuren auch bei veränderter Bandgeschwindigkeit störungsfrei abzutasten (Dynamic Tracking, s. u.). Darüber hinaus sollen professionelle Geräte für den Schnittbetrieb geeignet sein, was bedeutet, dass anhand eines aufgezeichneten Timecodes jedes Bild exakt lokalisiert werden kann, und dass Zuspieler und Recorder von einem externen Gerät fernsteuerbar sind (s. Kap. 9.3). Moderne Systeme lassen sich universell einsetzen, sie nutzen Schnittstellen, bei denen die Laufwerksfunktionen durch digital codierte Befehle ausgelöst werden, die seriell über-

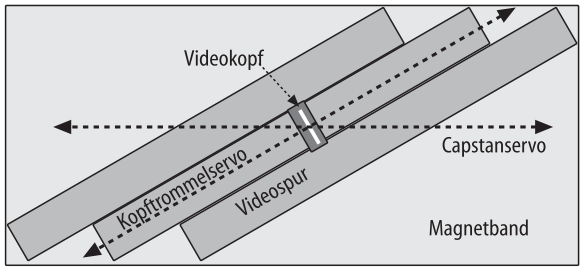


Abb. 8.26. Kopfposition in Abhängigkeit von den Servosystemen

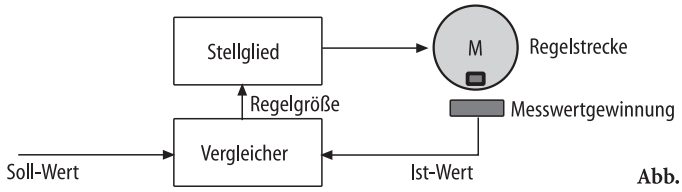


Abb. 8.27. Servo-Regelkreis

tragen werden. Als »Quasi-Standard« hat sich hier die sogenannte Sony 9-Pin-Schnittstelle auf Basis der RS-422A-Computerschnittstelle durchgesetzt.

Die wichtigste Aufgabe der Servosysteme ist die Geschwindigkeits- und Phasenregelung für die Kopftrommel- und die Bandgeschwindigkeit (Abb. 8.26). Weitere Servosysteme in MAZ-Geräten dienen der Kontrolle eines konstanten Band-Aufwickelzuges. Der Regelkreis dient in allen Fällen der Nachregelung einer Motordrehzahl (Regelstrecke). Die gegenwärtige Drehzahl wird als Ist-Größe mit einer vorgegebenen Soll-Größe verglichen und das Ergebnis (Regelgröße) beeinflusst über ein Stellglied die Drehzahl (Abb. 8.27).

Zur Erfüllung der genannten Aufgaben und Bereitstellung der Ist-Größen brauchen die Servosysteme Hilfssignale, die bei fast allen Aufzeichnungsformaten auf einer Steuer-, Kontroll- oder CTL-Spur im Längsspurverfahren aufgezeichnet werden (Abb. 8.28). Ein stationärer CTL-Kopf zeichnet Scanner-Referenz-Pulse auf, die mit der Kopftrommel-Umdrehung verkoppelt sind. Die Signale auf der CTL-Spur ergeben die mechanisch festgelegte Referenz für die Videospurlage. Sie entsprechen quasi einer elektronischen Perforation, die es ermöglicht, dass bei der Wiedergabe jede Spur genau getroffen wird.

Um die Kopftrommelrotation auf dem Sollwert zu halten, wird der Kopftrommel-Servo eingesetzt. Bei den meisten Formaten, bei denen mit zwei Köpfen ein Halbbild pro Spur aufgezeichnet wird, beträgt der Sollwert genau 25 Umdrehungen pro Sekunde. Mit der Kopftrommel-Phasenregelung wird zusätzlich die Position der Videoköpfe relativ zum Referenzsignal berücksichtigt. Die Position muss so geregelt sein, dass sich der Kopf bei der Aufnahme genau dann am Spuranfang befindet wenn der V-Synchronimpuls vorliegt. Die Kopfposition bestimmt auch den Zeitpunkt der Kopfumschaltung. Sie muss kurz

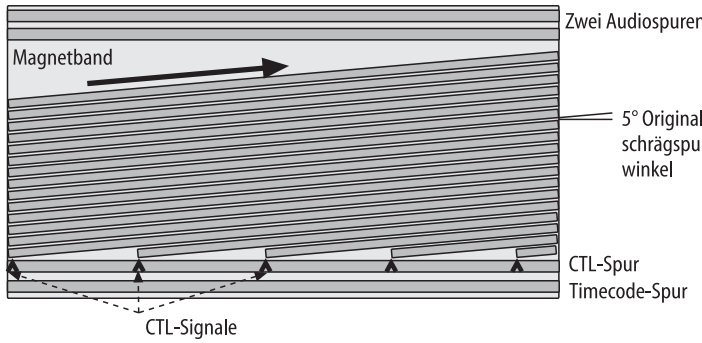


Abb. 8.28. Spurlage und CTL-Impulse

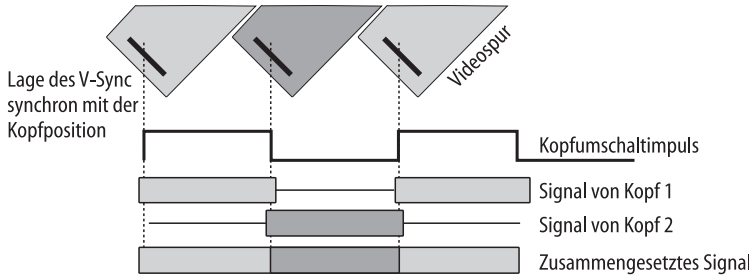


Abb. 8.29. V-Sync und Kopfumschaltung

vor dem V-Sync vorgenommen werden, damit ein Wiedergabemonitor sich auf ein ungestörtes V-Sync-Signal einstellen kann (Abb. 8.29). Bei professionellen Videogeräten liegt die Kopfumschaltung ca. zwei Zeilen vor dem V-Sync, bei Amateurgeräten ca. sechs Zeilen. Der letztgenannte Betrag bewirkt, dass bei der Wiedergabe mit Amateurgeräten am unteren Bildrand auf dem Monitor im Underscan-Modus oft einige gestörte Zeilen erscheinen.

Der Capstan-Regelkreis kontrolliert die Bandgeschwindigkeit, von der die Lage der Schrägspuren und der Schrägspurwinkel abhängig ist. Bei der Aufnahme wird der Soll-Wert aus einem stabilen Quarzoszillator abgeleitet, dabei werden auch die Signale für die CTL-Spur generiert. Bei der Wiedergabe orientiert sich der Regelkreis an eben diesen Signalen auf der CTL-Spur. Beim Video 8-Format, bei DV u. a. werden die CTL-Signale nicht auf Längsspuren, sondern zusammen mit dem Videosignal auf den Schrägspuren aufgezeichnet. Der Capstan-Regelkreis wird über Pegel- und Frequenzerkennung dieser als ATF (Automatic Track Following) bezeichneten Signale gesteuert.

Die Position des CTL-Kopfes relativ zur Kopftrommel ist entscheidend für eine genormte Spurlage. Bei Bändern, die auf verschiedenen Geräten bespielt wurden, kann sich aufgrund unterschiedlicher Positionen des CTL-Kopfes zwischen Aufnahme- und Wiedergabegerät eine Spurlagenabweichung ergeben, die bei der Wiedergabe durch die Veränderung des Tracking, d. h. durch eine Verzögerung des CTL-Signals, ausgeglichen werden kann. Diese wirkt wie eine Lageänderung des CTL-Kopfes (Abb. 8.30). Die Tracking-Einstellung kann manuell erfolgen. An professionellen Geräten steht dazu als Hilfsmittel ein Anzeigeinstrument zur Verfügung, das den maximalen Videopegel anzeigt, wenn die Köpfe genau die Videospurmitte abtasten. Bei Geräten mit Auto-Tracking wird anhand des Signalpegels das Servosystem automatisch immer so nachgeführt, dass sich stets ein maximaler Pegel ergibt.

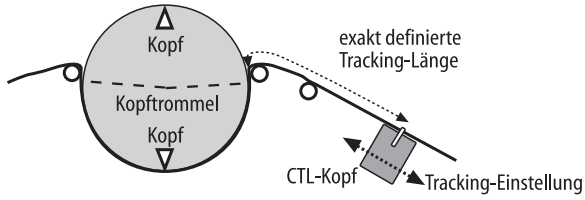


Abb. 8.30. Tracking

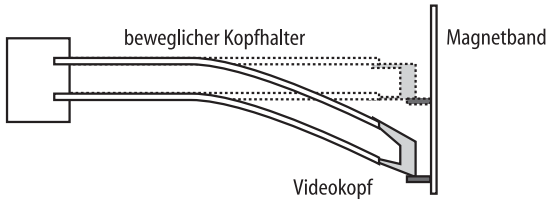


Abb. 8.31. Bewegliche Videoköpfe

Die bisher verwendeten Begriffe Tracking und Auto-Tracking beziehen sich nur auf die Regelung des Bandantriebs (Capstan-Tracking). Sie werden aber auch für eine andere Art der genauen Spurabtastung verwendet, die meist mit Dynamic Tracking (DT) (auch Automatic Scan Tracking, AST) bezeichnet wird [33]. Beim Dynamic Tracking wird der Videokopf beweglich und in der Bewegung steuerbar auf der Kopftrommel montiert, damit er auch bei Abweichungen von der Standard-Spurlage immer auf der Spurmitte geführt werden kann (Abb. 8.32). Dies ist insbesondere bei der Wiedergabe mit veränderter Bandgeschwindigkeit erforderlich. Die dabei erforderliche Spurlagenveränderung ist gering, die entsprechend minimale Verbiegung des Kopfhalters kann daher mit Hilfe von kleinen Hubmagneten oder piezoelektrischen Materialien erreicht werden. Ein Piezokristall ändert seine Länge unter dem Einfluss einer elektrischen Spannung. Die Halterungen des Videokopfes werden so gefertigt, dass zwei entgegengesetzt polarisierte Piezo-Schichten verwendet werden, von denen sich unter der Steuerspannung eine dehnt und eine zusammenzieht, sodass der Kopfhalter verbogen wird (Abb. 8.31).

8.2.3.6 Zeitlupe und Zeitraffer

Diese Effekte entstehen bei der Wiedergabe mit einer Bildfolge-Geschwindigkeit, die von der der Aufnahme abweicht. Die Bildfolgefrequenzen bei der Aufnahme f_A und Wiedergabe f_W lassen sich über einen Faktor k für Zeitlupe bzw. -raffer verknüpfen: $f_W = k \cdot f_A$. Die Festlegung bezieht sich auf die Bildfolgen, die Bildfrequenz des Videosignals wird natürlich beibehalten. Die Kopftrommelumdrehungszahl bleibt konstant, und der Bandvorschub wird entsprechend dem Faktor k verändert. Der Wert k = 1 entspricht dem Normalbetrieb, bei k = 0 wird immer das selbe Bild wiederholt. Bei einem k zwischen 0 und 1 ergibt sich eine Zeitlupenwiedergabe. Sie kann durch zyklischen Wechsel zwischen der Standbildabtastung und der Normalwiedergabe (z. B. Im Verhältnis 2:1) erreicht werden (Abb. 8.32). Bei k > 1 entsteht der Zeitraffereffekt [156]. Zur Erzielung der doppelten Bildfolgegeschwindigkeit wird z. B. jede zweite

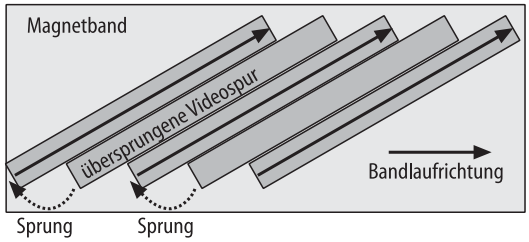


Abb. 8.32. Zeitraffer 2:1 durch Überspringen jeder 2. Spur

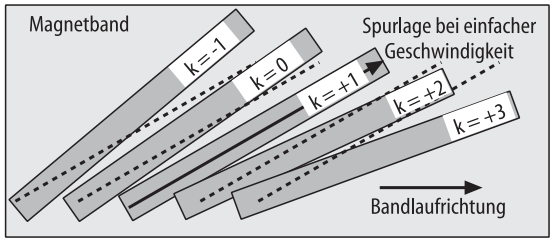


Abb. 8.33. Veränderte Spurlage bei Zeitlupe- und Zeitraffer

Spur einfach übersprungen. Professionelle MAZ-Geräte ermöglichen Werte für k zwischen – 1 und + 3, d. h. erlaubte Bandgeschwindigkeiten liegen zwischen einfach rückwärts bis dreifach vorwärts.

Beim Zeitlupen- und Zeitrafferbetrieb tritt das Problem auf, dass sich der Spurwinkel verändert. Abbildung 8.33 zeigt die Spurlage bei Normalbetrieb im Vergleich zum Betrieb mit veränderten Geschwindigkeiten. Ohne Korrekturmaßnahmen ist das Videosignal bei der Zeitlupen- oder Zeitraffer-Wiedergabe also gestört. Mit Hilfe der Dynamic Tracking-Funktion wird aber bei professionellen Geräten der Kopf auch bei verändertem Spurwinkel auf der Spur gehalten, die somit korrekt abgetastet werden kann. Bei Normalgeschwindigkeit kann beim Dynamic-Tracking-Betrieb die Bildqualität geringfügig verschlechtert sein, daher sollten bei Normalgeschwindigkeit die stationären statt der beweglichen (DT-) Köpfe benutzt werden.

8.3 Analoge Magnetbandaufzeichnung

Das große Problem bei der Videoaufzeichnung, die Verarbeitung der hohen Bandbreite, wird durch den Einsatz rotierender Videoköpfe gelöst. Ein zweites Problem besteht darin, dass die untere Grenzfrequenz bis nahe an 0 Hz heranreicht, und der gesamte Frequenzbereich damit mehr als 20 Oktaven (Frequenzverdoppelungen) umfasst. Da die Wiedergabespannung linear von der Frequenz abhängt (ω-Gang), wird bei Signalfrequenzen nahe Null eine extrem kleine Spannung induziert, und der Signal-Rausch-Abstand wird entsprechend schlecht. Eine dritte Schwierigkeit resultiert aus der nichtlinearen Magnetisierungskennline.

Die genannten Probleme werden gelöst, indem bei allen analogen MAZ-Formaten wenigstens das Luminanzsignal vor der Aufzeichnung frequenzmoduliert wird, womit die Information nicht mehr in der Signalamplitude, sondern in der Frequenz verschlüsselt ist.

Die Grundlagen der Frequenzmodulation sind in Kap. 4.1 beschrieben. Jedem Videosignalwert wird eine Frequenz zugeordnet. Die Amplitude des FM-Signals ist im Idealfall konstant und enthält keine Information. Das Basisbandsignal wird in den Frequenzen des Hubbereichs verschlüsselt, ober- und unterhalb des Hubbereichs treten Seitenbänder auf (Abb. 8.34). Da die Amplitude keine Information trägt, können Amplitudenschwankungen durch Begrenzung eliminiert werden, und das Band lässt sich immer bis zur Sättigung aussteu-

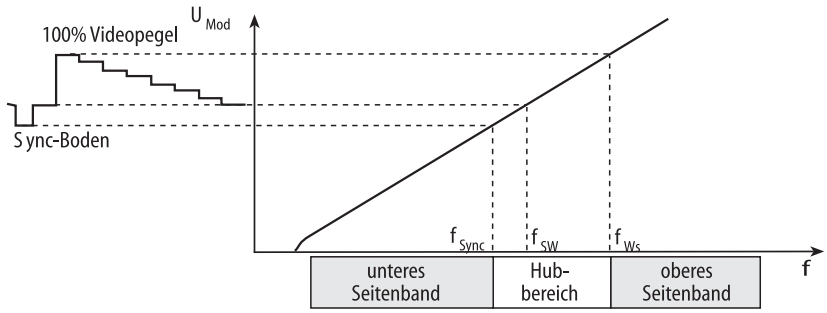


Abb. 8.34. FM bei Aussteuerung mit einem Videosignal

ern. Die Nichtlinearität der Magnetisierungskennlinie spielt fast keine Rolle mehr, und eine Vormagnetisierung ist nicht erforderlich. Allerdings ergibt sich bei FM gegenüber der nichttransformierten Aufzeichnung eine erhöhte Bandbreite. Die bei der Magnetbandaufzeichnung unter Beachtung ökonomischer Gesichtspunkte (Bandverbrauch, etc.) maximal nutzbare Grenzfrequenz beträgt ca. 16 MHz. Bei der professionellen FM-Videosignalaufzeichnung wird das Basisband mit den Grenzfrequenzen 0 Hz und 5 MHz in einen Bereich zwischen etwa 2 MHz und 16 MHz transformiert. Der Frequenzbereich umfasst nur noch 3–4 Oktaven, die im Wiedergabekopf frequenzabhängig induzierten Spannungen unterscheiden sich damit erheblich weniger als im Basisband.

Aufgrund der aus dem ω-Gang und den Verlusten resultierenden Kopfkurve können neben dem Hubbereich nur die direkt benachbarten Seitenbänder genutzt werden (Abb. 8.35). Es wird mit der Schmalband-FM, mit einem Modulationsindex < 1 gearbeitet. Der Hubbereich muss in einigem Abstand vom Video-Basisband liegen, damit sich nach der Demodulation Träger- und Videosignal ausreichend trennen lassen. Der FM-Hub beträgt bei professionellen Geräten ca. 2 MHz. Bei nicht professionellen Geräten wird der Hub auf ca. 1,2 MHz begrenzt und die Gesamtbandbreite beträgt etwa 8 MHz. Der Hubbereich ist durch die Frequenz f_{sync} für den Synchronboden und durch die Frequenz f_{ws} für den Weißwert festgelegt. Wegen der großen Bedeutung des Synchronsignals liegt f_{sync} am niederfrequenten Ende des Hubbereichs, hier sind die frequenzabhängigen Verzerrungen am geringsten.

Bei vielen Systemen wurde der genutzte Frequenzbereich im Laufe der Weiterentwicklung des Formates (Verfügbarkeit verbesserter Magnetbänder und -köpfe) ausgedehnt, die Frequenz für den Synchronboden wurde erhöht

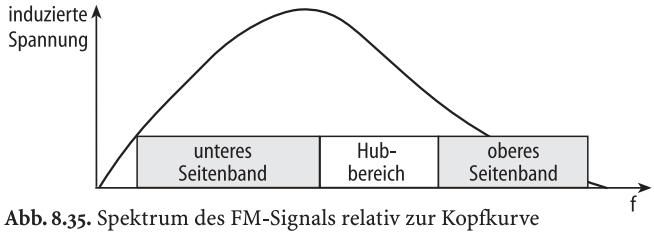


Abb. 8.35. Spektrum des FM-Signals relativ zur Kopfkurve

und der Hubbereich vergrößert. Diese weiterentwickelten Formate werden mit High-Band (HB), bzw. Super High-Band oder auch nur Super und High bezeichnet, z. B. U-Matic HB, Super-VHS, Hi 8.

Bei jeder Signalübertragung und -aufzeichnung werden störende Rauscheinflüsse wirksam. Bei Videosignalen trifft dies insbesondere auf hochfrequente Anteile zu, denn das Spektrum eines durchschnittlichen Videosignals zeigt, dass die Amplituden zu hohen Frequenzen hin abnehmen, womit sich bei hohen Frequenzen ein kleiner Signal-Rauschabstand ergibt. Zur Verbesserung des Signal-Rauschabstands werden die hochfrequenten Signalanteile vor der Aufnahme verstärkt. Die absichtlich herbeigeführte Verzerrung wird bei der Wiedergabe entsprechend ausgeglichen. Auf der Aufnahmeseite wird eine definierte Höhenanhebung (Preemphasis) vorgenommen, bei der Wiedergabe eine entsprechende Höhenabsenkung (Deemphasis), bei der dann auch die Rauschanteile mit abgesenkt werden.

Der Grad der Höhenanhebung ist durch passive Bauteile festgelegt und für jedes Aufzeichnungsformat genau spezifiziert. Abbildung 8.36 zeigt die Prinzipschaltung einer Preemphasis-Stufe und die Auswirkungen der Preemphasis im Zeit- und Frequenzbereich. Tiefe Frequenzen werden durch den in Reihe liegenden Kondensator abgeschwächt. Die Zeitkonstante aus C und R_1 bestimmt die Übergangsfrequenz, und mit R_1/R_2 wird der Grad der Pegelanhebung festgelegt. Neben der dargestellten Haupt- oder linearen Preemphasis wird auch mit nichtlinearer bzw. Neben-Preemphasis gearbeitet. Hier werden die Kennlinien durch pegel- und frequenzabhängige Widerstände (Dioden) verändert. Die Preemphasis wird dann bei kleinen Pegeln besonders groß, womit sich die Auflösung feiner Strukturen verbessert. Die Hochfrequenzanhebung bewirkt starke Überschwinger an steilen Signalübergängen. Je steiler der Signal- oder Schwarz-Weiß-Übergang, desto größer wird der Überschwinganteil. Abbildung 8.36 zeigt die Überschwinger bei einem Rechtecksignal. Der Maximalwert der Überschwinger muss auf einen festen Betrag begrenzt werden (Clipping), damit die Systembandbreite nicht überschritten wird. Der Begrenzungswert ist vom Aufzeichnungsformat abhängig und kann in Richtung Schwarz und Weiß unterschiedlich sein. Typische Werte sind z. B. 150% der Gesamtamplitude als Schwarz- und 200% als Weiß-Cliplevel.

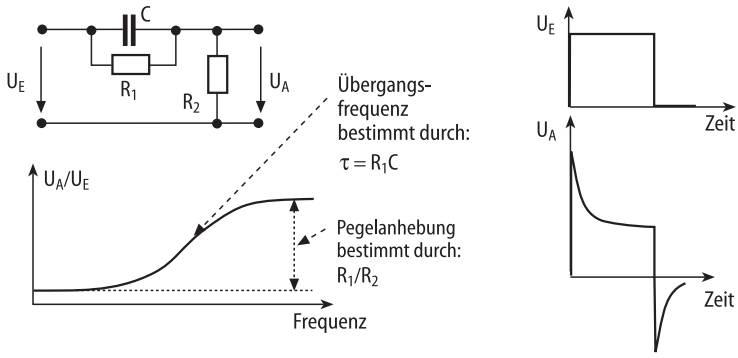


Abb. 8.36. Preemphasis-Schaltung, Wirkung im Frequenz- und Zeitbereich

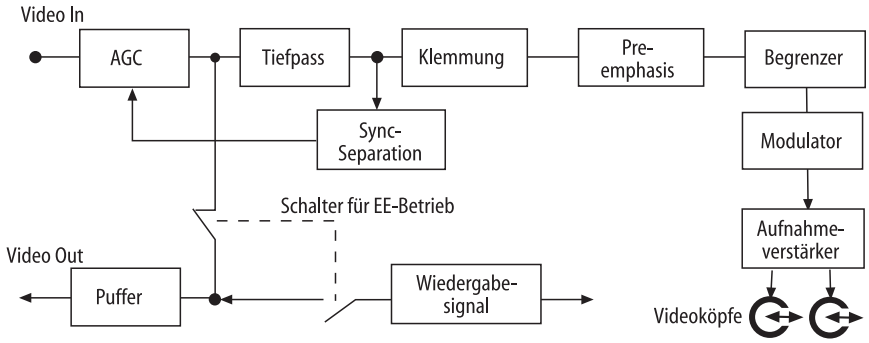


Abb. 8.37. Blockschaltbild der Aufnahmeelektronik

8.3.1 Signalverarbeitung bei der FM-Aufzeichnung

Ein Blockschaltbild der Aufnahmeelektronik ist in Abb. 8.37 dargestellt. Zunächst wird das Signal mit Hilfe eines automatisch geregelten Verstärkers (Automatic Gain Control, AGC) auf konstantem Pegel gehalten. Die Regelung arbeitet mit der H-Sync-Amplitude als Referenzwert. Das Synchron-Signal wird abgetrennt und außer zur AGC-Stufe zur Klemmstufe und zu den Band- und Kopf-Servosystemen geführt. Hinter der AGC-Stufe kann das Signal gleich wieder zur Ausgangsbuchse geschickt werden. Damit wird der sogenannte Durchschleif- oder E-E-Betrieb (Electronics to Electronics) ermöglicht, d. h. das Signal wird zur Kontrolle an den Ausgang geführt, wenn das Gerät nicht auf Wiedergabe-Betriebsart geschaltet ist.

Das geregelte Signal gelangt zu einem Tiefpass-Filter mit 5 MHz Grenzfrequenz und weiter zur Klemmstufe, die die Aufgabe hat, den Signalpegel für den Schwarzwert unabhängig vom wechselnden Videosignal auf einem konstanten Wert zu halten. Nach Durchlaufen der Preemphasis-Stufe und den Begrenzern für Schwarz- und Weißwert gelangt das Signal zum Frequenzmodulator. Dies ist ein vom Videosignal spannungsgesteuerter Oszillator (Voltage Controlled Oscillator, VCO), der im Aussteuerungsbereich eine lineare Kennlinie aufwei-

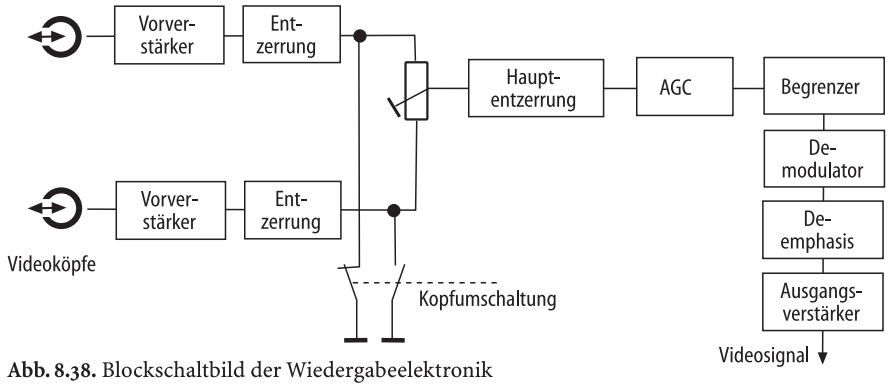


Abb. 8.38. Blockschaltbild der Wiedergabeelektronik

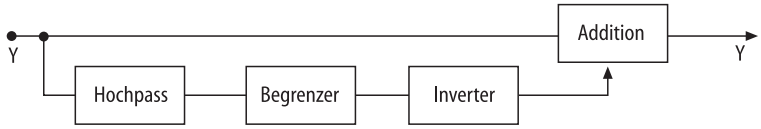


Abb. 8.39. Prinzip der Rauschunterdrückung

sen muss. Über ein Hochpass-Filter, das der Begrenzung des unteren FM-Seitenbandes dient, gelangt das Signal auf den Aufnahmeverstärker und schließlich wird es über die rotierenden Übertrager zu den Videoköpfen geführt.

Abbildung 8.38 zeigt ein Blockschaltbild der Wiedergabeelektronik. Für die Wiedergabe können die gleichen Köpfe wie für die Aufnahme benutzt werden. Im Kopf wird eine sehr kleine Spannung erzeugt, die erheblich verstärkt werden muss. Das Signal gelangt über die rotierenden Übertrager zu den Vorverstärkern Jedem Videokopf wird ein eigener Verstärker zugeordnet, um die nötigen Pegel- und Entzerrungs-Einstellungen separat vornehmen und die Signale einander angleichen zu können.

Nach der Vorverstärkung werden die Signale der einzelnen Videoköpfe zu einem kontinuierlichen Signalfluss vereinigt. Dieses gelangt dann zu einer Wiedergabe-AGC-Stufe und ggf. zu einem analogen Drop Out Detektor (s. Abschn. 8.4.3). Das Signal wird schließlich begrenzt und zum FM-Demodulator geführt.

Zur Demodulation werden in der Videotechnik Zähldiskriminatoren eingesetzt. Die Häufigkeit der Nulldurchgänge ist der FM-Frequenz proportional, die Integration ergibt dann bei hoher einen großen und bei geringer Häufigkeit einen kleinen Wert, und damit ist das Basisbandsignal zurückgewonnen. Das Signal wird anschließend einer Deemphasis unterzogen, mit der die bei der Aufnahme vorgenommene Höhenanhebung wieder rückgängig gemacht wird.

Im Basisband werden noch Maßnahmen zur Erhöhung des Schärfeeindrucks im Bild durchgeführt. Das Arbeitsprinzip ist das gleiche wie bei den Detailing-

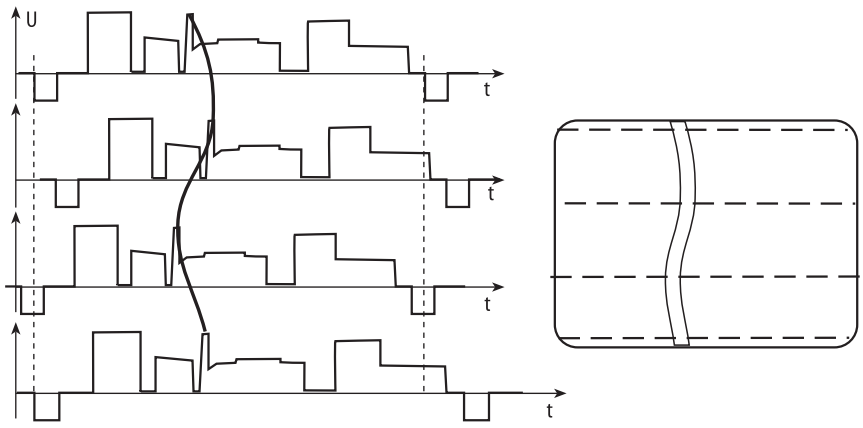


Abb. 8.40. Zeitbasisfehler und Auswirkung im Bild

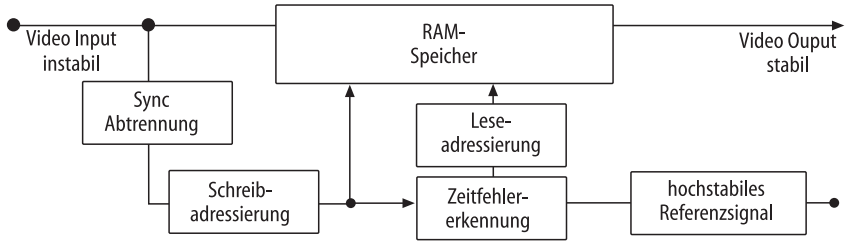


Abb. 8.41. Funktionsprinzip des Time Base Correctors

Schaltungen zur Konturanhebung in einer Kamera (s. Kap. 6.2) Häufig wird in den Wiedergabeweg noch eine Rauschunterdrückung geschaltet. Im einfachen Fall wird mit Hilfe einer Signaldifferenzierung der Rauschanteil extrahiert und mit inverser Phasenlage dem Original wieder hinzuaddiert, wodurch das Rauschen kompensiert wird (Abb. 8.39).

8.3.2 Zeitbasiskorrektur (TBC)

Trotz bester Servo- und Trackingsysteme lässt sich eine mechanische Restungenauigkeit bei der Abtastung der Spuren nicht verhindern. Die Instabilität führt zu Zeitfehlern (Jitter), die auch die Synchronsignale und damit die Zeitbasis des Signals betreffen. In vielen Fällen wird der Fehler im Bild nicht sichtbar, z. B. bei einfacher Überspielung zwischen Recordern oder bei Monitor-Wiedergabe, da sich die Geräte auf die Schwankung einstellen. Bei Fernsehgeräten wird dafür oft eine Umschaltung der Regelzeitkonstante vorgenommen (automatische Umschaltung bei Wahl des A/V-Eingang für Videorecorder).

Wenn dagegen mehrere Videosignalquellen gemischt werden sollen (Studiobetrieb), so ist deren Synchronität unbedingt erforderlich, da sonst die Bilder horizontal gegeneinander verschoben erscheinen oder in horizontaler Richtung schwanken (Abb. 8.40). Das FBAS-Signal darf einen Zeitfehler von maximal 3 ns enthalten, beim professionellen Einsatz von MAZ-Maschinen ist daher eine Zeitbasiskorrektur notwenig. Das Gerät, das diese Funktion übernimmt, heißt Time Base Corrector (TBC), es ist ein externes Gerät oder bereits fest in das MAZ-Gerät eingebaut (Abb. 8.42). Die Aufgabe des TBC ist die Stabilisierung des Signals, es wird dabei an einen zentralen Studiotakt bzw. an einen konstanten Taktmittelwert, als sogenannte Eigenführung, angepasst (Abb. 8.41). Die Zeitbasiskorrektur beruht auf einer steuerbaren Verzögerungsleitung, in der ei-

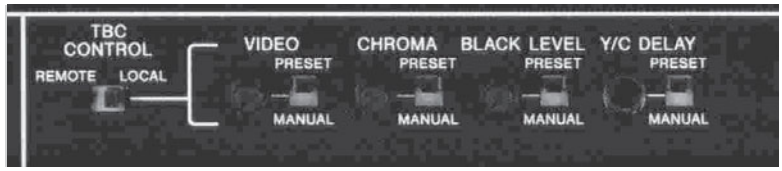


Abb. 8.42. TBC-Bedienelemente an einer MAZ-MAschine [208]

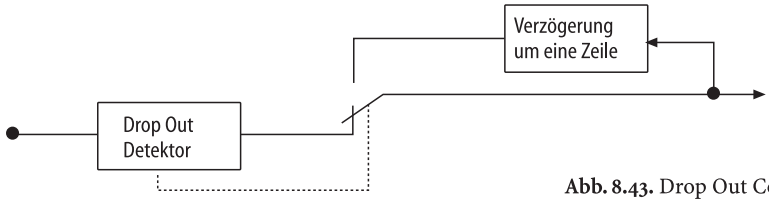


Abb. 8.43. Drop Out Compensation

nige Zeilen (z. B. 16) des Videobildes oder ein komplettes Halb- oder Vollbild zwischengespeichert werden. Zur Korrektur wird die Verzögerungszeit von einem Signal abhängig gemacht, das aus dem schwankenden Signal gewonnen wird (Zeitfehlererkennung). Ist die Zeilendauer in Bezug auf das Referenzsignal zu kurz, so wird die Verzögerungszeit verlängert und umgekehrt.

Wegen des in digitaler Form vorliegenden Signals kann der TBC ohne großen Aufwand zusätzliche Funktionen erfüllen, so werden hier häufig Signalveränderungen ermöglicht. Meistens sind Video- und Chromapegel und die Horizontal-Phase einstellbar. Die zugehörigen Einsteller (Abb. 8.41) finden sich am MAZ-Gerät oft etwas versteckt. Da sie an Schnittplätzen zum Angleichen der Pegel von verschiedenen Aufzeichnungsgeräten aber häufig gebraucht werden sind Fernbedieneinheiten (TBC Remote) verfügbar.

Die Digitalebene des TBC bietet sich auch zum Einsatz von digitalen Filtern an, mit denen eine Rauschminderung erzielt werden kann. Weiterhin wird der TBC häufig zur sogenannten Drop Out Compensation eingesetzt (Abb. 8.43). Damit durch den kurzzeitigen Ausfall des Signals (Drop Out) aufgrund von Kopfverschmutzung, Bandfehlern etc. keine Bildstörung (Spratzer) entsteht (Abb. 8.44), wird im Drop Out Compensator vor der Demodulation der Signalfehler festgestellt und das gestörte Signal durch ein fehlerfreies Nachbarsignal ersetzt. Die Verfälschung fällt kaum auf, da benachbarte Bildbereiche in der Regel sehr ähnlich sind. In einfachen Geräten wird die Drop Out Compensation auch auf analoger Ebene realisiert. Hier wird eine Zeile des Videosignals gespeichert, die dann ausgelesen wird, wenn die aktuelle Zeile einen Drop Out aufweist.

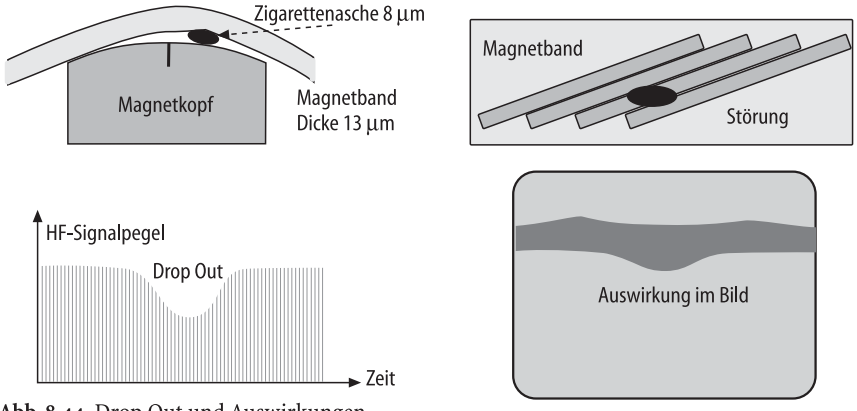


Abb. 8.44. Drop Out und Auswirkungen

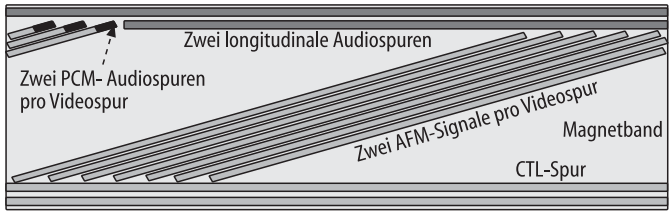


Abb. 8.45. Audioaufzeichnungsmöglichkeiten in einem analogen Videorecorder

8.4 Audioaufzeichnung in Analog-Recordern

Audiosignale haben im Vergleich zu Videosignalen eine viel geringere Grenzfrequenz und können auf Längsspuren aufgezeichnet werden, wobei allerdings die Tonqualität insbesondere bei Videorecordern für den Heimbereich (VHS) aufgrund der geringen Bandgeschwindigkeit sehr mäßig sein kann. Um auch bei langsam laufendem Band eine gute Tonqualität zu erreichen, wird auch für die Audioaufzeichnung die Frequenzmodulation angewendet. Eine weitere Alternative ist die digitale PCM-Aufzeichnung (Abb. 8.45).

Die Längsspur-Aufzeichnung ist die am häufigsten benutzte Art der analogen Tonaufzeichnung. Das Videoband zieht an einem feststehenden Tonkopf vorbei, wobei bei professionellen Systemen mehrere durchgehende Längsspuren am Rand des Videobandes entstehen. Das Band wird hier in Abhängigkeit von der Audio-Signalamplitude mehr oder weniger stark magnetisiert.

Die nichtlineare Magnetisierungskennlinie (s. Abschn. 8.2.1) bewirkt Signalverzerrungen, auf die das Gehör sehr empfindlich reagiert. Um die Signal-Verzerrungen bei der Audioaufzeichnung zu mindern, werden mit der HF-Vormagnetisierung technische Maßnahmen ergriffen, die dazu führen, dass nur die linearen Teile der Magnetisierungskennlinie wirksam werden. Das Audiosignal wird dazu einem HF-Signal (z. B. 100 kHz) aufgeprägt, wobei die Vormagnetisierung so gewählt wird, dass das Audiosignal immer im Bereich des linearen Kennlinienteils bleibt (Abb. 8.46). Die Größe des Vormagnetisierungsstromes bestimmt entscheidend den Verzerrungsgrad. Die Festlegung wird dafür häufig

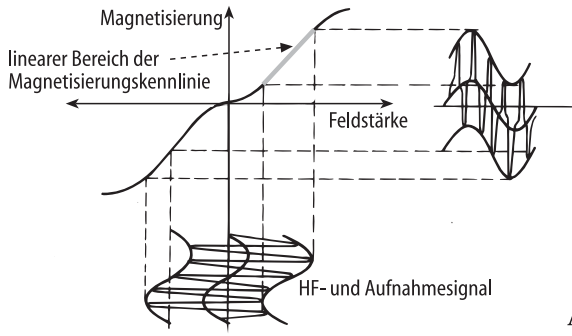


Abb. 8.46. HF-Vormagnetisierung

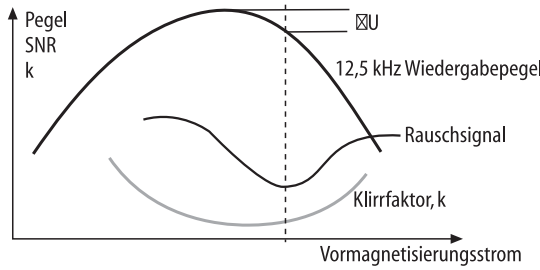


Abb. 8.47. Einstellung der Vormagnetisierung

so vorgenommen, dass der Wiedergabepegel bei 12,5 kHz nach Überschreitung des Pegelmaximums um einen definierten Wert ΔU abfällt. In diesem Bereich ergibt sich ein großer Rauschspannungsabstand bei kleinem Klirrfaktor (Abb. 8.47). Natürlich werden die Verzerrungen auch durch die NF-Signalamplitude beeinflusst. Diesbezüglich wird die Aussteuerungsgrenze meistens so festgelegt, dass ein Klirrfaktor von 1% oder 3% nicht überschritten wird.

Die maximale Systemdynamik, d. h. das geräteabhängige Verhältnis von leisestem zu lautestem Signal, wird vom Signal-Rauschabstand bestimmt. Leise Passagen sollen nicht vom Rauschen verdeckt werden. Bei professionellen Geräten werden mehr als 52 dB Dynamik erreicht. Der Signal-Rauschabstand kann verbessert werden, indem auch beim Audiosignal bei der Aufnahme eine frequenzbereichs- und pegelabhängige Signalanhebung für hohe Frequenzen vorgenommen wird, die bei der Wiedergabe entsprechend rückgängig gemacht wird. Das Signal wird aufnahmeseitig komprimiert (verringerte Dynamik) und wiedergabeseitig entsprechend expandiert (Abb. 8.48), wobei das Rauschen abgesenkt wird. Häufig eingesetzte Compander-Verfahren dieser Art sind die Systeme Dolby B und C, im deutschsprachigen Raum (B-Format) wird auch Telcom C4 verwendet. Rauschunterdrückungssysteme für Audiosignale sind im Gegensatz zur Video-Preemphasis meistens abschaltbar.

Die Längsspuraufzeichnung ermöglicht in professionellen Geräten die Aufzeichnung von Frequenzen zwischen ca. 50 Hz und 15 kHz. Der Frequenzbereich und auch die Dynamik werden aber bei jeder erneuten Kopie des Materials eingeschränkt, sodass mehr als vier Generationen nicht vertretbar sind.

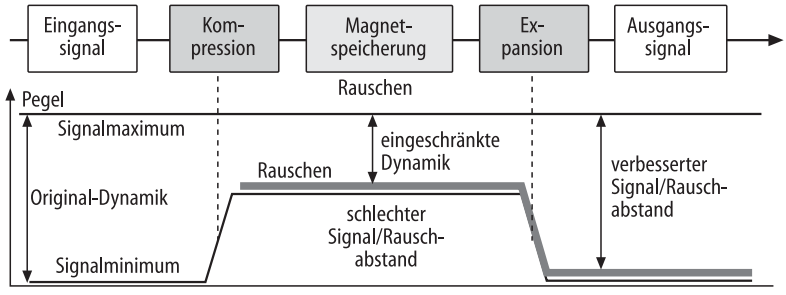


Abb. 8.48. Funktionsprinzip der Rauschunterdrückung

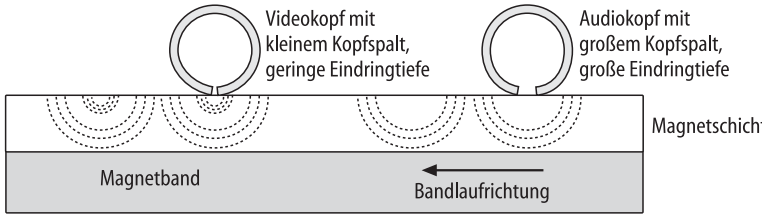


Abb. 8.49. Video- und Audio-FM-Aufzeichnung

Die Audiosignalqualität kann gesteigert werden, wenn die Aufzeichnung wie beim Videosignal auch mit rotierenden Köpfen vorgenommen wird. Bei diesem Verfahren wird mit Frequenzmodulation gearbeitet. Die Tonaufzeichnungsart wird mit Audio-FM (AFM), im Amateurbereich auch mit HiFi-Ton bezeichnet. Das AFM-Signal wird zusammen mit dem Videosignal auf den Videospuren aufgezeichnet. Das AFM-Signal kann also nicht verändert werden, ohne das Videosignal zu beeinflussen, das Bildsignal und der AFM-Ton sind nicht unabhängig voneinander editierbar.

Die AFM-Aufzeichnung arbeitet simultan für zwei Tonkanäle. Zur Aufzeichnung werden die Videoköpfe oder separate, den Videoköpfen voreilende, AFM-Köpfe benutzt (Abb. 8.49). Den beiden Tonkanälen werden eigene Trägerfrequenzen unterhalb des unteren FM-Seitenbandes zugeordnet. Beim Betacam-Format haben sie z. B. die Werte 310 kHz und 540 kHz, bei VHS 1,4 MHz und 1,8 MHz (Abb. 8.50). Die nutzbare Bandbreite liegt formatabhängig pro Kanal bei etwa ±100 kHz und der Hub bei ca. 50 kHz. Die Trennung von Audio- und Videosignalen wird durch Filter vorgenommen und beim Einsatz von separaten AFM-Köpfen durch die unterschiedlichen Kopfspaltbreiten und Azimutwinkel der AFM- und Videoköpfe unterstützt. Der Azimutversatz (s. Abschn. 8.1) beträgt bei den Audioköpfen z. B. ± 30°, während er für die Videoköpfe z. B. ± 6° beträgt. Damit ergibt sich eine Übersprechdämpfung nicht nur gegenüber den Nachbarspuren, sondern auch zwischen den Audio- und Videosignalen. Der AFM-Kopf zeichnet relativ tiefe Frequenzen auf und kann dementsprechend einen breiteren Kopfspalt als der Videokopf haben. Das Magnetfeld kann hier tief in die Magnetschicht des Bandes eindringen. Der nacheilende Videokopf magnetisiert dagegen nur die Bandoberfläche und löscht das AFM-Signal. Das AFM-Signal befindet sich damit vorwiegend in der Tiefe der Magnetschicht (ca. 0,8 µm), das Videosignal eher an der Oberfläche (ca. 0,3 µm) (Abb. 8.49). Bei der Abtastung wird das hochfrequente Videosignal im Audiokopf weitgehend kurzgeschlossen, während der Videokopf nur die Signale aus der Bandoberfläche erfasst.

Abb. 8.50. AFM-Frequenzmultiplex

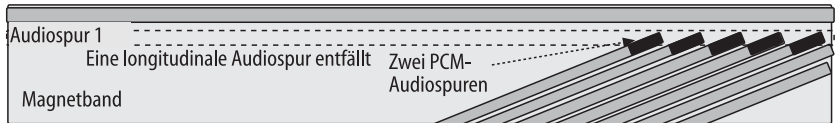


Abb. 8.51. Lage der PCM-Audiospuren bei Betacam SP und MII

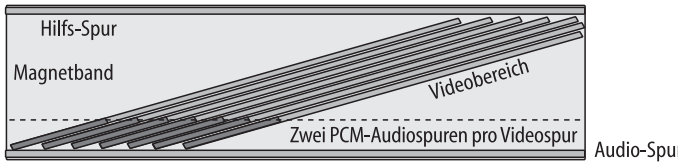


Abb. 8.52. PCM-Spuren bei Video8

Als dritte Möglichkeit für eine Audioaufzeichnung in analogen Videorecordern gibt es die PCM-Aufzeichnung von digitalisierten Audiosignalen. In Analogrecordern, die mit PCM-Audioaufzeichnung ausgestattet sind, wird das Signal für zwei Kanäle digital gewandelt und direkt in einem separaten Bereich auf dem Band gespeichert. Es ist damit, wie das Audiosignal auf der Longitudinalspur, unabhängig vom Videosignal editierbar. Die hohe Signalfrequenz erfordert eine Aufzeichnung auf Schrägspuren mit Magnetköpfen, die sich auf der Kopftrommel befinden. Da bei der Konzeption der Formate kein separater Bandbereich reserviert wurde, wird für die PCM-Aufzeichnung eine longitudinale Tonspur geopfert. Abbildung 8.51 zeigt die Veränderung der Spurlage beim Format Betacam SP.

Nach der ggf. erforderlichen A/D-Wandlung werden die Abtastwerte umgeordnet, mit Zusatzdaten zum Fehlerschutz versehen und schließlich zeitkomprimiert, da die kleinen PCM-Bereiche bei jeder Kopftrommelumdrehung nur kurz abgetastet werden. Bei der Wiedergabe findet entsprechend eine zeitliche Expansion statt. Es wird mit Digitalisierungsparametern für höchste Qualität, d. h. mit einer Abtastrate von 44,1 kHz oder 48 kHz und einer Amplitudenauflösung von 16 oder 20 Bit gearbeitet. Beim Heimformat Video 8 kann der PCM-Ton unabhängig vom Bild editiert werden, ein eigener Bereich wurde dafür bereits bei der Konzeption des Formates berücksichtigt. Der PCM-Datenbereich liegt unterhalb der Videospuren (Abb. 8.52), wenn er genutzt werden soll, muss die Kopftrommelumschlingung von 185° auf 221° erhöht werden.

Bei der Nutzung der Magnetbandaufzeichnung ist die richtige Aussteuerung sehr wichtig. Zur Erzielung eines guten Signal-Rauschabstandes muss

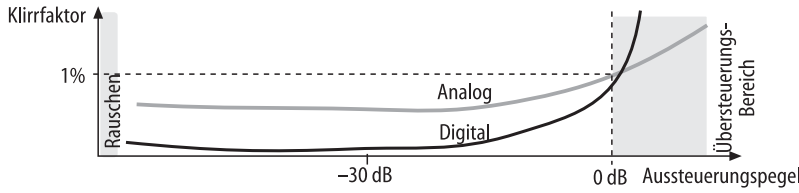


Abb. 8.53. Aussteuerungsverhalten von analogem und digitalem Signal

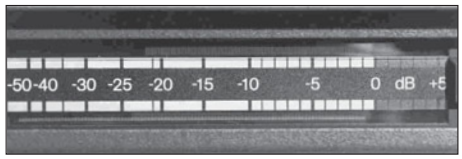


Abb. 8.54. Display des Peak Program Meters

Abb. 8.55. Display des VU-Meters

mit möglichst hohem Pegel aufgezeichnet werden. Mit steigendem Pegel steigen aber auch die Verzerrungen, und damit der Klirrfaktor. In Deutschland wird der Maximalpegel so festgelegt, dass ein Klirrfaktor von 1 % THD (Total Harmonic Distortion) nicht überschritten wird, international findet sich diesbezüglich auch der Wert 3 % THD. Analogsysteme verhalten sich bei Spitzenwertüberschreitung so, dass die Verzerrungen relativ langsam weiter ansteigen. Abbildung 8.53 zeigt im Vergleich, dass sich Digitalsysteme ganz anders verhalten die Verzerrungen steigen bei Pegelüberscheitung sehr stark und kaum kontrollierbar an. Aus diesem Grund arbeitet man bei Digitalsystemen mit einem Sicherheitsbereich, dem Headroom. Ein Headroom von 10 dB bedeutet, dass der erlaubte Maximalpegel so gewählt wird, dass er 10 dB unterhalb des Pegels liegt, bei dem die Verzerrungen tatsächlich sprunghaft ansteigen. Um verschiedene Geräte möglichst einfach miteinander verbinden zu können, wird festgelegt, auf welche Spannung ein Sinussignal verstärkt wird, also mit welchem Arbeitspegel gearbeitet wird. In Deutschland beträgt der Audio-Arbeitspegel für professionelle System +6 dBu, was 1,55 V entspricht. International ist der Wert +4 dBu gebräuchlich.

Zur Aussteuerungskontrolle werden Spannungsmessgeräte mit logarithmischer, und damit der Lautstärkeempfindung angepasster Skala eingesetzt. Die Instrumente werden in PPM-Anzeigen (Peak Program Meter) und VU-Anzeigen (Volume Unit) unterschieden. Zur effektiven Verhinderung von Übersteuerung ist nur die PPM-Anzeige geeignet (Abb. 8.54), denn sie reagiert mit einer Ansprechzeit von 5 ms so schnell, dass die Signalspitzen sichtbar werden.

VU-Anzeigen (Abb. 8.55) sind zu träge, um kurze Spitzenwerte anzuzeigen. Sie haben aber den Vorteil, dass sie in etwa das mittlere Aussteuerungsniveau wiedergeben. Sie sind damit besser als das PPM an die menschliche Lautstärkeempfindung angepasst, die vom Signalmittelwert abhängt.

Die Audiosignalführung ist bei professionellen Videorecordern symmetrisch, sowohl für die digitale wie für die analoge Übertragung werden fast ausschließlich XLR-Stecker eingesetzt (Abb. 8.56). Im Amateurbereich kommen Cinch-, und auch Klinkenstecker bei unsymmetrischer Signalführung zum Einsatz, außerdem wird hier das Audiosignal auch über SCART-Stecker geführt.

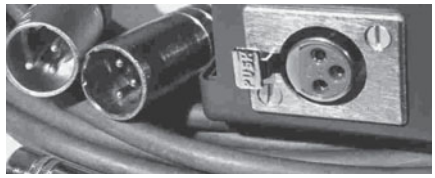


Abb. 8.56. Audio-Steckverbindungen

8.5 FBAS-Direktaufzeichnung

Zu Beginn der Entwicklung von farbtüchtigen Videorecordern wurde das gesamte FBAS-Signal direkt aufgezeichnet. Das erste studiotaugliche Format beruhte auf 2“- breiten Bändern und Querspuraufzeichnung. Mitte der 70er Jahre wurde die Schrägspuraufzeichnung auf 1“ breiten Bändern eingeführt, die offenen Spulen der entsprechenden Formate B und C sind noch heute in den Archiven der Rundfunkanstalten zu finden.

Die FBAS-Direktaufzeichnung erfordert sowohl bei der Aufnahme als auch bei der Wiedergabe wenig elektronischen Aufwand, allerdings müssen die Schaltungen genau dimensioniert sein, um höchste Qualität zu erreichen. Der Nachteil des Verfahrens ist die erforderliche hohe Video-Bandbreite, die Geräte sind daher aufwendig gebaut, groß und wenig mobil. FBAS direkt aufzeichnende Geräte benötigen einen TBC, denn das Farbsignal wird mit dem Farbhilfsträger aufgezeichnet, der einen Phasenfehler von maximal 3 ns aufweisen darf.

Die im Folgenden beschriebenen Aufzeichnungsformate sind für die US-Norm mit 525 Zeilen und 60 Hz vielfach anders spezifiziert als für das europäische TV-System mit 625 Zeilen und 25 Hz Bildwechselfrequenz. Die Parameter werden hier im Folgenden bezüglich der europäischen Norm dargestellt.

8.5.1 Das 2“-Quadruplex-System

Das Quadruplex-System war das erste brauchbare Videoaufzeichnungsverfahren. Es wurde 1956 von der Firma Ampex zunächst als S/W-Aufzeichnungsformat eingeführt und war in weiterentwickelter, farbtauglicher Form (High-Band, 1964) bis in die 80er Jahre Studiostandard (Abb. 8.57). Heute wird es höchstens noch zur Wiedergabe von Archivbändern eingesetzt.

Das Quadruplex-Verfahren hat seine Bezeichnung von den vier Videoköpfen, die symmetrisch auf der Kopfscheibe angebracht sind. Die Kopfscheibe rotiert quer zum Band, daher auch die Bezeichnung Querspuraufzeichnung. Das Band hat eine Breite von 50,8 mm (2“) und wird über ein Vakuumsystem der runden Kopfscheibenform angepasst (Abb. 8.58). Der Kopfscheibendurch-

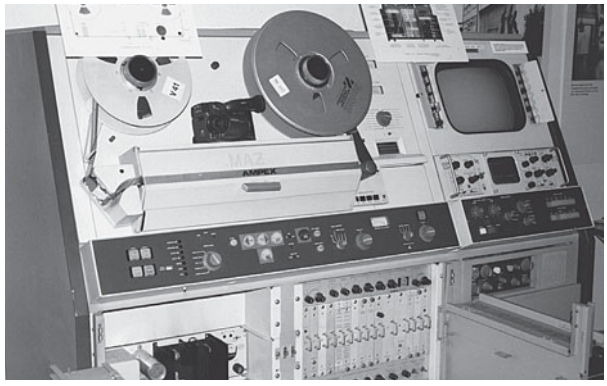


Abb. 8.57. 2“-Quadruplex-MAZ

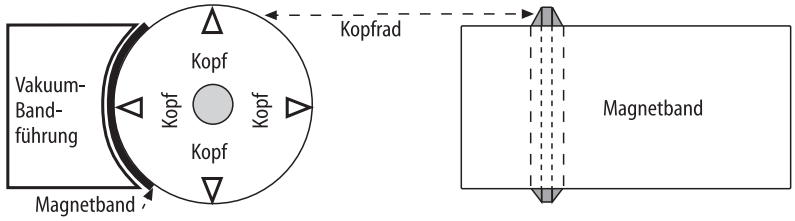


Abb. 8.58. Videokopfanordnung beim Quadruplex-Format

messer beträgt 52,5 mm, die Kopfscheibe rotiert mit 250 U/s. Es werden 1000 Spuren pro Sekunde geschrieben und die Schreibgeschwindigkeit beträgt ca. 40 m/s. Eine Sekunde Videosignal umfasst im europäischen TV-System 15625 Zeilen, d. h. jede Spur beinhaltet mit Überlappung ca. 18 Zeilen. Es findet also eine segmentierte Aufzeichnung statt, ein Halbbild wird in 20 Spuren gespeichert. Die Bandlaufgeschwindigkeit betrug anfangs 39 cm/s, später konnte auch mit 19,5 cm/s aufgezeichnet werden.

Die Spurbreite und der Abstand der Spuren (Rasen) hängt von der Bandgeschwindigkeit ab. Tabelle 8.3 zeigt eine Übersicht über die wichtigsten geometrischen Parameter, Tabelle 8.4 gibt eine Übersicht über die Parameter der Signalverarbeitung auf der farbtauglichen Entwicklungsstufe (High Band).

Eine Längsspur von 1,8 mm Breite oberhalb des Videoaufzeichnungsbereichs wird zur Tonaufzeichnung in konventioneller Technik genutzt. Aufgrund der hohen Bandgeschwindigkeit wird eine gute Tonqualität erzielt. Durch die räumliche Trennung von Ton- und Videokopf ergibt sich ein Ton- Videosignalversatz von 0,6 s. Mechanische Schnitte, die am Anfang der Entwicklung noch durchgeführt wurden, sind daher problematisch und nur in längeren Tonpausen ohne Störung möglich. Unterhalb des Videobereichs befinden sich die Cue-Spur und eine Steuerspur. Die Cue-Spur dient zur Aufzeichnung von Regieanweisungen oder Timecode. Auf der Steuerspur (CTL) wird in Abhängigkeit von der Kopfdrehzahl und -phase ein Kontrollsignal aufgezeichnet, das bei der Wiedergabe dazu dient, die Bandgeschwindigkeit so zu steuern, dass die Videospuren genau abgetastet werden. Abbildung 8.59 zeigt die relative Lage der Magnetköpfe.

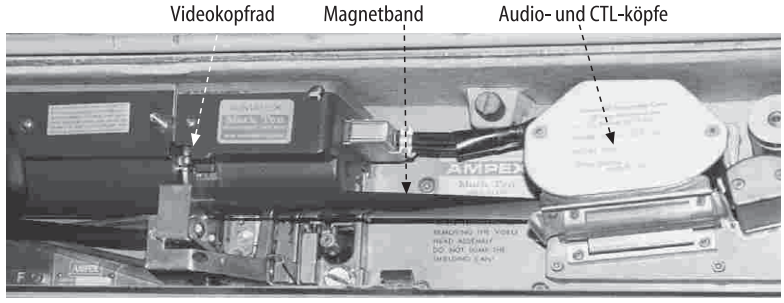


Abb. 8.59. Bandführung beim Quadruplex-Format

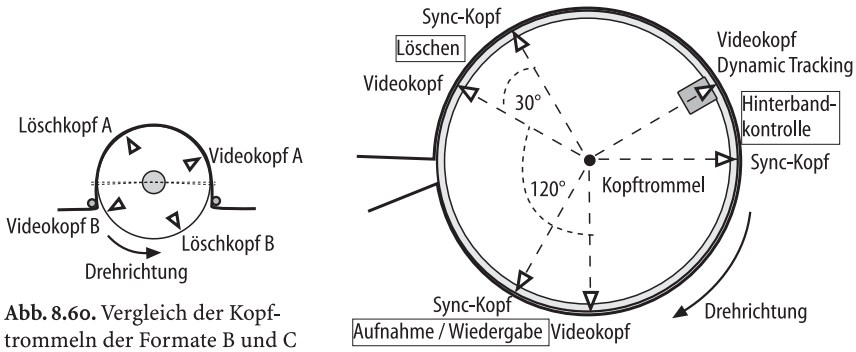


Abb. 8.60. Vergleich der Kopftrommeln der Formate B und C

8.5.2 Das 1"-A- und C-Format

Mitte der 70er Jahre wurden mit dem A-Format die ersten studiotauglichen Geräte auf den Markt gebracht, die mit Schrägspuraufzeichnung arbeiteten. Der Spurwinkel beträgt nur etwa 3°, und das Band wird in Ω-Umschlingung im 350°-Bogen um eine große Kopftrommel mit 13 cm Durchmesser herumgeführt. Damit wird eine Videospurlänge von ca. 40 cm erreicht, die auch bei der hohen Schreibgeschwindigkeit von mehr als 20 m/s ausreicht, um ein ganzes Halbbild aufzunehmen. Das Signal wird mit nur einem Kopf geschrieben, daher kann auf der Videospur nicht das gesamte Videosignal aufgezeichnet werden. Etwa 16 Zeilen pro Halbbild fallen weg oder werden mit separaten Sync-Köpfen aufgezeichnet, die den Videoköpfen um 30° vorauseilen. Den Video- und Sync-Köpfen für die Aufnahme (Record, Rec) werden im Abstand von 120° noch je ein Lösch- und Wiedergabekopf hinzugefügt (Playback, PB). Abbildung 8.60 zeigt die Anordnung der Köpfe auf der Trommel. Die Videoaufzeichnung erfolgt ohne gekippten Kopfazimut, dafür mit einem Sicherheitsabstand zwischen den Spuren.

Das A-Format wurde schnell zum C-Format weiterentwickelt, da es nur zwei Tonspuren beinhaltet, von denen eine im professionellen Einsatz als Timecode-Spur genutzt werden muss [229]. Beim Übergang auf das C-Format wurden die Spurlagen-Parameter geändert (Abb. 8.61). Der Spurwinkel wurde auf ca. 2,5° verkleinert und damit Platz für fünf Longitudinalspuren geschaffen. Ne-

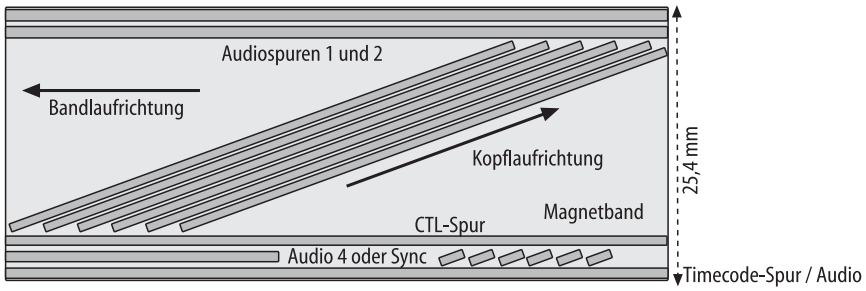


Abb. 8.61. Spurlage beim C-Format

Tabelle 8.2. Geometrische Parameter der Formate Quadruplex, B und C

	Quadruplex	B-Format	C-Format
Magnetbandbreite	50,8 mm	25,4 mm	25,4 mm
Bandgeschwindigkeit	39,7 cm/s	24,3 cm/s	24 cm/s
Relativgeschwindigkeit	41,2 m/s	24 m/s	21,4 m/s
Videospurlänge	46,1 mm	80 mm	411,5 mm
Videospurbreite	254 µm	160 µm	160 µm
Spurzwischenraum	143 µm	40 µm	54 µm
Spurwinkel	quer	14,3°	2,5°
Kopftrommeldurchmesser	52,5 mm	50,3 mm	134,6 mm
Kopftrommeldrehzahl	250 U/s	150 U/s	50 U/s

ben der Kontrollspur sind drei Audiospuren nutzbar. Die fünfte Spur kann als weitere Audiospur genutzt werden, oder es wird in diesem Bereich mit separaten rotierenden Köpfen der Teil des Synchronsignals aufgezeichnet, der auf den Videospuren wegen der nur 350° betragenden Kopftrommelumschlingung keinen Platz findet. Die Audioqualität ist gut, denn das Band läuft mit einer Geschwindigkeit von 24 cm/s. Die Videosignalverarbeitung entspricht der in Abschn. 8.3 beschriebenen. Die FM-Trägerfrequenz für den Schwarzwert beträgt 7,68 MHz bei 1,74 MHz Hub. Abbildung 8.62 zeigt ein 1"-C-MAZ-Gerät. Die Tabellen 8.2 und 8.3 geben im Vergleich zum B-Format die mechanischen Spezifikationen sowie einen Überblick über die Signal-Spezifikationen an.

Geräte nach dem C-Standard können mit beweglichen Köpfen ausgestattet sein, die eine genaue Spurabtastung auch bei veränderter Bandgeschwindigkeit ermöglichen, sodass eine störungsfreie Zeitlupen- und Zeitrafferwiedergabe möglich wird. Für besonders hochwertige Zeitlupenwiedergabe, wie sie häufig bei der Sportübertragung erforderlich ist, stehen beim C-Format sogenannte

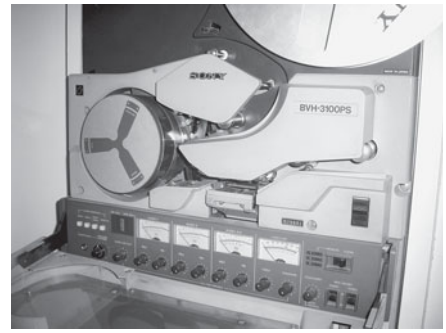


Abb. 8.62. 1"-C-MAZ , rechts mit geöffneter Klappe

Tabelle 8.3. Signalparameter der Formate Quadruplex, B und C

	Quadruplex	B-Format	C-Format
Bandbreite (–3 dB)	6 MHz	5,5 MHz	5,5 MHz
Signal/Rauschabstand	> 42 dB	> 48 dB	> 45 dB
FM-Träger (Blanking)	7,8 MHz	7,4 MHz	7,68 MHz
FM-Hub	2,14 MHz	2,14 MHz	1,74 MHz
Audiobandbreite (–3 dB)	30 Hz–15 kHz	50 Hz–15 kHz	50 Hz–15 kHz

Super-Slow-Motion-Geräte zur Verfügung, mit denen in Verbindung mit einer Spezialkamera Signale mit dreifacher zeitlicher Auflösung, also mit 90 statt 30 Bildern/s (NTSC-System), aufgezeichnet werden können. Das Super-Slow-Motion-System weist damit gegenüber herkömmlicher Zeitlupenwiedergabe eine wesentlich feinere Bewegungsauflösung auf. Geräte nach dem C-Format bieten, gemessen an Standards der 90er Jahre, eine sehr gute Qualität und werden im Archivbereich noch eingesetzt. Sie waren im Ausland und in privaten deutschen Produktionshäusern häufiger anzutreffen als die qualitativ vergleichbaren Geräte im B-Format, die vorwiegend in den öffentlich-rechtlichen Anstalten Deutschlands verwendet wurden.

8.5.3 Das 1"-B-Format

1"-B-Geräte für geschlossen codierte FBAS-Signale waren in deutschsprachigen Ländern die Standard-MAZ-Geräte für hochwertige Videoproduktionen, bzw. der Sendestandard. Sie wurden von der Firma Bosch/BTS als Typenreihe BCN hergestellt. Abbildung 8.63 zeigt ein portables Gerät. Das B-Format erreicht mit etwa gleichen Parametern wie beim C-Format, nämlich einer Bandgeschwindigkeit von 24 cm/s und einer Schreibgeschwindigkeit von 24 m/s, die gleiche Signalqualität, d. h. eine Videobandbreite von 5,5 MHz bei einem Signal-Rauschabstand von mehr als 46 dB. Die FM-Trägerfrequenz für den Schwarzwert beträgt 7,4 MHz, der Hub 2,14 MHz. Tabelle 8.3 gibt eine Übersicht über die Signal-Spezifikationen.

Abb. 8.63. Portabler Recorder im B-Format

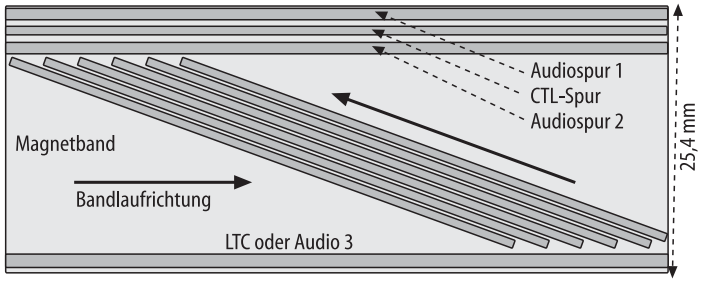


Abb. 8.64. Spurlage beim B-Format

Der wesentliche Unterschied zum C-Format ist der völlig andere mechanische Aufbau, der dadurch bestimmt ist, dass eine viel kleinere Kopftrommel mit nur 5 cm Durchmesser eingesetzt wird (Abb. 8.60 und 8.65). Um Banddehnung und Bandverschleiß zu reduzieren, wird das Band nur zu 190° um die Kopftrommel herumgeführt. Das Signal wird mit zwei um 180° versetzten Köpfen abwechselnd geschrieben. Aufgrund der geringen Kopftrommel-Umschlingungsstrecke ergibt sich ein vergleichsweise großer Spurwinkel von 14,4° und eine Spurlänge von 8 cm, in der ein ganzes Halbbild nicht aufgezeichnet werden kann. Das B-Format arbeitet daher mit segmentierter Aufzeichnung. Jede Spur beinhaltet ca. 50 Zeilen, damit ist ein Halbbild in sechs Segmente aufgeteilt. Um die erforderliche Relativgeschwindigkeit zu erreichen, rotiert das Kopfrad mit 150 U/s. Abbildung 8.64 zeigt die Spurlage des B-Formates. Neben den Videospuren befinden sich an der oberen Bandkante zwei Audiospuren und dazwischen die Kontrollspur. An der unteren Bandkante liegt die Timecode-Spur. Tabelle 8.2 gibt eine Übersicht über die Spurbreiten und weitere mechanische Spezifikationen. Die für professionelle Systeme geforderte Signalstabilität wird mit einem TBC erreicht, der auch bei der Standbild- und Zeitlupenwiedergabe benutzt wird. Die Realisierung dieser Funktionen ist beim B-Format aufgrund der segmentierten Aufzeichnung recht aufwendig. Die Bildteile müssen digitalisiert und in einem Gesamtbildspeicher zusammengesetzt werden, aus dem dann das Signal in verschiedenen Geschwindigkeiten ausgelesen wird. Um die dafür erforderlichen Bildteile in den Speicher einzulesen, ist eine aufwendige mechanische Steuerung nötig, die ein schnelles Hin- und Zurückspringen an bestimmte Bandpositionen ermöglicht, woraus das typische Rattern bei dieser Betriebsart resultiert.

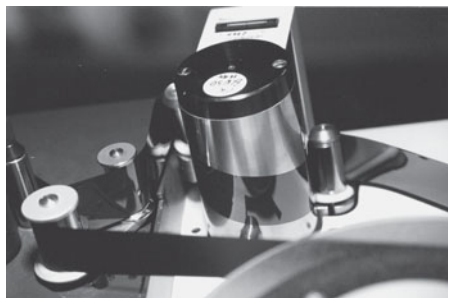


Abb. 8.65. Kopftrommel eines Recorders im B-Format

8.6 Colour-Under-Aufzeichnung

Parallel zur Entwicklung der professionellen MAZ-Geräte mit Schrägspuraufzeichnung wurden auch preiswerte Schrägspur-Geräte für den Massenmarkt entwickelt. Wichtig waren hier vor allem geringe Kosten und einfache Handhabung der Geräte, die durch die Einführung von Cassetten erreicht wurde. Die Kostenreduktion gegenüber professionellen Geräten bezieht sich sowohl auf die Vereinfachung der Geräte selbst, als auch auf die Reduzierung des Bandverbrauchs. Der letzte Punkt erfordert eine Verminderung der Band- und der Band-Kopf-Relativgeschwindigkeit, woraus eine Verminderung der nutzbaren Bandbreite resultiert. Die Bandgeschwindigkeit wurde zunächst auf 9,5 cm/s (U-Matic, 1968), und später, mit der Verfügbarkeit hochwertigerer Magnetbänder, über VHS (1976), mit ca. 2,3 cm/s, bis hinab zu 2 cm/s (Video 8, 1984) gesenkt. Die Kopf-Band-Relativgeschwindigkeit beträgt zwischen 8,5 m/s (U-Matic) und 3,1 m/s (Video 8). Aufgrund dieser Parameter kann bei Standardsystemen eine Bandbreite von nur etwa 3 MHz genutzt werden.

Die verringerte Bandbreite begrenzt die erreichbare Auflösung der Helligkeitsinformation auf ca. 250 Linien, was für ein Heimformat hingenommen werden kann, sie macht aber die Direkt-Aufzeichnung des Chromasignals, das am oberen Ende des Frequenzspektrums liegt, völlig unmöglich. Die Verarbeitung des Farbsignals geschieht daher mit dem Colour Under-Verfahren, bei dem das Farbsignal vom oberen Ende des Frequenzspektrums an das untere verlegt wird. Am unteren Ende des Spektrums ergibt sich aufgrund der Frequenzmodulation des Luminanzsignals eine Lücke. Das FM-Spektrum reicht nur bis zu ca. 1,5 MHz hinab, sodass der Bereich zwischen 0 und 2 MHz für ein in den unteren Frequenzbereich versetztes Chromasignal genutzt werden kann (Abb. 8.66). Das Luminanz- und das konvertierte Chrominanzsignal werden auf diese Weise im Frequenzmultiplex aufgezeichnet, ohne sich zu überlagern.

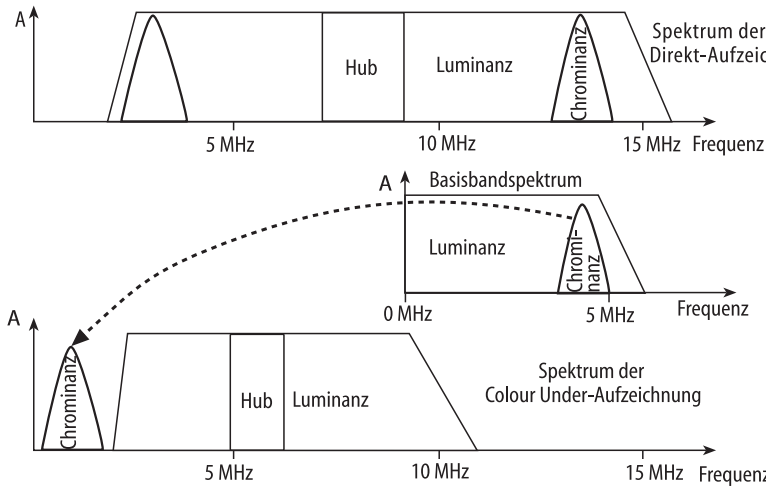


Abb. 8.66. Vergleich der Spektren von FBAS- und Colour-Under-Aufzeichnung

8.6.1 Signalverarbeitung

Das Blockschaltbild in Abb. 8.67 zeigt die Signalverarbeitung für die Aufnahme beim Colour-Under-Verfahren. Das Luminanzsignal durchläuft die in Abb. 8.67 oben abgebildeten Stufen. Für das Chromasignal existiert ein zweiter Signalweg, der im unteren Teil der Abbildung dargestellt ist. Die wesentliche Einheit ist hier die Mischstufe zur Herabsetzung des Farbsignals. Hier wird der Farbträger (bei PAL 4,43 MHz) je nach Format auf eine Frequenz zwischen ca. 500 kHz und 1 MHz herabgesetzt. Das geschieht durch Mischung mit einer Trägerfrequenz derart, dass sich durch die Differenz die gewünschte, herabgesetzte Frequenz ergibt. Die Mischung erfolgt nach folgendem Schema: Chroma-Frequenz f_{SC} (z. B. 4,43 MHz), gemischt mit einer Trägerfrequenz f_T (z. B. 5,06 MHz für VHS), ergibt $f_T + f_{SC} = 9{,}49$ MHz und $f_T - f_{SC} = f_{CU} = 0{,}63$ MHz. Letztere ist die gewünschte herabgesetzte Colour-Under-Frequenz.

Das herabgesetzte Chromasignal bleibt unverändert ein QAM-Signal, trägt also weiterhin die Information über den Farbton in der Phase und über die Farbsättigung in der Amplitude. Da der Farbträger bei ca. 700 kHz liegt, steht für die Seitenbänder nur eine Bandbreite von ca. ± 600 kHz zur Verfügung. Der herabgesetzte Farbträger wird mit dem frequenzmodulierten Luminanzsignal gemischt. Das hochfrequente FM-Signal wirkt dabei wie bei der Audio-Aufzeichnung als HF-Vormagnetisierung, die das Chromasignal in den linearen Teil der Kennlinie verschiebt. Voraussetzung für eine gute Funktion ist eine optimale Einstellung der Pegelverhältnisse für das Chroma- und das FM-Signal. Hier ist es wichtig, die Frequenz des herabgesetzten Chroma-Trägers in einer AFC-Schaltung (Automatic Frequency Control) fest mit der Zeilenfrequenz f_H zu verkoppeln, damit die Vormagnetisierung durch das FM-Signal möglichst wenige Störungen hervorruft. Mit der Zeilenzahl n gilt für das PAL-System: $f_{CU} = (n + 1/8) \cdot f_H$. Beim Format VHS folgt daraus mit n = 40 die Colour Under-Frequenz: $f_{CU} = 626{,}953$ kHz, bei U-Matic ergibt sich mit n = 59: $f_{CU} = 923{,}828$ kHz, und für Video 8 folgt mit n = 46,7: $f_{CU} = 732{,}421$ kHz.

Da die Colour Under-Systeme mit sehr eng nebeneinander liegenden Spuren oder ganz ohne Spurzwischenraum arbeiten, ergibt sich bei der Abtastung

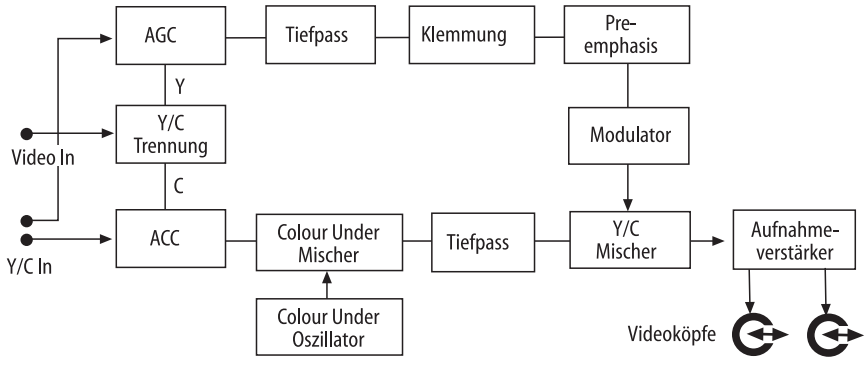


Abb. 8.67. Colour-Under-Signalverarbeitung bei der Aufnahme

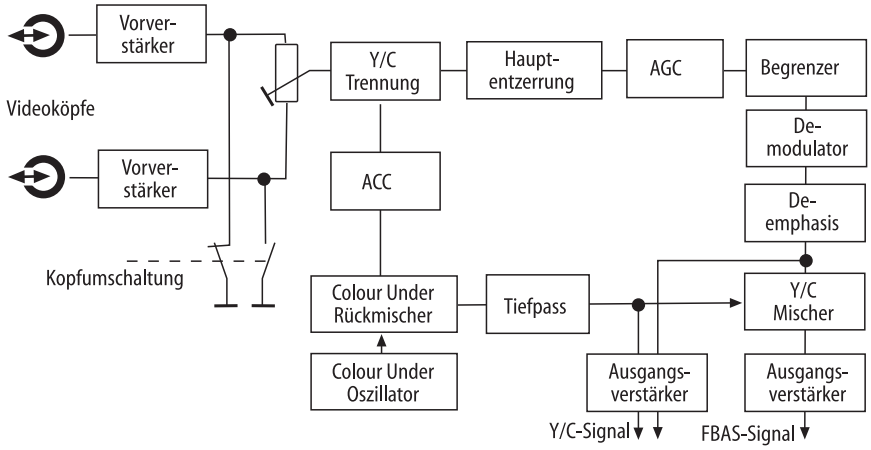


Abb. 8.68. Colour-Under-Signalverarbeitung bei der Wiedergabe

das Übersprechproblem mit den Signalen, die auf benachbarten Spuren aufgezeichnet werden. Für das Chroma-Signal wird daher mit der Phasenfortschaltung (Colour Under Phase Shift) dafür gesorgt, dass die Phasenlage in jeder Zeile um 90° gedreht wird, damit das Nutz- und das Nachbarspur-Störsignal immer eine entgegengesetzte Phasenlage aufweisen.

Bei der Signalwiedergabe hat die Chroma-Signalverarbeitung die Aufgabe, die Phasenfortschaltung zurückzunehmen und das konvertierte Farbsignal in die Ausgangslage zurück zu transformieren (Abb. 8.68). Die Aufwärtskonvertierung wird wie im Aufnahmeweg durch die Mischung mit einer Hilfsfrequenz erreicht, die um 4,43 MHz höher liegt als die herabkonvertierte Farbträgerfrequenz, sodass sich durch die Subtraktion die PAL-Farbträgerfrequenz ergibt. Schließlich gelangt das Chromasignal über den automatischen Farbabschalter (ACK) zur Ausgangsstufe, und weiter zur Y/C- oder zur FBAS-Ausgangsbuchse.

Die getrennte Signalführung für die Luminanz- und Chrominanzsignale sollte soweit wie möglich aufrechterhalten werden, da jede Zusammenführung der Komponenten in den meisten angeschlossenen Geräten wieder rückgängig gemacht werden muss (z. B. bei der Kopie auf einen zweiten Recorder) und damit qualitätsmindernd wirkt. S-VHS und Hi 8-Systeme sind deshalb meist mit Y/C 443-Anschlussmöglichkeiten versehen. Im Professional-Bereich gibt es auch Bildmischer und Peripheriegeräte, die speziell für diese Signalform ausgelegt sind. Die Bezeichnung YC 443 steht für eine getrennte Signalführung, wobei die Zahl 443 bedeutet, dass das Chroma-Signal mit einem Farbhilfsträger der Frequenz 4,43 MHz übertragen wird. Im Gegensatz dazu gibt es andere Bezeichnung wie YC 686, YC 623 oder YC 924. Hier stehen die Zahlen für die konvertierten Farbträgerfrequenzen, d. h., dass hier das Chromasignal nicht wieder heraufgesetzt wird. Bei einer Überspielung auf einen Recorder des gleichen Formates kann damit die Chroma-Konvertierungsstufe umgangen werden, sodass die Qualität dadurch nicht gemindert wird (Abb. 8.69). Eine Geräteverbindung auf der Ebene der konvertierten Chroma-Signale wird über sogenannte DUB-Anschlüsse realisiert.

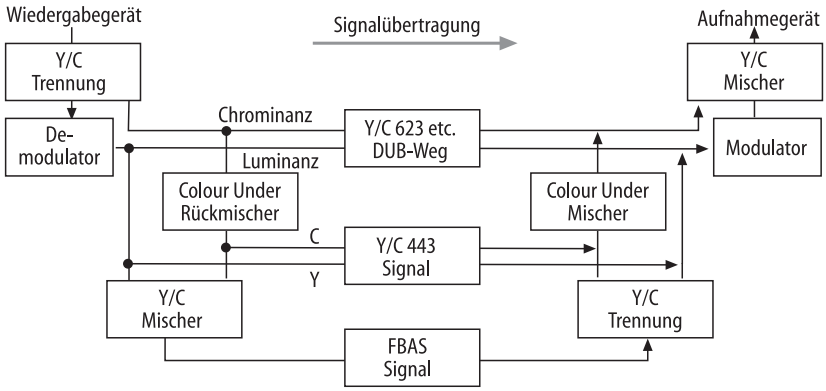


Abb. 8.69. Geräteverbindungen beim Colour Under-Verfahren

8.6.2 Colour-Under-Formate

8.6.2.1 U-Matic

U-Matic ist ein Anfang der 70er Jahre von Sony eingeführtes Format mit Cassetten und 3/4"- (19 mm) breiten Bändern. Es wurde als ein für damalige Verhältnisse preiswertes, leicht zu handhabendes System hoher Aufzeichnungsdichte entwickelt. Die maximale Spieldauer beträgt eine Stunde. Der Name resultiert daraus, dass das automatisch aus der Cassette gezogene Band in U-förmiger 180°-Umschlingung um die Kopftrommel gezogen wird (U-Loading). U-Matic wurde im Laufe der Zeit zu U-Matic High Band und U-Matic SP (Superior Performance) weiterentwickelt und konnte sich als Standardformat im Professional-Bereich, also für Industriefilm, Präsentationen etc. durchsetzen. Das Format ist sehr robust und ermöglicht die Herstellung portabler Geräte. U-Matic High Band war das erste Cassettenformat, das den Ansprüchen der Fernsehanstalten an ein Format zur elektronischen Berichterstattung (EB) genügte. Bis zur Einführung von Betacam und Betacam SP war U-Matic in diesem Sektor sehr verbreitet, daher existieren die Bänder noch in Archiven.

Abbildung 8.70 zeigt die Spurlage. Die Bandlauf- und die Kopfgeschwindigkeit ist gegenüber den 1"-MAZ-Maschinen etwa halbiert. Die Bandgeschwindigkeit beträgt 9,5 cm/s, die Schreibgeschwindigkeit 8,54 m/s. Auf der Spurlän-

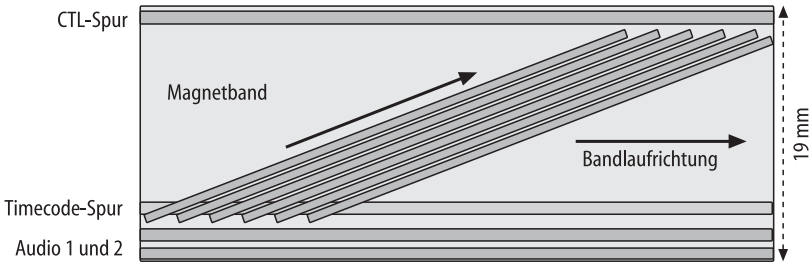


Abb. 8.70. Spurlage bei U-Matic

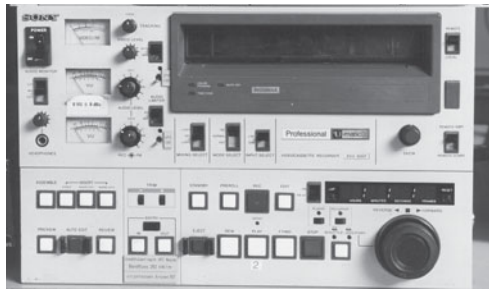


Abb. 8.71. U-Matic-Recorder

ge von 17,1 cm wird je ein Halbbild von einem der beiden um 180° versetzten Köpfe geschrieben. An der oberen Bandkante befindet sich die Steuerspur, an der unteren Kante liegen zwei Audio-Längsspuren, die aufgrund der recht hohen Bandgeschwindigkeit eine gute Audioqualität mit einer Audio-Grenzfrequenz von 15 kHz ermöglichen. Abbildung 8.71 zeigt einen U-Matic-Recorder.

U-Matic High Band und U-Matic SP sind auch deshalb professionell nutzbar, weil sie die Möglichkeit bieten, Timecode auf einer eigenen Spur aufzuzeichnen, sodass also keine Tonspur geopfert werden muss. Da aus Kompatibilitätsgründen das Spurschema von U-Matic Low Band übernommen wurde, fand die Timecode-Längsspur (LTC-Spur) nur noch im Bereich der Videospuren Platz (Abb. 8.70). Das niederfrequente Timecode-Signal kann durch Filter vom Videosignal getrennt werden und liegt in dem Videospur-Bereich, in dem die vertikalen Synchronsignale aufgezeichnet sind. Die Spuranordnung hat den Nachteil, dass der LTC nur bei der ersten Bespielung zusammen mit dem Videosignal aufgezeichnet werden kann.

8.6.2.2 Betamax

Das Betamax-Format der Firma Sony wurde 1975 vorgestellt und war das erste Recorder-Format, das einen wirklichen Massenmarkt erschließen konnte. Kurz darauf erschien das qualitativ vergleichbare Konkurrenzprodukt VHS von JVC. Aufgrund der besseren Marktstrategie fand nach einiger Zeit VHS die größere Verbreitung, sodass Betamax keine große Marktbedeutung erlangen konnte. Die Aufzeichnung geschieht nach dem Colour Under Verfahren mit zwei um

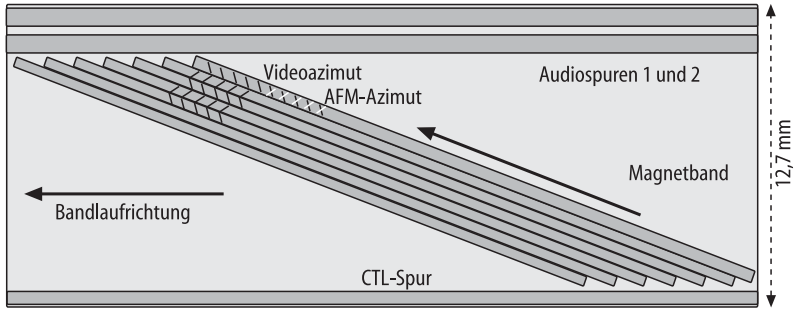


Abb. 8.72. Spurlage bei Betamax und VHS

Tabelle 8.4. Signalparameter der Formate Betamax und VHS

	Betamax	VHS
Y-Bandbreite (–3 dB)	3 MHz	3 MHz
Colour Under Träger	686/689 kHz	627 kHz
Signal/Rauschabstand	> 42 dB	> 42 dB
FM-Träger (Blanking)	4,2 MHz	4,1 MHz
FM-Hub	1,4 MHz	1 MHz
Audiobandbreite (–3 dB)	50 Hz–12 kHz	50 Hz–12 kHz
FM-Audio (Option)	20 Hz–20 kHz	20 Hz–20 kHz

180° versetzten Köpfen auf Cassetten (mit etwas kleineren Abmessungen als bei VHS), die ein Band von 1/2"-Breite enthalten. Die maximale Spieldauer beträgt 180 Minuten. Jede Spur enthält ein Halbbild, zur Übersprechdämpfung wird mit einem Azimut von ± 7° gearbeitet. Tabelle 8.4 zeigt die Signal-Spezifikationen und Abb. 8.72 das Spurlagenschema, das dem von VHS gleicht. Die beiden Audiospuren sind bei einfachen Geräten auch zu einer Monospur zusammengefasst, eine Time Code Spur existiert nicht. Die Luminanz-Bandbreite ist bei Betamax auf 3 MHz begrenzt, der Luminanzsignal/Rauschabstand beträgt 42 dB. Die Konvertierung des Chroma-Signals geschieht mit zwei verschiedenen Frequenzen, sodass der herabgesetzte Farbträger für jede Spur abwechselnd bei 689,45 kHz und 685,55 kHz liegt. Tabelle 8.5 zeigt die mechanischen Spezifikationen. Das Format wurde zu Betamax ED ausgebaut, die Auflösung konnte dabei so gesteigert werden, dass es semiprofessionellen Ansprüchen genügt.

8.6.2.3 Video Home System, VHS

Das von der Firma JVC entwickelte Format VHS wurde 1976 vorgestellt. Es konnte sich gegen das Konkurrenzprodukt Betamax, sowie auch gegen die europäischen Konkurrenzformate VCR und Video 2000 durchsetzen. Es war dann das mit Abstand bedeutendste analoge Heimanwenderformat und als Massenprodukt preiswerter Distributionsstandard von Industrie- und Schulungsfilmen. VHS-Geräte wurden von sehr vielen Herstellern gefertigt.

Das VHS-Format ist Betamax sehr ähnlich. Auch hier findet mit zwei Videoköpfen eine Colour-Under-Aufzeichnung auf Cassetten mit 1/2"-Band statt.

Tabelle 8.5. Geometrische Parameter der Formate Betamax und VHS

	Betamax	VHS
Magnetbandbreite	12,7 mm	12,7 mm
Bandgeschwindigkeit	1,9 cm/s	2,3 cm/s
Relativgeschwindigkeit	5,8 m/s	4,9 m/s
Videospurlänge	117 mm	97 mm
Videospurbreite	33 µm	49 µm
Spurzwischenraum	0	0
Spurwinkel	5°	6°
Azimut	± 7°	± 6°
Kopftrommeldurchmesser	74 mm	62 mm (41 mm (C))
max. Spieldauer	180 Min.	300 Min.

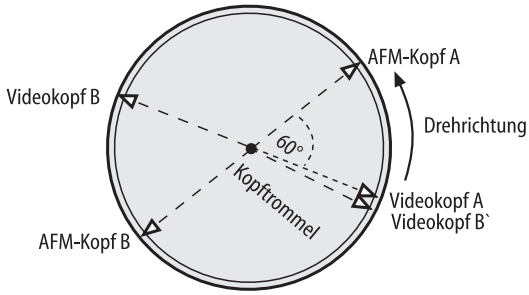


Abb. 8.73. Kopftrommel des VHS-Recorders (Dreikopf mit AFM-Ton)

Die Cassette hat die Abmessungen 188 mm x 104 mm x 25 mm. Die Luminanz-Auflösung beträgt wie bei Betamax wegen der auf 3 MHz beschränkten Bandbreite 240 Linien. Die Unterschiede zwischen Betamax und VHS werden in den in Tabelle 8.4 dargestellten Signal-Spezifikationen und den mechanischen Spezifikationen deutlich: VHS arbeitet mit einer etwas kleineren Kopftrommel und die Bandgeschwindigkeit ist mit 2,34 cm/s etwas größer als bei Betamax (Tabelle 8.5). Das Magnetband wird bei VHS meist im M-Loading-Verfahren um die Kopftrommel geführt. Abbildung 8.73 zeigt die Kopftrommel und Abb. 8.72 die Spurlage beim VHS-Format. Bei Standard-Recordern gibt es neben der Kontrollspur nur eine Audio-Längsspur, bei neueren Geräten können zwei Spuren genutzt werden, indem die Monospur in zwei Teile geteilt wird. Die Audiobandbreite ist wegen der geringen Bandgeschwindigkeit auf ca. 12 kHz beschränkt. Eine hochwertige Audioaufzeichnung zweier weiterer Audiospuren ist optional mit AFM möglich. Bei VHS werden separate AFM-Köpfe benutzt, die AFM-Trägerfrequenzen liegen bei 1,4 MHz für den linken und 1,8 MHz für den rechten Kanal im Bereich des unteren FM-Seitenbandes (Abb. 8.74).

Im VHS-System werden keine steuerbaren Videoköpfe eingesetzt, daher ist die Wiedergabe bei nicht normgerechter Geschwindigkeit gestört. Zur Reduktion der Standbildfehler arbeiten aufwendigere VHS-Geräte daher mit mehreren Wiedergabeköpfen. Ein Dreikopf-Recorder enthält neben den Köpfen 1 und 2 zusätzlich den Kopf 2‘, der ganz in der Nähe von Kopf 1 angeordnet ist, aber einen Azimutwinkel aufweist, der dem von Kopf 2 entspricht (Abb. 8.73). Bei der Standbildwiedergabe wird das Band so gesteuert, dass mit den Köpfen 2 und 2‘ dieselbe Spur abgetastet wird (Field Still).

Das VHS-System wurde sehr vielfältig weiterentwickelt. Zunächst wurde eine Long-Play-Funktion eingeführt, mit der die maximale Spieldauer verdoppelt werden kann. Die Bezeichnung VHS-HQ (High Quality) steht für eine weiterentwickelte Signalverarbeitung auf der Basis verbesserter Bandmaterialien.

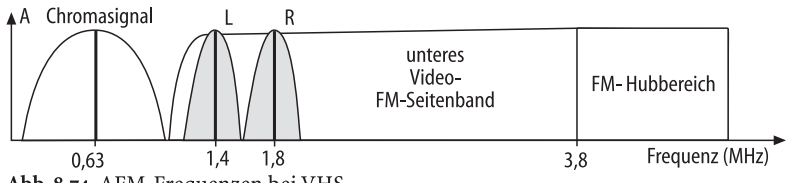


Abb. 8.74. AFM-Frequenzen bei VHS

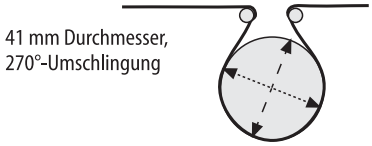


Abb. 8.75. Kleine Kopftrommel bei VHS

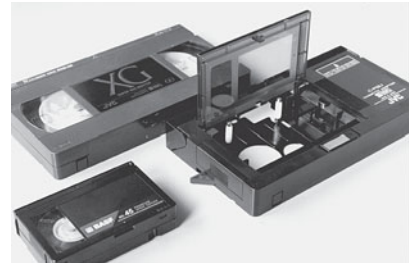


Abb. 8.76. VHS-C-Cassette mit Adapter

Im Zuge der Entwicklung kleiner portabler Camcorder und als Konkurrenz zu Video 8 wurde das VHS-System durch die Einführung von verkleinerten Compact-Cassetten (VHS-C) erweitert. Die maximale Spieldauer einer VHS-C-Cassette beträgt 45 Minuten. Das VHS-C-Format ist kompatibel zu VHS, die Kompatibilität ist aber evtl. eingeschränkt, da in den portablen Geräten häufig kleinere Kopftrommeln mit nur 41 mm Durchmesser eingesetzt werden, die eine Erhöhung des Kopftrommelumschlingungswinkels und den Einsatz von vier Köpfen erfordern (Abb. 8.75). VHS-C-Cassetten können mit einem Adapter in Standgeräten abgespielt werden (Abb. 8.76).

Eine wesentliche qualitative Weiterentwicklung erfuhr das VHS-Format durch die Einführung von Super-VHS. Mit Hilfe verbesserter Videobänder und -köpfe konnte die Luminanzbandbreite auf mehr als 4 MHz gesteigert werden. Damit sind Kopien über die zweite Generation hinaus möglich, die bei VHS qualitativ praktisch nicht vertretbar sind. S-VHS ist abwärtskompatibel zu VHS, die mechanischen Parameter bleiben unverändert. S-VHS-Bänder können mit VHS-Geräten nicht wiedergegeben werden. Die Veränderungen betreffen vor allem eine Vergrößerung des FM-Hubes von 1 MHz auf 1,6 MHz sowie eine Verlagerung des Gesamtspektrums zu höheren Frequenzen hin. Die unterste FM-Frequenz liegt nun bei 5,4 MHz.. Die Chrominanz-Signalverarbeitung bleibt im Wesentlichen unverändert, denn die Colour Under-Frequenz liegt aus Kompatibilitätsgründen weiterhin bei 627 kHz (Abb. 8.77).

Durch den Übergang zu S-VHS konnte die Video-Qualität des Formates so gesteigert werden, dass es semiprofessionellen Ansprüchen genügt. Es wurden

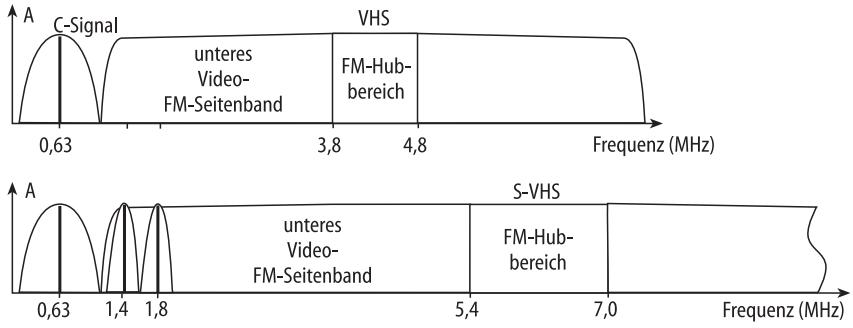


Abb. 8.77. Erweitertes FM-Spektrum bei S-VHS

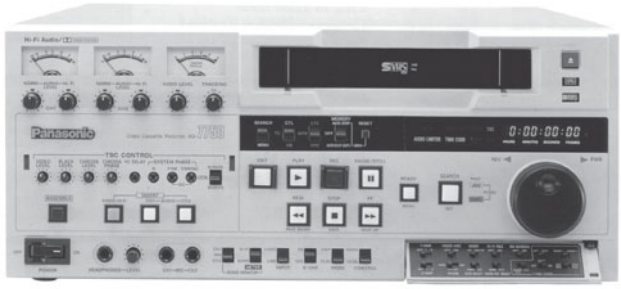


Abb. 8.78. S-VHS-Professional Recorder [164]

mit der Professional-S-Reihe besonders stabil gefertigte Geräte angeboten, bei denen ein hoher Schaltungsaufwand betrieben wurde, und die sich auch in professionelle Schnittsysteme einbinden lassen (Abb. 8.78). Ein Problem bleibt die schlechte Farbauflösung. Noch schwerwiegender für den professionellen Einsatz ist die problematische Audiosignalqualität. Hochwertige Qualität kann nur mit den AFM-Spuren erzielt werden. Diese sind aber nicht unabhängig vom Bild editierbar, sodass bei der Nachbearbeitung immer die Längsspuren benutzt werden müssen. Schließlich ist auch die für die professionelle Nachbearbeitung unerlässliche Anwendung des Timecode-Verfahrens eingeschränkt, da dafür eine Tonspur geopfert werden muss.

Unter der Bezeichnung W-VHS gibt es eine zu VHS bzw. S-VHS kompatible Weiterentwicklung zur Aufzeichnung von HDTV-Signalen. W-VHS gehört nicht mehr zu den Colour Under-Verfahren, denn die Farbdifferenzkomponenten werden im Zeitmultiplex verarbeitet. Abbildung 8.79 zeigt das Arbeitsprinzip von W-VHS. Die VHS-Spurbreite ist gedrittelt, zwei Teilspuren dienen der analogen Videoaufzeichnung, auf der dritten werden im HD-Modus zwei Audiokanäle digital aufgezeichnet. Mit der auf 3,33 cm/s erhöhten Bandgeschwindigkeit können ein 1125/60-HD-Signal oder wahlweise zwei Standard NTSC-Signale mit hoher Qualität aufgezeichnet werden. Das Format lässt sich somit gut für die Aufzeichnungen von 3D- oder Widescreen Bildern mit doppelter Bildbreite einsetzen. Zur Audioaufzeichnung stehen weiterhin zwei AFM-Kanäle zur Verfügung. Die Erweiterungen des VHS-Formates für die digitale Aufzeichnung werden mit D-VHS und Digital S bezeichnet. Näheres dazu findet sich in Abschnitt 8.8.

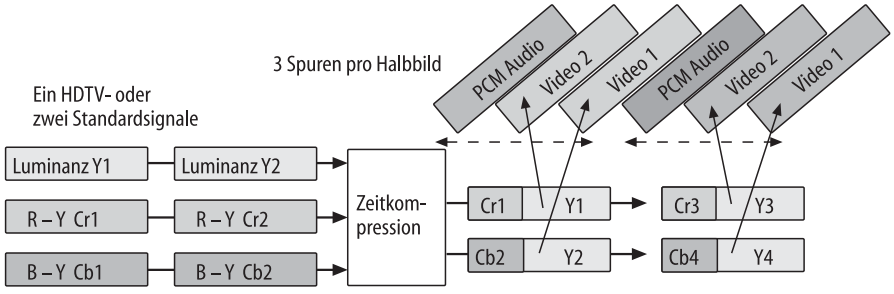


Abb. 8.79. Signalverarbeitung bei W-VHS

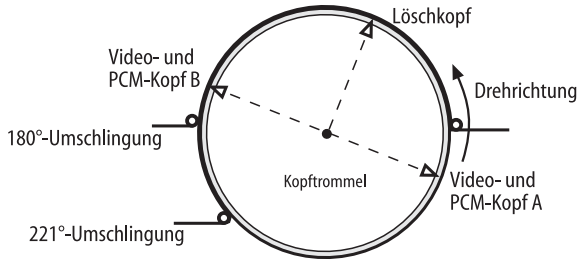


Abb. 8.80. Video 8 Kopftrommel

8.6.2.4 Video 8

Um Videorecorder mit einer Kamera zu einem kompakten Camcorder für den Massenmarkt zusammenfassen zu können, entwickelten eine Vielzahl von Herstellern das Format Video 8, das 1984 vorgestellt wurde. Die Bezeichnung rührt vom 8 mm breiten Bandmaterial her, das die Verwendung sehr kleiner Cassetten mit den Abmessungen 95 mm x 62,5 mm x 15 mm ermöglicht. Video 8 war das dominante analoge Camcorder-Format im Amateurbereich. Die maximale Spieldauer einer Video 8-Cassette beträgt 90 Minuten.

Die Aufnahme erfolgt wie bei VHS mit zwei auf der Kopftrommel gegenüberliegenden Videoköpfen mit der Slanted Azimut-Technik zur Übersprechdämpfung. Die Kopftrommel (Abb. 8.80) hat standardmäßig einen Durchmesser von 40 mm. Die Tabellen 8.6 und 8.7 zeigen die mechanischen und die Signalspezifikationen. Die Bandgeschwindigkeit beträgt 2 cm/s (SP). Bei dieser Geschwindigkeit ergibt sich ein Sicherheitsabstand (Rasen) zwischen den Videospuren. Bei der im Long-Play-Modus (LP) halbierten Bandgeschwindigkeit verschwindet dieser und die Spuren werden überlappend aufgezeichnet. Bei

Tabelle 8.6. Geometrische Parameter des Formats Video 8

Magnetbandbreite	8 mm
Bandgeschwindigkeit	2 cm/s
Relativgeschwindigkeit	3,1 m/s
Videospurlänge	62 mm
Videospurbreite	27 µm
Spurzwischenraum	7 µm
Spurwinkel	5°
Azimut	± 10°
Kopftrommeldurchmesser	40 mm

Tabelle 8.7. Signalparameter der Formate Video 8 und Hi 8

	Video 8	Hi 8
Y-Bandbreite (–3 dB)	3 MHz	4 MHz
C-Bandbreite (–6 dB)	1 MHz	1 MHz
Colour Under Träger	732 kHz	732 kHz
Signal/Rauschabstand	> 44 dB	> 45 dB
FM-Träger (Blanking)	4,6 MHz	6,3 MHz
FM-Hub	1,2 MHz	2 MHz
Audiobandbreite (AFM)	30 Hz–15 kHz	30 Hz–15 kHz

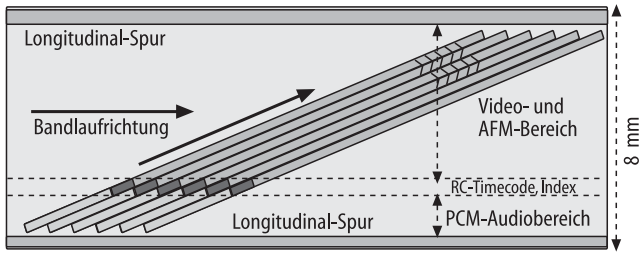


Abb. 8.81. Spurlage bei Video 8

der Spuranordnung zeigen sich wesentliche Unterschiede zu VHS. Abbildung 8.81 zeigt nur eine Cue-Hilfsspur an der oberen und eine Audio-Hilfsspur an der unteren Bandkante, eine Kontrollspur ist nicht vorhanden. Der Schrägspurbereich enthält dagegen neben den Videosignalen auch Gebiete für Audio- und Timecode-Daten. Der Unterschied zu VHS zeigt sich auch im Signalspektrum (Abb. 8.82). Zwischen den Y- und C-Frequenzbereichen befinden sich das Audiosignalspektrum (AFM), und darunter liegen die Bandsteuersignale (ATF s. u.). Das Audiosignal wird standardmäßig frequenzmoduliert zusammen mit dem Videosignal in der Schrägspur aufgezeichnet (Abb. 8.81).

Bei der Definition von AFM als Audio-Standardaufzeichnung entsteht das Problem, dass das Tonsignal nicht unabhängig vom Bildsignal bearbeitet werden kann. Aus diesem Grund ist optional mit der PCM-Aufzeichnung eine zweite Audioaufzeichnungsmöglichkeit vorgesehen. Der für die PCM-Aufzeichnung reservierte Bereich auf dem Band wird durch eine Erhöhung der Kopftrommelumschlingung von 180° auf 221° erreicht. Zwischen den Video- und PCM-Bereichen befindet sich ein Zwischenraum, der für die Aufzeichnung eines Rewritable Consumer Time Code (RCTC) vorgesehen ist.

Eine weitere Besonderheit von Video 8 ist der Verzicht auf eine CTL-Spur. Statt des Kontrollsignals von der Längsspur kommen bei Video 8 die

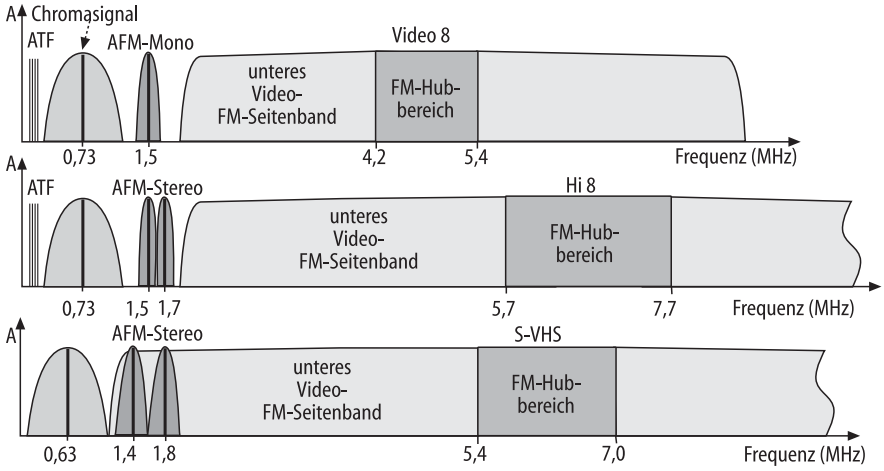


Abb. 8.82. Vergleich der Signalspektren von Video 8, Hi 8 und S-VHS

Abb. 8.83. Hi 8-Professional Recorder [208]

Steuerinformationen aus der Auswertung von vier ATF-Frequenzen (Automatic Track Following), die unterhalb des Chrominanzbereichs bei liegen. In jeder Spur wird eine dieser Frequenzen mit geringem Pegel aufgezeichnet, sie wiederholen sich zyklisch. Die Information über die Genauigkeit der Spurabtastung wird dann gewonnen, indem die Pilotsignale der Hauptspur und der mit abgetasteten Nachbarspuren überlagert werden. So entstehen für jede Kombination Differenzfrequenzen. Durch Auswertung der Intensitäten wird die Information für die Capstan-Regelung gewonnen. Wenn beide Intensitäten gleich sind, befindet sich der Kopf genau auf der Mitte der Spur.

Video 8 wurde durch Erweiterung des Luminanzfrequenzbereichs zu Hi 8 verbessert. Hi 8 ist abwärtskompatibel zu Video 8. Die mechanischen Spezifikationen bleiben unverändert, ebenso wie die Colour-Under-Trägerfrequenz. Die Signalspezifikationen für das Y-Signal (Tabelle 8.7) sind noch besser als bei S-VHS, denn der Hub beträgt 2 MHz und der Synchronboden liegt bei 5,7 MHz, sodass eine Luminanzsignal-Bandbreite von mehr als 4 MHz genutzt werden kann (Abb. 8.82). Hi 8 war das verbreitetste analoge Camcorder-Format für anspruchsvolle Amateuraufnahmen. Auch hier stehen Hi 8-Professional-Geräte zur Verfügung, die in Schnittplätze eingebunden werden können (Abb. 8.83).

8.6.2.5 Video 2000 und VCR

Die in Europa entwickelten Colour-Under-Systeme VCR und Video 2000 seien hier nur kurz erwähnt, sie haben keine große Verbreitung gefunden. Beide Formate arbeiten mit Cassetten und 1/2"-Band. Bei VCR liegen die Bandwickel übereinander. Video 2000 weist die Besonderheit auf, dass die Cassette gewendet und zweiseitig benutzt werden kann. Auf jeder Bandhälfte liegen außen die Audiospuren und in der Mitte die Kontrollspuren. Die Videospurbreite beträgt nur 23 μm (Abb 8.84). Zur Kopfnachführung wird das DTF-Verfahren angewandt, das als Vorläufer des Video 8-ATF-Systems betrachtet werden kann.

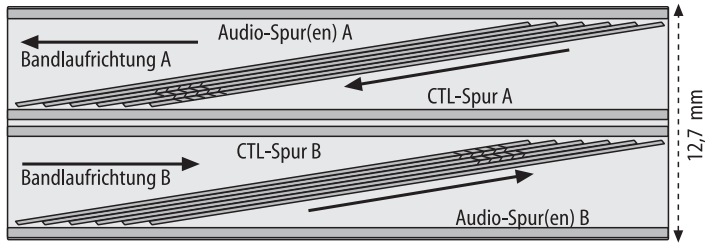


Abb. 8.84. Spurlage bei Video 2000

8.7 Komponentenaufzeichnung

Zu Beginn der 80er Jahre wurde versucht, die Qualität professioneller MAZ-Geräte unter Beibehaltung der Kompaktheit der Amateurgeräte zu erreichen. Die Weiterentwicklung beruhte bei Sony auf dem Betamax-Recorder und wird mit Betacam bezeichnet, bei Panasonic entstand das M-Verfahren auf der Basis des VHS-Recorders. Bei beiden Formaten wird eine Cassette mit 1/2"-breitem Band verwendet. Die Formate sollten qualitativ in eine ganz andere Klasse vorstoßen als die Amateurformate, sodass Betacam- und M-Geräte trotz mechanischer Ähnlichkeiten sowohl elektrisch als auch mechanisch wesentlich aufwendiger und teurer gefertigt werden als Betamax und VHS-Recorder.

Das M-Verfahren wird auch als Chromatrack-M-Verfahren bezeichnet. Damit wird der wesentliche Schritt zur verbesserten Chromasignalverarbeitung deutlich: für das Chrominanzsignal steht eine eigene Spur zu Verfügung. Eine weitere qualitätssteigernde Maßnahme war die Erhöhung der Bandgeschwindigkeit um etwa den Faktor 3 bzw. 5 gegenüber den Amateurformaten, womit bei den zu Beginn der 80er Jahre zur Verfügung stehenden Magnetbändern eine Grenzfrequenz von 4 MHz für das Luminanzsignal erreicht werden konnte. Ein drittes Aufzeichnungsverfahren mit getrennten Spuren für Y und C beruht auf 1/4"-Cassetten und wurde unter der Bezeichnung QuarterCam entwickelt. Das Verfahren ist bedeutungslos, es konnte nie Marktrelevanz erreichen.

Chromatrack-M und Betacam konnten sich dagegen im professionellen Bereich etablieren, als sie Ende der 80er Jahre unter Nutzung hochwertiger Magnetbänder und erweiterter FM-Frequenzbereiche zu MII und Betacam SP ausgebaut wurden. Bei Betacam SP und MII wird eine Grenzfrequenz von 5,5 MHz für das Y-Signal erreicht. Ein sichtbarer Qualitätsverlust ist erst bei Kopien ab der fünften Generation zu bemerken. Das 1988 vorgestellte Format Betacam SP setzte sich als das weltweit dominierende professionelle analoge MAZ-Format durch. MII ist qualitativ gleichwertig, hat aber trotz eines erheblich geringeren Verkaufspreises einen wesentlich kleineren Marktanteil (ca. 10% relativ zu Betacam SP).

Der wesentliche Vorteil von Betacam SP und MII ist dadurch gegeben, dass die Signalkomponenten Y, C_R und C_B separat aufgezeichnet werden und auch separat ausgegeben werden können. Daraus resultiert die Bezeichnung Komponentenaufzeichnung.

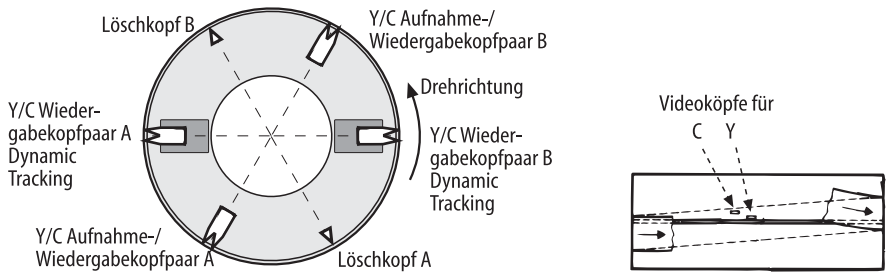


Abb. 8.85. Betacam SP-Kopftrommel in Auf- und Seitenansicht

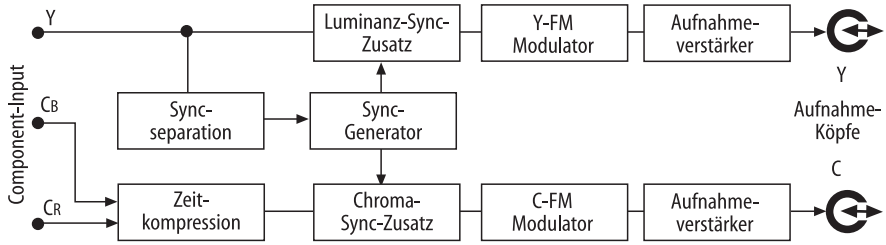


Abb. 8.86. Signalverarbeitung bei der Komponentenaufzeichnung

MII und Betacam SP-Geräte unterstützen die Timecode-Varianten LTC und VITC (s. Kap. 9). Sie sind mit hochwertigen TBC und mit Fernsteueranschlüssen nach RS-422 ausgerüstet. Beide Formate genügen damit allen Anforderungen eines professionellen Schnittbetriebs (sehr häufiges Umspulen und exaktes Auffinden bestimmter Bildsequenzen im Dauerbetrieb). Das Komponentensignal hielt mit dem Betacam SP-System schnell Einzug in den professionellen Bereich. Es wurden Kreuzschienen, Mischer etc. mit entsprechenden Anschlüssen entwickelt, und in den 90er Jahren wurden viele Studios für die Komponentensignalbearbeitung umgebaut, was vielfach mit einer Verdreifachung des Verkabelungsaufwands einherging.

8.7.1 Signalverarbeitung

Auch bei der Komponentenaufzeichnung geschieht die Verarbeitung des Luminanzsignals im Wesentlichen wie bei der FM-Direktaufzeichnung, das Chrominanzsignal wird jedoch getrennt verarbeitet. Soweit noch nicht vorliegend, müssen durch Filterung und Demodulation zunächst die Farbdifferenzsignale gewonnen werden, die dann mit separaten Chromaköpfen auf einer eigenen Chromaspur aufgezeichnet werden. Die Spur für das Chrominanzsignal ist geringfügig schmaler als die für das Luminanzsignal. Mit zwei um 180° versetzten Kopfpaaren wird statt einer Videospur pro Halbbild ein Spurpaar geschrieben. Die Videoköpfe für Y und C haben dafür einem kleinen Höhenversatz (< 0,1 mm), der bei der Seitenansicht der in Abb. 8.85 dargestellten Betacam-Kopftrommel deutlich wird. Die Abbildung zeigt auch die Bestückung der Trommel mit zwei Löschköpfen, zwei Kopfpaaren für Aufnahme/Wiedergabe und zwei beweglichen Paaren für die Dynamic Tracking-Wiedergabe.

Abbildung 8.86 gibt eine Übersicht über die Signalverarbeitung bei der Komponentenaufzeichnung. Die Chromakomponenten werden ebenso wie das Y-Signal frequenzmoduliert aufgezeichnet, es liegen zwei fast identische Signalwege mit einer Bandbreite von 5,5 MHz und 4 MHz vor. Der wesentliche Unterschied zu anderen Aufzeichnungsverfahren ist der Einsatz der Kompressionsstufe für die Chromakomponenten. Diese ist erforderlich, weil sich die beiden Farbdifferenzkomponenten die Chromaspur teilen. Dabei wird das Zeitmultiplex-Verfahren angewendet, das gegenüber dem Frequenzmultiplex eine bessere Trennung der Komponenten erlaubt.

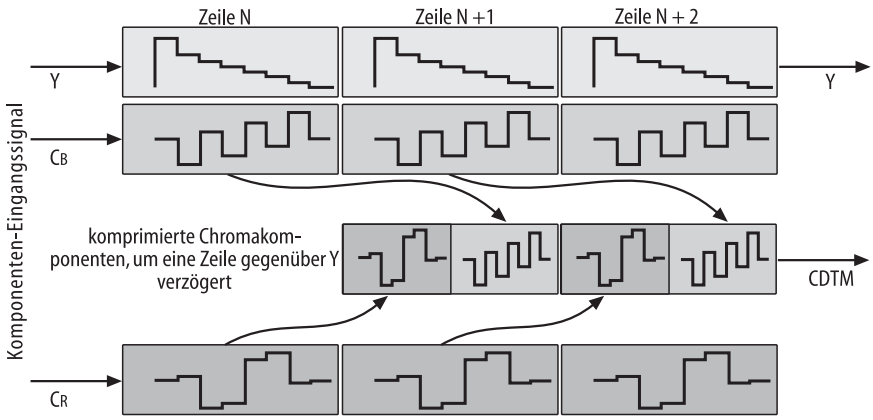


Abb. 8.87. Chroma-Zeitkompression CTDM

Die gleichzeitig vorliegenden Farbdifferenzsignale werden jeweils zeitlich um den Faktor zwei komprimiert und nacheinander aufgezeichnet. Bei der Wiedergabe werden die Signale entsprechend expandiert. Das Verfahren heißt bei Betacam Compressed Time Division Multiplex (CTDM) und bei MII Chroma Time Compressed Multiplex (CTCM) (Abb. 8.87). Da ein Übertragungskanal mit fester Bandbreite von zwei Signalen genutzt wird, können beide nur mit halber Grenzfrequenz aufgezeichnet werden. In der Praxis wird der Chrominanzkanal für eine Bandbreite von 4 MHz ausgelegt. Die Farbdifferenzsignale sind daher auf 2 MHz bandbegrenzt. Das ist ein guter Wert, verglichen mit dem Standard-PAL-Signal, bei dem die Bandbreite des Chrominanzkanals auf 1,3 MHz eingeschränkt ist. Die Zeitkompression kann in einem Schieberegister vorgenommen werden, wobei das Signal in Echtzeit ein- und mit doppelter Geschwindigkeit wieder ausgelesen wird. Bei diesem Vorgang werden die Chroma-Signalanteile gegenüber dem Y-Signal um eine Zeile verzögert, das gleiche passiert bei der zur Wiedergabe nötigen Expansion. Die Gesamtverzögerung von zwei Zeilen wird elektronisch ausgeglichen.

Aufgrund der mechanischen Trennung von Y- und C-Signalwegen kann ein relativer Zeitfehler auch zwischen den Kanälen auftreten. Damit dieser im TBC ausgeglichen werden kann, wird beiden Signalwegen ein speziell generierter Hilfssynchronimpuls zugesetzt (Abb. 8.88). Außerdem wird zur Kennzeichnung

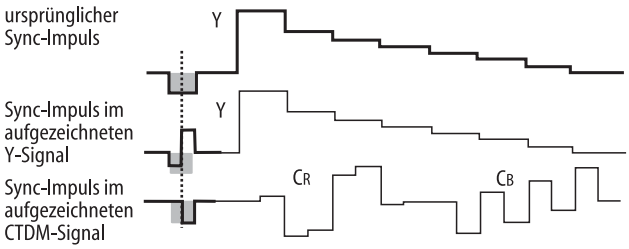


Abb. 8.88. Referenzimpulse bei Betacam SP

Abb. 8.89. Betacam SP-Recorder der BVW-Serie [208]

der PAL-8V-Sequenz bei Betacam SP in jeder 12. und 325. Zeile den Chrominanzkomponenten ein 8V-Ident-Impuls zugesetzt, und in der 8. und 312. Zeile des Y-Signals wird zusätzlich der Vertical Interval Subcarrier (VISC) eingetastet [33]. Bei der Signalverarbeitung zur Wiedergabe eines Komponentensignals werden im Wesentlichen die Luminanz- und Chrominanzkomponenten demoduliert und anschließend eine Zeitbasiskorrektur vorgenommen, wobei sich die Zeitfehlererkennung an den speziellen Synchronsignalen orientiert. Im TBC liegt das Signal in digitaler Form vor, daher kann hier mit wenig Aufwand auch die Zeitkompression rückgängig gemacht werden.

8.7.2 Komponentenaufzeichnungsformate

8.7.2.1 Betacam SP

Das verbreitetste Videoformat mit analoger Komponentenaufzeichnung ist Betacam SP. Abbildung 8.89 zeigt ein Gerät der BVW-Serie. Für das Y-Signal steht eine Bandbreite von 5,5 MHz und für die Farbdifferenzkomponenten je eine Bandbreite von 2 MHz zur Verfügung, Tabelle 8.8 zeigt die Signalspezifikationen. Die Spurlage auf dem 1/2“ breiten Magnetband ist in Abb. 8.90 dargestellt. An der unteren Bandkante befinden sich zwei Longitudinalspuren für CTL- und Timecode-Signale, an der oberen Bandkante zwei Audio-Längsspuren mit 0,6 mm Breite und 0,4 mm Abstand. In der Mitte liegen die Video-Schrägspurpaare, die mit Slanted Azimut und einem zusätzlichen Sicherheitsabstand aufgezeichnet werden. Die mechanischen Spezifikationen finden sich

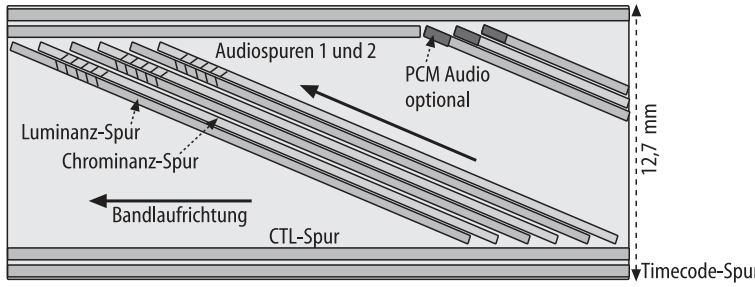


Abb. 8.90. Spurlage bei Betacam SP

Tabelle 8.8. Signalparameter der Formate Betacam SP und M II

	Betacam SP	M II
Y-Bandbreite (–3 dB)	5,5 MHz	5,5 MHz
C-Bandbreite (–6 dB)	2 MHz	2 MHz
Signal/Rauschabstand (Y)	> 48 dB	> 47 dB
FM-Träger (Blanking Y)	7,4 MHz	7,4 MHz
FM-Hub (Y)	2 MHz	2,6 MHz
Audiobandbreite (Lng)	50 Hz–15 kHz	40 Hz–15 kHz
AFM- und PCM-Audio	20 Hz–20 kHz	20 Hz–20 kHz

Tabelle 8.9. Geometrische Parameter der Formate Betacam SP und M II

	Betacam SP	M II
Magnetbandbreite	12,7 mm	12,7 mm
Bandgeschwindigkeit	10,2 cm/s	6,6 cm/s
Relativgeschwindigkeit	5,8 m/s	5,9 m/s
Videospurlänge	115 mm	118 mm
Videospurbreite (Y, C)	86 µm, 73 µm	56 µm, 36 µm
Spurzwischenraum (Y, C)	7 µm, 2 µm	4 µm, 3 µm
Spurwinkel	4,7°	4,3°
Azimut	± 15°	± 15°
Kopftrommeldurchmesser	74,5 mm	76 mm
max. Spieldauer (S, L)	32 Min., 108 Min.	23 Min., 97 Min.

in Tabelle 8.9. Die Magnetbänder befinden sich in Cassetten, es können zwei Cassettengrößen genutzt werden. Die kleine S-Cassette mit den Maßen 188 mm x 106 mm x 30 mm erlaubt eine maximale Spieldauer von 32 Minuten, die große L-Cassette (254 mm x 145 mm x 30 mm) maximal 108 Minuten. Das 1/2“-Format ermöglicht die für den EB-Einsatz nötige Verbindung von Kamera und Recorder zu einem Camcorder. Für diese Cassetten stehen auch computergesteuerte Cassetten-Wechselautomaten (Betacart) zur Verfügung, mit denen unter Nutzung mehrerer Betacam SP-Player ein automatischer Sendeablauf realisiert werden kann.

Die Bandgeschwindigkeit beträgt bei Betacam SP 10,1 cm/s. Damit lässt sich auch auf den Längsspuren ein Audiosignal recht hoher Qualität aufzeichnen,

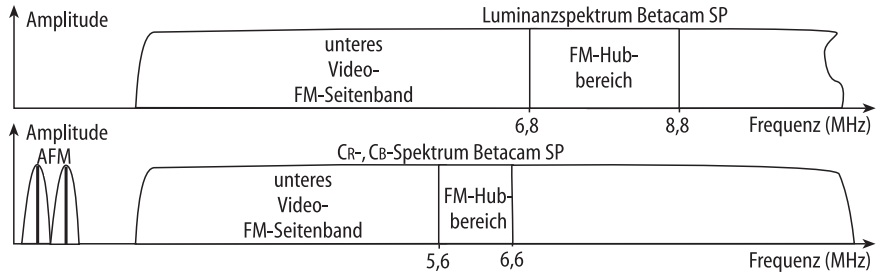


Abb. 8.91. Luminanz- und Chrominanz-Spektren bei Betacam SP

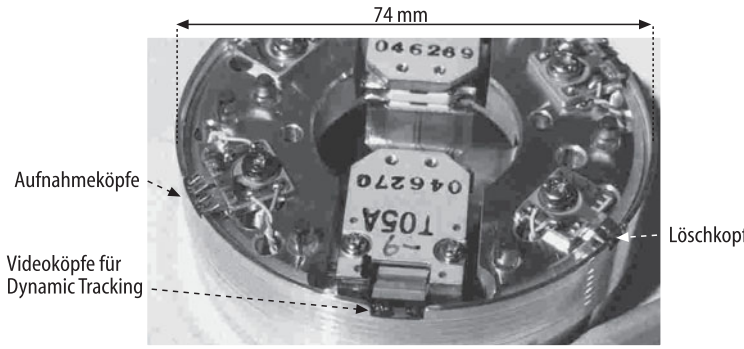


Abb. 8.92. Kopftrommel eines Betacam SP-Recorders

der Audiofrequenzbereich reicht von 50 Hz bis 15 kHz. Abbildung 8.91 zeigt die Spektren von Luminanz- und Chrominanz-Kanälen. Optional kann auch ein zweikanaliges, unabhängig vom Bild bearbeitbares PCM-Audiosignal aufgezeichnet werden, dabei muss allerdings die Audiospur zwei geopfert werden (Abb. 8.90). Die optionale Funktion Dynamic Tracking ermöglicht ein ungestörtes Stand- und Zeitlupenbild. Dynamic Tracking erfordert die Montage von beweglichen Wiedergabeköpfen. Abbildung 8.92 zeigt die entsprechend bestückte Kopftrommel eines Betacam SP-Recorders.

Betacam SP-Geräte sind mit einem digitalen TBC ausgestattet, mit dem das Signal stabilisiert und die Zeitexpansion für die Farbdifferenzsignale vorgenommen wird. Mit dem TBC kann das Ausgangssignal der Maschinen bezüglich des Schwarzwertes (0 ... 100 mV), Weißwertes (± 3 dB) und der Chromaamplitude (± 3 dB) verändert werden.

Außer vom Systementwickler Sony wurden Betacam SP-Geräte auch von Ampex, Thomson und BTS hergestellt. Die Geräte sind äußerlich einander sehr ähnlich. Professionelle Geräte trugen die Typbezeichnung BVW. Bereits zu Beginn der 90er Jahre brachte Sony eine Typenreihe mit der Bezeichnung PVW heraus, mit der Betacam SP auch im semiprofessionellen Bereich etabliert werden sollte. Entsprechende Geräte werden etwa zum halben Preis eines BVW-Gerätes angeboten. Die Ausstattung der PVW-Geräte unterscheidet sich kaum von denen der Broadcast-Serie, sie sind aber einfacher gefertigt und die Signalspezifikationen haben größere Toleranzen (Abb. 8.93).

Abb. 8.93. Betacam SP-Recorder PVW [208]

Abb. 8.94. Betacam SP-Recorder UVW [208]

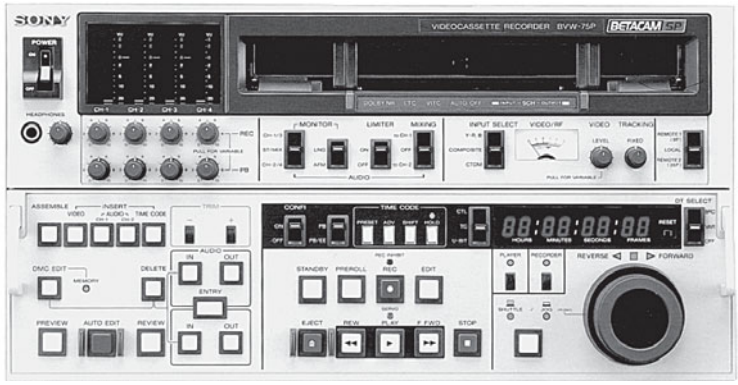


Abb. 8.95. Bedienelemente auf der Frontseite eines Betacam SP-Recorders [208]

Eine nochmalige Vereinfachung der Geräte wurde 1995 mit der UVW-Serie vorgenommen, für die die Preise gegenüber PVW wiederum etwa halbiert sind. Diese Geräte sind äußerlich deutlich anders gestaltet (Abb. 8.94), bieten aber trotzdem alle Standardfunktionen und lassen sich problemlos in professionelle Schnittsysteme einbinden. UVW-Geräte arbeiten jedoch nur mit dem LTC.

Die Geräte der Betacam-Serie werden hier so genau beschrieben, weil sich die Geräteausstattung und die Bedienelemente in ähnlicher Form an vielen weiteren MAZ-Geräten, auch aus dem Digitalbereich wiederfinden. Abbildung 8.95 zeigt aus diesem Grunde in großer Darstellung die Vorderseite eines Betacam SP-Recorders der BVW-Serie und die verfügbaren Bedienelemente. In der oberen Hälfte befindet sich der Cassetten-Schacht, der zur Aufnahme beider Cassetten-Größen geeignet ist. Links daneben liegen die vier Audio-Aussteuerungsanzeigen und die zugehörigen Steller für die In- und Output-Level (Rec, PB). Unter dem Cassettenschacht liegen die Schalter zur Video- und Audio-Eingangswahl und zur Auswahl der Abhörkanäle. Über den Schalter ganz rechts wird festgelegt, ob der Recorder ferngesteuert (Remote) oder über die frontseitigen Bedienelemente gesteuert wird (Local). In der Mitte befinden sich die üblichen Tasten für Start, Stop, etc. und rechts das Jog/Shuttle-Rad zur bequemen Steuerung des Bandlaufs. Weitere Bedienelemente werden zugänglich, wenn die untere Hälfte der Frontplatte nach vorn gezogen, bzw. hoch geklappt wird (Abb. 8.96).

Die Rückwand eines Betacam SP-Recorders und die verfügbaren Anschlüsse zeigt Abb. 8.97. Das Videosignal wird in Komponentenform zu den BNC-

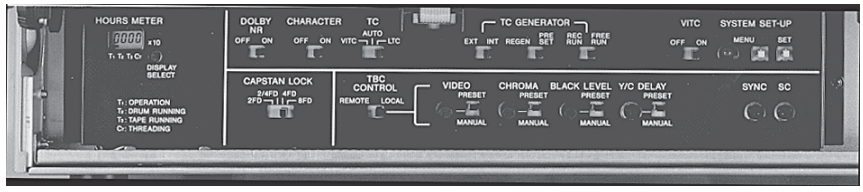


Abb. 8.96. Bedienelemente unter der Frontplatte eines Betacam SP-Recorders [208]

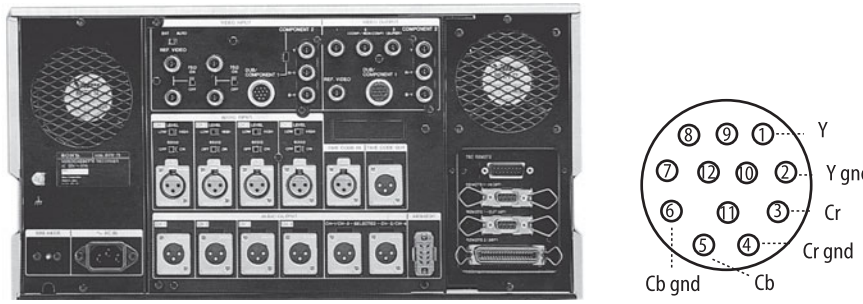


Abb. 8.97. Rückseite eines Betacam SP-Recorders und Dub-Steckerbelegung [208]

Buchsen für Y, C_R und C_B (Component 2) geführt. Für FBAS-Signale steht der Video-Eingang zur Verfügung. Eine dritte Eingangsmöglichkeit besteht über den Dub-Weg (Component 1) (Abb. 8.97). Hier wird das Komponentensignal im zeitkomprimierten Zustand belassen. Für die gleichen Signalformen sind Ausgänge vorhanden. Das FBAS-Signal steht dreifach zur Verfügung. Einer dieser Anschlüsse ist mit »Super« gekennzeichnet. Damit können Timecode- oder Laufwerksdaten in das Ausgangsbild eingeblendet werden.

8.7.2.2 M II

MII ähnelt Betacam SP qualitativ in fast allen Punkten, allerdings sind die Cassetten nicht austauschbar. Abb. 8.98 zeigt ein Gerät dieses Formats. Die Tabellen 8.8 und 8.9 zeigen die mechanischen und elektrischen Spezifikationen im Vergleich zu Betacam SP. MII nutzt das hier als CTCM (Chrominance Time Compression Multiplex) bezeichnete Zeitkompressionsverfahren. Auch bei MII gibt es je zwei Longitudinal-, FM- und optional PCM-Audiospuren.

Geräte des Formats M II wurden von Panasonic und JVC hergestellt. Es war wie Betacam SP universell für den Broadcastbereich verwendbar. Trotz der Tatsache, dass es die gleiche Leistungsfähigkeit wie Betacam SP bei halbiertem Preis bot konnte es sich aber nur bei sehr wenigen Rundfunkanstalten, wie z. B. beim ORF, etablieren.

Abb. 8.98. MII-Recorder [164]

8.8 Digitale Magnetbandaufzeichnung

Die Vorteile und Grundlagen der digitalen Signalverarbeitung wurden bereits in Kapitel 3 vorgestellt. Die Möglichkeit der Vermeidung von Qualitätsverlusten hat bei Magnetbandaufzeichnung besondere Bedeutung, da hier im Vergleich zu anderen Signalverarbeitungsstufen das Signal recht stark durch schlechten Signal/Rauschabstand und Drop Out etc. beeinträchtigt wird. Hinzu kommen Besonderheiten der Magnetaufzeichnung allgemein und des mechanischen Bandtransports mit Servosystemen und Spurabtastung, die in den Abschnitten 8.2 und 8.3 dargestellt sind und auch bei digitalen MAZ-Systemen gelten. Die Digitalaufzeichnung kann diese Probleme nicht umgehen, aber die Daten können mit einem Fehlerschutz versehen werden, sodass in den meisten Fällen eine vollständige Korrektur möglich wird, auch wenn z. B. eine Spur nicht optimal abgetastet wird.

Die digitale Aufzeichnung ist den Eigenschaften der magnetischen Aufzeichnung sehr gut angepasst, denn den zwei logischen Signalzuständen Null und Eins werden einfach die magnetischen Polarisationsrichtungen Nord und Süd zugeordnet. Innerhalb des Bereichs einer Bandwellenlänge können zwei Bit aufgezeichnet werden. Es werden Bandwellenlängen von ca. 0,5 µm erreicht.

Das Zeitraster des Analogsignals (Zeilendauer) wird beim Digitalsignal irrelevant, da das gesamte Signal durch die Abtastung einem neuen, extrem feinen Raster unterliegt. Damit besteht die Möglichkeit, die Daten fast beliebig umordnen zu können. Die zeitliche Reihenfolge muss nicht mehr beachtet werden, es muss nur bekannt sein, wie die Daten vor der Wiedergabe zurück geordnet werden müssen. Die Daten eines Halbbildes werden daher in Segmente aufgeteilt, die wiederum verschiedenen Spuren zugeordnet werden, die mit mehreren Köpfen abwechselnd geschrieben werden. Bei den Digital-Formaten werden zwischen 6 und 16, beim HDTV-Format D6 sogar 48 Schrägspuren verwendet, um ein Halbbild aufzuzeichnen. Abbildung 8.99 zeigt die Halbbildsegmentierung beim Format D1. Die aufgezeichneten 300 Zeilen pro Halbbild werden in sechs Segmente mit jeweils 50 Zeilen aufgeteilt. Die Segmente sind wiederum

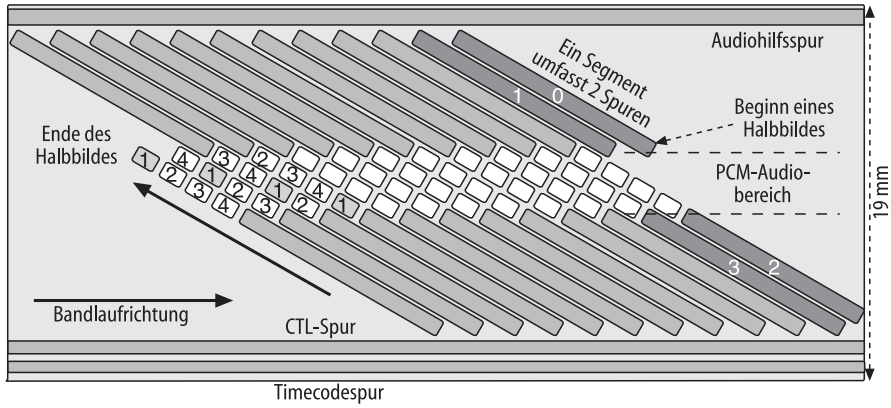


Abb. 8.99. Spurzuordnung und Halbbildsegmentierung beim D1-Format

Tabelle 8.10. Übersicht über die Digitalformate im SD-Bereich

Format	D1	D2	D3	D5	DCT	D-Beta	DV	D9	B-SX	D10
Videosignalart	Komp	FBAS	FBAS	Komp	Komp	Komp	Komp	Komp	Komp	Komp
Quantisierung (Bit)	8	8	8	10	8	10	8	8	8	8
Datenreduktion	–	–	–	–	2:1	2:1	5:1	3,3:1	10:1	3,3:1
Audiokanäle	4	4	4	4	4	4	2	2/4	4	8/4
Bandbreite (mm)	19	19	12,7	12,7	19	12,7	6,3	12,7	12,7	12,7
Spieldauer (Min.)	94	207	245	123	187	124	270	104	184	220
Hersteller	Sony BTS	Ampex Sony	Panas. JVC	Panas. JVC	Ampex	Sony	div.	JVC	Sony	Sony

auf zwölf Spuren verteilt. Abbildung 8.100 zeigt, dass die D1-Kopftrommel mit vier Köpfen bestückt ist. Sie muss mit 150 Umdrehungen/s rotieren, damit die zwölf Spuren in einer Halbbilddauer von 20 ms geschrieben werden können.

Durch die Freiheit bei der Umordnung der Daten ist es möglich, digitale Videorecorder durch einfache Umschaltung sowohl für Videosignale im europäischen 625/50-Format als auch in US-Format mit 525 Zeilen und 59,94 Hz zu nutzen. Für das 525-System werden für jedes Halbbild 5/6 der Spuren des 625-Systems verwendet. (z. B. 6 Spuren in Europa, 5 in USA). Die meisten digitalen Videorecorder können beide Formate verarbeiten, was allerdings nicht bedeutet, dass sie zur Formatkonversion eingesetzt werden können.

Tabelle 8.10 bietet eine Übersicht über die Digitalformate für Standardauflösung. Die sechs für den Broadcast-Einsatz konzipierten digitalen Magnetbandaufzeichnungsformate haben die Bezeichnungen D1–D5, D10, DCT, DVCPro, und Digital Betacam. Die Bezeichnung D4 wird nicht verwendet. Die Digitalformate D1–D5 arbeiten ohne Datenreduktion, die Formate D1, und D5 zeichnen dabei ein digitales Komponentensignal auf. Die Formate D2 und D3 sind für die Aufzeichnung eines digitalen Composite-Signals vorgesehen.

DCT, Digital Betacam und alle neueren Formate arbeiten mit Komponentensignalen, dabei wird für DCT und Digital Betacam eine Datenreduktion von 2:1 eingesetzt. Die DV-Varianten DVCPro und DVCAM reduzieren mit 5:1, DVCPro50, D9 und D10 mit 3,3:1. Abgesehen von einigen selten eingesetzten digitalen HDTV-Recordern arbeiten alle digitalen Magnetbandaufzeichnungsformate mit Cassetten. Die Breite des Bandes beträgt 19 mm bzw. 3/4“ (D1, D2, D6, DCT), 1/2“ (D3, D5, Digital Betacam, Betacam SX, D9, D10) oder 1/4“ (DV).

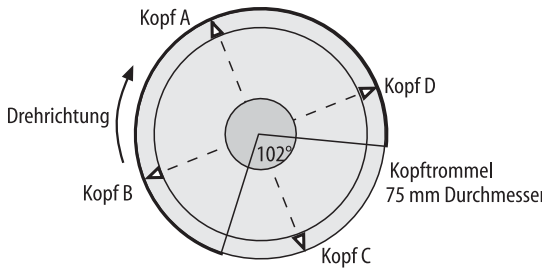


Abb. 8.100. Kopftrommel und Umschlingung beim D1-Format

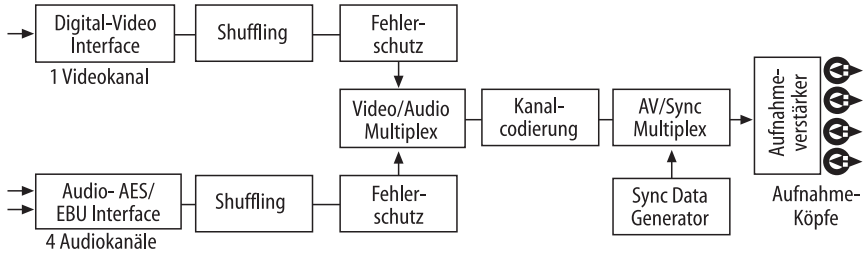


Abb. 8.101. Signalverarbeitung bei der Aufnahme

8.8.1 Signalverarbeitung

Fast alle digitalen Videorecorder enthalten Analog/Digital- und D/A-Umsetzer, damit sie in eine analoge Umgebung eingebunden werden können. Die Daten werden ggf. noch einer Verarbeitungsstufe zur Datenreduktion zugeführt und liegen dann als paralleler oder serieller Bitstrom vor.

Bevor dann die Video- und die gleichwertig behandelten Audiodaten den Magnetkopf erreichen, ist es erforderlich, sie den besonderen Eigenschaften des Übertragungskanals anzupassen. Die Nichtlinearität der Magnetisierungskennlinie ist hier kein Problem, da ja nur zwei Zustände unterschieden werden müssen, aber die Grenzfrequenzen des Kanals müssen berücksichtigt werden. Der zweite wichtige Punkt ist die Berücksichtigung der typischen Fehler die der Übertragungskanal hervorruft. Diese können durch die Forward Error Correction (FEC) bereits vor der Aufzeichnung berücksichtigt werden, sodass sie bei der Wiedergabe korrigiert werden können. Bei Digitalsystemen die datenreduzierte Signale aufzeichnen, muss die Signalverarbeitung ggf. auch noch die entsprechende Quellencodierung umfassen.

Zunächst besteht die Aufgabe der Signalverarbeitung darin, die logisch zusammengehörigen Daten im Hinblick auf einen effektiven Fehlerschutz zu erweitern und umzuordnen. Im Multiplexer werden dann der Video- und Audiodatenstrom zusammengefasst (Abb. 8.101). Darüber hinaus hat die Signalverarbeitung noch die Aufgabe, die Daten der Kanalcodierung entsprechend umzuordnen, wobei die Wahl der Codierung von dem jeweils entstehenden Signalspektrum und dem Fehlerschutz bestimmt ist.

Bevor die Daten schließlich zu den Aufnahmeverstärkern gelangen, wird der serielle Datenstrom durch Sync-Bytes erweitert, damit die umgeordneten und mit Fehlerschutz versehenen Daten eine Struktur erhalten, die es ermöglicht, die Daten zu erkennen und richtig zurückzuordnen. Am Beispiel des D1-Formats zeigt Abb. 8.102, dass den Nutzdaten mehrere Sync-, ID- und Prüf-Bytes zugesetzt werden, sodass ein Sync Block hier aus 134 Bytes besteht. Damit

Abb. 8.102. Sync-Block beim D1-Format

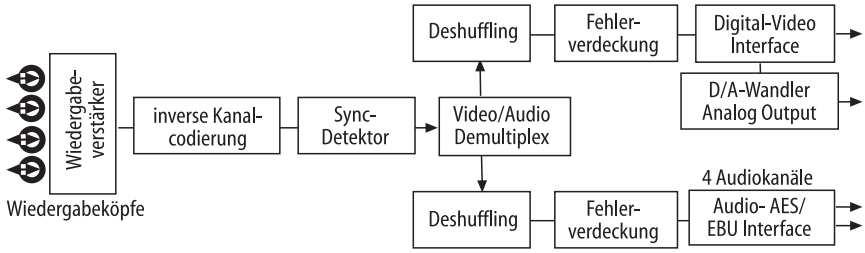


Abb. 8.103. Signalverarbeitung bei der Wiedergabe

wird ermöglicht, einerseits die Position des Sync-Blocks selbst zu erkennen und andererseits die Daten innerhalb des Blocks den richtigen Bildpunkten zuzuordnen, sodass auch bei veränderter Wiedergabegeschwindigkeit (Suchlauf) ein erkennbares Bild erzeugt werden kann.

Im Wiedergabeweg müssen die Daten zurück geordnet werden und durchlaufen die oben beschriebenen Stufen mit inverser Funktion, bis sie schließlich den Video/Audio-Demultiplexer und ggf. die D/A-Wandler erreichen. Bei der Einbindung digitaler Recorder in Studiosysteme ist zu bedenken, dass im Gegensatz zu Analogrecordern die Signalverarbeitung mit einer Zeitverzögerung von ca. einer Bilddauer verbunden ist. Die Signalverarbeitung bei der Wiedergabe ist in Abbildung 8.103 dargestellt.

Das einfache digitale Signal ist häufig dem Übertragungskanal schlecht angepasst. Es wird daher umgeformt und unter Berücksichtigung des Fehlerschutzes wird eine Kanalcodierung vorgenommen. Codierung bedeutet in diesem Zusammenhang Veränderung des Datenstroms durch Umordnung und Einfügen von Bits derart, dass sich der Originaldatenstrom eindeutig rekonstruieren lässt, hinsichtlich der Möglichkeit der Taktgewinnung aus dem Signal selbst, der spektralen Verteilung mit ausgeglichenem Frequenzgang sowie der Gleichwertanteile und der Polarität.

Ein einfaches Beispiel ist der RZ-Code. Hier wird nach jeder 1 das Bit auf Null zurückgesetzt (Return to Zero). Die höchsten Spektralanteile ergeben sich bei einem laufenden Wechsel zwischen 0 und 1 (Abb. 8.104). Wird das Signal dagegen bei aufeinanderfolgenden High-Zuständen nicht zurückgesetzt, so er-

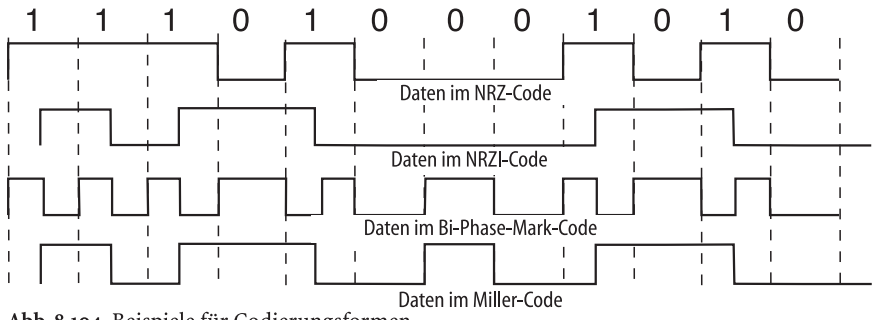


Abb. 8.104. Beispiele für Codierungsformen

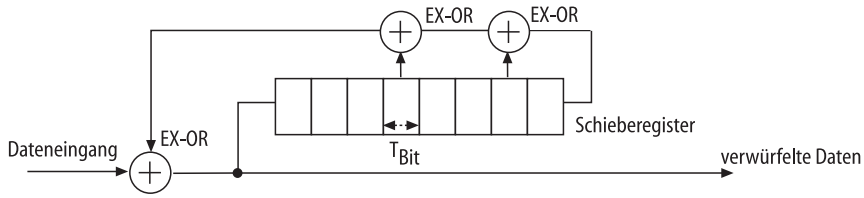


Abb. 8.105. Scrambling-Verfahren

gibt sich der NRZ-Code (Non Return to Zero) mit der halben oberen Grenzfrequenz, verglichen mit RZ. NRZ ist ein sehr simpler Code, aber weder selbsttaktend noch gleichspannungsfrei. Im Spektrum sind niederfrequente Anteile stark vertreten, was bei der Wiedergabe wegen des Induktionsgesetzes zu geringen Spannungen führt. Die spektrale Verteilung kann durch Verwürfelung (Scrambling) geändert werden. Daten sind NRZI (Non Return to Zero Inverse) codiert, wenn bei jeder 1 der Signalzustand zwischen 0 und 1 bzw. 1 und 0 umgeschaltet wird. Jede 1 entspricht einem Pegelwechsel und damit wird die Codierung polaritätsunabhängig.

Das NRZI-Spektrum ähnelt dem NRZ-Spektrum, weder NRZ noch NRZI sind selbsttaktend, sondern müssen dazu verwürfelt werden. Die Verwürfelung wird erreicht, indem der NRZ-Datenstrom über Exclusiv Oder mit einer Pseudo-Zufallsfolge verknüpft wird (Abb. 8.105). Ein verwürfelter Synchronized Scrambled NRZ-Code (SSNRZ) wird bei der Aufzeichnung im digitalen MAZ-Format D1 verwendet, ein verwürfelter NRZI-Code beim seriell-digitalen Videosignal (SDI). Eine andere Art von Codes, die selbsttaktend und polaritätsunempfindlich sind, sind die Miller-Codes (Abb. 8.104). Eine logische 1 entspricht hier einem Pegelsprung in der Taktperiode, aufeinander folgende Nullen einem Pegelsprung nach der Taktperiode [33]. Um den Miller-Code schließlich gleichspannungsfrei zu machen, wird der letzte Pegelsprung einer geraden Anzahl von Einsen zwischen zwei Nullen weggelassen (Miller²). Miller-Codes erreichen das Ziel, relativ schmalbandig zu sein. Miller² wird bei der D2-MAZ-Aufzeichnung verwendet.

Als weitere Codierungsart wäre das Bitmapping zu nennen, häufig werden dabei 8-Bit-Blöcke unter Berücksichtigung des Fehlerschutzes in größere Einheiten umgesetzt (z. B. 8/14-Modulation), wobei die Datenrate entsprechend

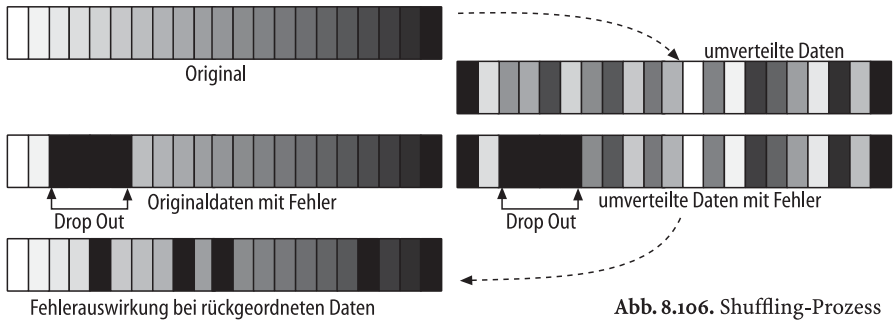


Abb. 8.106. Shuffling-Prozess

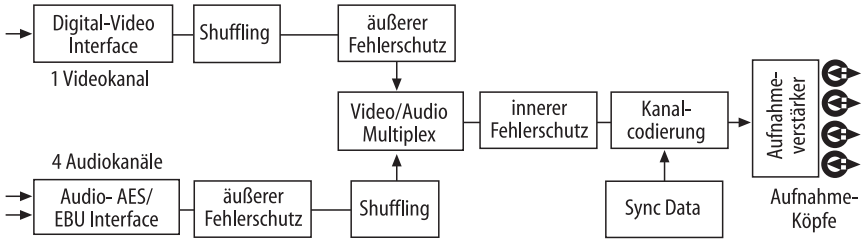


Abb. 8.107. Innerer und äußerer Fehlerschutz

steigt. Die Zuordnung geschieht mit Hilfe von Tabellen. Diese Variante der Codierung wird bei den Formaten D3 und D5 angewandt und zwar so, dass die maximale Lauflänge gleicher Bits im 14-Bit-Wort gleich sieben und die minimale Lauflänge gleich 2 ist. Damit ergibt sich eine schmale spektrale Verteilung.

8.8.1.1 Fehlerschutz

Die Magnetbandaufzeichnung ist besonders fehleranfällig. Zu dem auch in anderen Signalsverarbeitungsstufen auftretenden Rauschen kommen Signaleinbrüche (Drop-Outs) aufgrund von Fehlstellen an der Bandoberfläche hinzu und als Kopfzusetzer bezeichnete, durch Schmutz verursachte, kurzzeitige Kopfausfälle. Es können dabei kurze Bitfehler auftreten oder aber z. B. auch zusammenhängende Bereiche von einigen hundert Bits betroffen sein. Die Digitaltechnik bietet hier den Vorteil, dass die Daten durch Fehlerschutzbits erweitert und damit die meisten Fehler, unterstützt durch eine gezielte Datenumverteilung (Shuffling), korrigiert werden können.

Bei der Digitalaufzeichnung kommen meist Blockcodes zum Einsatz, wie sie in Kapitel 4 beschrieben sind. Dafür ist es wichtig, die logisch zusammengehörigen Daten physikalisch weit entfernt voneinander zu speichern. Es wird eine Datenumverteilung vorgenommen, die als Shuffling oder Interleaving bezeichnet wird (Abb. 8.106). Oft werden zwei Stufen zur Fehlerkorrektur eingesetzt. Dabei ist das Coder/Decoderpaar, das näher an der Signalquelle bzw. -senke liegt, für den sogenannten äußeren Fehlerschutz zuständig und das näher am Übertragungskanal liegende für den inneren (Abb. 8.107).

Die unvermeidlichen Restfehler bei Audio- und Videodaten werden meistens verdeckt, indem die vorhergehenden Werte gehalten oder Ersatzdaten aus der Interpolation zwischen Nachbarwerten berechnet werden (Abb. 8.108).

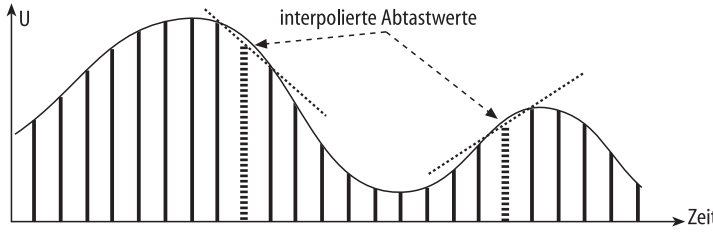


Abb. 8.108. Interpolation fehlender Abtastwerte

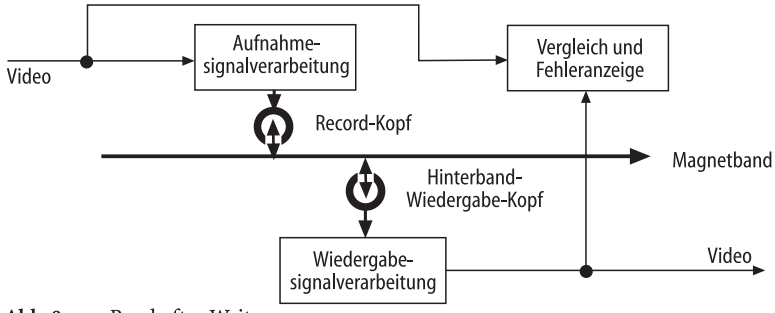


Abb. 8.109. Read after Write

8.8.1.2 Sonderfunktionen

Die für den professionellen Einsatz vorgesehenen Digitalformate sollten mindestens die gleichen Ausstattungsmerkmale wie Analogformate aufweisen. So sollte es auch möglich sein, im Suchlauf den Ton abzuhören. Diese Funktion wird häufig mit Hilfe einer einkanaligen analogen Audio-Hilfsspur realisiert, indem beim schnellen Suchlauf von Digital- auf Hilfsspurwiedergabe umgeschaltet wird. Moderne Geräte erlauben auch die Wiedergabe von Digitalton beim Suchlauf. Digitale Videorecorder sollen außerdem ungestörte Stand-, Zeitlupen- und Zeitrafferbilder liefern können. Zu diesem Zweck sind, wie bei analogen Systemen, viele Geräte mit separaten, beweglichen Videoköpfen für die Wiedergabe ausgestattet, die in den meisten Fällen den üblichen variablen Geschwindigkeitsbereich mit $k = -1$ bis $k = +3$ abdecken.

Der Einsatz separater Wiedergabeköpfe ermöglicht, dass die Daten direkt nach der Aufnahme wiedergegeben und kontrolliert werden können (Direct Read after Write, Abb. 8.109), sodass eine objektive Qualitätsüberwachung möglich wird. Weiterhin können separate Wiedergabeköpfe auch so eingesetzt werden, dass die damit gelesenen Daten in modifizierter Form direkt wieder aufgezeichnet werden (Read Modify Write, Abb. 8.110). Diese Funktion erlaubt eine Signalbearbeitung (z. B. Farbkorrektur) mit nur einer Maschine oder einen eingeschränkten A/B-Roll-Schnitt mit zwei Geräten. Es ist allerdings zu bedenken, dass die Originalaufnahme dabei gelöscht wird.

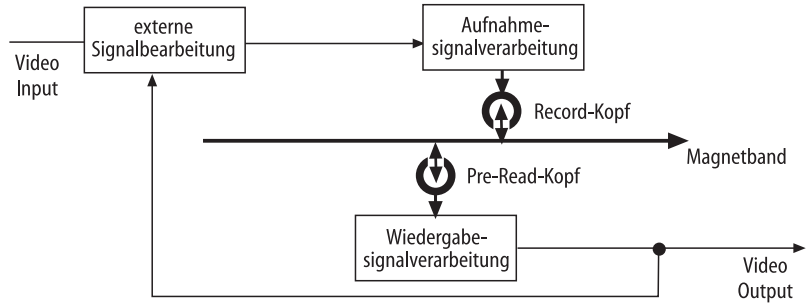


Abb. 8.110. Read Modify Write

Tabelle 8.11. Geometrische Parameter der Formate D1, D5 und D6

	D1	D5	D6
Magnetbandbreite	19 mm	12,7 mm	19 mm
Bandgeschwindigkeit	28,7 cm/s	16,7 cm/s	49,7 cm/s
Relativgeschwindigkeit	35,6 m/s	23,8 m/s	46 m/s
Videospurbreite	40 µm	18 µm	21 µm
Spurzwischenraum	5 µm	0	1 µm
Kopftrommeldrehzahl	150 Hz	100 Hz	150 Hz
Kopftrommeldurchmesser	75 mm	76 mm	96 mm
max. Spieldauer (S, M, L)	12 / 34 / 94 Min.	32 / 62 / 123 Min.	8 / 28 / 64 Min.

8.8.2 Digitale Komponentenformate ohne Datenreduktion

Die Formate dieser Klasse erfüllen höchste Qualitätsansprüche in Bereich der SD-Aufzeichnung. Bereits das erste, als D1 bezeichnete Format zur digitalen Videosignalaufzeichnung arbeitete mit Komponentensignalen nach dem Standard ITU-R 601 (Abb. 8.111). Das Format D5 bietet darüber hinaus 10-Bit-Signalverarbeitung und D6 eine HDTV-Aufzeichnung.

8.8.2.1 Das D1-Format

Das D1-Format gehört bis heute zu den Magnetbandaufzeichnungsverfahren mit der höchsten Qualität, allerdings erfüllt es nicht mehr die Ansprüche an ein kostengünstiges Produktionsmittel. Drei Cassetten (S, M, L) mit 3/4"-Band erlauben Spielzeiten bis zu 12, 34 oder 94 Minuten. Die Bandlaufgeschwindigkeit beträgt 28,7 cm/s, die Kopf-Band-Relativgeschwindigkeit 35,6 m/s (Tabelle 8.11). Die Kopftrommel dreht sich mit 150 U/s, das Signal wird mit vier Köpfen geschrieben. Die Aufzeichnung eines Halbbildes im 625/50-System erfordert 12 Spuren, ein 525/60-Halbbild wird in zehn Spuren aufgezeichnet. Wenn die Videoköpfe im 90°-Winkel angeordnet sind, beträgt die Kopftrommelumschlingung 270°. Die Kopftrommel und das Spurbild wurden bereits im vorigen Abschnitt vorgestellt (Abb. 8.99 und 8.100). Das Spurbild zeigt eine longitudinale Audio-Hilfspur am oberen Bandrand und eine CTL und Timecode-Spur am unteren. Die Hilfsspur ist erforderlich, da bei schnellem Vorlauf die Audiodaten nicht ausgewertet werden können. Zwischen den Längsspuren befindet sich der Schrägspurbereich. Hier ist jede Spur in zwei Videosektoren und dazwi-

Abb. 8.111. D1-Recorder [32]

Tabelle 8.12. Signalparameter der Formate D1, D5 und D6

	D1	D5	D6
Y-Bandbreite (–0,5 dB)	5,75 MHz	5,75 / 7,67 MHz	30 MHz
C-Bandbreite (–0,5 dB)	2,75 MHz	2,75 / 3,67 MHz	15 MHz
Pegelauflösung	8 Bit	10 Bit / 8 Bit	8 Bit
Signal/Rauschabstand (Y)	> 56 dB	> 62 / 56 dB	> 56 dB
aufgezeichnete Datenrate	227 Mbit/s	303 Mbit/s	1,2 Gbit/s
aufgez. Zeilenzahl/Halbbild	300	304	> 1150
kürzeste Wellenlänge	0,9 µm	0,71 µm	

schen liegende vier Audiosektoren aufgeteilt. Damit können vier digitale Audiokanäle hoher Qualität aufgezeichnet und unabhängig editiert werden.

Tabelle 8.12 zeigt die D1-Signalspezifikationen. Jede Signalkomponente wird mit 8 Bit aufgelöst und mit SSNRZ-Codierung (Scrambled Synchronized Non Return to Zero) aufgezeichnet. Die Gesamtdatenrate beträgt 227 Mbit/s. Als besondere Funktion können zwei D1-Recorder synchron betrieben werden.

8.8.2.2 Das D5-Format

Abbildung 8.112 zeigt einen D5-Studiorecorder von Panasonic. Das D5-Format (D14) wurde von Panasonic entwickelt, es ermöglicht als einziges Format eine unkomprimierte SD-Komponentenaufzeichnung auf 1/2“-Bändern und ist für Standardsignale das Aufzeichnungsformat mit der höchsten heute möglichen Qualität. Die D1-Qualität wird noch übertroffen, da D5 mit 10 Bit statt 8 Bit Amplitudenauflösung arbeiten kann (Tabelle 8.12). Mit vier digitalen Audiokanälen und Fehlerschutzbits wird eine Gesamtdatenrate von 303 Mbit/s mit 8/14-Modulation aufgezeichnet. Aufgrund der hohen Datenrate ist es bei D5 möglich, nicht nur Digitaldaten nach ITU-R 601 mit einer Abtastrate von 13,5/6,75 MHz und einer Auflösung von 10 Bit aufzuzeichnen, sondern auch mit erhöhter Horizontalauflösung bei einer Abtastrate von 18 MHz zu arbeiten. Die aktive Videozeile enthält dann 960 statt 720 Abtastwerte, die Auflösung der Abtastwerte muss dabei allerdings auf 8 Bit beschränkt werden. Dieser Modus ist vor allem für eine hochwertige Signalverarbeitung für das Bildformat 16:9 nützlich. Das D5-Format wurde als HD-D5 unter Nutzung von Datenreduktionsverfahren auch für die HDTV-Aufzeichnung eingesetzt. Aufgrund der hohen Datenrate ist dabei nur eine Kompression um den Faktor 4 erforderlich.

Abb. 8.112. D5-Recorder [164]

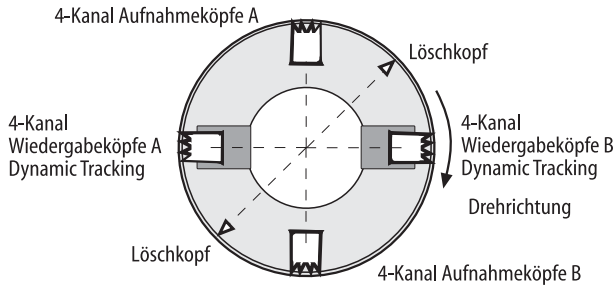


Abb. 8.113. Kopftrommel des D5-Recorders

Drei Cassettengrößen (S, L, XL) erlauben Spielzeiten von maximal 32, 62 oder 123 Minuten, allerdings werden an das Bandmaterial hohe Anforderungen gestellt und die Cassetten sind daher sehr teuer. Die hohe Aufzeichnungsdichte wird erreicht, indem sehr schmale Schrägspuren von nur 18 μm Breite bei einer Bandgeschwindigkeit von 16,7 cm/s aufgezeichnet werden (Tabelle 8.11). Abbildung 8.113 zeigt die D5-Kopftrommel. Das Spurbild (Abb. 8.114) ähnelt dem von D1, es zeigt eine longitudinale Audio-Hilfspur am oberen Bandrand und eine CTL- und Timecode-Spur am unteren. Dazwischen befindet sich wieder der Schrägspurbereich, der in zwei Videosektoren und dazwischen liegende vier Audiosektoren für die Aufzeichnung von vier digitalen Audiokanälen hoher Qualität aufgeteilt ist. Im Gegensatz zum D1-Format sind die Videospuren in Richtung des Bandlaufs geneigt und werden mit der Slanted Azimut-Technik ohne Sicherheitsabstand geschrieben.

D5-Recorder sind auf der Basis des D3-Formates von Panasonic entstanden. Beide Formate weisen ein sehr ähnliches Spurbild auf, die Bandgeschwindigkeit und die Datenrate sind bei D5 doppelt so hoch wie bei D3. Daher ist es möglich, dass D5-Recorder optional auch D3-Material wiedergeben können.

8.8.3 Digitale Compositesignal-Aufzeichnung

Das D1-Format wurde bereits zu einer Zeit entwickelt, als in den Studios noch die FBAS-Geräte dominierten. Als direkter Ersatz für 1“ B- und C-MAZ-Geräte

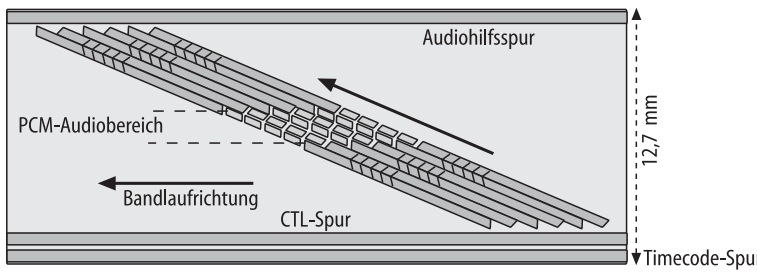


Abb. 8.114. Spurlage beim D5-Format (und DCT bei umgekehrter Laufrichtung)

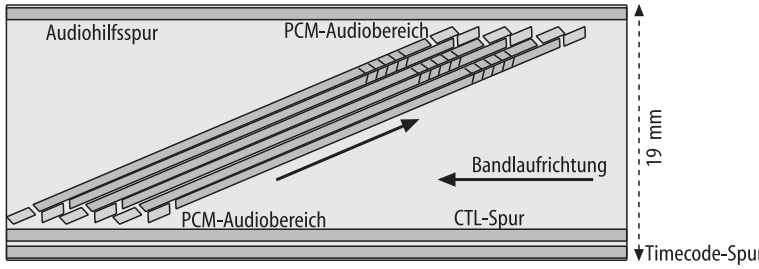


Abb. 8.115. Spurlage beim D6 und D2-Format (und D3 bei umgekehrter Laufrichtung)

waren D1-Geräte zu aufwendig. Hier war ein Format gefragt, das bei geringeren Kosten den wesentlichen Vorteil der Digitaltechnik, nämlich die gute Multigenerationsfähigkeit, bietet.

8.8.3.1 Das D2-Format

Das zweite Digitalformat (Ampex, Sony) arbeitet mit Composite-Signalen, die mit der 4-fachen Farbträgerfrequenz (17,7 MHz) bei einer Auflösung von 8 Bit abgetastet werden. Unter Einbeziehung der Fehlerschutzdaten und der vier digitalen Audiokanäle (max. 20 Bit/48 kHz) ergibt sich eine aufgezeichnete Datenrate von 152 Mbit/s. Pro Halbbild werden 304 Zeilen auf 8 Spuren aufgezeichnet, die ohne Sicherheitsabstand (Rasen) geschrieben werden. Die aktuelle Spur überdeckt dabei einen Teil der zuvor aufgezeichneten, das Übersprechen wird durch die wechselweise gekippten Kopfspalte vermindert. Bei einer Bandgeschwindigkeit von 13,2 cm/s betragen die Cassettenspieldauern 32, 94 bzw. 204 Minuten (Tabelle 8.13). Die Aufzeichnung erfolgt mit zwei Kopfpaaren, Abb. 8.115 zeigt die Spurlage und Abb. 8.116 die Kopftrommel.

Durch die Digitalaufzeichnung mit einer Grenzfrequenz von 6,5 MHz wird ein gutes Multigenerationsverhalten bis zur 20. Generation erreicht. Die Composite-Signalaufzeichnung weist aber, insbesondere wegen der geringen Chrominanz-Bandbreite, erhebliche Nachteile auf. Abbildung 8.117 zeigt einen D2-Recorder. Das Format ist veraltet.

8.8.3.2 Das D3-Format

Das digitale Composite-Format D3 hat bezüglich der Signalverarbeitung ähnliche Merkmale wie D2. Es gibt auch hier vier separat editierbare Audiokanäle, Hinterbandkontrolle sowie eine Wiedergabemöglichkeit mit –1- bis 3-facher

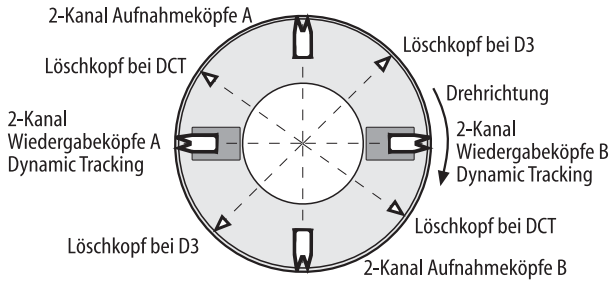


Abb. 8.116. Kopftrommel bei D2, D3 und DCT

Tabelle 8.13. Geometrische Parameter der Formate D2 und D3

	D2	D3
Magnetbandbreite	19 mm	12,7 mm
Bandgeschwindigkeit	13,2 cm/s	8,4 cm/s
Relativgeschwindigkeit	30,4 m/s	23,8 m/s
Videospurbreite	35 µm	18 µm
Spurzwischenraum	0	0
Kopftrommeldurchmesser	96,5 mm	76 mm
max. Spieldauer (S, M, L)	32 / 94 / 204Min.	64 / 125 / 245 Min.

Normalgeschwindigkeit. Auch das Spurbild ist ähnlich (Abb. 8.115), allerdings sind die Schrägspuren in Bandlaufrichtung geneigt. D3 ist wie D2 bei gleicher Gesamtdatenrate für die Aufzeichnung von 625 Zeilen bei 50 Hz oder 525/60 geeignet. Ein D3-Recorder ist in Abb. 8.118 dargestellt. Der wesentliche Unterschied zu D2 ist die Verwendung von 1/2"-Bändern. Aufgrund der gleichen Spurbreite wie bei D5 (18 µm) konnte die Aufzeichnungsdichte bei D3 gegenüber D2 von 7,2 auf 17,9 Mbit/cm² gesteigert werden. Damit werden bei einer Bandgeschwindigkeit von 8,4 cm/s Cassettenspieldauern von 64, 125 bzw. 245 Minuten erreicht (Tabelle 8.13). Das D3-Format eignet sich gut zur digitalen Aufzeichnung von altem FBAS-Material (z. B. 1"-B und -C oder U-Matic).

8.8.4 Digitale Komponentenformate mit geringer Datenreduktion

Um eine digitale Komponentenaufzeichnung zu geringeren Kosten zu ermöglichen, gleichzeitig aber auch ein robustes Format zu erhalten, bei dem die Aufzeichnungsdichte nicht bis an die Grenze des Machbaren getrieben wurde, entschied man sich bei Ampex und Sony zum Einsatz einer milden Datenreduktion. Dieser Schritt wurde bei der Einführung der Systeme als revolutionär angesehen, die Bildqualität der beiden Formate ist aber visuell auch nach vielen Generationen nicht zu unterscheiden. Außerdem stieg die Akzeptanz der Datenreduktion im Produktionsbereich generell, sodass sich DigiBeta gegen D5 durchsetzen konnte, da es als großen Vorteil die optionale Kompatibilität zum analogen Aufzeichnungsstandard Betacam SP bietet und geringere Folge-

Abb. 8.117. D2-Recorder

Abb. 8.118. D3-Recorder [164]

Abb. 8.119. DCT-Recorder [6]

kosten für das Bandmaterial mit sich bringt. Digital Betacam ist bis heute ein verbreitetes Format im Bereich anspruchsvoller digitaler SD-Produktion, während DCT, das zweite Format in diesem Kapitel, fast bedeutungslos ist.

8.8.4.1 Das DCT-Format

Mit DCT sollte ein besonders robustes Aufzeichnungsformat mit geringer Fehlerrate entwickelt werden. Mit den gleichen Cassettengrößen wie bei D2 werden Spieldauern von 29, 85 oder 187 Minuten erreicht. Inklusive der vier Audiokanäle, deren Sektoren wie bei D1 in der Bandmitte liegen (Abb. 8.114), ergibt sich bei DCT eine Gesamtdatenrate von 130 Mbit/s (Tabelle 8.14).

DCT steht für Digital Component Technology. Die Abkürzung kennzeichnet gleichzeitig auch das eingesetzte Datenreduktionsverfahren DCT. Bei der verwendeten Kompression um den Faktor 2 kann man davon ausgehen, dass bei vielen Bildvorlagen eine Redundanzreduktion möglich ist, andererseits tritt bei kritischen Vorlagen tatsächlich ein (unsichtbarer) Informationsverlust auf.

Aufgrund ihrer Zuverlässigkeit und Robustheit wurden DCT-Geräte vor allem im Post-Produktions-Bereich eingesetzt. Das Band wird mit einem Luftfilm anstelle von Reibrädern sehr schonend geführt, die Reaktionszeiten der Mechanik sind sehr gering. Beim Editing mit DCT-Maschinen (Abb. 8.119) werden Pre Roll-Zeiten (s. Abschn. 9.3) von weniger als zwei Sekunden erreicht.

8.8.4.2 Digital Betacam

Digital Betacam stammt vom Betacam-Systementwickler Sony und etablierte sich mit Hilfe der Kompatibilität zum analogen Format zu Beginn des neuen

Abb. 8.120. Digital Betacam-Recorder [208]

Tabelle 8.14. Geometrische Parameter der Formate DCT, Digital Betacam und Betacam SP

	DCT	Digital Betacam	Betacam SP
Magnetbandbreite	19 mm	12,7 mm	12,7 mm
Bandgeschwindigkeit	14,6 cm/s	9,7 cm/s	10,2 cm/s
Relativgeschwindigkeit	30,4 m/s	19 m/s	5,8 m/s
Schrägspurlänge	151 mm	123 mm	115 mm
Videospurbreite	35 µm	24 µm	86 µm
Datenreduktionsfaktor	2:1	2:1	
Kopftrommeldurchmesser	96,5 mm	81,4 mm	74,5 mm
max. Spieldauer (S, M, L)	29 / 85 / 187 Min.	40 / - / 124 Min.	32 / - / 108 Min.

Jahrhunderts als Standard im Broadcast-Bereich. Es werden Cassetten mit den gleichen mechanischen Abmessungen wie bei Betacam verwendet, sie ermöglichen hier Spieldauern von 40 bzw. 124 Minuten. Das 1/2"-Cassettenformat erlaubt den Bau von Camcordern, die ähnlich kompakt sind wie die üblichen Betacam SP-Camcorder. Für das Format sind Studio-Recorder (Abb. 8.120) und -Player ebenso verfügbar wie portable Field-Recorder. Das Komponentensignal wird mit 10 Bit und Abtastraten nach CCIR 601 digitalisiert, die Gesamtdatenrate beträgt 125,6 Mbit/s (Tabelle 8.14).

Digital Betacam arbeitet mit einer DCT-Variante von Sony, um eine Datenreduktion zu erreichen. Die Verwendung von Datenreduktionstechniken ist bei diesem Format darin begründet, dass eine Komponentensignalaufzeichnung mit 1/2"-Bändern möglich werden sollte, ohne die Anforderungen an Spurbreite und Bandqualität so hoch zu treiben wie beim Konkurrenzformat D5. Außerdem war bei der Entwicklung Rücksicht auf die Abwärts-Kompatibilität zum Betacam SP-Format zu nehmen.

Das Datenreduktionsverfahren musste so ausgelegt sein, dass eine konstante Datenrate am Ausgang entsteht, hier hat die Netto-Datenrate den Wert 109 Mbit/s. Da bei der Abtastung von 608 Zeilen mit je 720 Bildpunkten eine Datenrate von 218 Mbit/s entsteht, hat der Datenreduktionsfaktor hat damit den Wert zwei. Durch geschickte Wahl der Quantisierung der DCT-Ausgangsdaten, kann erreicht werden, dass die Schwelle zwischen Redundanz- und Irrelevanzreduktion nur in seltenen Fällen, bei sehr kritischen Bildvorlagen überschritten wird und damit eine exzellente Bildqualität zur Verfügung steht. Als Kanalcodierung

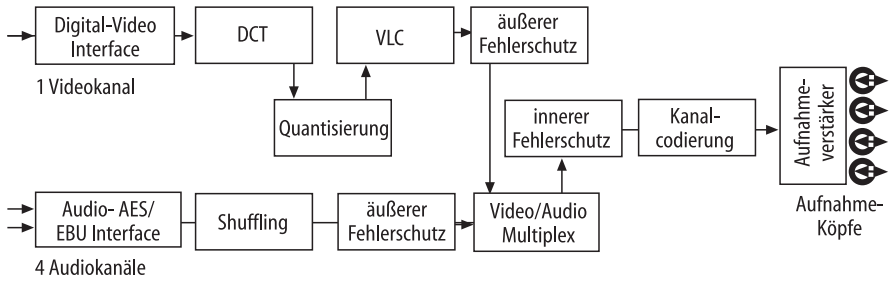


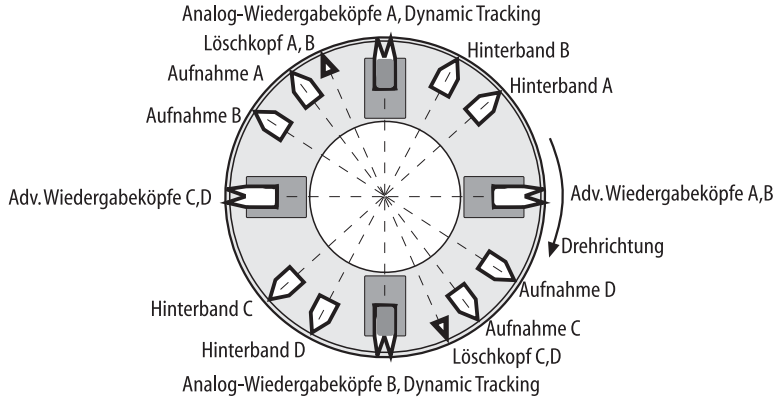
Abb. 8.121. Signalverarbeitung bei Digital Betacam

Sync-Bytes (2)	ID-Bytes (2)	Daten-Bytes (162)	Fehlercode Bytes (14)

Abb. 8.122. Datenblöcke bei Digital Betacam

wird eine Form des SNRZI-Codes eingesetzt, die ein enges Frequenzspektrum erzeugt und mit Partial Response IV bezeichnet wird. Abbildung 8.121 zeigt die Signalverarbeitung bei der Aufnahme, Abbildung 8.122 die Struktur des Sync-Blocks der 180 Byte enthält, von denen 162 Nutzbytes sind.

Digital Betacam ermöglicht optional die Wiedergabe analoger Betacam-Bänder. Bei entsprechenden Geräten müssen dafür separate Köpfe montiert sein (Abb. 8.123). Das in Abbildung 8.124 dargestellte Spurbild ist ähnlich wie bei D5, es weist aber die Besonderheit auf, dass zwischen den Audio- und Videosektoren Pilottonfrequenzen aufgezeichnet werden, die zusammen mit den üblichen CTL-Signalen auf der Längsspur ein besonders gutes Tracking-Verhalten ermöglichen. Die Längsspuren an den Rändern stimmen beim Digital- und Analogformat überein.



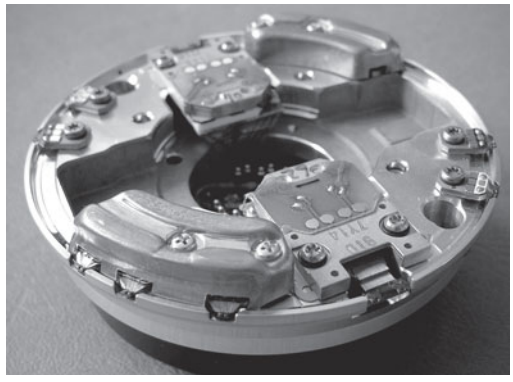


Abb. 8.123. Digital Betacam-Kopfrad mit (oben) und ohne (unten) Analogköpfen

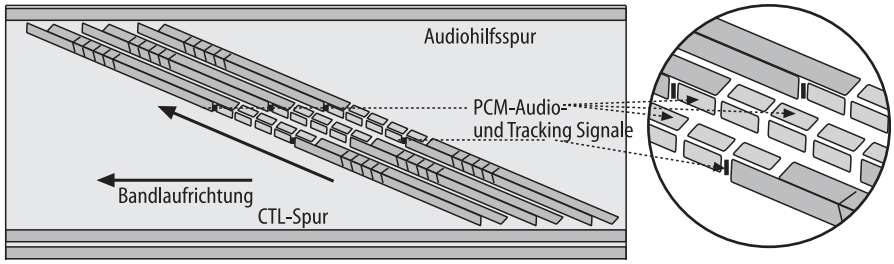


Abb. 8.124. Spurlage bei Digital Betacam

8.8.5 Digitale Komponentenaufzeichnung mit DV-Datenreduktion

Nachdem die Diskussionen um die Verwendbarkeit von Datenreduktionsverfahren im Produktions- und Postproduktionsbereich abgeklungen waren, wurden viele weitere Formate entwickelt, die mit höheren Datenreduktionsfaktoren arbeiten und für den Corporate und News-Bereich, sowie für einfache Produktionen konzipiert wurden. Die Akzeptanz der Datenreduktion wurde durch die relativ gute Bildqualität gesteigert, die das erste Digitalaufzeichnungsformat für den Heimanwenderbereich, nämlich DV, bietet.

Die ersten Formate aus dieser Klasse basierten denn auch auf dem DV-Format und wurden geringfügig modifiziert. Diese Entwicklungen wurden DV-CAM und DVCPro (D7) genannt und stammen von Sony und Panasonic. Von JVC folgte das Format Digital S, das auch als D9 bezeichnet wird. Hier wird der DV-Datenreduktionsalgorithmus eingesetzt, aber mit doppelter Datenrate, in gleicher Weise wie bei der Weiterentwicklung von DVCPro zu DVCPro50.

8.8.5.1 Das DV-Format

Bereits Ende der 80er Jahre bemühten sich viele Geräteentwickler um ein Digitalformat für den Heimanwenderbereich. Ein Konsortium verschiedener großer Gerätehersteller wie Thomson, Sony, Philips und Panasonic konnte daraufhin 1994 einen einheitlichen Standard definieren, der zunächst DVC (Digital Video Cassette) später nur DV genannt wurde. Das Format wurde im Hinblick auf die Entwicklung sehr kompakter Geräte konzipiert. Die Kopftrommel hat einen Durchmesser von nur 21,7 mm. Das Band läuft mit einer Geschwindigkeit von 1,88 cm/s. Es werden sehr kleine Cassetten mit 1/4“-Band (6,35 mm) und den Abmessungen 66 mm x 48 mm x 12 mm verwendet. Die Spieldauer beträgt hier 60 Minuten, größere Cassetten (125 mm x 78 mm x 15 mm) ermöglichen Spieldauern bis zu 270 Minuten. Die Cassetten enthalten einen Speicher in Form eines ID-Board in dem Basiseigenschaften wie z. B. Bandtype (ME oder MP) festgehalten sind, oder einen Halbleiterspeicher. Der nicht flüchtige MIC-Halbleiterspeicher gestattet die Aufzeichnung zusätzlicher Informationen.

Die Signalspezifikationen entsprechen fast professionellen Maßstäben. Es wird ein 8-Bit-Komponentensignal aufgezeichnet, wobei das Luminanzsignal Y mit 13,5 MHz abgetastet wird. Das Format ist in Europa ebenso einsetzbar wie in den USA, allerdings gibt es einige Unterschiede. Im 625/50-System wer-

Tabelle 8.15. Geometrische Parameter der diversen DV-Formate

	DV	DVCam	DVCPro	DVCPro50
Magnetbandbreite	6,3 mm	6,3 mm	6,3 mm	6,3 mm
Bandgeschwindigkeit	1,88 cm/s	2,82 cm/s	3,38 cm/s	6,76 cm/s
Relativgeschwindigkeit	10,2 m/s	10,2 m/s	10,2 m/s	10,2 m/s
Schrägspurlänge	34 mm	34 mm	34 mm	34 mm
Videospurbreite	10 µm	15 µm	18 µm	18 µm
Kopftrommeldurchmesser	21,7 mm	21,7 mm	21,7 mm	21,7 mm
max. Spieldauer (S, M, L)	60/ - /270 Min.	40/ - /180 Min.	33/63/126 Min.	16/30/60 Min.

den die Chrominanzanteile C_R und C_B mit 6,75 MHz abgetastet, jeder Anteil steht aber nur jede zweite Zeile zur Verfügung (4:2:0). Im 525/60-System wird dagegen statt der vertikalen die horizontale Auflösung reduziert, d. h. die Abtastrate für die Farbdifferenzkomponenten wird halbiert (4:1:1). Die Videoausgangsdatenrate beträgt damit 125 Mbit/s (s. Abschn. 3.5.1), sie wird durch eine DCT-Datenkompression um den Faktor 5 auf 25 Mbit/s reduziert. Hinzu kommen Fehlerschutz, Zusatz- und Audiodaten, sodass die aufgezeichnete Datenrate insgesamt ca. 42 Mbit/s beträgt. Die Daten werden mit zwei Köpfen auf das Band geschrieben. Vor der Aufzeichnung werden die Daten 24/25-moduliert und ein ATF-System mit vier Pilotfrequenzen realisiert. Schließlich werden sie mit einem NRZI-Kanalcode versehen und auf das Band geschrieben.

Alle Signale, inklusive der Steuerdaten, werden auf Schrägspuren aufgezeichnet, dafür sind im 625/50-System pro Bild 12 Spuren und im 525/60-System 10 Spuren erforderlich. Längsspuren sind nur optional vorgesehen, Abb. 8.125 zeigt die Spurlage. Die Steuersignale befinden sich im ITI-Sektor (Insert Track Information) am Spuranfang, der bei Insert-Aufnahmen nicht überschrieben wird. Es folgen die Audio- und Videosektoren, und am Schluss der Schrägspur liegt ein Subcode-Bereich der zur Aufzeichnung der Uhrzeit und von Timecode-Daten genutzt wird. Alle Sektoren sind durch Zwischenräume (Gaps) getrennt und separat editierbar. Die Schrägspurlänge beträgt 34 mm, die Spurbreite 10 µm. Tabelle 8.15 zeigt weitere mechanische Parameter.

Das bei DV angewandte Datenreduktionsverfahren ist in Kapitel 3 genauer dargestellt. Als herausragendes Merkmal ist dabei die reine Intraframe-Codierung zu nennen, die aber auch bei MPEG-2 mit I-Frame-only erreicht wird. Bei DV wird darüber hinaus die konstante Datenrate nicht dadurch gewonnen, dass quasi rückwärtsgewandt die Quantisierung vergröbert wird, wenn

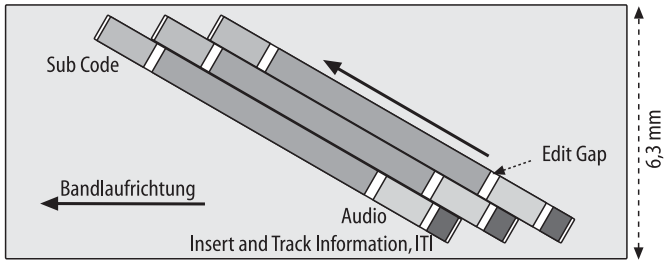


Abb. 8.125. Spurlage beim DV-Format

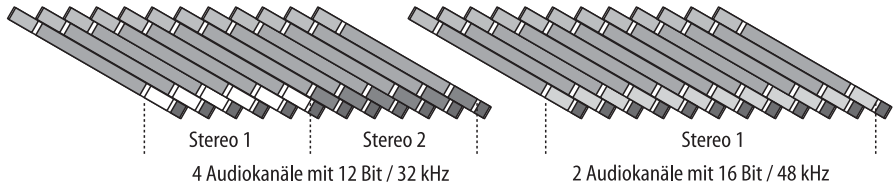


Abb. 8.126. Audio Record Modes bei DV

der Datenpuffer am Ausgang zu voll wird, sondern vorwärtsgerichtet, indem vor der Datenreduktion mit Hilfe einer Bildanalyse eine Quantisierungstabelle bestimmt wird, die für eine gegebene Bitrate qualitativ optimal ist. Abbildung 8.127 zeigt ein frühes DV-Amateurgerät.

Zur Audioaufzeichnung stehen wahlweise zwei Kanäle mit 16 Bit Quantisierung und 44,1/48 kHz Abtastrate oder vier Kanäle mit 12 Bit/32 kHz zur Verfügung, die separat bearbeitet werden können (Abb. 8.126). Die Audio- und Videodaten können synchronisiert oder unsynchronisiert aufgezeichnet werden (lock/unlock). Die Synchronisation ist bei professionellem Einsatz unverzichtbar, in diesem Fall muss auf 44,1 kHz als Abtastrate verzichtet werden [231].

8.8.5.2 DVCAM

Dieses Format ist die Weiterentwicklung des DV-Formats für anspruchsvollere Anwendungen (Professional Sector) von Sony. Die Datenverarbeitung bleibt die gleiche wie bei DV, aber die Geräteausstattung und die Anschlüsse entsprechen den Anforderungen des professionellen Bereichs. DV und DVCAM sind nicht in beiden Richtungen, sondern nur abwärts kompatibel. Die Tabellen 8.21 und 8.22 zeigen die Spezifikationen. Der wesentliche Unterschied besteht in der Erhöhung der Bandgeschwindigkeit von 18,8 mm/s auf 28,2 mm/s und damit einer Erhöhung der Spurbreite von 10 μm auf 15 μm. Auf diese Weise soll die Aufzeichnung robuster werden. Die maximale Aufzeichnungsdauer auf einer Mini-DV-Cassette sinkt damit von 60 auf 40 Minuten und bei Standard-Cassetten von 270 auf 180 Minuten. Die Audioverarbeitung erfolgt im Lock-Mode, damit ist wahlweise die Aufzeichnung von 2 Kanälen mit 48 kHz/16 Bit oder von 4 Kanälen mit 32 kHz/12 Bit möglich. Das Format bietet optional eine Timecode-Verarbeitung. Abbildung 8.128 zeigt einen DVCAM-Recorder.

Abb. 8.127. DV-Camcorder [208]

Abb. 8.128. DVCAM-Recorder [208]

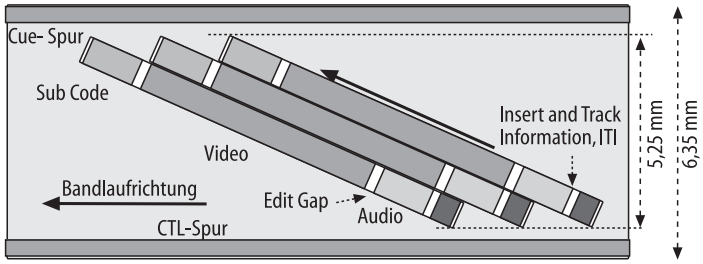


Abb. 8.129. Spurlage bei DVCPro

8.8.5.3 DVCPro

Diese Bezeichnung, bzw. auch die Bezeichnung D7, steht für die Weiterentwicklung des DV-Formates für den Professional Sektor von Panasonic. Die Spurbreite wird noch weiter als bei DVCAM, nämlich von 10 µm auf 18 µm erhöht. Das geschieht durch eine auf 33,8 mm/s gesteigerte Geschwindigkeit. Die Geräte werden auch hier professionellen Ansprüchen entsprechend ausgestattet. Zur Audioaufnahme ist nur der 2-Kanalbetrieb mit 48 kHz/16 Bit vorgesehen, vierspurig bespielte Cassetten können aber wiedergegeben werden. Zusätzlich wird das Audiosignal auf einer longitudinal-Cue-Spur aufgezeichnet um ein einfaches Monitoring beim schnellen Umspulen zu erreichen. Hinzu kommt die Aufzeichnung einer CTL-Spur, mit deren Hilfe eine schnellere Servoreaktion erzielt werden soll. Ebenso wie die Cue-Spur ist auch diese Spur nur bei DVCPro zu finden. Die Längsspuren liegen an den Bandrändern (Abb. 8.129).

Die Signalverarbeitungvon DVCPro unterscheidet sich in Europa von der beim Standard-DV-Format, da hier, ebenso wie in NTSC-Ländern, mit 4:1:1-Abtastung gearbeitet wird. Nach der Abtastung entstehen für jedes Bild 720 x 576 Luminanzwerte und für jedes Farbdifferenzsignal 180 x 576 Werte.

Bei DVCPro kann die Standard-Cassettengrößen verwendet werden, die aufgrund der erhöhten Bandgeschwindigkeit eine Spielzeit von maximal 123 Minuten erlaubt. Darüber hinaus können nur bei diesem Format Midsize-Cassetten eingesetzt werden, deren Abmessungen zwischen denen des Standard- und Mini-DV-Formats liegen und damit eine Spielzeit von 63 Minuten bieten (Tabelle 8.15). Der Cassettenaustausch zwischen DVCAM und DVCPro war anfänglich nur für DVCPro möglich und für DVCam nicht. Neuere DVCAM-Geräte erlauben auch die Wiedergabe von DVCPro. Beide Formate sind abwärtskompatibel zu DV.

Abb. 8.130. DVCPro50-Recorder [164]

8.8.5.4 DVCPro50

Das Format DVCPro war im Bereich der elektronischen Berichterstattung (EB) erfolgreich, für den Einsatz im Produktionsbereich gab es einige Hindernisse: Die Chrominanzauflösung ist insbesondere für Key-Effekte zu gering, da sie nicht einer 4:2:2-Abtasung entspricht, die Datenreduktionsfaktor ist zu hoch und es fehlen zwei Audiokanäle. Um diese Nachteile auszuräumen, wurde DVCPro zu DVCPro50 weiterentwickelt. Der Name rührt daher, dass statt einer Video-Nettodatenrate von 25 Mbit/s jetzt 50 Mbit/s aufgezeichnet werden. Die Bandgeschwindigkeit wurde gegenüber DVCPro verdoppelt, um die höhere Datenrate aufzeichnen zu können. Aus Kompatibilitätsgründen werden zwei DVCPro-Schaltungen parallel betrieben, die Datenwortbreite von 8 Bit wird beibehalten. Aus der Addition von zweimal 4:2:0 ergibt sich quasi ein 4:2:2:4-Abtastmuster. Statt aber die Y-Samples doppelt aufzuzeichnen wird der Speicherplatz genutzt, um die Daten aufzunehmen, die aus der Verringerung des Datenreduktionsfaktors entstehen. Der ursprüngliche Faktor war 5:1. Die neue 4:2:2-Abtaststruktur erfordert 2/3 der Werte des 4:2:2:4-Abtastmusters, sodass der Reduktionsfaktor auf 2/3 von 5:1, also auf 3,3:1 gesenkt werden kann. Die Verdopplung der Speicherkapazität ergibt dabei auch die gewünschte Verdopplung der Audiokanäle. Abbildung 8.130 zeigt einen DVCPro50-Recorder. Mit diesen technischen Parametern genügt DVCPro50 den Ansprüchen die an MAZ-Formate im SD-Produktionsbereich gestellt werden.

8.8.5.5 D9

Das als D9 bezeichnete Format von JVC wurde unter der Bezeichnung Digital S auf den Markt gebracht. Damit wird auf eine Brücke zum S-VHS-Format verwiesen, auf dessen mechanischen Eigenschaften das Digitalformat beruht und zu dem es abwärtskompatibel ist. Es wird ein 1/2“-breites Magnetband in einer Cassette verwendet, die äußerlich der S-VHS-Cassette entspricht.

Bei D9 wird ein 4:2:2 abgetastetes Signal mit 8 Bit aufgezeichnet. Dabei wird die gleiche Signalverarbeitung wie bei DVCPro50, also mit zwei parallel arbeitenden DV-Codern, eingesetzt. Die Spurlage unterscheidet sich natürlich von der bei DV, in Abb. 8.131 wird deutlich, dass sie aus Kompatibilitätsgründen wie das S-VHS-Format zwei Audiospuren am oberen und eine am unteren Rand des Bandes enthält. Dazwischen liegen die für die Digitaldaten genutzten Schrägspuren, die so aufgeteilt sind, dass der Audio- und Subcode-Bereich in der Bandmitte liegen. Hinzu kommt das ITI-Segment, das nicht nur am Spurbeginn sondern auch in der Mitte und am Ende aufgezeichnet wird.

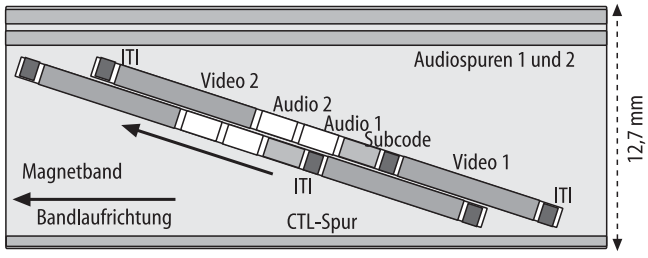


Abb. 8.131. Spurlage beim D9-Format

Abbildung 8.132 zeigt einen D9-Recorder, Tabelle 8.16 die mechanischen Parameter. Das D9-Format bot zu günstigen Preisen eine hohe Signalqualität, die in vielen Fällen auch für hochwertige Produktionen ausreicht, allerdings war das Format aus Markt- und Kompatibilitätsgründen nur wenig verbreitet.

8.8.5.6 Digital 8

Ein weiteres Beispiel für die Verwendung der DV-Datenreduktion ist das zweite Digitalformat von Sony, das für den Heimanwenderbereich gedacht ist. Es wird in Anlehnung an das Video8- bzw. Hi8-Format mit Digital8 oder D8 bezeichnet. Das Format arbeitet DV-gemäß mit 4:2:0-Abtastung bei 8 Bit und wurde entwickelt, um eine Abwärtskompatibiltät zu dem im analogen Camcorder-Markt weit verbreiteten Format Hi8 zu erreichen. D8-Geräte zeichnen auf Hi8-kompatiblen Cassetten nur digital auf, sind aber in der Lage, analoge Video8- oder Hi8-Aufzeichnungen wiederzugeben, wobei die Signale sogar gewandelt und über i-Link (IEEE 1394) digital ausgegeben werden können. Eine 90-Minuten Hi8-Cassette reicht für 60 Minuten Digitalaufzeichnung aus. Das Format hatte nur geringe Bedeutung.

8.8.6 Digitale Komponentenaufzeichnung mit MPEG-Datenreduktion

Vor allem die Firma Sony favorisierte den Einsatz von MPEG-2 im professionellen Umfeld. Sie setzte im Markt für preiswertere Produktionen auf DVCAM und wollte für Broadcastzwecke die MPEG-2-Codierung in allen Bereichen der Produktion etablieren.

8.8.6.1 Betacam SX

Betacam SX war das erste professionelle MAZ-Format, das mit einer Datenreduktion nach MPEG-2 arbeitete (Abb. 8.133). Es nutzt den Umstand, dass hier auch die zeitliche Dimension zur Datenreduktion genutzt werden kann die bei gleicher Qualität größere Reduktionsfaktoren als bei DV ermöglichen. Es ist für den professionellen Bereich konzipiert, weist Ähnlichkeiten mit dem Format Digital Betacam auf und ist abwärtskompatibel zu Betacam SP. Als typisches Einsatzgebiet wird der News-Bereich angegeben, wo es auf kostengünstige Produktion ankommt. Geringe Kosten werden durch preiswerte Bandmaterialien und geringen Bandverbrauch durch einen hohen Datenreduktionsfaktor erreicht. Die Datenreduktion nach MPEG-2 bietet dabei die Möglichkeit der Verwendung eines Reduktionsfaktors von 10:1 durch Nutzung von Groups of

Abb. 8.132. Digital S (D9)-Recorder [118]

Tabelle 8.16. Geometrische Parameter der Formate D9, S-VHS und Betacam SX

	D9	S-VHS	Betacam SX	D10
Magnetbandbreite	12,7 mm	12,7 mm	12,7 mm	12,7 mm
Bandgeschwindigkeit	5,77 cm/s	2,3 cm/s	5,96 cm/s	5,37 cm/s
Relativgeschwindigkeit	14,5 m/s	4,85 m/s		12,7 m/s
Videospurlänge	88 mm	100 mm		
Videospurbreite	20 µm	49 µm	32 µm	21,7 µm
Schrägspuren/Bild	12 (bei 625Z)	2	12/GOP	8 (bei 625Z)
max. Spieldauer (S, M, L)	34/64/104 Min.	240 Min.	60/ - /184 Min.	71/ - /220 Min.

Pictures in der IB-Form. Vor der Datenreduktion liegt ein digitales Komponentensignal mit Abtastraten im Verhältnis 4:2:2 und 8-Bit-Auflösung vor. Die MPEG-Reduktion wird nach dem professional Profile @ Main Level durchgeführt, d. h. das 4:2:2-Verhältnis bleibt erhalten und es werden auch Zeilen der vertikalen Austastlücke aufgezeichnet (insgesamt 608).

Es wird auf 1/2"-Metallpartikelband aufgezeichnet, das preisgünstig ist und mit einer Geschwindigkeit von ca. 6 cm/s läuft (Tabelle 8.16). Das Spurbild ähnelt dem von Digital Betacam. Beide Formate sind kompatibel zu Betacam SP. Das Format zeichnet einen MPEG-Elementarstrom nach 4:2:2P@ML auf, intern muss dieser aber, wie bei jedem Bandformat, durch einen Umordnungsprozess für die Bandaufzeichnung optimiert werden [34].

Betacam SX bietet eine bessere Bildqualität als Betacam SP und ist aufgrund der 4:2:2-Abtastung besser für den Produktionsbereich geeignet als DVCPro, das für ein ähnliches Marktsegment konzipiert ist. Problematisch kann allerdings die Tatsache sein, dass aufgrund der IB-Bildgruppen im Prinzip nur eine Schnittgenauigkeit von zwei Bildern erreicht wird. Dieser Umstand wurde als nicht mehr akzeptabel angesehen und daher wurde eine Möglichkeit für den bildgenauen Schnitt auch für die IB-Bildfolgen geschaffen. Maßgabe ist dabei, dass die IBIB-Reihenfolge nicht unterbrochen wird. Das Problem ist, dass ein B-Frame am Schnittpunkt eines seiner Referenzbilder verliert. Zur Lösung des Problems werden die Frames am Schnittpunkt mit Hilfe von speziellen Leseköpfen vor dem Einfügen einer neuen Szene gelesen und die kritischen Bilder werden von B- in B_u-Bilder umcodiert, die nur noch unidirektional von den Nachbarbildern abhängen.

Abb. 8.133. Betacam SX-Recorder [208]

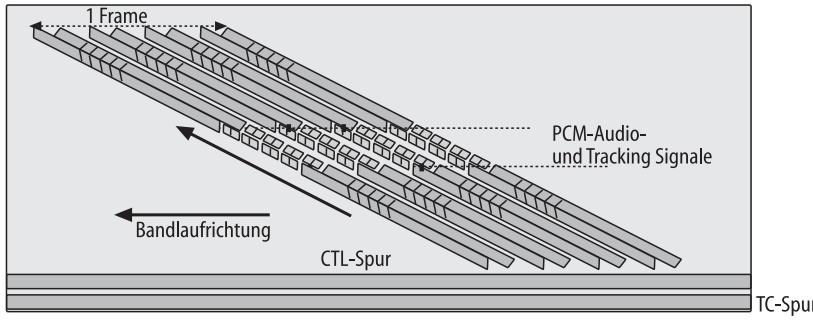


Abb. 8.134. Spurlage beim D10-Recorder

8.8.6.2 D10 (IMX)

Betacam SX wurde, insbesondere in Europa, nicht sehr gut am Markt angenommen, da für ein hochwertiges professionelles Format heute eine Datenrate von 50 Mbit/s mit bildgenauem Schnitt gefordert wird. Ein solches wurde von Sony dann unter der Bezeichnung IMX-Recorder entwickelt und als D10-Format standardisiert. Der IMX-Recorder bietet ähnliche Merkmale wie ein SX-Recorder, nutzt aber die MPEG-I-Frame-only-Codierung bei Aufzeichnung eines Signals im 4:2:2P@ML-Format mit 50 Mbit/s (Tabelle 8.24). Es gibt damit, im Gegensatz zu Betacam SX, keinen zeitlichen Bezug der Bilder untereinander und der Zugriff auf jedes Einzelbild ist problemlos möglich. Das Format bietet acht Audiokanäle PCM, 16 Bit/48 kHz oder vier Kanäle mit 24 Bit.

Die Nutzung des 1/2“-Bandes erlaubt eine optionale Abspielkompatibilität zu Digital Betacam, Betacam SX und Betacam SP. Die Bandgeschwindigkeit liegt mit weniger als 6 cm/s noch unter dem Niveau von Betacam SX und da die Spurbreite verkleinert werden konnte (Tabelle 8.16) wird mit der großen Cassette eine Spieldauer von 220 Minuten erreicht, was eine Verdopplung gegenüber Betacam SP darstellt. Abbildung 8.134 zeigt die Spurlage auf dem Band Neben der Digitaldatenaufzeichnung im Schrägspurbereich gibt es zwei Längsspuren für CTL- und Timecode-Daten.

Ein IMX-Recoder ist in Abbildung 8.135 dargestellt. Er bietet alle Funktionalitäten eines Studiorecorders und eine einfache Integration in die bestehende Studioumgebung. Die weitgehende Abspielkompatibilität, hat viele Rundfunkanstalten bewogen, D10 als digitales Nachfolgeformat für Betacam SP zu wählen.

Abb. 8.135. D10-Recorder [208]

8.8.7 Digitale HD-MAZ

Die unkomprimierte Aufzeichnung eines HD-Signals ist anspruchsvoll, denn sie erfordert die Verarbeitung einer Datenrate von mehr als 1 Gbit/s wenn ein 4:2:2-HD-Komponentensignal bei 10-Bit-Auflösung verwendet wird. Tatsächlich ist nur ein MAZ-Format in der Lage, diese Datenrate unkomprimiert aufzuzeichnen, nämlich das erste HD-Cassettenformat mit der Bezeichnung D6.

Alle anderen HD-Formate beherrschen nur eine erheblich geringere Datenrate und arbeiten daher in den meisten Fällen mit Chromaunterabtastung und stärkerer Datenreduktion. Wenn keine Unterabtastung der Farbdifferenzkomponenten verwendet wird, wird die Bezeichnung 4:4:4 beibehalten, die aus der Digitalisierung von SD-Signalen entlehnt ist. Während mit 4:4:4 bei SD-Signalen die Abtastung der drei Kanäle des RGB-Signals mit je 13,5 MHz gemeint ist, bezieht sie sich bei HD-Signalen auf drei RGB-Kanäle, die jeweils mit 74,25 MHz abgetastet werden. Diese qualitativ anspruchsvollste Variante der Videodatenspeicherung wird nur von wenigen Speichersystemen wie z. B. HDCAM SR unterstützt. In der Regel wird mit Y, Cr, Cb-Komponenten und Chromaunterabtastung gearbeitet, d. h. mit 4:2:2, also 74,25 MHz für die Y-Abtastung und 37,125 MHz für C_R und C_B. Da die HD-Signale auch in diesem Fall noch sehr hohe Datenraten erzeugen, wird bei vielen Formaten noch eine weitere Unterabtastung verwendet, wobei, im Gegensatz zur SD-Aufzeichnung, auch die Y-Komponente einbezogen wird.

In diesem Zusammenhang existieren Bezeichnungen wie 3:1:1 oder 3:1,5:1,5. Damit ist gemeint, dass in der Zahlenrelation zu 4:2:2 die Abtastraten und entsprechend die horizontalen Bildpunktzahlen für alle drei Komponenten reduziert werden. So steht 3:1:1 dafür, dass die Luminanzinformation eines Bildes von 1920 x 1080 auf 1440 x 1080 Bildpunkte reduziert wird und bei den Chrominanzanteilen nur eine Auflösung von 480 x 1080 zur Verfügung steht. Abbildung 8.136 zeigt eine Übersicht, in der die Unterabtastung am Beispiel des HD-Formats mit 1080 Zeilen veranschaulicht ist.

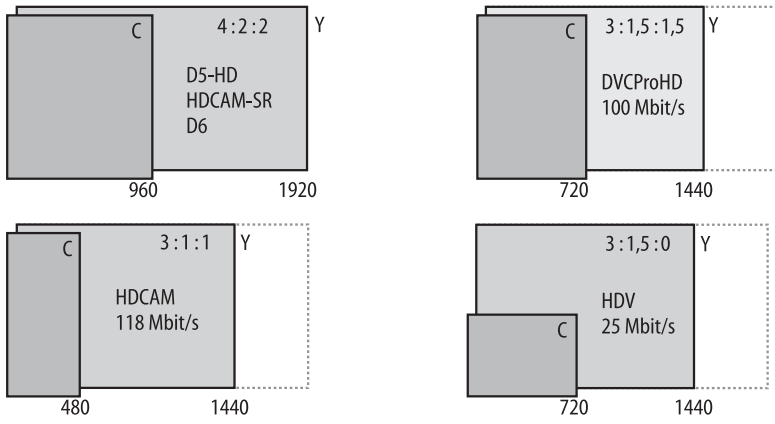


Abb. 8.136. Veranschaulichung der Unterabtastungen bei verschiedenen HD-MAZ-Formaten

Tabelle 8.17. Parameter der HD-MAZ-Formate

Format	Sampl.	Bit	Quellrate	aufgezeich.	1080p/50	1080i/25	720p/50
HDCAM SR	4 : 4 : 4	10	3110	880 Mbit/s	3,5 : 1		
HDCAM SR	4 : 4 : 4	10	1555	440 Mbit/s		3,5 : 1	
HDCAM SR	4 : 2 : 2	10	1037	440 Mbit/s		2,4 : 1	2,1 : 1 (921)
D5 HD	4 : 2 : 2	10	1037	232 Mbit/s		4,5 : 1	
D5 HD	4 : 2 : 2	8	737	232 Mbit/s			3,2 : 1
D6/Voodoo	4 : 2 : 2	8	829	829 Mbit/s		1 : 1	
HDCAM	3 : 1 : 1	8	518	118 Mbit/s		4,4 : 1	
DVCPro HD	3 : 1,5 : 1,5	8	622	100 Mbit/s		6,2 : 1	5,5 : 1 (553)
HDV	3 : 1,5 : 0	8	466	25 Mbit/s		18,6 : 1	
HDV	4 : 2 : 0	8	553	19,7 Mbit/s			28 : 1

Tabelle 8.17 zeigt eine Übersicht über die bandbasierten HD-Aufzeichnungsverfahren und die sich ergebenden Datenreduktionsfaktoren für unterschiedliche Betriebsparameter der einzelnen Formate. Die Werte in Klammern geben die veränderten Quelldatenraten in Mbit/s an.

8.8.7.1 Das D6-Format/Voodoo

Mit dem D6-Format existiert bereits zu Beginn der 90er Jahre ein HDTV-Aufzeichnungsformat, das ohne Datenreduktion auskommt (Abb. 8.137). Bei D6 werden ein Videosignal und 12 Audiokanäle auf 3/4"-Band mit einer Gesamtdatenrate von 1,2 Gbit/s aufgezeichnet. Die Bandgeschwindigkeit beträgt dabei ca. 50 cm/s. Mit den gleichen drei Cassettengrößen wie bei D1 werden Spieldauern von 8, 28, oder 64 Minuten erreicht. Tabelle 8.11 zeigt weitere mechanische Spezifikationen. Die Spurlage ähnelt der von D1, allerdings befinden sich die Audiosektoren am Schrägspurrand (Abb. 8.115).

Das Format wurde im Laufe der Zeit so weiterentwickelt, dass es verschiedene Bildraten einschließlich der 24p-Aufzeichnung beherrscht und trug dann die Bezeichnung Voodoo. Um dieses Format im Bereich des digitalen Films einsetzen zu können, wurde darüber hinaus eine 10 Bit-RGBA-Aufzeichnungsoption (4:4:4:4) entwickelt, die sich einer Datenreduktion um den Faktor 3 bedient, die auf dem Wavelet-Algorithmus beruht. Das System kann auch einfach als schneller und hochkapazitiver Datenspeicher für Filmdaten, z. B. im DPX-Format, eingesetzt werden. D6 war teuer und wenig verbreitet.

Abb. 8.137. D6-Recorder [219]

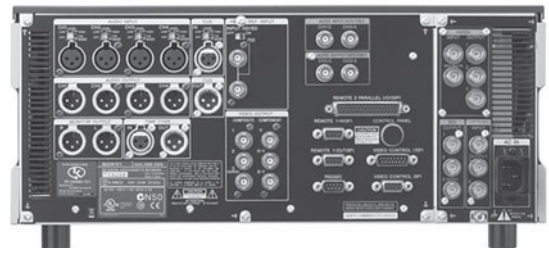


Abb. 8.138. HDCAM-Recorder-Rückseite [208]

8.8.7.2 HDCAM und HDCAM SR

Das erste komprimiert arbeitenden HD-MAZ-System, dass zum Ende des letzten Jahrhunderts von der Firma Sony unter der Bezeichnung HDCAM (D11) entwickelt wurde, war lange Zeit ohne Konkurrenz. Beim Format HDCAM, das zur Betacam-Familie gehört, steht nur eine Datenrate von 183 Mbit/s bzw. eine Netto-Videodatenrate von 118 Mbit/s zur Verfügung. Um damit einen HD-Aufzeichnung zu ermöglichen, wird, wie beschrieben, die Eingangsdatenrate durch Filter mit 3:1:1-Unterabtastung auf 518 Mbit/s für das aktive Bild reduziert. Zusätzlich werden die 1080i-Signale mit einem Reduktionsfaktor von 4,4 komprimiert. Nach außen werden 10-Bit-Signale bereitgestellt, während der interne Reduktionsprozess mit 8 Bit arbeitet. HDCAM-Recorder sind separat verfügbar, können aber auch so kompakt gebaut werden, dass sie mit Videokameras kombiniert werden können. Diese Geräte sind dann mit den Datenreduktions- und Subsampling-Nachteilen der Recorder behaftet. Wenn das Signal an einer HD-SDI-Schnittstelle ausgegeben wird, kann es unter Umgehung des internen Recorders auch ggf. unkomprimiert auf ein anderes Speichermedium aufgezeichnet werden. Abbildung 8.138 zeigt die Rückseite eines HDCAM-Recorders, Die Frontseite ähnelt sehr der des IMX-Gerätes (Abb. 8.135).

Das HDCAM-Format war erfolgreich, aber recht teuer und wegen der 3:1:1-Unterabtastung nicht für anspruchsvolle Produktionen geeignet. Daher entwickelte Sony weitere HD-MAZ-Formate: HDCAM wurde zum hochwertigeren, Format HDCAM SR ausgebaut und für den semiprofessionellen und Heimanwenderbereich stehen die preisgünstigeren Formate HDV und XDCAM HD zur Verfügung, wobei letzteres eine bandlose Aufzeichnung ist (s. u).

HDCAM SR ist im Bereich der komprimiert arbeitenden Systeme z. Zt. das qualitativ hochwertigste HD-MAZ-Format, denn es stellt eine Videonettodatenrate von 440 Mbit/s zur Verfügung, die im High-Quality-Modus, bei verdoppelter Bandgeschwindigkeit, auf 880 Mbit/s gesteigert werden kann. Die verwendete Kompression ist nach dem MPEG-4 Studio Profile standardisiert. Es werden 1080i/25-Signale wahlweise mit 4:2:2 oder 4:4:4 aufgezeichnet, wobei zwischen den Datenreduktionsfaktoren 2,4:1 oder 3,5:1 umgeschaltet wird. Die Datenreduktion geschieht einzelbildweise. Durch die mögliche Verdopplung der Datenrate ist die Verarbeitung von 1080p/50-Signalen möglich, sodass das Format sehr gut für hochqualitative Anwendungen, auch im Bereich des digitalen Films, geeignet ist. Abbildung 8.139 zeigt einen stationären und einen portablen Recorder.

Abb. 8.139. Stationärer und mobiler HDCAM SR-Recorder [208]

8.8.7.3 HD-D5

Sonys Konkurrent Panasonic entwickelte ein komprimiert aufzeichnendes MAZ-System auf der Basis des Formats D5, das entsprechend als HD-D5 oder D15 bezeichnet wird (Abb. 8.112). Das D5-Standardformat bietet bereits eine Videobitrate von netto 232 Mbit/s. Daher können HD-Komponentensignale mit hoher Auflösung (4:2:2) bei 10 bit aufgezeichnet werden und trotzdem braucht nur ein Reduktionsfaktor von 4,5 eingesetzt zu werden, wenn als Signalform 1080i/25 vorliegt. HD-D5 stellt damit ein HD-Format mit vergleichsweise hoher Qualität dar, wurde in der Praxis aber selten verwendet.

8.8.7.4 DVCProHD

Als zweites HD-MAZ-Format von Panasonic wurde das DVCProHD-Format (D12) angeboten, bei dem allerdings nur eine Datenrate von 100 Mbit/s zur Verfügung steht. Sie wird über die Verdopplung der Bandgeschwindigkeit eines DVCPro50-Systems erreicht. Das Format basiert also auf DV und zeichnet auf das schmale 1/4"-Band nur 8-bit-Signale auf. Dabei ist es möglich, mit 1080i/25- ebenso zu arbeiten wie mit 720p-Signalen. Ähnlich wie bei HDCAM, wird das Signal vor der Datenreduktion mit 3:1,5:1,5 unterabgetastet. Die Datenreduktion arbeitet dann mit ca. 6,2:1. Abbildung 8.140 zeigt einen Recorder dieses Formates der in der Lage ist, große DV-Cassetten aufzunehmen,.

8.8.7.5 HDV

HDV wurde von Sony und anderen Firmen als HD-MAZ-System für den Heimanwenderbereich entwickelt. Die Grundidee ist, die Geräte des erfolgreichen DV-Formates HD-tauglich zu machen. DV erlaubt die Herstellung preisgünstiger Geräte, die DV-Videodatenrate von 25 Mbit/s wurde für HDV beibehalten.

Da HD-Signale im Vergleich zu SD das etwa 5-fache Datenvolumen haben, muss für HDV eine effektivere Datenreduktion verwendet werden als bei DV. Es wird MPEG-2 Long GOP verwendet, d. h., dass auch die zeitliche Abhängigkeit der Bilder in einer Group of Pictures zur Datenreduktion ausgenutzt wird. Zusätzlich werden die Signale vor der Reduktion unterabgetastet.

HDV existiert in zwei Varianten, beide für 8-bit-Signale: Die Variante HD1 arbeitet mit 720p, mit 25 oder 50 Hz bei 4:2:0. Die GOP-Länge beträgt 14 und es werden 19,7 Mbit/s ausgenutzt, der Reduktionsfaktor bei 25p beträgt 14:1. HD2 ist für 1080i/25 konzipiert. Die Kompression und die Bildauflösung richtet sich nach MPEG-2 im MP@H14, also dem High Level mit 1440 Bildpunkten horizontal mit 8 Bit und 4.2:0-Chromaunterabtastung. Anders ausgedrückt ist das eine Unterabtastung von 3:1,5:0. Die GOP-Länge ist 15 und es wird eine

Abb. 8.140. DVCProHD-Recorder [164]

Tabelle 8.18. HDV-Spezifikationen

	HD1 720p	HD2 1080i	1080p
Videosignal	720p /24, 25, 30, 50, 60 Hz	1080i / 25, 30 Hz	24, 25, 30Hz
Bildpunktzahl	1280 x 720	1440 x 1080	
Samplingfrequenz	74,25 MHz	55,7 MHz	
Unterabtastung	4:2:0		
Quantsierung	8 Bit		
Datenreduktion	MPEG-2 MP@HL	MPEG-2 MP@HL-14	
Datenrate	19,7 Mbit/s	25 Mbit/s	
Audiokompression	MPEG-1 Layer 2		
Audiodigitalisierung	48 kHz/16 bit (2 Kanäle)		
Audiodatenrate	384 kbit/s		
Datenformat	MPEG-Systems TS	MPEG-Systems PES	

Videodatenrate von 25 Mbit/s bei einer Datenreduktion von 19:1 erzielt. Bei HD1 wird der MPEG-Transportstrom aufgezeichnet, bei HD2 der PES. Für die Audiokompression wird MPEG1 Layer 2 verwendet mit 384 kbit/s für zwei Audiokanäle (Tabelle 8.18). Als Datenstrominterface ist Firewire vorgesehen.

Für die Recorder können alle DV-Bänder zum Einsatz kommen, es werden jedoch besonders hochwertige Materialien empfohlen. Neben HDV-Bändern sind auch DVCAM-Bänder gut geeignet. Diese bieten Spieldauern, die das 1,5-fache des aufgedruckten Wertes betragen (z. B. 60 Min. statt 40 Min. bei Mini-DV-Kassetten), da das Band bei HDV genauso schnell läuft wie bei DV und nicht 1,5-fach schneller wie bei DVCAM. HDV bietet eine relativ gute Qualität und hat eine hohe Verbreitung, da ähnlich wie bei DV oft versucht wird, die preisgünstigen Geräte auch im professionellen Sektor einzusetzen.

8.8.7.6 D-VHS

Diese Bezeichnung steht für ein Heimanwenderformat, das einen MPEG-2-Datenstrom auf VHS-Cassetten aufzeichnet. Damit VHS- und S-VHS-Aufzeichnungen wiedergegeben werden können, entspricht die Spurlage der von VHS, wobei in den Videoschrägspuren Digitaldaten gespeichert werden. Auf einem D-VHS-Band können bis zu 44 GByte gespeichert werden. Die maximale Datenrate beträgt 14,1 Mbit/s, mit höherer Geschwindigkeit stehen 28,2 Mbit/s für HDTV (1080i oder 720p) zur Verfügung. D-VHS hatte kaum Verbreitung.

Abb. 8.141. Stationärer Recorder und Camcorder im HDV-Format [208]

8.9 Bandlose digitale Signalaufzeichnung

Aufgrund der hohen Datenraten, die bei Videosignalen auftreten, waren lange Zeit bandbasierte Speicher unumgänglich. Im Laufe der Zeit wurden jedoch die Speicher der IT-Technologie so groß und schnell, dass sie mit den Magnetbandverfahren konkurrieren konnten. Es ist zu erwarten, dass zukünftig keine bandbasierten Speicherverfahren mehr entwickelt werden, denn die IT-Technologie bietet die Vorteile eines Massenmarktes mit geringen Preisen und die Unabhängigkeit von den Spezifika des aufgezeichneten Signals. D. h. es ist prinzipiell unerheblich, ob Audio- oder Videodaten in SD oder in HD aufgezeichnet werden oder welcher Codec benutzt wird. Damit signalisiert die Etablierung der bandlosen Aufzeichnung den Beginn der zweiten Generation der Digitalisierung, die durch diese Entkopplung der Aufzeichnungs- und Verarbeitungssysteme von den Besonderheiten des Videosignals gekennzeichnet ist.

Als Datenträger kommen vorwiegend magnetische und optische Platten zum Einsatz. Mit dem Erfolg der DVD, aber besonders durch die Weiterentwicklung zur Blu-ray Disc, kamen dann auch immer öfter optische Speicher zum Einsatz. Daneben werden aber neuerdings immer mehr Festspeicherkarten verwendet, die eine Aufzeichnung ohne die Verwendung mechanischer Bauteile ermöglicht. Festspeicher bieten die größte Zugriffsgeschwindigkeit, sind allerdings teuer und es können dementsprechend bisher nur relativ geringe Video-Speicherkapazitäten realisiert werden. Magnetplatten haben eine höhere Kapazität bei einer Zugriffsgeschwindigkeit, die bei Nutzung intelligenter Verbindungen mehrerer Platten auch für den HD-Videobereich geeignet ist.

Ein weiterer großer Vorteil bandloser gegenüber den klassischen MAZ-Verfahren ist der Umstand, dass ohne Umspulen in sehr kurzer Zeit ein wahlfreier Zugriff auf beliebige Abschnitte des aufgezeichneten Videomaterials möglich ist. Damit bieten die bandlosen Speicher Vorteile, die vor allem für das nonlineare Editing unverzichtbar sind. Nonlineares Editing bezeichnet einen Schnittbetrieb bei dem es möglich ist, wie beim Filmschnitt, innerhalb einer bereits aufgezeichneten Sequenz eine weitere einzufügen, ohne die bereits aufgenommene dabei zu überschreiben. Ein zweites Beispiel, bei dem die bandlose Speicherung einen besonders großen Vorteil bietet, ist die automatische

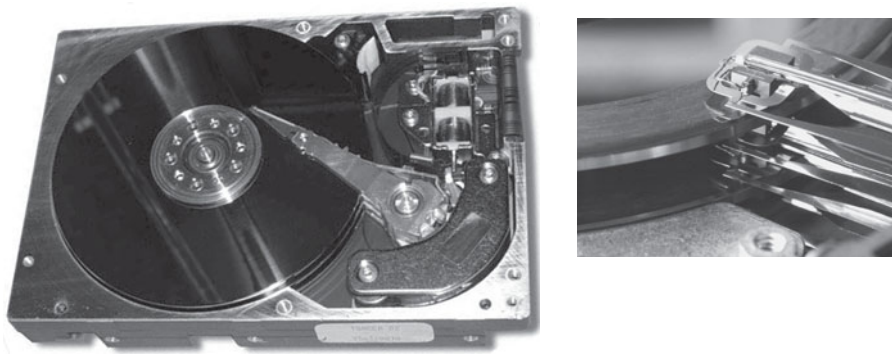


Abb. 8.142. Festplatte geöffnet und Magnetköpfe an der Festplatte

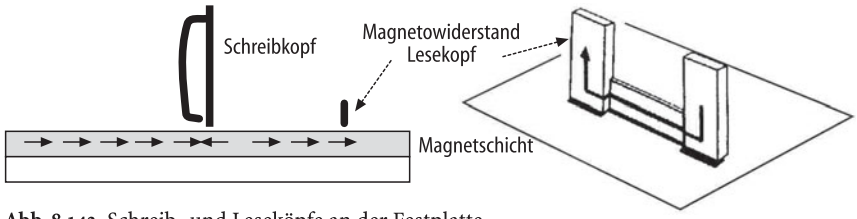


Abb. 8.143. Schreib- und Leseköpfe an der Festplatte

Sendeabwicklung. Der schnelle Zugriff auf häufig wiederkehrende Beiträge wird hier wesentlich vereinfacht, außerdem besteht die Möglichkeit, die Daten komprimiert zu speichern. Über digitale Netze können die Beiträge ohne Kopiervorgang schnell eingespielt und ausgetauscht werden.

8.9.1 Magnetplattensysteme

Diese für die Computertechnik wichtigsten Massenspeicher haben auch beim bandlosen Videoeinsatz die größte Bedeutung. Magnetplattenspeicher (Hard Disks) werden als mechanisch abgeschlossene Festplatten realisiert. Abbildung 8.142 zeigt sie ohne Abdeckung. Das Grundprinzip der Magnetspeicherung ist das gleiche wie bei der digitalen Magnetbandaufzeichnung. Die Information wird als magnetischer Flusswechsel mit einer angepassten Kanalcodierung in einer Magnetschicht aus Kobaltlegierung gespeichert und lässt sich fast beliebig oft schreiben und lesen. Unterschiede zu Bandsystemen gibt es vor allem im mechanischen Bereich. Fest- und Wechselplattensysteme bestehen aus einem schnell rotierenden Stapel von Aluminium-Platten mit den Standard-Durchmessern 3,5“ und 2,5“. Die Platten sind meist beidseitig mit einer Magnetschicht versehen. Die Abtastköpfe umfassen zangenartig die Platte (Abb. 8.142) und schweben auf einem Luftpolster im Abstand von ca. 50 nm darüber. Die Köpfe werden radial bewegt, aufgrund der Plattenrotation können so alle Bereiche der Platte erfasst werden. Die Plattenumdrehungszeit und die Antriebsmechanik bestimmen entscheidend die Zugriffsgeschwindigkeit.

Die Informationen werden in konzentrischen Spuren gespeichert, übereinander liegende Spuren werden zu Zylindern zusammengefasst. Eine Spur wird in Sektoren geteilt, die die kleinsten Dateneinheiten bilden und typische Größen von 512, 1024 oder 4096 Byte aufweisen. Die Nutzdaten werden mit Header und Prüfsummen versehen, die je nach Format pro Sektor 40 bis 100 Byte umfassen.

Die Schreib/Leseköpfe in Festplatten können wie MAZ-Magnetköpfe induktiv arbeiten, d. h. zum Schreiben einer Information wird ein Strompuls durch eine Spule geschickt und das am Schreibkopf austretende Magnetfeld reicht zur Magnetisierung der Platte aus (Abb. 8.143). Der Lesevorgang kann auch nach dem Induktionsgesetz erfolgen, die Verkleinerung der Köpfe zur Erzielung höherer Speicherdichten ist aber bezüglich des Lesevorgangs an ihrer technologischen Grenze angelangt. Daher wird heute statt des Schreibkopfs selbst, ein dort integriertes Magnetowiderstandselement für den Lesevorgang

verwendet (Abb. 8.143). Durch dieses Element fließt ein kleiner Strom, über den die Widerstandsänderung registriert wird, die durch das variable Magnetfeld hervorgerufen wird [5].

Wichtige Parameter für die Leistung einer Festplatte sind Kapazität, Zugriffszeit und Datentransferrate. Die Transferrate ergibt sich aus der Rotationsgeschwindigkeit und der Anzahl von Bits pro Umdrehung. Die mittlere Zugriffszeit setzt sich aus der mittleren Suchzeit und dem Zeitraum zusammen, der gebraucht wird, bis sich der gewünschte Sektor unter den Schreib-Lesekopf gedreht hat. Die übertragbare Datenrate, die bei mehr als 50 MByte/s liegt, wird neben der Zugriffszeit wesentlich durch die Schnittstellen für den Datentransport bestimmt. Die Kapazitäten der Festplatten konnten im Laufe der Zeit auf mehr als ein TByte gesteigert und die mittleren Zugriffszeiten konnten auf unter 10 ms gesenkt werden. Die letzten Steigerungen gelangen insbesondere durch die Entwicklung von Magnetköpfen für das „Perpendicular Recording", bei dem sich die magnetischen Bereiche in die Tiefe statt in die Breite der Magnetschicht erstrecken und somit erheblich dichter liegen.

Das Datenvolumen der Festplatte erscheint nach außen hin als Einheit, eine Logik blendet fehlerhafte Sektoren aus. Die Platte kann durch Partitionen in mehrere virtuelle Festplatten aufgeteilt werden, die dann wiederum mit unterschiedlichen Dateisystemen formatiert werden können. Zur Verbesserung der Zugriffszeit kann die Datenverarbeitung mittels des Nativ Command Queing (NCQ) dafür sorgen, dass bei vielen Anfragen an die Platte unnötige Bewegungen des Schreib-/Lesekopfes vermieden werden, indem das System eigenständig bestimmt, wie die Abarbeitung vorgenommen werden soll.

8.9.1.1 Schnittstellen

Die für den Computerbereich entwickelten Standardfestplatten weisen auch die Schnittstellen der Computertechnik auf, die wichtigsten Typen sind hier IDE und SCSI mit ihren Nachfolgern SATA und SAS. Bei beiden befinden sich die Steuerelektronik und der Festplattencontroller auf der Harddisk, damit werden unbekannte Laufwerke erkannt und ein Management defekter Bereiche auf der Platte einbezogen. Bei der (E)IDE-Schnittstelle ist zusätzlich der Hostadapter für die direkte Anbindung an den Computer auf die Platte verlegt, wodurch die Festplattencontroller sehr preiswert wurden, aber die Flexibilität eingeschränkt ist, da das Hostsystem auf IBM-kompatible PC festgelegt ist. Das Controller-System übernimmt die Pufferung der Daten, die erforderlich ist, weil ein Sektor immer am Stück gelesen oder geschrieben werden muss. Weiterhin das Caching, die Zwischenspeicherung von Daten, auf die besonders häufig zugegriffen wird, und die Erstellung von Defektlisten, in denen fehlerhafte Blöcke registriert werden, damit den Anwendern nur der fehlerfreie Restbereich zugänglich bleibt.

IDE steht für Integrated Device Electronics und ist eine parallele Schnittstelle mit 16 Datenleitungen. Das zugehörige Softwareprotokoll, nach dem manchmal auch die Schnittstelle benannt ist, heißt ATA (auch Advanced Technology Attachment with Packet Interface, ATAPI).

SCSI bedeutet Small Computers System Interface und war die flexiblere Schnittstelle für den professionellen Sektor. Der SCSI-Standard beschreibt elek-

trische Eigenschaften und ein Protokoll für ein geräteunabhängiges Ein- und Ausgabeverfahren. Der Standard wurde bereits 1986 etabliert und ständig weiterentwickelt. SCSI-Anschlusskabel sind 50- oder 68-polig (Wide), es werden entsprechend breite Stecker (Centronics) verwendet. Eine Standardverbindung ist single ended und erlaubt Buslängen von 6 m, bei Fast-SCSI mit mehr als 5 MByte/s ist die Entfernung auf 3 m beschränkt. Jede Datenleitung erfordert am Busabschluss einen Terminator mit 220 Ω gegen 5 V und 330 Ω gegen Masse (Abb. 8.144).

SCSI erlaubt eine Geräteidentifikation mit Hilfe von ID-Nummern. Damit lassen sich durch einfaches Zusammenstecken bis zu sieben Festplatten an einem Computersystem betreiben. Die Elektronik beinhaltet für jedes Gerät einen Controller, der es dem Gerät ermöglicht, als Initiator oder Target zu agieren. Zur Steuerung sind acht Kommandogruppen gebildet worden. Vor einer Verbindung wird zunächst eine Identifizierung der Geräte als Plattenlaufwerk, CD-ROM, etc. durchgeführt. Die SCSI-3 Definitionen unterstützen dabei die Bildung von RAID-Arrays.

Die Steigerung der Transferraten von Parallelschnittstellen wie IDE und SCSI auf Werte, die wesentlich über 320 MByte/s liegen, scheitert an dem Umstand, dass es immer schwieriger wird, den Datenstrom auf allen Leitungen synchron zu halten bzw. das Übersprechen zwischen den einzelnen Leitungen zu unterdrücken. Daher sind die Nachfolger seriell ausgelegt, als Serial ATA (SATA) und als Serial SCSI (SAS). Neben der höheren Datentransferrate bieten sich hier auch Vorteile durch kleinere Steckverbindungen, die zudem so ausgelegt wurden, dass die Festplatten im laufenden Betrieb gewechselt werden können (Hot Plug). Bei SATA betrug die Transferrate zunächst 150 MByte/s, sie konnte im Laufe der Zeit auf 600 MByte/s gesteigert werden.

SAS benutzt die gleiche Steckverbindung wie SATA, die Datenraten erreichen gegenüber SATA die doppelten Werte. SATA 2 definiert inzwischen auch Kabel und Stecker für den externen Betrieb, die als External Serial ATA (eSATA) bezeichnet werden. Abbildung 8.144 zeigt die SATA-Steckverbindung im Vergleich zu IDE- und SCSI.

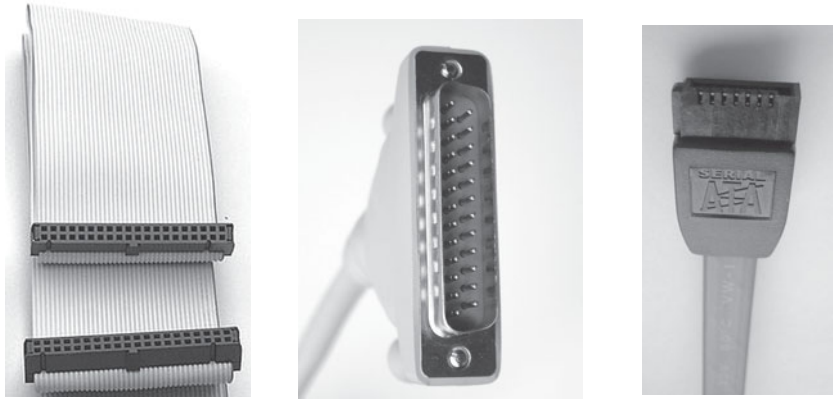


Abb. 8.144. Steckverbindungen für IDE-, SCSI- und SATA-Festplatten

8.9.1.2 RAID-Systeme

Große Speicherkapazitäten werden heute vor allem durch den intelligenten Zusammenschluss vieler preiswerter Standard-Festplatten zu einem Redundant Array of Independent (Inexpensive) Disks (RAID) erreicht [81]. Das Array erscheint nach außen hin als ein großes Speichersystem bei dem eine hohe Übertragungsgeschwindigkeit durch parallelen Zugriff auf die einzelnen Platten erreicht wird. Der Begriff Redundant steht dabei für eine mehrfache Verteilung der Daten auf die einzelnen Platten, womit sich sowohl die Datenrate, als auch die Datensicherheit erhöhen lässt. Der letzte Aspekt ist für den Verbund von besonderer Bedeutung, da sich die Fehlerwahrscheinlichkeiten der einzelnen Platten summieren. Ein einfacher Festplattenzusammenschluss, ohne Fehlerschutz, wird als JBOD (Just a bunch of disks) bezeichnet.

Die kleinste Dateneinheit einer RAID-Konfiguration ist ein Chunk. Kleine Chunks (1 Bit bis 16 Byte) bieten eine hohe Transfergeschwindigkeit, denn die Datei ist auf mehrere Chunks verteilt, auf die parallel zugegriffen wird. Große Chunks (bis einige kB) unterstützen eher die gleichzeitige Bearbeitung mehrerer Daten-Anforderungen. Für die verschiedenen Anwendungsbereiche wurden sieben RAID-Levels definiert, mit verschiedener Berücksichtigung von Datenübertragungsrate und -sicherheit (Abb. 8.146). Die einfachste Konfiguration (RAID 0) steht für eine simple Datenverteilung auf mehrere Platten ohne erhöhte Datensicherheit (Striping). RAID 1 beschreibt eine Spiegelung der Daten, also schlicht die doppelte Aufzeichnung des Materials. RAID 2 schützt die Daten durch einen fehlerkorrigierenden Hamming-Code, der über alle Platten verteilt wird, sodass beim Ausfall einer Platte kein Datenverlust auftritt. Das Verfahren ist aufgrund des großen Anteils an Redundanz nicht effizient und wird kaum eingesetzt. Bei RAID 3 werden kleine Chunks in Byte-Größe gespeichert, wobei das Fehlersicherungssystem des einzelnen Laufwerks ausgenutzt wird. Zusätzlich werden Paritätsdaten ermittelt und auf einem dafür vorgesehenen Laufwerk abgelegt. Damit kann der Ausfall einer Datenplatte ausgeglichen werden. RAID 3 ist aber nicht sehr gut für den Mehrkanalbetrieb geeignet, da dafür große Zwischenspeicher erforderlich werden.

RAID 4 arbeitet wie Variante 3, allerdings mit großen Chunks, es wird aufgrund des unsymmetrischen Schreib- Leseverhaltens selten verwendet. RAID 5 macht das Verhalten symmetrisch, indem die Parity-Daten nicht auf ein sepa-

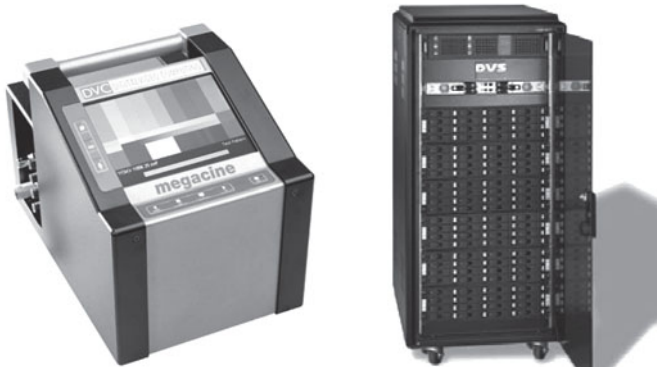


Abb. 8.145. RAID-Festplatten-Arrays

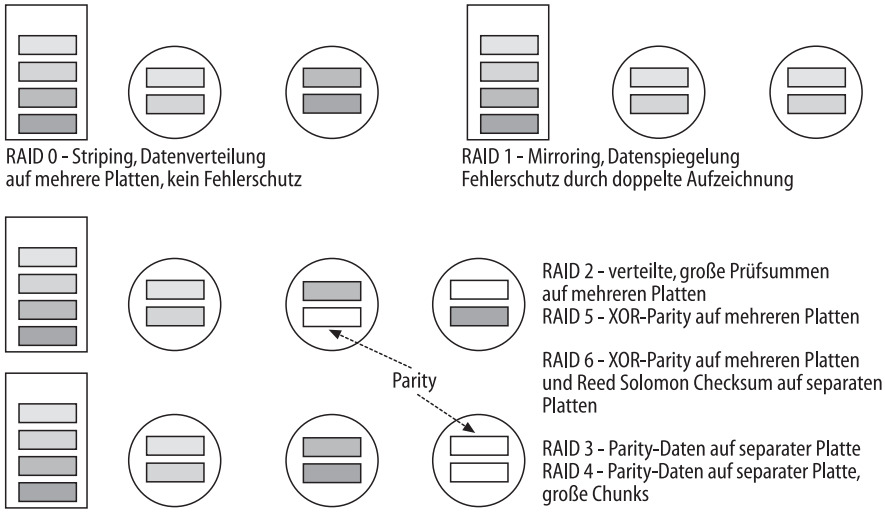


Abb. 8.146. RAID-Level

rates Laufwerk, sondern über alle Platten verteilt werden. Es ist damit besonders gut für Mehrkanalanwendungen geeignet und entsprechend weit verbreitet. RAID 6 entspricht etwa RAID 5, jedoch ist die Datensicherheit durch ein zusätzliches Fehlerschutzsystem noch einmal gesteigert, sodass das RAID auch den Ausfall von zwei Platten verkraftet. Bei vielen Arrays können die defekten Platten in laufendem Betrieb ausgewechselt werden (Hot Swap). Abbildung 8.145 zeigt ein mobiles System, das speziell für den Filmbereich ausgelegt wurde um unkomprimierte HD-Daten aufzuzeichnen. Es arbeitet intern mit 2,5"-Festplatten, die im RAID-Verbund vor Datenverlusten geschützt sind. Daneben ist ein großer RAID-Festplattenverbund für HD-Daten zu sehen.

8.9.1.3 Harddisk-Recorder

Festplatten als Speichermedien wurden bereits Ende der 80er Jahre in Harddisk-Recordern eingesetzt. Bekannte Typen stammten von Quantel und von Abekas, die beide zu ihrer Zeit sehr teure Geräte herstellten, die eine Speicherkapazität von ca. einer Minute für unkomprimierte Daten aufwiesen. Die Einsatzbereiche lagen dort, wo der Random Access trotz des hohen Preises große Vorteile brachte, wie z. B. die Einspielung häufig wiederkehrender kurzer Videosequenzen oder von Standbildern, ohne die Kopfabnutzung von MAZ-Geräten in Kauf nehmen zu müssen.

Moderne Harddiskrecorder sind als RAID-Array aufgebaut. Sie werden oft zur Aufzeichnung unkomprimierter Daten mit 8 oder 10 Bit und Abtastraten von 4:2:2 bis 4:4:4:4, also für RGBA-Signale eingesetzt, sie können aber auch mit datenreduzierten Signalen betrieben werden. Der Plattenverbund ist insbesondere für HD-Anwendungen meist unentbehrlich, denn eine einzelne Festplatte liefert in der Praxis selten mehr als 50 MB/s. Dabei ist vor allem zu bedenken, dass eine einzelne Platte nur dann einen relativ großen Datenstrom

aufrechterhalten kann, wenn dieser kontinuierlich gelesen wird. Eine Average Access Time in der Größenordnung von 10 ms ist aber nicht ausreichend kurz, um häufige Szenenwechsel und die damit verbundenen schnellen Kopfpositionswechsel zuzulassen. Das Problem ist von nichtlinearen Editingsystemen her bekannt. Dort wird empfohlen, die Platte vor der Bearbeitung zu defragmentieren und damit große zusammenhängende Gebiete für die Aufzeichnung zu schaffen. Dieses Vorgehen ist für Harddiskrecorder nicht praktikabel, hier muss eine Funktion ohne Vorbereitung realisiert werden können. Es wird der sogenannte True Random Access gefordert, d. h. durch den Parallelzugriff auf die Daten im RAID-Verbund muss jederzeit der uneingeschränkte Zugriff auf jedes Einzelbild möglich sein. Dafür muss die Zugriffszeit wiederum kürzer sein als die Dauer der vertikalen Austastlücke von 1,6 ms.

Harddiskrecorder unterstützen manchmal eine Vielzahl verschiedener Signalformen. Gute Geräte ermöglichen auch eine mehrkanalige Audioaufzeichnung, bieten Timecode und die serielle Schnittstelle nach RS 422, sodass sie wie eine MAZ angesteuert und in ein Schnittsteuersystem integriert werden können. Zunehmend gewinnen auch Vernetzungsmöglichkeiten an Bedeutung. Abbildung 8.147 zeigt einen größeren Disk-Recoder von Sony. Auch dieses Gerät arbeitet mehrkanalig und ist sowohl in einer HD-Ausführung als auch als Slow-Motion-Variante erhältlich. Auch hier sind MAZ-ähnliche Bedienelemente an der Front zu finden.

8.9.1.4 Mediaserver

Media- oder Videoserver sind hochleistungsfähige bandlose Speichersysteme für digitale Audio- und Videosignale auf RAID-Basis, meist Level 5 oder 6. Der wesentliche Unterschied zu Harddiskrecordern ist die Möglichkeit, eine Vielzahl von Kanälen gleichzeitig ein- und ausspielen zu können. Während Harddiskrecorder als separate Stand-Alone-Geräte oft für hochwertige Produktionen ohne oder mit geringer Datenreduktion arbeiten, werden Videoserver im News-Bereich und für die Distribution eingesetzt und arbeiten meist mit stärkerer Datenreduktion. Oft wird hier das MPEG-Verfahren verwandt, da es eine einfache Änderung der Datenrate erlaubt.

Abb. 8.147. Disk-Recorder [186]

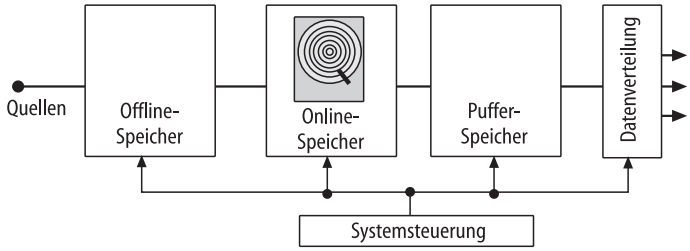


Abb. 8.148. Server-Grundstruktur

Zu den Leistungsmerkmalen zählen neben der Anzahl der Mediaströme, die im simultaneous true random access verarbeitet werden, die Qualität des Datenmanagements und die Möglichkeit der Verwendung verschiedener Datenraten. Auch hier werden Timecode und Steuerprotokolle aus dem MAZ-Bereich unterstützt. Besonders wichtig ist die Möglichkeit zur Vernetzung mit anderen Systemen über Fibre Channel, ATM, SDTI (s. Kap. 10) oder wenigstens SCSI. Abbildung 8.148 zeigt das Blockschaltbild eines Servers. Basis ist ein High Speed Datenbus mit einer Leistung von mehreren Gbit/s. Hinzu kommen Pufferspeicher (RAM), die einen kontinuierlichen Datenfluss sicherstellen, Schnittstellen und ggf. Signalwandler. Für die Funktionssteuerung steht meist ein eigener Datenbus zur Verfügung.

Durch die Verbindung digitaler Produktionsbereiche mit Mediaservern über schnelle Netzwerke lassen sich komplette bandlose Videosysteme von der Akquisition bis zur Sendung realisieren. Recht früh wurden für diesen Zweck die Profile-Serie von GrassValley bekannt. Abbildung 8.149 zeigt die Rückseite eines Gerätes, das eine vierkanalige Ein- oder Ausspielung von HD-Signalen über HD-SDI ermöglicht. Intern arbeitet das Gerät mit einer Datenreduktion nach MPEG-2@HL bei einer Datenrate von bis zu 80 Mbit/s pro Kanal und Echtzeitcodierung. Generell wird versucht, die Einbindung der Server möglichst flexibel zu halten. Meist ist daher die Nutzung von Fibre Channel-Interfaces ebenso möglich wie die der videospezifischen SDI-Anschlüsse nebst Timecode-, GPI- und RS-422-Verbindungen (s. Kap. 9). Moderne Videoserver sind SD- und HD tauglich und darauf ausgelegt, nur mit starker Datenreduktion zu arbeiten. Sie erreichen damit Speicherkapazitäten, die für viele Tage ausreichen.

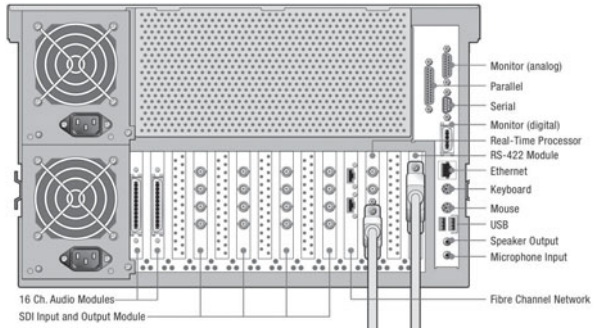


Abb. 8.149. Anschlussfeld eines Profile-Servers

8.9.2 Optische Speichermedien

Zu den optischen Speichersystemen gehören die Compact Disc (CD), die Digital Versatile Disc (DVD) und die Blu-ray Disc (BD). Das früher vorhandene BD-Konkurrenzformat namens HD-DVD wurde von der BD verdrängt. Bereits die DVD und später die BD entstanden auf Basis der CD, die 1982 als neuer digitaler Datenträger für den Digital Audio-Bereich vorgestellt wurde. Die Einführung war äußerst erfolgreich, die CD konnte die Schallplatte fast völlig verdrängen. Das System wurde mit geringfügigen Änderungen dann als CD-ROM für allgemeine Daten der Computertechnik fortentwickelt und zu Beginn der 90er Jahre auch zur Speicherung von Videosignalen eingesetzt. Aufgrund der auf 650 MB beschränkten Speicherkapazität und der geringen Datentransferrate von weniger als 0,5 MByte/s mussten die Videodaten dafür stark reduziert werden. Ein diesbezüglicher Versuch der Einführung eines videofähigen Systems war die CDi (Compact Disc interactive), die mit einer Datenreduktion nach MPEG-1 arbeitete. Dieses System war jedoch nicht leistungsfähig genug, um sich zu etablieren.

Bei der DVD konnten dann sowohl die Übertragungsgeschwindigkeit als auch die Speicherkapazität so gesteigert werden, dass mit einer Datenreduktion nach MPEG-2 die Wiedergabe eines hochqualitativen Programms von mehr als zwei Stunden Dauer möglich wurde. Die DVD-Video konnte sich schnell etablieren und verdrängte die VHS-Bänder im Heimbereich. Die BD soll die Funktion der DVD im Bereich HD-Video übernehmen.

8.9.2.1 Funktionsprinzip

Die Grundfunktion ist bei CD, DVD und BD gleich, sie sei hier am Beispiel der CD erläutert. Bei der CD befinden sich die Daten auf einer Scheibe aus einem hochwertigen Polycarbonat-Kunststoff, der einen sehr gleichmäßigen Brechungsindex aufweist. Der Scheibendurchmesser beträgt 12 cm bei einer Dicke von 1,2 mm (Abb. 8.150). Die Daten befinden sich in einer spiralförmigen Spur, die mit konstanter Geschwindigkeit von innen nach außen gelesen wird (Constant Linear Velocity CLV, im Gegensatz zu Constant Angular Velocity CAV). Die Rotationsfrequenz variiert dabei für die CD zwischen 8,1 und 3,3 Hz.

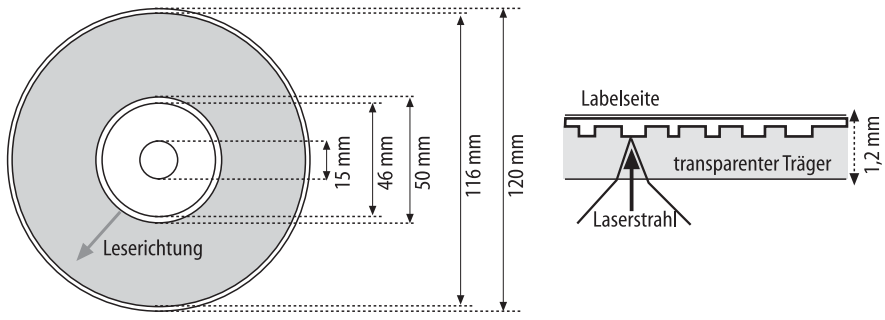


Abb. 8.150. Abmessungen der CD und Lage der Datenschicht

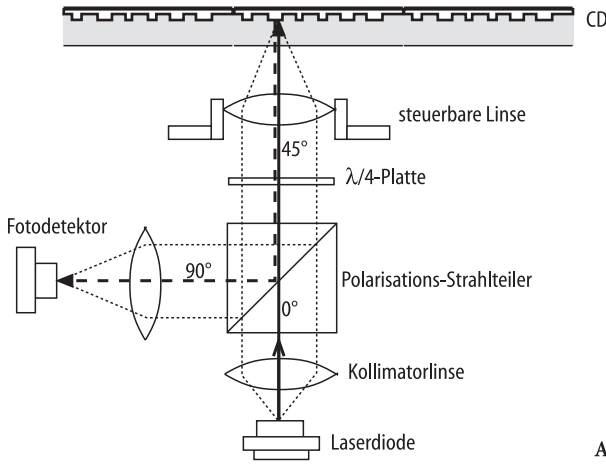


Abb. 8.151. CD-Arbeitsprinzip

Daraus resultiert eine Geschwindigkeit von 1,2 m/s, die zu einer Spieldauer der Audio-CD von 75 Minuten führt.

Den beiden Zuständen des gespeicherten Digitalsignals sind Vertiefungen und Erhöhungen (Pits und Lands) zugeordnet, die sich zusammen mit einer dünnen Reflexionsschicht aus Aluminium innerhalb des Mediums, an der Plattenoberfläche, der Labelseite befinden (Abb. 8.150). Zum Lesen der Informationen wird ein Laserstrahl aus einem kleinen Halbleiterlaser von unten durch die Scheibe geschickt. Beim Übergang in das dichtere Medium wird der Strahl gebrochen, sodass insgesamt ein Fokussierung auftritt. Für die CD ist festgelegt, dass der Brechungsindex n = 1,46 beträgt und der Laserstrahl so fokussiert wird, dass er mit einem Durchmesser von 0,8 mm auf die Scheibe fällt. Daher ist die Gleichmäßigkeit der Brechungseigenschaften und die Einhaltung der Plattendicke wichtig.

Bei der CD hat der Laserstrahl eine Infrarot-Wellenlänge von 780 nm, die sich im Medium mit n = 1,46 auf ca. 530 nm reduziert. Der Laserstrahl wird an der Datenschicht reflektiert und die gewünschte Information wird aus der Änderung der Reflexionseigenschaften beim Übergang von Pit zu Land gewonnen. Bei diesem Wechsel wird die Strahlintensität moduliert, sodass als Empfänger ein einfacher Fotodetektor ausreicht, dessen Ausgangsspannung registriert wird (Abb. 8.151).

Für die Modulation ist das kohärente Licht des Lasers erforderlich, das je nach Phasenlage konstruktiv oder destruktiv interferieren kann [226]. Die destruktive Interferenz folgt aus einer Phasenverschiebung von der halben Wellenlänge $\lambda/2$. Bei der Reflexion wird das Pit zweimal durchlaufen. Die Tiefe der Pits wird mit $\lambda/4 = 0{,}12\ \mu m$ daher so gewählt, dass das in den Pits reflektierte Licht mit dem von den Lands reflektierten Licht destruktiv interferiert und damit weniger Lichtintensität den Detektor erreicht als bei der Reflexion am Land allein. Der Lichtfleck ist deshalb breiter als die Vertiefung. Der Effekt wird durch eine Intensitätsschwächung des reflektierten Licht durch Streuung an den Pits verstärkt.

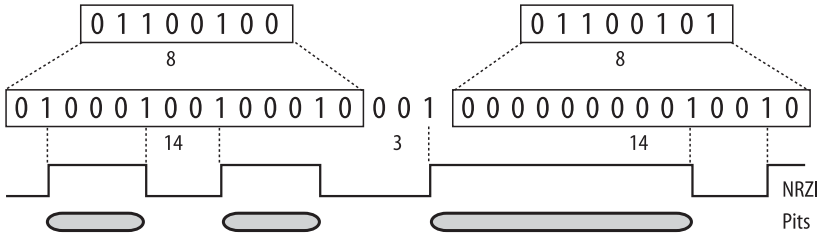


Abb. 8.152. EFM bei der CD

Damit der Detektor nur vom reflektierten Licht getroffen wird, arbeitet das Abtastsystem mit einem Strahlteiler aus zwei verbundenen Prismen, die so gewählt sind, dass die Grenzfläche für Licht einer bestimmten Polarisationsrichtung fast transparent ist, während sie Licht mit einer um 90° dagegen verdrehten Polarisationsrichtung spiegelt. Abbildung 8.151 zeigt das Funktionsprinzip. Das Laserlicht wird zu einem Strahl gebündelt und fällt dann auf den Strahlteiler, der nur für eine lineare Polarisationsrichtung transparent ist. Darauf folgt eine $\lambda/4$-Platte, die eine Phasenverschiebung von 45° bewirkt, sodass der Strahl nicht mehr linear, sondern zirkular polarisiert ist. Das Licht dringt in die Platte ein, wird mehr oder weniger stark reflektiert und durchläuft die $\lambda/4$-Platte erneut. Es ist dann um insgesamt 90° verdreht, sodass es am Strahlteiler in Richtung Detektor reflektiert wird. Die ganze Einheit kann sehr kompakt aufgebaut und so gelagert werden, dass eine seitliche Verschiebung zur automatischen Korrektur der Spurlage ebenso möglich wird, wie eine vertikale Versetzung mit der Fokussierungsfehler ausgeglichen werden.

Die Daten auf der CD werden in Blöcke und Frames zusammengefasst und mit einem CIRC-Fehlerschutz versehen. Dabei kommt die Eight to Fourteen Modulation (EFM) zum Einsatz, bei der 8-Bit in 14-Bit-Worte umgesetzt werden. Zusätzlich werden jeweils 3 sogenannte Mergin Bits angehängt, sodass zwischen zwei logischen 1-Zuständen minimal zwei und maximal zehn 0-Zustände auftreten (Abb. 8.152). Die Kanaldatenrate inklusive Fehlerschutz beträgt 4,32 Mbit/s, die Signaldatenrate 2,03 Mbit/s.

8.9.2.2 DVD

Bei der Entwicklung der DVD wurde auf die Abwärtskompatibilität zur CD geachtet. Daher wird der gleiche Datenträger, die Scheibe mit 120 mm Durchmesser und 1,2 mm Dicke verwendet. Die kleine CD-Größe mit 8 cm Durchmesser wird auch unterstützt. Der innere Radius des Spurbereichs beträgt

Tabelle 8.19. DVD-Standardisierung

Book	A	B	C	D	E
Anwendung	DVD-ROM	DVD-Video	DVD-Audio	DVD-R	DVD-RW
physikalisch	Nur lesen	Nur lesen	Nur lesen	einmal beschreibbar	mehrfach beschreibbar
Dateisystem	UDF-Bridge (M-UDF + ISO 9660)			UDF	UDF

Tabelle 8.20. Vergleich der Parameter von CD und DVD

	CD	DVD	BD
Durchmesser	120 mm	120 mm	120 mm
Dicke der Disk	1,2 mm	1,2 mm	1,2 mm
Substratdicke	1,2 mm	2 x 0,6 mm	1,2 mm
Laserwellenlänge	780 nm	650 und 635 nm	405 nm
numerische Apertur der Linse	0,45	0,6	0,85
Spurabstand	1,6 µm	0,74 µm	0,32 µm
minimale Pitlänge	0,83 µm	0,4 µm	0,16 µm
Fehlerkorrektur	CIRC	Reed Solomon Product Code	
Kanalcodierung	8/14 EFM	8/16 EFM+	

24 mm, der äußere 58 mm, dazwischen befindet sich die Datenspur, aus der die Daten mit konstanter Lineargeschwindigkeit ausgelesen werden. Die DVD wurde gleich für Multiformatanwendungen ausgelegt, als DVD-ROM, d. h. als Speicherformat für Computeranwendungen, als DVD-Video, DVD-Audio und in beschreibbarer Form als DVD-R, DVD-RW, DVD+RW und DVD-RAM. Die Spezifikationen sind in sogenannten Books festgelegt (Tabelle 8.19).

Im Unterschied zur CD besteht die DVD-Scheibe aus zwei Substratschichten von je 0,6 mm Dicke. Eine wesentliche Veränderung gegenüber der CD ist zudem die Verwendung eines Lasers mit kürzerer Wellenlänge, sie beträgt 635 nm. Damit kann die Aufzeichnungsdichte gesteigert werden. Eine weitere Steigerung wird möglich, da anstatt eines Spurabstands von 1,6 µm bei der CD bei der DVD 0,74 µm benutzt wird. Abbildung 8.153 zeigt die Größen im Vergleich zur CD und BD und die Lage der Datenschichten, die bei der DVD in der Mitte sind. Auch die Pitlänge ist kürzer und variiert bei der DVD zwischen 0,4 und 2 µm, während sie bei der CD zwischen 0,8 und ca. 3 µm beträgt. Als dritte Maßnahme wird ein größerer Anteil an Nutzdaten verwendet. Schließlich wurde noch die Rotationsgeschwindigkeit gegenüber der CD etwa verdreifacht, die Abtastgeschwindigkeit beträgt bei der DVD 3,8 m/s, sodass sich insgesamt auf einer Datenschicht eine effektiv nutzbare Kapazität von 4,7 GByte ergibt. Tabelle 8.20 zeigt den Vergleich der Spezifikationen von DVD und CD.

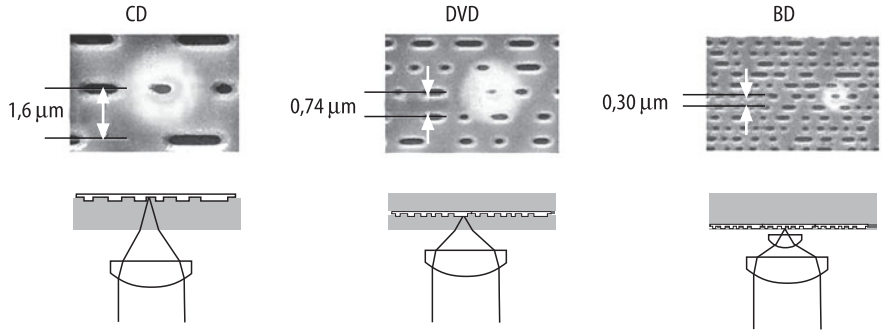


Abb. 8.153. Strukturgrößen und Lage der Datenschichten bei CD, DVD und BD

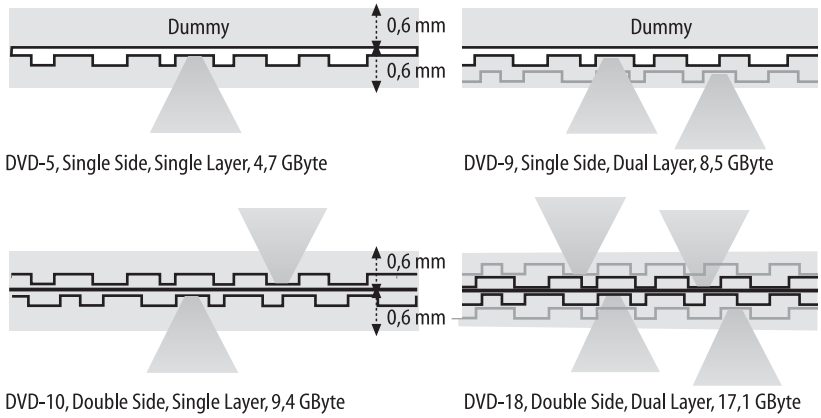


Abb. 8.154. Layer-Struktur der DVD

Die als DVD-5 spezifizierte Speicherkapazität von 4,7 GByte kann durch Verwendung zweier Daten-Layer nahezu verdoppelt werden. Die zwei Schichten sind durch eine ca. 50 µm dicke Zwischenlage getrennt, wobei die dem Laser zugewandte Schicht halb durchlässig ist, damit der Strahl auch die zweite Schicht erreichen kann. Welche der beiden Schichten gelesen wird, hängt vom Fokussierungszustand des Lasers ab. Diese Dual-Layer-Version bietet eine Kapazität von 8,5 GByte und ist als DVD-9 spezifiziert. Weitere Steigerungen ergeben sich aus der Verwendung zweier Seiten, da die Scheibe aus zwei Teilen von je 0,6 mm Dicke besteht. Die Datenrate steigt damit auf 9,4 GB (single Layer DVD-10) bzw. 17 GByte (Dual Layer, DVD-18). Wenn die Platte nicht umgedreht werden soll, sind für die Auslesung zweier Seiten auch zwei Abtastsysteme erforderlich (Abb. 8.154) [212].

Die Daten auf der DVD werden in einer spiralförmigen Spur (Track) geschrieben, die von Ein- und Ausgangsbereichen (Lead in und out) begrenzt ist. Lead-in enthält Informationen wie beispielsweise die Anzahl der Layer, Durchmesser, Spurführung, Nummer des ersten Sektors und Kopierschutzschlüssel. Middle bzw. Lead-out signalisieren, dass die Mitte einer Dual-Layer-Disc bzw. das Ende erreicht ist. Die erste Spur für den Layer 0 führt vom inneren Teil zum äußeren Rand der Disc. Bei Vorliegen zweier Layer gibt es zwei Möglichkeiten der Spurführung. Die Variante Opposite Track Path erlaubt einen weitgehend unterbrechungsfreien Datenstrom. Der Laser beginnt dabei im Zentrum und läuft über die Lead-in Zone bis zum äußeren Rand. Dann fokussiert der Laser auf den zweiten Layer und läuft in der entgegen gesetzten Richtung bis zum Lead-out-Bereich in der Mitte zurück. Wenn die Spuren andererseits parallel von innen nach außen verlaufen, spricht man von Parallel Track Path. Bei diesem Verfahren besitzt jeder Layer sowohl einen Lead-in- als auch einen Lead-out-Bereich.

Die spiralförmige Spur ist in Sektoren aufgeteilt. Jeder dieser Sektoren hat eine nutzbare Größe von 2048 Bytes und ist einem Bit-Shifting-Prozess unterworfen, um eine möglichst effektive Fehlerkorrektur vornehmen zu können.

Dem Sektor werden zusätzlich 12 Start- und 4 End-Byte hinzugefügt. Die Daten lassen sich in 12 Reihen darstellen, darin dienen 4 Byte zur Sektoridentifikation (ID). Falls diese fehlerhaft übertragen werden, sind zwei Byte zur Fehlererkennung vorgesehen (ID Error Detection). Eine 6 Byte umfassende Information gibt zusätzlich Auskunft über die Kopierfähigkeit des Sektors. Am Ende der 12 Reihen befindet sich ein 4 Byte-Fehlerschutzcode für den gesamten Sektor. Insgesamt ergibt sich somit also eine logische Sektorgröße von 2064 Byte.

Die Daten werden dann weiter zu einem Error Correction Code-Block (ECC-Block) umgeordnet. In diesem befinden sich untereinander 192 Reihen mit einer Blockbreite von 172 Byte. Für jede der 172 Spalten wird ein 16-Byte großer, äußerer Umrechnungscode (Reed-Solomon) erstellt, der wiederum zu einem, 172 Spalten beinhaltenden und 16 Reihen hohen, Block zusammengefasst wird. Für die nunmehr 208 Spalten des ECC-Blocks, wird ein innerer 10-Byte-breiter Umrechnungscode erstellt und angefügt. Der Datenblock wird im Anschluss wieder in Sektoren aufgeteilt, die aufgezeichnet werden. Ein Sektor besteht dann aus 13 Reihen (12 + 1) a´ 182 Byte (172 +10). Durch Hinzunahme von 52 Sync-Byte entsteht schließlich das Recording Frame mit 2418 Byte.

Zur Anpassung an die Eigenschaften des Mediums wird eine 8/16-Modulation (EFM Plus), vorgenommen. Das bedeutet, dass mit Hilfe einer Tabelle jedem Byte 16 Bits zu geordnet werden, damit zwischen zwei Einsen mindestens 2 und max. 10 Nullen stehen. Die modulierten Daten werden dann mit dem NRZI-Code (Non Return to Zero Inverse) auf die Platte geschrieben.

8.9.2.3 DVD-Video

Die Grundlage aller DVD-Varianten ist die DVD-ROM. Das zugehörige Filesystem nennt sich UDF. DVD-Video und DVD-Audio stellen speziell formatierte Varianten dar. Die Spezifikationen der DVD-Video sehen vor, dass das Dateisystem auf der obersten Ebene einen Ordner mit der Bezeichnung »Video_TS« (bzw. »Audio_TS« für die DVD-Audio) enthalten muss. TS bedeutet hier Title Set und hat nichst mit dem MPEG-Transportstrom zu tun. Die Files innerhalb eines solchen Ordners sind fest strukturiert. Dateien mit der Endung ».VOB« enthalten die eigentlichen Bild- und Tondaten sowie auch die Untertitel und Still-image-Menüs. Dateien mit der Endung ».IFO« und ».BUP« enthalten die Programminformationen bzw. eine Sicherheitskopie davon.

Die DVD liefert direkt eine Datenrate von 26 Mbit/s. Nach der Demodulation stehen davon noch 13 Mbit/s und nach der Fehlerkorrektur 10,08 Mbit/s zur Verfügung. Mit Hilfe des sogenannten Bitbudgeting können diese als eigentliche Nutzdaten flexibel zugeteilt werden. Für jeden Teildatenstrom der DVD-Video, also Video, Audio und Subtitel, können höchsten 9,8 Mbit/s genutzt werden. Es ist festgelegt, dass ausschließlich datenreduzierte Videosignale verwendet werden dürfen, die nach MPEG-1 oder MPEG-2 codiert sind.

Im Gegensatz zur Datenübertragung bei DVB wird der MPEG-Elementarstrom (PES) für CD- und DVD-Anwendungen nicht in einen Transportstrom (TS) sondern in einen Programmstrom (PS) umgesetzt (s. Kap. 3.8). Alle Elementarströme des Programmstroms gehören zu einem einzigen Programm. Eine MPEG-GOP darf nach der DVD-Spezifikation maximal 15 (PAL) bzw. 18 Bilder (NTSC) umfassen. Es werden die Bildfrequenzen 25 Hz und 29,97 Hz

sowie die TV-Systeme 625/50 und 525/60, jedoch keine HD-Signale, unterstützt. Weiterhin wird mit Hilfe von Pan Scan oder Letterbox die Verwendung der Bildseitenverhältnisse 4:3 und 16:9 ermöglicht (s. Kap. 4.4).

Die Audiospezifikationen der DVD-Video sind vielfältig, ein Videostrom kann von bis zu acht Audiokanälen begleitet sein. Es werden die Datenreduktionsverfahren von MPEG-Layer II unterstützt, mit Datenraten zwischen 32 und 192 kbit/s für Mono- bzw. 64 bis 384 kbit/s für Stereosignale. Bei MPEG-2 kann zusätzlich ein Extensionstream mit maximal 528 kbit/s für Mehrkanalanwendungen genutzt werden. Die Abtastfrequenz beträgt 48 kHz. Dieser Wert gilt auch für das alternative Datenreduktionsverfahren Dolby AC3. Für diese Version werden alle Modi, nämlich 1 bis 3 Frontkanäle ohne Surroundsound und 3 Frontkanäle mit mehreren Surroundkanälen unterstützt. Die Datenrate liegt zwischen 64 und 448 kbit/s. Bei typischen Videofilmen wird eines dieser Audioformate, meist AC3, eingesetzt, die DVD kann aber auch mit PCM-Signalen ohne Datenreduktion versehen sein. Weiterhin werden auch Kinotonformate wie DTS und SDDS (s. Kap. 5) unterstützt.

Neben der Vielfalt bei der Wahl der Audiokanäle bietet die DVD-Video weitere Flexibilität und Möglichkeiten des interaktiven Umgangs. Die Benutzer werden mit Hilfe von Menüs geführt, die auf dem Bildschirm erscheinen. Darüber können Filme gestartet, gestoppt und gewechselt werden. Dabei gibt es die Möglichkeit zur Unterbrechungsfreien Verzweigung (seamless branching) um z. B. bestimmte Szenen ungestört überspringen zu können. Es kann einer von acht Begleittönen oder einer von bis zu 32 Untertiteln gewählt. Eine gegenüber gewöhnlichen Videorecordern völlig neue Funktionalität ist die Wahl verschiedener Kameraperspektiven (Angle) für einen Videobeitrag (der dann natürlich entsprechend aufwendig produziert sein muss). Zu den weiteren Merkmalen der DVD-Video gehören Kopierschutzverfahren und eine Kindersicherung. Außerdem wurde ein Ländercode eingeführt, sodass DVD nur abgespielt werden, wenn sie den gleichen Code aufweisen wie das Abspielgerät. Damit kann international unterschiedlichen Verwertungsrechten, z. B. von Spielfilmen, Rechnung getragen werden. Die Welt wird dabei in sechs Zonen aufgeteilt, in Nord/Südamerika, Europa, verschiedene Teile Asiens etc.

Für den Umgang mit dem Medium ist in den Spezifikationen der DVD-Video ein Minimalset von Bedienfunktionen festgelegt, die als Fernbedienung von DVD-Playern oder als Menü von Player-Applikationen in Computern realisiert sind (Abb. 8.155). Der interaktiven Umgang bei der DVD erfordert, dass die Funktionen in geeigneter Weise angesprochen werden und die Benutzer mit Hilfe der Bildschirmmenüs geführt werden können. Die dafür nötigen didaktisch/gestalterischem Tätigkeiten werden als DVD-Authoring bezeichnet.

Abb. 8.155. Standardbedienelemente zur Steuerung von DVD-Playern

8.9.2.4 DVD-Aufzeichnung

Für die Aufnahme auf DVD-Medien stehen sowohl einmal beschreibbare DVD (R) als auch mehrere mehrfach verwendbare DVD (RW) zur Verfügung. Die einmal beschreibbare DVD-R nutzt einen organischer Farbstoff auf der Speicherschicht dessen Reflexionseigenschaft durch Laserbestrahlung punktuell verändert werden kann. Die DVD-R hat eine Kapazität von 4,7 GB pro Seite und ist kompatibel zur DVD-ROM, DVD-Video und DVD-Audio.

Beim Schreibvorgang folgt der Laserstrahl einer auf dem DVD-Rohling bereits vorgeprägten Schlingerbahn mit einer festen Frequenz von 22,05 KHz (Wobbled Groove). Beim Abweichen von der Mittellinie wird es durch PLL-Regelung möglich, innerhalb der Grenzen der Spur zu bleiben und auch die Rotationsgeschwindigkeit zu steuern. Die DVD-R wird in die Varainte A (für Authoring-Anwendung mit Aufzeichnung des Disc Description Protokolls, DDP) und G (General) unterscheiden. Die erste arbeitet mit einer Wellenlänge von 650 nm, die zweite mit 635 nm. Als Alternative zur DVD-R wird die DVD+R angeboten, die auf der DVD+RW basiert und auch 4,7 Gbyte bietet.

Auch die mehrfach beschreibbaren Varianten DVD-RW und DVD+RW erfordern ein vorkonfiguriertes Medium. Die Aufzeichnung erfolgt wieder in Phase Change-Technologie mit Einstrahl-Überschreibtechnik (Abb. 8.156). Durch Erhitzung und schnelle Abkühlung werden mit verschiedenen Laserleistungen Temperaturen von ca. 200 °C oder von über 500 °C erzeugt, womit die Beschichtung eine kristalline oder amorphe Phase erreicht. Die kristalline Phase, die bei der niedrigeren Temperatur entsteht, hat beim Lesevorgang einen wesentliche höheren Reflexionsgrad als die amorphe, sodass die zwei Zustände unterschieden werden können Die DVD+RW erlaubt das direkte Editieren der Daten auf der Platte und nutzt mit Zoned CLV eine Kombination aus CLV und CAV. Da dazu ein veränderter Befehlssatz erforderlich ist, ergibt sich im Gegensatz zur DVD-RW eine nicht vollständige Kompatibilität und daher ist diese Form nicht offiziell anerkannt. Die Kapazität beträgt bei beiden Varianten 4,7 GB pro Seite, genau wie bei der letzten zu erwähnenden Variante, nämlich DVD-RAM, die aber für Videoanwendungen unbedeutender ist.

Das DVD-Recording geschieht via PC aber auch in Standalone-Recordern, die den VHS-Recorder im Heimbereich ersetzen. Gute Geräte unterstützen als Recorder und Player alle zuvor genannten DVD R(W)-Varianten und bieten auch eine Schnittstelle für die digitale Datenübertragung.

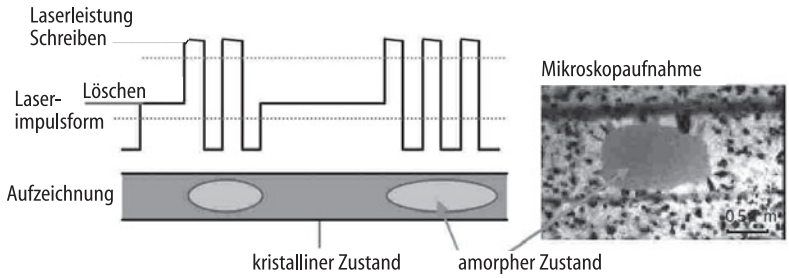


Abb. 8.156. Aufzeichnungs- und Löschvorgang bei der DVD und Blu-Ray Disc [122]

8.9.2.5 HD-DVD und Blu-ray Disc

Die Speicherkapazität der DVD kann durch den Einsatz eines blauen Lasers mit kürzerer Wellenlänge (405 nm) noch weiter gesteigert werden. Als Nachfolgesystem, insbesondere für die Verbreitung von Filmen mit HD-Auflösung, waren zunächst zwei Formate vorgesehen, die HD-DVD und die Blu-ray Disc (BD). Letztere hat sich schließlich am Markt allein durchgesetzt.

Die Blu-ray Disc erreicht eine Datenrate von 36 Mbit/s und eine Speicherkapazität von 25 GB [18]. Die Datenrate kann durch den Einsatz von zwei Lasern verdoppelt werden. Die Disc war zunächst durch ein Gehäuse geschützt (Abb. 8.157 rechts), ist heute aber ohne verwendbar. Sie kann, nach dem gleichen Prinzip wie bei der DVD, auch als Dual Layer-Variante verwendet werden, womit sich die Kapazität auf 50 GB verdoppelt. Zukünftig sollen Blu-ray Discs mit 4 bzw. 8 Datenschichten bis zu 200 GB Speicherkapazität aufweisen.

Mit dem Einsatz von Datenreduktionsverfahren können mehr als zwei Stunden HD-Material gespeichert werden. Die Spezifikationen erlauben hier drei Kompressionsverfahren, nämlich MPEG-2@HL, H.264/AVC und Windows Media VC1 (s. Kap. 3). Es sind folgende Blu-ray Disc-Varianten erhältlich: die BD-ROM als vorbespielte Disc für Filme, Spiele, Software, etc., die BD-R als einmal beschreibbare und die BD-RE wiederbeschreibbare Speicher. Bei der BD-RE gleicht der Aufzeichnungsvorgang dem von HD-DVD und DVD (RW): Es wird mit einem Laser umschaltbarer Leistung gearbeitet, der bei kleiner Leistung das Medium löschen kann. Das bedeutet, dass sich die Oberfläche nach dem Löschvorgang in einem kristallinen Zustand befindet, während durch bestimmte Impulse bei hoher Leistung ein amorpher Zustand erreicht wird, der für den Laserstrahl andere Reflexionseigenschaften aufweist als der kristalline Zustand (Abb. 8.156).

Die BD findet vor allem im Heimbereich für die Distribution von Kinofilmen Verwendung, die meist mit einer Auflösung von 1080p produziert werden. Im Tonbereich werden dafür viele moderne Codecs unterstützt. Moderne Verfahren wie Dolby TrueHD oder DTS-HD bieten eine sehr gute Audioqualität für bis zu 8 Surround-Kanäle. Die meisten Blu-ray-Geräte lassen sich auch als DVD-Player nutzen. Bezüglich der interaktiven Funktionen gibt es gegenüber

Abb. 8.157. XDCAM-Recorder und Blu-Ray Disc-Medium [186]

der DVD Erweiterungen, wie überganglose Menüsteuerung, Unterstützung von Picture-in-Picture-Funktionen und interaktive Bonusinhalte mittels BD-Java.

8.9.2.6 Professional Disc

Die Blu-ray Disc wurde unter der Bezeichnung Professional Disc für den professionellen Bereich weiterentwickelt. Entsprechende Geräte, die unter der Bezeichnung XDCAM gehandelt werden, sind seit 2004 verfügbar. Abbildung 8.157 zeigt einen Recorder der sich in gewöhnliche Studioumgebungen und auch Netzwerke einbinden lässt. Abbildung 8.158 macht deutlich, dass das Aufzeichnungsverfahren so robust ist, dass es auch in Camcordern eingesetzt werden kann. Damit gewinnt man, unter dem Stichwort Random Access, die gleichen Vorteile, z. B. die Möglichkeit des Retro Loop-Recording, wie sie bei Camcordern mit Festplatten oder Festspeichern gegeben sind.

Die maximale Datentransferrate der BD wurde für XDCAM auf 72 Mbit/s verdoppelt, um zu erreichen, dass die 50 Mbit/s des IMX-Codecs aufgezeichnet werden können. Die Daten werden als Files abgelegt und lassen sich unter IMX auch mit 30 Mbit/s oder 40 Mbit/s skalieren. Alternativ lässt sich meist ein DVCAM-Datenstrom aufzeichnen.

Die Kapazität der Professional Disc beträgt 23,3 GB, womit Spieldauern von 85 Minuten bei DVCAM-Aufzeichnung, bzw. 75, 60 oder 45 Min. bei IMX-Aufzeichnung mit 30, 40 oder 50 Mbit/s erreicht werden. Im Hinblick auf eine effiziente Anbindung an IT-Umgebungen wird im Gesamtsystem die Aufzeichnung von Metadaten und UMID (s. Kap. 10) sowie von Proxy AV-Daten unterstützt, die die eigentlichen AV-Daten mit geringer Auflösung repräsentieren und eine schnelle Übersicht bei Netzwerkzugriff ermöglichen.

Die SD-Aufzeichnung auf XDCAM-Medien beruht sowohl beim Einsatz von IMX als auch bei DVCAM auf reiner Intraframekompression. Nach der erfolgreichen Einführung des Formates, bemühte sich Sony, es auch für den HD-Bereich einsatzfähig zu machen [58]. Da die maximale Datenrate bei XDCAM 50 Mbit/s beträgt, musste die Kompression effektiver werden. Die Lösung ist, mit MPEG-2 codierten Daten nach dem Long-GOP-Verfahren, also mit zeitlich abhängigen Bildern zu arbeiten. Zunächst wurde XDCAM HD entwickelt, das Signale verarbeitet, die ähnlich wie bei HDV, örtlich mit 3:1,5:0 unterabgeta-

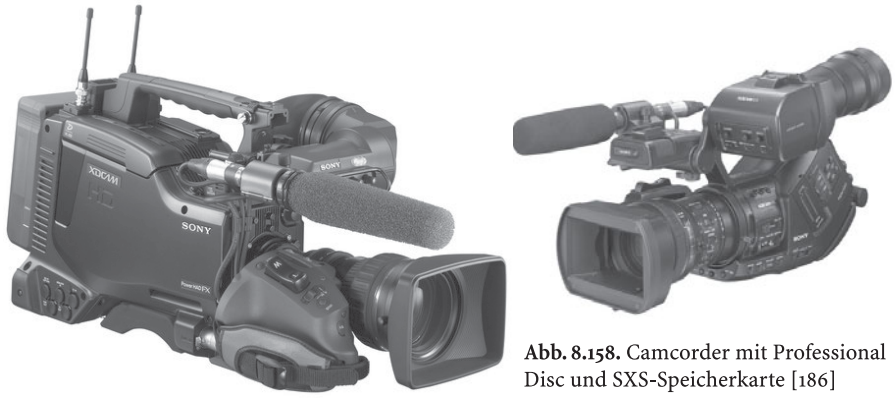


Abb. 8.158. Camcorder mit Professional Disc und SXS-Speicherkarte [186]

Tabelle 8.21. Aufnahme- und Wiedergabespezifikationen von XDCAM-HD422-Geräten

Bezeichn.	Codec	Auflösung	Bitrate	Subsampl	Audiokan.
XDCAM HD422	MPEG-2 4:2:2P@HL	1920 x 1080	50 Mbit/s	4:2:2	8
		1280 x 720	50 Mbit/s	4:2:2	8
XDCAM HD	MPEG-2 MP@HL14	1440 x 1080	18/25/35 Mbit/s	4:2:0	4
		1280 x 720	25/35 Mbit/s	4:2:0	4
IMX	MPEG-2 4:2:2P@ML	720 x 576	30/40/50 Mbit/s	4:2:2	8/4
DVCAM	DV	720 x 576	25 Mbit/s	4:2:0/4:1:1	4

stet sind. Als Datenraten können 18, 25 oder 35 Mbit/s gewählt werden. Mit der Einführung des Nachfolgeformates XDCAM HD422 wurde dieser Mangel behoben, d. h. dass mit 4:2:2-Signalen bei 8 Bit gearbeitet wird. Mit der letztgenannten Variante, die inzwischen bei vielen Rundfunkanstalten Akzeptanz findet, werden 1080i- und 720p-HD-Daten also auf 50 Mbit/s reduziert, womit Aufzeichnungsdauern von über 45 Minuten erreicht werden. 1080i-HD-Signale werden dabei um den Faktor 16,6:1 und 720p-Signale mit 14,7:1 reduziert.

Tabelle 8.21 zeigt, dass bei XDCAM HD422-Geräten auch XDCAM HD sowie für die SD-Aufzeichnung auch IMX und DVCAM unterstützt werden.

8.9.2.7 Weitere optische Speicher

Bei den optischen Speichermedien wird zur Weiterentwicklung an der Fluorescent Multilayer Disk (FMD) gearbeitet, bei der anstelle der zwei Schichten, auf die der Laser bei der DVD fokussiert werden kann, mit 10 bis 20 Schichten von fluoreszierenden Stoffen gearbeitet wird. Die Scheibe ist völlig transparent und soll eine Speicherkapazität von 140 GByte aufweisen [200].

Noch größeres Potenzial bietet die holografische Speichertechnik, bei der nicht einzelne Bits auf einer Fläche, sondern umfassende Bitkombinationen als Laser-Interferenz-Muster in der Tiefe einer Datenschicht gespeichert werden [135]. Die Interferenzmuster entstehen durch Überlagerung zweier Laserstrahlen. Die Information kann mit flächigen Sensoren ausgelesen werden.

Bei der Ausführung als Holographic Versatile Disc (HVD) werden zwei Laser verwendet. Ein roter Laser liest Daten von einer CD-ähnlichen Schicht, die der Ortsbestimmung und der Synchronisation dienen. Die eigentlichen Nutzdaten werden mit einem blau-grünen Laser (526 nm) von einer holografischen Schicht im oberen Bereich des Speichermediums ausgelesen. Eine HVD soll eine Kapazität von bis zu 3,9 TByte und eine Transferrate von 1 Gbit/s erreichen. Diese Leistungsdaten wurden im Laufe der Zeit aber schon von Festwertspeichern erreicht die heute immer dominanter werden.

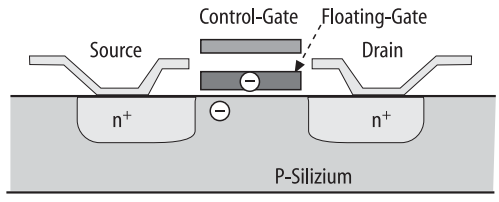


Abb. 8.159. Flash-Speicherzelle

8.9.3 Festwertspeicher

Die neueste Entwicklung bei den Video-Speichersystemen nutzt Festwertspeicherkarten (Flash-RAM), die als SD- (Secure Digital) oder CF- (CompactFlash)-Cards u. a. auch für die digitale Fototechnik eingesetzt werden. Da diese als Heimanwenderprodukt eine weite Verbreitung gefunden haben, gibt es eine starke Entwicklungskraft, die die Speicherkapazitäten schnell erhöht und die Preise senkt. Die SD-Card hatte eine Speicherkapazität von maximal 2 GB, die Nachfolger SDHC und SDXC erreichen bis 32 GB bzw. 104 GB bei Transferraten von 6 MB/s bzw. 104 MB/s. Flash-Speicher erfordern keine beweglichen Teile und stellen somit eine sehr robuste Technologie dar. Sie sollen die Informationen für bis zu 10 Jahre speichern und 100000-mal beschreibbar sein.

Abbildung 8.159 zeigt den Aufbau einer Flash-Speicherzelle. Sie zeichnet sich durch einen isoliert angebrachtes, sogenanntes Floating Gate aus. Durch Anlegen einer Steuerspannung an das darüber befindliche Controlgate werden Ladungsträger dazu angeregt, in das Floating Gate überzutreten. Aufgrund der Isolierung bleiben sie auch nach Abschalten der Spannung in diesem Bereich, damit ist die Information gespeichert und kann beim Auslesen abgerufen werden. Vor einer Wiederbeschreibung muss der Speicher gelöscht werden.

Die Flash-Speicher wurden im professionellen Videosektor zuerst von Panasonic, unter der Bezeichnung P2, angeboten. Hier werden SD-Speicherbausteine zu je 4 Einheiten in eine schmale Karte nach dem PCMCIA-Standard der PC-Technik eingebaut, sodass die Karte direkt in einen Laptop-PC eingesteckt werden kann. Dabei werden Datentransferraten von bis zu 640 Mbit/s erreicht. Bei einer Speicherkapazität von 32 GB pro Karte können 128 Min. DVCPro-Material, 64 Min. DVCPro50- oder 32 Min. HD-Signale im DVCProHD-Standard aufgezeichnet werden. Zur Erhöhung der Aufzeichnungsdauer bieten verfügbare Geräte Slots für mehrere Karten. Der Hersteller spezifiziert die genannten Datenreduktionsstandards sowie auch den AVC-Intra-Codec (s. Kap. 3) [105].

Während P2 auf SD-Karten basiert, werden bei der Red-Kamera CF-Cards als Speicher genutzt. Von Sony war die Flashtechnologie zunächst unter der Bezeichnung SXS mit ExpressCards als Nachfolger der PCMCIA-Cards mit Datenraten bis zu 2,5 Gbit/s verfügbar (Abb. 8.160). Geräte und Camcorder mit diesen Speichern haben bei Sony die Bezeichnung XDCAM EX. Ein Camcorder ist in Abbildung 8.158 rechts dargestellt. Inzwischen haben sich die Speicherbausteine so schnell weiterentwickelt, dass sie auch für die hohen Datenraten einsetzbar sind, die im Bereich Digital Cinematography anfallen (s. Kap. 6). Unter der Bezeichnung SR-Memory bietet Sony hier Speicherkarten mit einer Kapazität bis zu 1 TB und einer Datenrate bis zu 5, 5 GBit/s, die auch den dort anfallenden 4k-RAW-Daten bei bis zu 120 fps gerecht werden sollen.

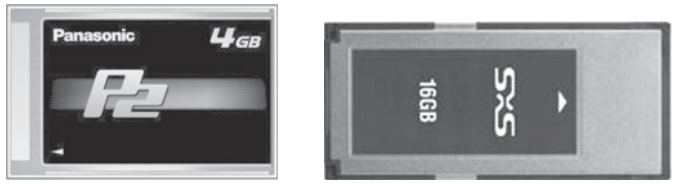


Abb. 8.160. P2- und SXS-Speicherkarten

9 Videosignalbearbeitung

Der Begriff Videosignalbearbeitung steht hier für die Verknüpfung mehrerer Signale durch Mischung, Stanzeffekte, digitale Tricks im Live-Betrieb sowie für die Umstellung aufgezeichneter Videosequenzen (Schnittbetrieb, Editing). In diesem Kapitel werden entsprechende Verfahren und Geräte für die Bereiche Produktion und Postproduktion beschrieben.

9.1 Bildmischer

Bildmischer dienen zur Umschaltung und Mischung von Videosignalen. Sie werden im englischen Sprachraum als Switcher bezeichnet. Mit dem Begriff Mischer werden die Funktionen getroffen, bei denen zwei Signale gleichzeitig vollflächig sichtbar sind. Im Laufe der Zeit hat der Funktionsumfang der Bildmischer immer weiter zugenommen, heute dienen sie zur Verbindung von Videosignalen mit Standard-Tricks und Stanzen u. v. m., woraus auch die Bezeichnung Trickmischer resultiert. Der Bildmischer ist die zentrale technische Einrichtung bei der Live-Produktion, wird als Einzelgerät also vor allem dort gebraucht, wo ein Echtzeitbetrieb erforderlich ist.

Eine wesentliche Voraussetzung für die fehlerfreie Mischung ist die absolute Synchronität der Bildgeneratoren. Dieser Gleichlauf wird im Studiobetrieb erreicht, indem alle angeschlossenen Bildquellen an einen zentralen Studiotakt angepasst werden (Abb. 9.1). Da nicht alle Geräte die gleiche Signalverarbei-

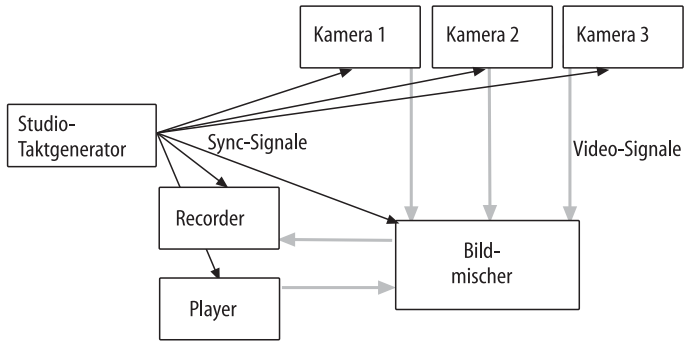


Abb. 9.1. Synchronisation aller Bildquellen durch einen Zentraltakt

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4_9, © Springer-Verlag Berlin Heidelberg 2013

tung aufweisen, und die zur Verbindung dienenden Leitungen unterschiedliche Längen haben, ist der Gleichlauf durch den zentralen Takt allein noch nicht gewährleistet. An professionellen Geräten ist deshalb eine Möglichkeit zum Feinabgleich der Phasenlage zwischen dem Videosignal und dem Studiotakt gegeben (s. Kap. 10.) Es reicht aus, die horizontalen Bildlagen (H-Phase) und bei PAL-Signalen auch die Farbträgerphasen (Subcarrier-Phase) in Übereinstimmung zu bringen, da die Zeitabweichungen maximal im µs-Bereich liegen. Auch bei Digitalmischern müssen die an der Mischung beteiligten Signale im Gleichtakt laufen, hier ist aber keine Anpassung und kein Abgleich erforderlich, denn Digitalmischer enthalten standardmäßig Synchronizer, die meistens einem Korrekturbereich von ca. 50 µs aufweisen.

9.1.1 Mischerkonzepte

Ein Bild- oder Trickmischer soll die Umschaltung und Mischung von mindestens zwei Videosignalen ermöglichen. Die als Mix bezeichnete Bildmischung ist kein Schaltvorgang, hier sind beide Bilder ganzflächig mit einer bestimmten Intensität sichtbar. Wenn die Mischung einfach additiv vorgenommen wird, so entsteht ein Videoüberpegel, der durch eine nachgeschaltete Regelstufe wieder reduziert werden muss (Abb. 9.2 und Abb. 9.5).

Da bei allen Mischzuständen ein Videopegel von 100% nicht überschritten werden soll, ist es einfacher, die Mischung über zwei gegenläufig gesteuerte Verstärker vorzunehmen. Mit einem Steuerhebel (Faderbar) wird dabei eine Steuerspannung eingestellt, die den Verstärkungsgrad des einen in der gleichen Weise herabsetzt wie den des anderen herauf [198]. Bei Mittelstellung des Faders tragen so beide Signale zu 50% zum Gesamtsignal bei (Abb. 9.4).

Die Schaltfunktion in einem Bildmischer wird in zweierlei Hinsicht erforderlich. Es muss ein Hartschnitt (Cut) möglich sein, d. h. eine Umschaltung zwischen zwei Signalen in der vertikalen Austastlücke. Des Weiteren soll die Umschaltung aber auch nach Ablauf einer bestimmten Anzahl von Zeilen oder eines bestimmten Prozentsatzes der aktiven Zeilendauer vorgenommen werden können. Diese Art führt dann zum Standardtrick (Wipe) und zu den

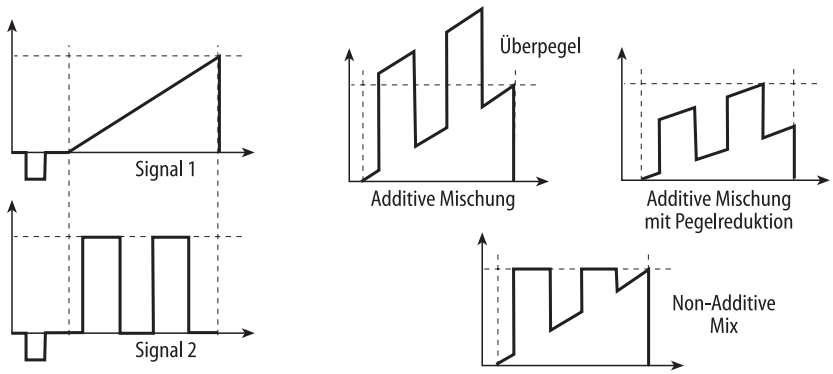


Abb. 9.2. Additive und nicht-additive Mischung

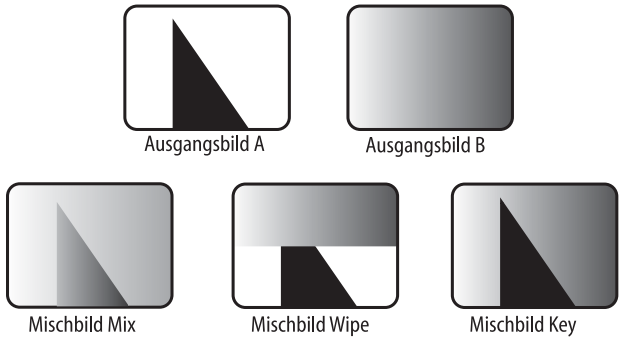


Abb. 9.3. Mischeffekte Mix, Wipe und Key

Stanzverfahren (Key) (Abb. 9.3). Eine besondere Form der Umschaltung ist die geschaltete Mischung, bei der immer das Signal mit dem größten Videopegel eingeschaltet wird (Non Additive Mix, Abb. 9.2).

Die sehr schnelle Ausführung der beschriebenen Steuerung der zur Mix-Funktion benutzten Verstärker entspricht einem Schaltvorgang. Dadurch lässt sich mit nur einem Arbeitsprinzip eine sogenannte Mix/Effektstufe aufbauen, die in Abhängigkeit von einer Steuerspannung die Funktionen Mix, Wipe und Key erfüllen kann, allerdings nur für zwei Signalquellen (Abb. 9.4).

9.1.1.2 Next Channel System

Sollen mehr als zwei Signale an der Mischung beteiligt werden, so kann einerseits eine additive Mischung mehrerer Quellen nach dem Prinzip der Audiomischer verwendet werden, bei dem jedem Kanal ein eigener Pegelsteller zugeordnet ist (Abb. 9.5). Der Nachteil dieses früher verwendeten Knob-a-channel-Systems ist aber, dass dabei eine aufwendige Herabregelung des Überpegels notwendig wird.

Heute werden daher praktisch alle Bildmischer so gebaut, dass das oben genannte Prinzip der gegenläufigen Steuerung zweier Verstärker (AB- oder Next Channel-Prinzip) nicht aufgegeben werden muss. Der Kern dieser Mischer ist die Mix/Effekt-Stufe, bei der kein Überpegel entstehen kann.

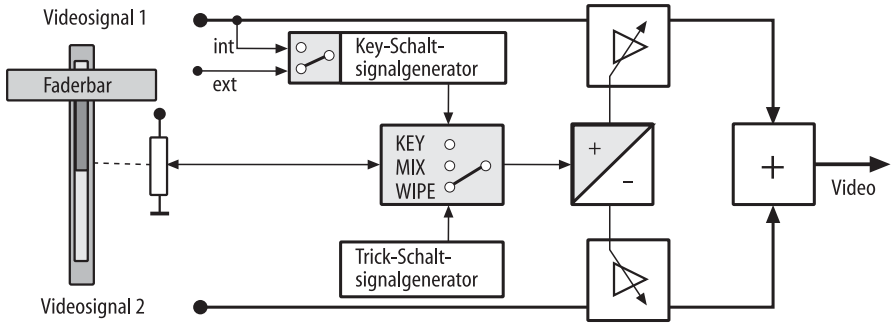


Abb. 9.4. Mix/Effekt-Stufe

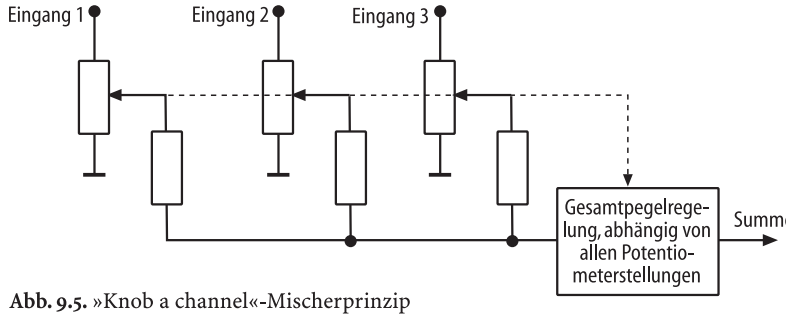


Abb. 9.5. »Knob a channel«-Mischerprinzip

Das Problem der Beschränkung auf zwei beteiligte Signale wird hier gelöst, indem mehrere Mischeffekt-Stufen benutzt werden, die je einer Ebene zugeordnet sind. Mehrere Ebenen werden dann kaskadiert, d. h, dass das Ausgangssignal der Mischstufe einer höheren Ebene als wählbarer Eingang der unteren Ebenen zur Verfügung steht (Abb. 9.6). Die zwei an der Mischung oder dem Effekt beteiligten Signale werden in jeder Ebene über eine Kreuzschiene ausgewählt, daher ergeben sich die großen Bedienoberflächen der Mischer.

9.1.1.1 Videokreuzschienen

Kreuzschienen werden eingesetzt, um den beiden Eingängen der Mix/Effektstufe und auch anderen Geräten auf einfache Weise verschiedene Signalquellen zuordnen zu können. Sie dienen allgemein der flexiblen Verteilung mehrerer Signalquellen an mehrere Videogeräte, wobei die Auswahl auf Tastendruck vorgenommen wird.

Eine einfache Umschaltung zwischen zwei Videosignalen führt in der Regel zu Bildstörungen, da die Synchronsignalstruktur gestört wird, wenn mitten im Halbbild auf ein anderes Signal gewechselt wird. Aus diesem Grund wird dafür gesorgt, dass mit Hilfe elektronisch verriegelter Schalter die Signalumschaltung nur in der vertikalen Austastlücke, also jeweils nach einem Bild, geschieht. Die Steuerlogik für die Schalter entspricht der eines UND-Gatters. Mit dem Tastendruck wird die Umschaltung nur vorbereitet. Erst wenn neben dem

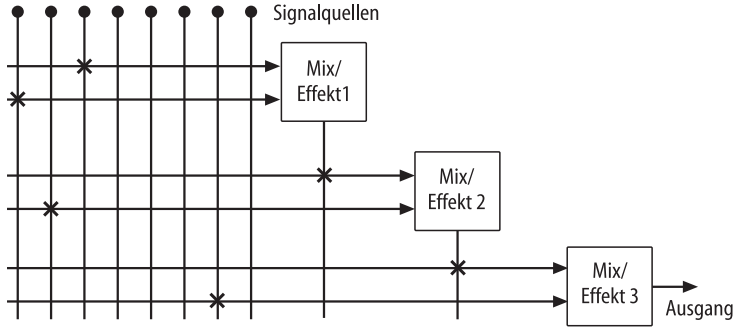


Abb. 9.6. Prinzip der kaskadierten Ebenen

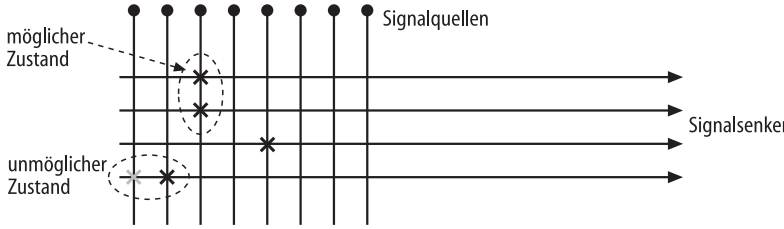


Abb. 9.7. Videokreuzschiene 8 x 4

Tasten- auch der V-Synchronimpuls vorliegt, wird der Schaltvorgang tatsächlich ausgelöst.

Die Ein- und Ausgänge der Kreuzschiene bilden eine Matrix (Abb. 9.7), die entsprechend der Anzahl, z. B. bei 10 Ein- und 5 Ausgängen, als 10 x 5-Kreuzschiene bezeichnet wird. Jedem Kreuzungspunkt ist für die beschriebene Funktion ein elektronischer kontaktloser Schalter zugeordnet. Auf Tastendruck kann ein Eingang auf ein oder mehrere Ausgänge geschaltet werden, es ist aber nicht möglich, mehrere Eingänge auf einen Ausgang zu legen. Abbildung 9.7 zeigt mögliche Schaltzustände.

Wie analoge Kreuzschienen bewirken auch digitale Typen eine Umschaltung zwischen digitalen Composite- oder Komponentensignalen, orientiert am Timing Reference Signal in der V-Austastlücke.

9.1.2 Bildübergänge

Im heute gebräuchlichen A/B-Mischertyp ist die zentrale Einheit die Misch/Effekt-Stufe (Abb. 9.4), die für die in fast allen Mischern möglichen Bildübergangs-Funktionen Cut, Mix und Wipe eingesetzt werden kann.

Der Begriff Cut bezeichnet den Hartschnitt, die direkte Umschaltung zwischen den Signalquellen. Die Quellen können entweder auf der Kreuzschiene geschaltet werden, oder sie werden dort vorgewählt, und der Wechsel erfolgt

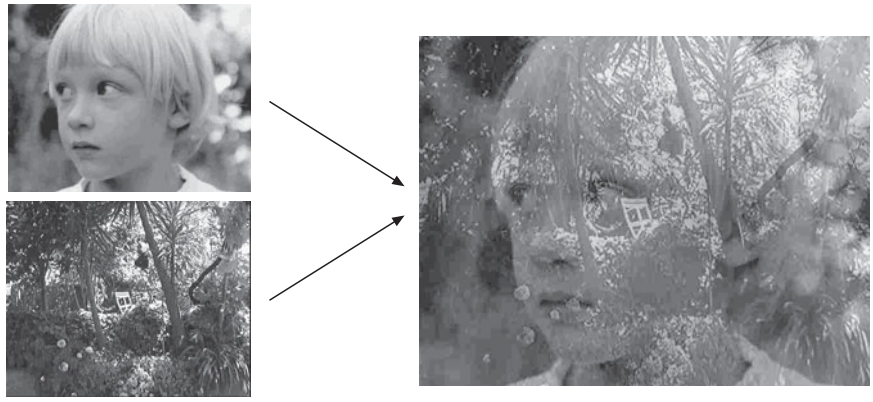


Abb. 9.8. Bildmischung

mit Hilfe der Taste Cut. Die eigentliche Mischung ist dadurch gekennzeichnet, dass beide Bilder ganzflächig sichtbar sind. Bei der Überblendung (Mix, Dissolve) wird das eine Bild mehr oder weniger schnell unsichtbar, während das andere ganzflächig umso mehr in Erscheinung tritt (Abb. 9.8). Der Ablauf kann von Hand gesteuert oder automatisch durchgeführt werden.

Mit der Trickblende (Wipe) wird ein Schaltvorgang zwischen den Signalen ausgelöst, der bewirkt, dass von beiden Bildern nur Teilflächen zu sehen sind. Die einfachste Form des Wipe ist eine waagerechte oder senkrechte Teilung des Bildschirms. Bei senkrechter Teilung kann z. B. im linken Teil die linke Hälfte des Bildes A zu sehen sein, während der rechte Teil die rechte Hälfte von Bild B zeigt. Mit Hilfe des Faderbar oder der Automatikfunktion kann der Teilungspunkt verschoben werden, sodass sich das eine Bild wie von einem Scheibenwischer getrennt über das andere schiebt (Abb. 9.9).

Die senkrechte Teilung in zwei Hälften kann erreicht werden, indem in jeder Zeile nach Ablauf eines bestimmten Teils der aktiven Zeilendauer zwischen den Signalen umgeschaltet wird. Dagegen ergibt sich eine waagerechte Teilung, wenn die Schaltung nicht innerhalb der Zeile, sondern nach Ablauf einer bestimmten Zeilenanzahl geschieht. Bei der bewegten Blende wird über den Faderbar oder eine Automatikstufe eine Steuerspannung eingestellt, die den Umschaltzeitpunkt innerhalb der Zeile oder innerhalb des Bildes variiert.

Durch die Kombination der Schaltvorgänge innerhalb der Zeile und innerhalb eines Bildes können eine Vielzahl verschiedener Trickmuster erzeugt werden. Abbildung 9.9 zeigt Beispiele. Außer durch die intern erzeugten Trickmustersignale kann die Umschaltsteuerung auch durch ein externes Schablonensignal vorgenommen werden. Die Trick-Grundmuster können oft vielfältig variiert werden. Es ist möglich, die Kante am Bildübergang weich einzustellen (Soft Edge), indem im Übergangsbereich nicht hart geschaltet wird, sondern

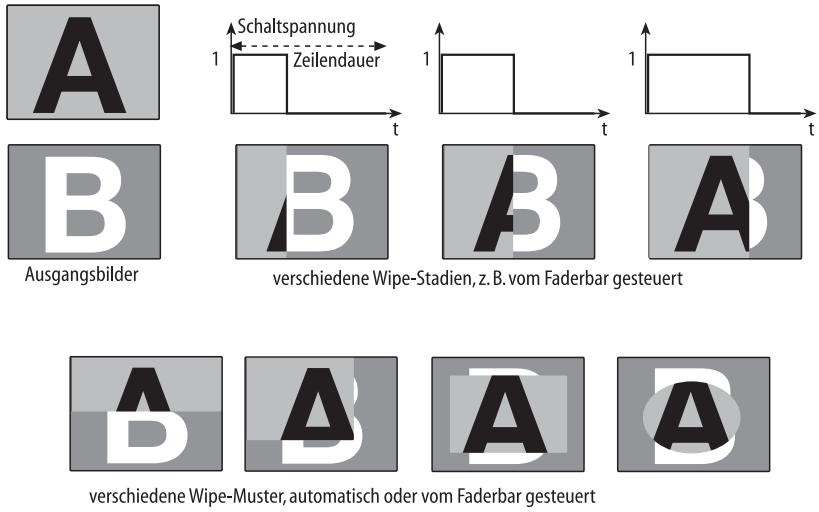


Abb. 9.9. Bildteilung, Wipe

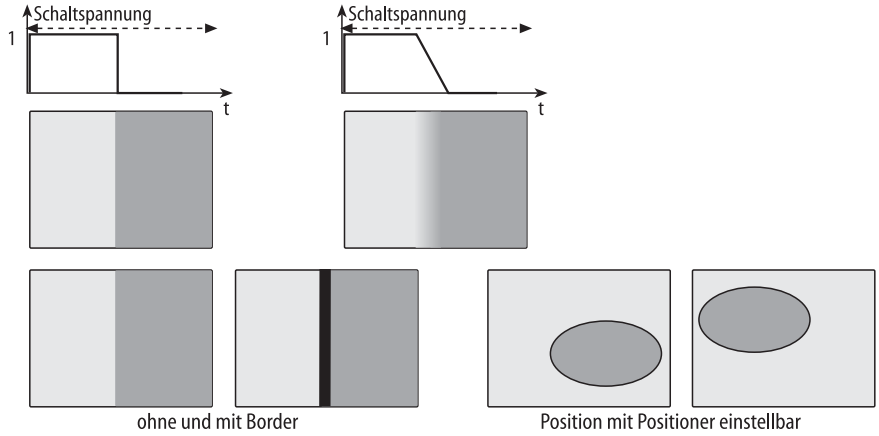


Abb. 9.10. Wipe mit Hard und Soft Edge, mit Border und verschiedenen Positionen

in einem einstellbaren Bereich um die Kante herum eine Überblendung vorgenommen wird (Abb. 9.10). Eine weitere Veränderung der Trickmuster besteht in der Modulation der Kante durch verschiedene Wellenformen. Die Kante kann auch mit einer Grenzlinie (Border) versehen werden, die mit Hilfe der im Bildmischer vorhandenen Farbflächengeneratoren eingefärbt und ihrerseits mit der Soft Edge-Funktion verändert werden kann. Bei Trickmustern, die einen Bereich von Bild A so begrenzen, dass er innerhalb von Bild B liegt (Rechteck, Kreis), ist häufig eine Veränderung des Seitenverhältnisses möglich, sodass z. B. der Kreis zum Oval wird. Die Größe des Ovals oder der anderen Muster wird durch den Faderbar festgelegt. Die Lage des Musters kann mit einem Positioner-Bedienhebel (Joy Stick) verschoben werden (Abb. 9.10).

9.1.3 Stanzverfahren

Auch das Stanzverfahren ist im Prinzip ein Schaltvorgang, allerdings kann die Umschaltung nicht nur nach einem vorgegebenen starren Muster vorgenommen sondern auch aus dem Bildsignal selbst abgeleitet werden. Damit ergibt sich der Vorteil, dass sich der Stanzbereich mit einem Bildobjekt verschiebt. Ein bewegter Bildbereich kann ein Hintergrundbild überlagern, sodass es z. B. aussieht, als bewege sich ein dunkler Skorpion, der im Vordergrundbild auf einem weißen Blatt Papier läuft, auf dem Hintergrundbild, das eine menschliche Hand zeigt. Der Stanzvorgang wird meist als Keying bezeichnet. Bei entsprechendem Umriss erscheint es, als sähe man durch ein Schlüsselloch (Abb. 9.11). Das Signal, aus dem die Form abgeleitet wird, ist das Vordergrundsignal (self key). Die Stanzform entspricht quasi einem Loch im Vordergrund, durch das der Hintergrund zu sehen ist.

Die Schwierigkeit bei der Keyfunktion ist, dass sich der auszustanzende Bereich im Vordergrundsignal deutlich vom Umfeld abheben muss, damit daraus eindeutig das Schaltsignal gewonnen werden kann. Es werden große Hellig-

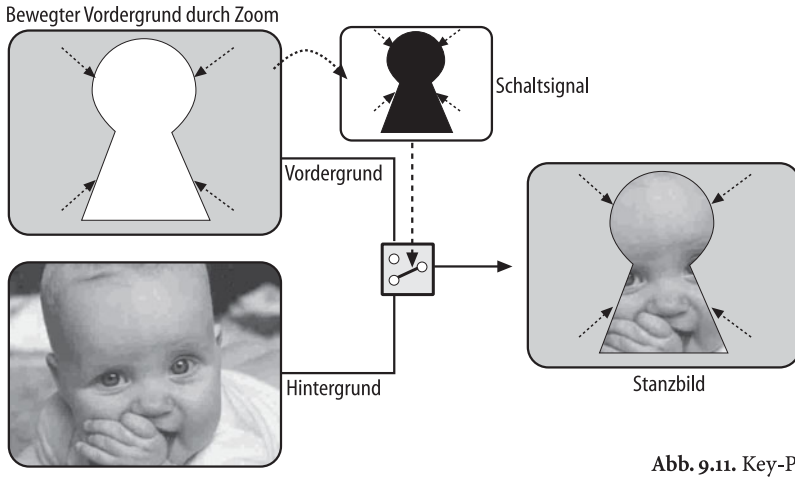


Abb. 9.11. Key-Prinzip

keitsunterschiede (Luminanz Key), oder besonders deutliche Farbunterschiede (Chroma Key) ausgenutzt.

Wenn sich der genannte dunkle Skorpion auf einem weißen Blatt Papier bewegt, kann das Schaltsignal (invers) z. B. per Luminanz Key gewonnen werden. Immer wenn ein Signalwechsel von weiß auf schwarz eintritt, wird auf die Bildquelle mit dem Skorpion umgeschaltet, andernfalls auf das Hintergrundsignal, das die Hand darstellt.

Ein Nachteil der gewöhnlichen Stanztechnik ist die erforderliche starre Einstellung von Hintergrundbild und Kameraposition, denn wenn sich z. B. die Kamera vom Skorpion im Vordergrundbild entfernen würde, erschiene er kleiner, und die Abbildung der Hand müsste sich zum Erhalt der Größenverhältnisse auch ändern. Eine derartige Anpassung des Hintergrundes an eine veränderte Kameraposition ist mit üblichen Key-Funktionen nicht möglich. Dazu sind zusätzlich besondere Hochleistungsrechner erforderlich, die in der Lage sind, das Hintergrundbild in Echtzeit neu zu berechnen (s. Abschn. 10.6).

9.1.3.1 Luminanz Key

Hier wird das Stanzsignal anhand der Helligkeit des Vordergrundbildes gewonnen. Mit Hilfe des Stellers für den Clip Level (Slice) wird ein bestimmter Videopegel definiert und dann festgestellt, an welchen Stellen dieser Wert überschritten wird (Abb. 9.12). Aus diesen so definierten Bereichen wird ein Schwarz-Weißsignal gewonnen, das die Umschaltsteuerung übernimmt. Alle Bildbereiche, in denen der Clip Level unterschritten wurde, zeigen das Hintergrundbild und alle Bereiche, in denen der Level überschritten wurde, das Vordergrundbild. Die Zuordnung kann auch umgekehrt werden (Key Invert). Das in den ausgestanzten Bereich eingesetzte Signal wird als Key Fill bezeichnet.

In der Realität wird nicht hart zwischen Vorder- und Hintergrund umgeschaltet, dafür wären extrem steile Flanken im Stanz- oder Schablonensignal erforderlich. Das Vordergrundbild fügt sich besser in das Hintergrundbild ein, wenn weiche Key-Grenzen benutzt werden. Dazu werden die im Schablonen-

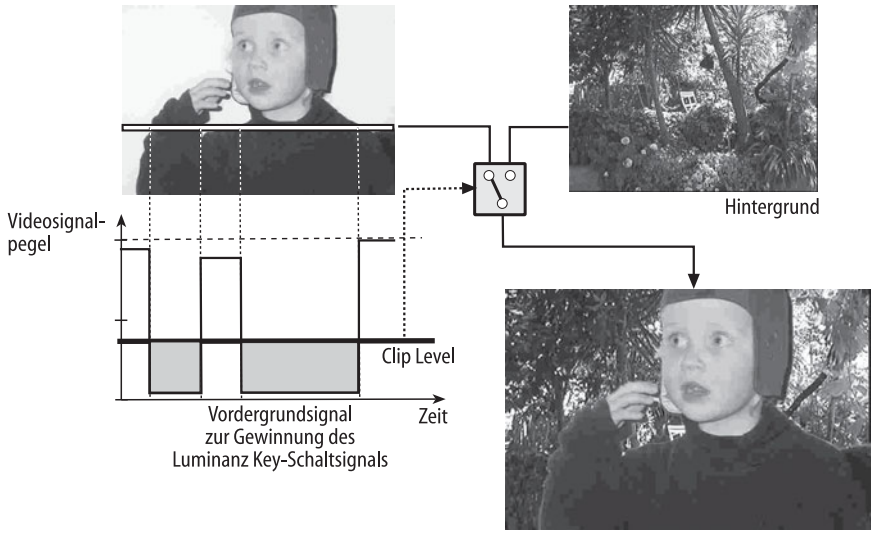


Abb. 9.12. Luminanz Key

oder Keysignal auftretenden Kantenverläufe benutzt, die begradigt werden und zur Steuerung der Mix/Effect-Stufe bzw. der Multiplier in einer separaten Keystufe eingesetzt werden. Genauso wie bei den weich eingestellten Wipe-Kanten wird hier eine Überblendung vorgenommen, und die Breite des Effekts um die Trickkante herum wird durch einen Softness- oder Gain-Steller festgelegt [213]. Die Softness sinkt mit der Steilheit des Schaltsignals (Abb. 9.14). Die Illusion der Einheit von Vorder- und Hintergrundbild beim Key-Trick hängt von der geschickten Einstellung der Flanken des Schablonensignals mit den Stellern Clip Level und Gain ab.

9.1.3.2 Linear Key

Diese Funktion kommt dem Bedürfnis entgegen, dass Bildbereiche, in denen der Hintergrund durchscheint, nicht nur in der Nähe der Stanzkanten sondern auch unabhängig vom Stanzbereich verfügbar sein sollten. Linear Key wird

»Bauchbinde«, erzeugt mit Linear Key

Abb. 9.13. Linear Key

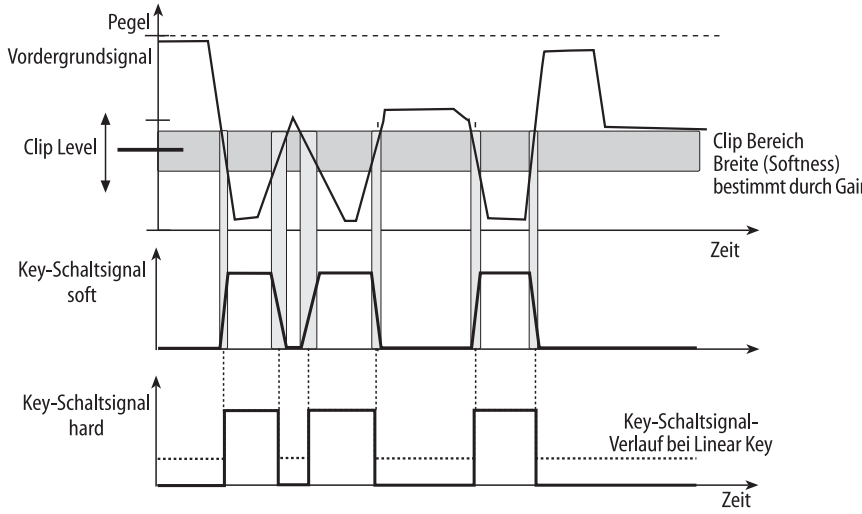


Abb. 9.14. Clipped Key, Soft Key und Linear Key

oft in Verbindung mit Schriftgeneratoren und digitalen Effektgeräten eingesetzt und ermöglicht z. B. die Einblendung von Untertiteln auf transparenten Flächen, sogenannten Bauchbinden (Abb. 9.13) oder transparente Schatten an Buchstaben. Für Linear Key wird ein Schablonensignal verwendet, das einen Signalpegel erzeugt, der zwischen den extremen Key-Schaltzuständen liegt, quasi ein Grausignal und nicht nur eine S/W-Schablone (Abb. 9.14). Dieses Signal steuert im gewünschten Bildbereich den Multiplizierer so, dass der Hintergrund durchscheint. Zusätzlich kann bestimmt werden, ob das Key-Fill-Signal auch transparent ist oder voll durchgesteuert wird. Die Bezeichnung Linear Key gilt, wenn die Transparenz im gestanzten Bild mit der im Original übereinstimmt.

9.1.3.3 Chroma Key

Ein häufig gewünschter Effekt ist das Einsetzen von Personen in künstliche Hintergründe. Für diese und viele andere Stanz-Funktionen ist das Luminanz-Key-Verfahren schlecht geeignet, da sich meist keine ausreichend ausgeprägten Helligkeitsunterschiede ergeben. Bessere Ergebnisse können hier durch das Chroma-Key-Verfahren erzielt werden, bei dem das Stanzsignal aus der Farbe und nicht aus der Helligkeit des Vordergrundbildes abgeleitet wird. Mit dem Chroma-Key-Verfahren kann also eine Person in einen fremden Hintergrund eingesetzt werden, wobei sie sich auch vor dem Hintergrund bewegen darf. Es ist dafür zu sorgen, dass ein großer Unterschied zwischen den Farben der Person und ihrem Hintergrund entsteht. Da der menschliche Hautton wenig Blaufärbung aufweist, ergibt sich dieser Kontrast meist am einfachsten, wenn die Person vor einem möglichst gleichmäßig beleuchteten blauen Hintergrund steht (Blue Screen oder Blue Box). Allgemein kann jede andere Farbe eingesetzt werden, meist werden aber blau und grün verwendet. Es gibt spezielle Chroma-Key-Farben, die sich für den Vorgang besonders gut eignen.

Das Schaltsignal wird gewonnen, indem das Vordergrundsignal auf die Blauwerte des Blue Screen untersucht wird, dazu sollte möglichst ein RGB-Signal vorliegen. Am Mischer wird mit dem Steller »Chroma Key Hue« der Farbton der blauen Wand möglichst genau eingestellt (Abb. 9.17). Mit den Stellern Clip Level und Softness wird wie beim Luminanz Key der Pegel bestimmt, an dem das Stanzsignal zwischen Schwarz und Weiß wechselt und die Breite der Überblendzone festgelegt. Das Verfahren ist kritischer als beim Luminanz Key, denn das Stanzsignal hängt von der Qualität des Videosignals ab. Dieses ist generell nicht frei von Rauschen, sodass die Ermittlung des Umschaltpunktes nicht konstant ist. Das Rauschen wird besonders kritisch, wenn die Signalbandbreite beschränkt ist, wie es bei den Chrominanzanteilen der Fall ist [195].

Ein bekanntes Beispiel für den Einsatz der Chroma-Key-Technik sind die Nachrichtensendungen. Die Moderatoren agieren in einer Blue Box, während die Hintergrundbilder elektronisch überblendet werden (Abb. 9.15). Die Moderatoren dürfen keine blauen Kleidungsstücke tragen, da sonst an diesen Stellen das Hintergrundbild sichtbar würde. Bei einem Hemd mit blauem Kragen erschiene ein Moderator halslos. Die Voraussetzung für ein sauberes Stanzverfahren ist ein hochwertiges Signal, daher sollte das Verfahren möglichst bei der Produktion mit dem Kamerasignal als Vordergrundsignal eingesetzt werden. Wenn das Chroma Key-Verfahren in der Nachbearbeitung verwendet werden muss, sollte bei der Produktion laufend darauf geachtet werden, dass sich das aufgezeichnete Signal auch als Stanzsignal eignet.

Die saubere Einstellung des Key ist eine kritische Angelegenheit und nur bei guter Ausleuchtung des blauen Hintergrundes möglich. Eine schlechte Einstellung führt zu ausreißenden Stanzkanten, die die Illusion eines zusammengehörigen Bildes zerstören (Abb. 9.16). Kritische Stellen sind Haare und Gegenstände, durch die der Hintergrund sichtbar sein soll.

Schatten, die eine Person auf den blauen Hintergrund wirft, können sichtbar bleiben, wenn eine Shadow-Keystufe benutzt wird, mit der der Schatten-

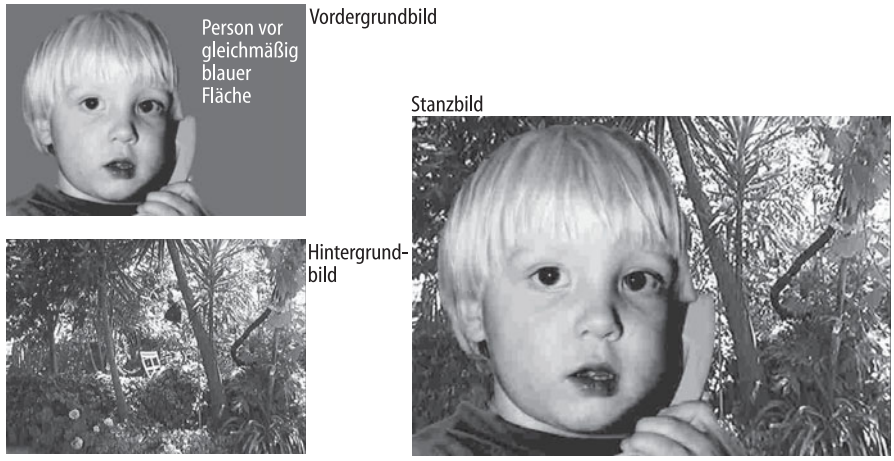


Abb. 9.15. Chroma Key

Abb. 9.16. Gute und schlechte Chroma-Key-Einstellung

bereich im Vordergrundbild registriert und das Hintergrundbild mit Schwarz abgedunkelt wird. Bei besonders kritischen Vordergrundbildern, z. B. wenn der Hintergrund durch Rauch oder Glas sichtbar ist, wird das Ultimatte Chroma Key-Verfahren eingesetzt (s. Abschn. 10.6), ein aufwendiges Verfahren, das durch mehrere Mischprozesse ein besonders realistisches Gesamtbild bietet.

9.1.3.4 Extern Key

Hier wird das Stanzsignal von außen zugeführt, sodass das Vordergrundbild mit extern vorgegebenen Umrissen, ggf. per DSK, über das Hintergrundbild gelegt werden kann (Abb. 9.17). Das externe Stanzsignal wird als Key Source und das in den Umriss eingesetzte Signal als Key Fill bezeichnet. Beim praktischen Einsatz ist der H-Phasenabgleich zwischen den Signalen zu beachten.

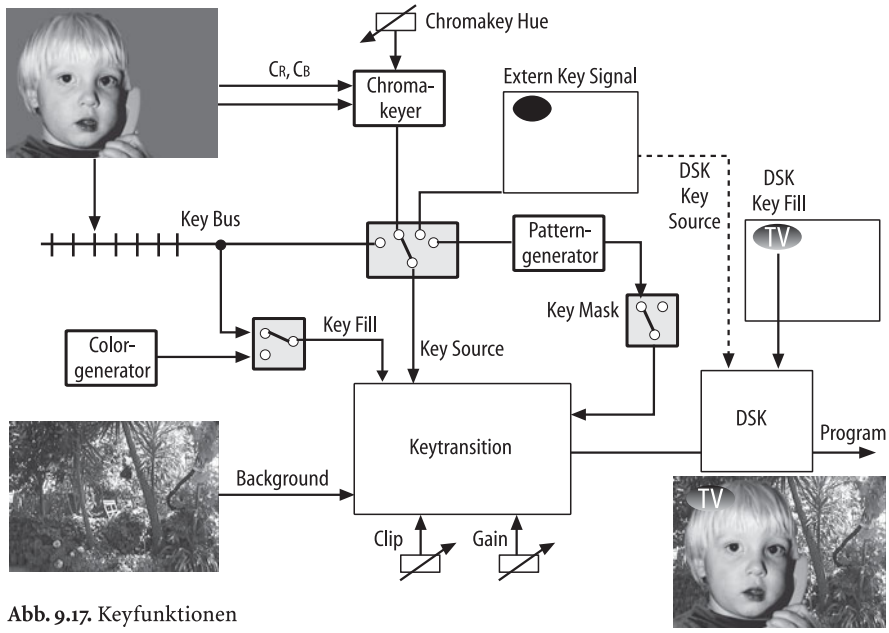


Abb. 9.17. Keyfunktionen

Sehr nützlich ist der Extern Key-Eingang für das Einstanzen von computergenerierter Grafik, wenn der Computer in der Lage ist, neben den eigentlichen, bildbestimmenden RGB-Kanälen mit einem separaten, sogenannten Alphakanal zu arbeiten, der ein Graustufen-Stanzsignal zur Verfügung stellt. Die Graustufe des Alphakanals bestimmt den Mischungsgrad zwischen Vorder- und Hintergrund. Wenn das Stanzsignal an den externen Key-Eingang angeschlossen wird, braucht bei der Bildgestaltung keine Rücksicht darauf genommen zu werden, dass die Grafik stanzfähig gestaltet wird, d. h. hohe Helligkeits- oder Farbkontraste aufweist.

9.1.3.5 Preset Pattern Key

Anstatt eines externen Stanzsignals können auch die für den Wipe-Effekt vorgegebenen Muster (Preset Pattern) eingesetzt werden, um daraus das Stanzsignal abzuleiten. Der ausgestanzte Bildbereich kann mit dem Vordergrundsignal gefüllt werden. Er lässt sich ggf. über den Bildschirm bewegen und mit den für die Wipe-Muster zur Verfügung stehenden Border-Effekten verändern. Im Unterschied zum normalen Wipe kann hier eine feste Bildteilung auf Tastendruck ausgelöst werden.

Der Wipe-Generator kann auch zur Maskierung des Vordergrundes eingesetzt werden (Key Mask), indem die Stanzfunktion auf die dem Muster entsprechenden Bereiche des Vordergrundbildes eingegrenzt wird. Soll z. B. ein Titel in ein Bild gestanzt werden, der in der Vorlage mit einem störenden Untertitel versehen ist, so lässt sich dieser mit einem richtig positionierten Rechteckmuster unterdrücken (Abb. 9.18). Anstatt durch das Vordergrundsignal kann der ausgestanzte Bereich auch mit einem Farbsignal aus den im Mischer vorhandenen Farbflächengeneratoren gefüllt werden (Matte Fill, Effect Matte). Weiterhin können die Grenzlinien zwischen Vorder- und Hintergrundbild besonders gestaltet werden, häufig wird z. B. eine eingestanzte Schrift mit einem andersfarbigen Schatten versehen, um auch bei unruhigem Hintergrund eine gute Lesbarkeit zu gewährleisten.

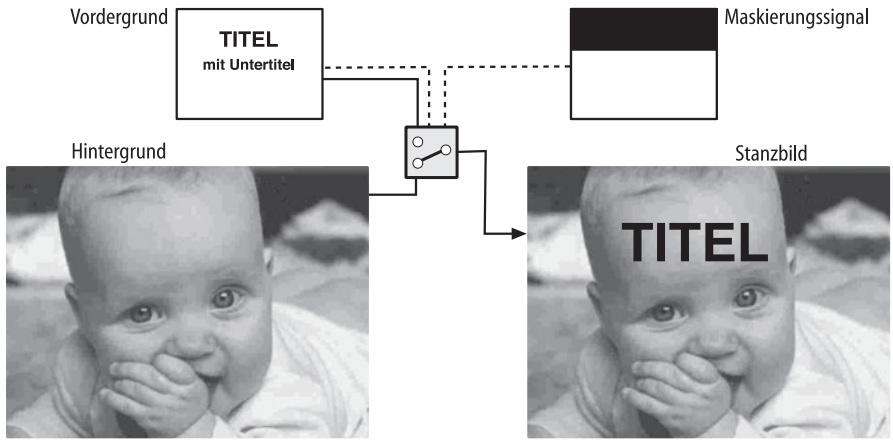


Abb. 9.18. Maskierter Stanzvorgang

9.1.4 Funktionen in Digitalmischern

Die generellen Funktionen sind in Analog- und Digitalmischern gleich, sie werden ähnlich realisiert und auch die Benutzeroberflächen der Geräte sind im Wesentlichen die Gleichen. Die Realisierung einer Mischfunktion entspricht einer Addition von Signalwerten. Im analogen Fall können dazu Spannungen überlagert werden, bei Digitalwerten werden Zahlen addiert (Abb. 9.19). Ähnlich wie bei Analogsignalen, bei denen die Übersteuerungen verhindert werden müssen, ist hier erforderlich, dass der erlaubte Zahlbereich nicht überschritten wird. Dazu wird die Rechengenauigkeit intern erhöht. Da das Ausgabesignal i. d. R. dieselbe Wortbreite wie die Eingangssignale haben wird, ist es erforderlich, die Wortbreiten vor der Ausgabe wieder zu reduzieren, also eine Requantisierung durchzuführen. Wenn bei diesem Prozess einfach nur Bits weggelassen werden, machen sich besonders bei geringen Pegeln Artefakte (Contouring) bemerkbar. Durch Rundungsprozesse oder Hinzufügen von Zufallssignalen (Dithering) können diese erheblich gemindert werden.

Bei der Mischung von nach ITU-R 601 standardisierten DSK-Signalen gibt es die Besonderheit, dass der minimale Analogwert nicht durch die Zahl Null repräsentiert wird, sondern im 8 Bit-Luminanzsignal durch den Wert 16. Daher muss dieser Wert zunächst von den zu mischenden Signalen subtrahiert und schließlich wieder zum Ausgangssignal hinzuaddiert werden (Abb. 9.19). Dabei wird meist nicht einfach der feste Wert 16 verwendet, sondern die Werte aus dem Beginn der Zeile, die den Werten der Austastlücke entsprechen, da die digitale Zeile ja geringfügig länger ist als die analoge. Auf diese Weise können auch kleine Aussteuerungsfehler berücksichtigt werden (digitale Klemmung).

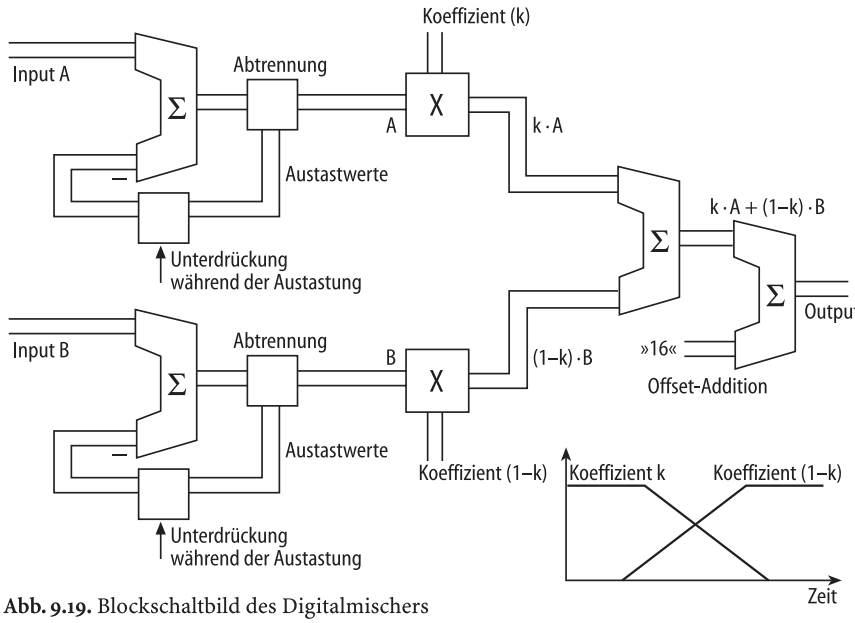


Abb. 9.19. Blockschaltbild des Digitalmischers

Derselbe Vorgang wird auch für die Farbdifferenzkomponenten durchgeführt, jedoch mit einem Offset von 128, bzw. 512 bei 10 Bit-Signalen. Vor der Addition wird hier dann das 2er-Komplement gebildet, sodass die Werte wieder positiv und negativ werden können, wie es auch im analogen Signal der Fall ist.

In Mix/Effect-Stufen wird nicht nur gemischt, sondern es werden auch Signale verstärkt und gedämpft. Diese Funktion wird durch Multiplikation mit bestimmten Koeffizienten realisiert. Eine per Faderbar gesteuerte Überblendung erfordert dabei zwei Multiplizierer, wobei der eine mit den Koeffizienten k der andere mit (1 – k) gesteuert wird [226]. Mit Addition und Mischung lassen sich die oben genannten Bildmischerfunktionen realisieren. Hinzu kommen Schaltvorgänge, die auf digitaler Ebene sehr leicht durchführbar sind. Die Digitaltechnik wird dabei aber nicht nur zur Kopie der analogen Funktionen eingesetzt. Sie ermöglicht eigene Lösungen, die oft im Funktionsumfang über die klassischen Mischerfunktionen hinausgehen und daher in Digital Video Effekt-Geräten (DVE) zu finden sind.

9.1.5 Bildmischeraufbau

Die gebräuchlichen Videomischer sind häufig auf eine Eingangs-Signalform festgelegt, einige Typen erlauben aber auch den gemischten Betrieb. Meist ist die eigentliche Elektronik eines Bildmischpultes vom Bedienpult getrennt (Abb. 9.20). Zur Verbindung dienen herstellerspezifische Systeme, die bei großen Mischern ähnlich arbeiten wie Computernetzwerke. Diese Verbindungen sind bei aufwendigen Geräten so leistungsfähig, dass ein Mischsystem über mehrere Bedienpulte gesteuert werden kann.

Die Elektronikeinheit mit den Ein- und Ausgängen kann zusammen mit der Elektronik der Bildquellen (Camera Control Unit) oder den Videorecordern in einem Maschinenraum untergebracht werden, sodass nur kurze Verbindungs-

Abb. 9.20. Bedienfeld und separate Elektronikeinheit eines Videomischers [186]

leitungen erforderlich sind. Die Elektronikeinheiten sind insbesondere bei analogen Komponentenmischern mit sehr vielen Buchsen bestückt.

Die manchmal komplex erscheinenden Benutzeroberflächen der Videomischer reduzieren sich bezüglich der Grundelemente auf die im wesentlichen gleich ausgestatteten Mischebenen mit den Mix/Effekt-Stufen und den zugeordneten Kreuzschienen, eine Einheit zur Anwahl der den Effekten zugeordneten Attribute (z. B. Wipe-Muster, Stanzfarbe, Keytype, Softedge etc.) und eine weitere Key-Stufe vor dem Signalausgang, die speziell dem Einstanzen von Logos und Schriften dient.

Die Bedienelemente sind dem Signalfluss entsprechend von links nach rechts angeordnet (Abb. 9.21). Die Kreuzschienen liegen links, die hier ausgewählten Signale werden über die in der Mitte liegenden Mix/Effekt-Stufen gemischt. Bevor sie zum Ausgang gelangen, durchlaufen sie die genannte separate Keystufe, die entsprechend der Vorstellung eines von links nach rechts verlaufenden Signalstromes als Down Stream Keyer (DSK) bezeichnet wird. Dieses Konzept ist an dem abgebildeten Produktionsmischer deutlich sichtbar.

Die zentrale Einheit der Ebene, die Mix/Effekt-Stufe, arbeitet im einfachen Fall mit zwei Eingängen. Zur Anwahl der Quellen ist entsprechend eine zweireihige Kreuzschiene erforderlich, bei der zwei Eingänge meist fest mit einem im Mischer erzeugten Schwarzbild und einem Farbbalken- bzw. einem Farbflächensignal belegt sind. An der oberen, mit Program (PGM) bezeichneten

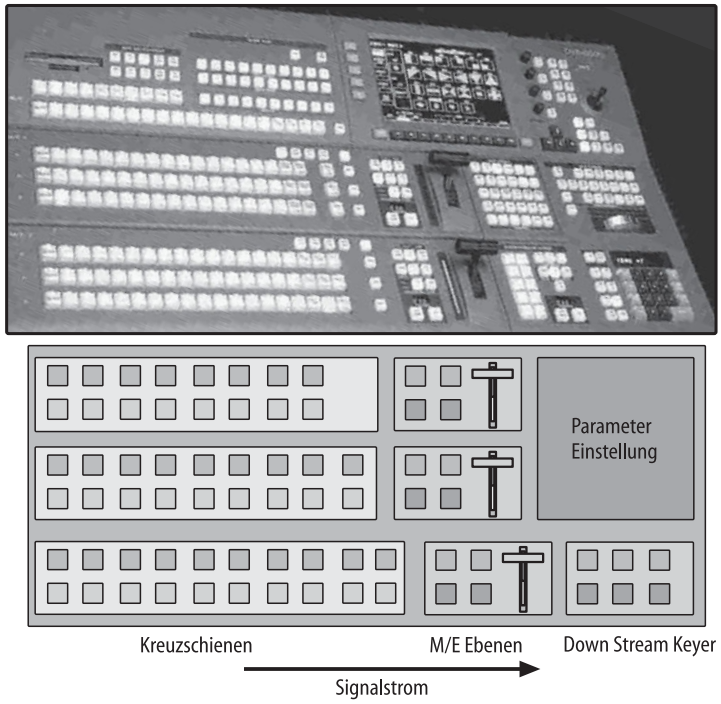


Abb. 9.21. Typische Bedienoberfläche eines Videomischers [186]

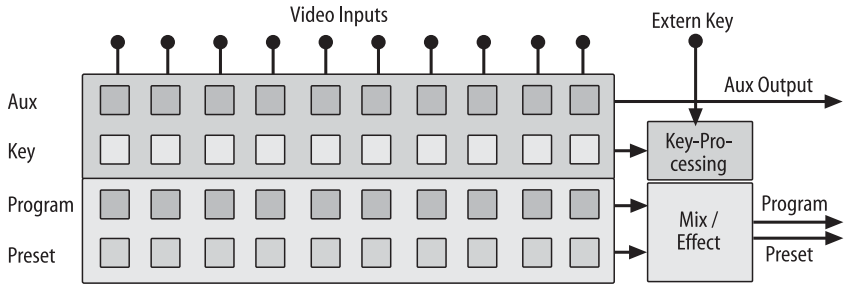


Abb. 9.22. Kreuzschiene mit AUX-, Key-, Program- und Preset-Bus

Tastenreihe (Bus) ist das zum Ausgang durchgeschaltete Signal angewählt, an der unteren, mit Preset (PST) bezeichneten, wird das Signal vorgewählt, das am nächsten Mischeffekt beteiligt sein soll. Im Flip-Flop-Betrieb wechselt das Signal auf der PST-Reihe nach Auslösung des Effekts zur PGM-Reihe und umgekehrt.

Die Effektstufe wird oft so erweitert, dass unabhängig von einem Wipe- oder Mix-Effekt ein Stanzvorgang vorgenommen werden kann. Da dazu zusätzlich die Anwahl einer Stanzsignalquelle erforderlich ist, sind die Kreuzschienen dreireihig aufgebaut. Die Program/Preset-Signale fungieren beim Stanzvorgang als Hintergrund und werden entsprechend auch als Background (BKGD) bezeichnet. Bei der dreireihigen Kreuzschiene liegt die Key-Reihe (Key Bus) oberhalb der Program-Reihe, die Tasten sind meist farblich abgehoben (Abb. 9.22). Anstelle des Key Bus, oder zusätzlich, wird häufig auch ein Universal-Bus (AUX Bus) in der Kreuzschiene verwendet. Hiermit kann unabhängig von allen Mischereinstellungen ein Signal zu einem separaten AUX-Ausgang geführt werden. Damit ist es möglich, alle Eingangssignale anzuwählen, um sie z. B. zu einem an den separaten Ausgang angeschlossenen Effektgerät zu schicken, dessen Ausgang wiederum mit einem Mischereingang verbunden ist. Auf diese Weise können alle Signale effektbehaftet in die Mischung aufgenommen werden, ohne dass die Signalverbindungen am Effektgerät verändert werden müssen.

Neben den Signalwahl-Tasten der Kreuzschiene sind in der Ebene noch die Tasten vorhanden, mit denen die Mix/Effekt-Stufe bedient wird. Bei einem

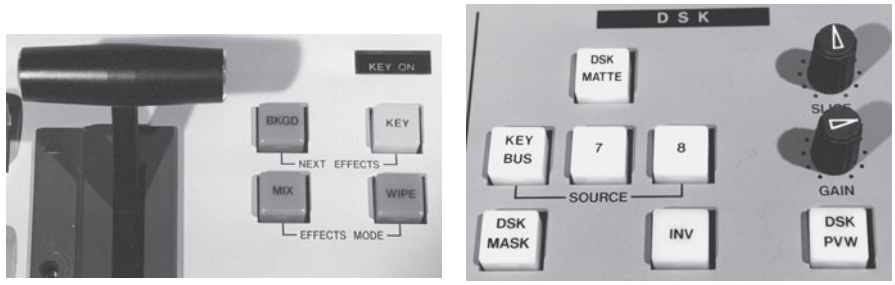


Abb. 9.23. Bedienelemente für den Transition Type und den DSK

einfachen Mischer ohne Key-Möglichkeit befinden sich hier mindestens vier Tasten und ein Hebel für manuelle Steuerung (Faderbar). An den in Abb. 9.23 unten dargestellten beiden Tasten wird der Effects Mode (Transition Type) ausgewählt, d. h. ob der Mischeffekt für die auf PST und PGM angewählten Signale eine Überblendung (Mix) oder ein Standardtrick (Wipe) sein soll. Die Form des Wipe wird über ein separates Bedienfeld festgelegt.

Bei Mischern, die einen Key-Vorgang erlauben, sind die zwei weiteren in Abb. 9.23 gezeigten Tasten (Next Effects) vorhanden. Die Taste BKGD ermöglicht die gerade beschriebenen Übergänge zwischen den Background-Signalen (PGM und PST), während die Taste Key bei der Effektauslösung die gestanzte Überlagerung des auf dem Key Bus angewählten Signals bewirkt. Um den jeweiligen Zustand anzuzeigen, werden selbstleuchtende Tasten verwendet. Eine gewählte Funktion oder ein Eingang wird hell leuchtend, ein vorgewählter Kanal durch geringere Intensität signalisiert. Die Bedientasten für den Keyer selbst sind in Abb. 9.24 dargestellt.

Weitere Tasten dienen der Auslösung des voreingestellten Effekts. Cut bewirkt einen sofortigen Übergang zum PST-Signal, einen harten Schnitt. Die Taste Auto löst dagegen einen langsamen Übergang aus, dessen Dauer meist als Anzahl von Bildern zwischen eins und einigen hundert festgelegt werden kann. Die dritte Möglichkeit zur Ausführung des voreingestellten Effekts besteht in der Nutzung des Faderbars, mit dem individuell z. B. eine Überblendung nur zur Hälfte vorgenommen werden kann, die dann langsam wieder zum Ausgangsbild zurückgeführt wird.

9.1.5.1 DSK

Wie bereits beschrieben, befindet sich vor dem Ausgang der meisten Bildmischer eine unabhängige Down Stream Key-Stufe, die vor allem dem Einstanzen von Logos und Schriften dient (Abb. 9.23). In vielen Fällen ist die Stanzart auf Luminanz Key eingeschränkt, dafür sind aber häufig aufwendige Border- und Shadow-Effekte möglich. Für den DSK sind separate Eingänge vorhanden. Die Down Stream Key-Stufe ist vor allem für Signale vorgesehen, die einem bereits bearbeiteten Bild zum Schluss überlagert werden sollen. Damit lässt sich auch mit einem einfachen Mischer, dessen einzige Key-Stufe durch eine

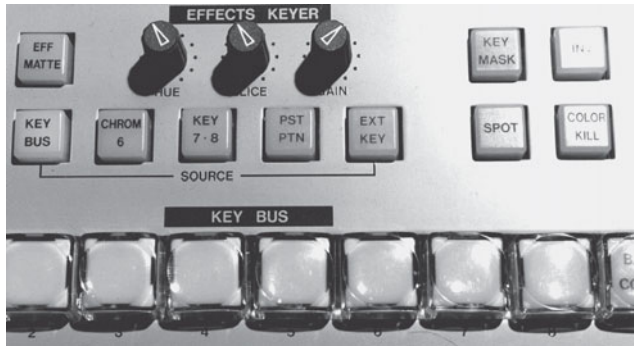


Abb. 9.24. Bedienelemente für die Key-Funktionen

Chroma-Key-Anwendung belegt ist, eine Einstellung realisieren, die häufig bei der Produktion von Nachrichtensendungen auftritt: Auf dem Keybus wird das Kamerasignal angewählt und mittels Chroma Key wird das Vordergrundbild über die mit PGM und PST austauschbaren Hintergrundbilder gelegt. In dieser Situation erlaubt es der DSK, z. B. einen Schriftzug zusätzlich über das Gesamtbild zu stanzen (Abb. 9.17). Noch hinter dem DSK ist bei vielen Mischern die Fade-to-Black-Einrichtung angeordnet, die auf Tastendruck eine Abblende nach Schwarz ermöglicht, ohne dass Veränderungen in der Mix/Effektstufe oder im DSK vorgenommen werden müssen.

9.1.5.2 Farbflächen

Alle modernen Bildmischer enthalten eine Anzahl unabhängiger Farbflächengeneratoren (Matte). Sie werden für die Erzeugung der Fill-Farben und zur Einfärbung von Schatten und Bordern etc. eingesetzt. Die Einstellung erfolgt mit den drei Einstellern »Hue« für den Farbton, »Sat« für die Farbsättigung und »Lum« für die Helligkeit des Farbsignals. Große Bildmischer erlauben auch die Erzeugung von Flächen mit Farbverläufen.

9.1.5.3 Tally

Dieser Begriff steht für die Signalisierung des Kamerazustandes über das Kamera-Rotlicht. Wenn der Bildmischer als Produktionsmischer zur Verarbeitung mehrerer Kamerasignale eingesetzt wird, so ist es erforderlich, den aufgenommenen Personen und den Kameraleuten anzuzeigen, welche Kamera gerade »geschnitten«, d. h. auf Sendung geschaltet ist. Da das gewünschte Signal in der Regie am Bildmischer bestimmt wird, müssen die Mischer in Abhängigkeit vom Kreuzschienenzustand Schaltsignale generieren, die zu den Kameras geführt werden. Große Mischer bieten dabei auch die Möglichkeit der Signalisierung der im Preset-Bereich angewählten Quellen (Preset-Tally).

9.1.5.4 Schnittstellen

Weitere Schaltsignale für den Bildmischer werden über das General Purpose Interface (GPI) geführt. Als Information werden hier nur die Zustände Ein/Aus unterschieden. Diese Verbindung wird bei Mischern als einfache Fernsteuerung vor allem dazu genutzt, bei Veränderung des GPI-Signalzustandes, z. B. durch ein Schnittsteuersystem, einen Mischeffekt auszulösen.

Wenn eine umfassendere Beeinflussung des Mischerzustandes mit einer Vielzahl verschiedener Befehle erforderlich ist, wird anstelle des GPI eine serielle Schnittstelle verwendet. Hier kommt z. B. die RS 422-Schnittstelle zum Einsatz, über die auch die Laufwerksfunktion von MAZ-Maschinen gesteuert werden (s. Abschn 9.3.5). Auch die serielle Schnittstelle wird meist zur Kommunikation mit Schnittsteuersystemen eingesetzt, allerdings ist es aufgrund der vielen verschiedenen Mischertypen schwierig, die Befehle eindeutig den Funktionen zuzuordnen. Damit beschränken sich die Steuerfunktionen auch bei seriellen Schnittstellen meist auf einfache Vorgänge wie die Bildquellenanwahl, die Bestimmung des Transition Type etc. Schon die Auswahl eines Wipemusters stößt auf Schwierigkeiten, da in verschiedenen Mischern unterschiedliche Muster implementiert sind.

9.1.5.5 Effects Memory

Die Auslösung einer sehr komplexen Funktion durch ein externes Steuersignal wird durch die Speicherung des gesamten Mischerzustandes möglich. In einem »Snapshot« werden hier alle aktuellen Betriebsdaten gespeichert und auf einem Speicherplatz im Effects Memory (E-MEM) oder einer externen Diskette bzw. Speicherkarte abgelegt. Durch einen einfachen Aufruf eines dieser Speicherplätze auf Knopfdruck oder durch ein externes Signal kann der entsprechende Mischerzustand sofort wiederhergestellt werden, und ein komplexer Mischvorgang kann automatisch ablaufen. An einigen Bildmischern können die Zustände zusätzlich in eine zeitliche Abfolge gebracht werden, sodass nach einer Definition der zugehörigen Zeiten ganze Sequenzen von Mischerzuständen ablaufen können.

9.1.6 Bildmischer-Typen

Bildmischer unterscheiden sich durch die verarbeitbaren Signalformen, durch digitale oder analoge Arbeitsweise, durch die Anzahl der Eingänge, die Anzahl der Ebenen, bezüglich der Signalverarbeitung und vor allem durch die Qualität der Chroma Key-Stufe. Moderne Bildmischer arbeiten heute digital und sind darüber hinaus sowohl für SD- als auch für HD geeignet. Auch bei Digitalmischern wird meistens das oben erläuterte Benutzerkonzept der Analogmischer übernommen. Einige Typen arbeiten dagegen mit einer sehr kleinen Benutzeroberfläche, über die sehr viele Funktionen erreichbar sind. Diese Geräte sind sehr kompakt, haben aber den Nachteil der Unübersichtlichkeit und langen Zugriffszeit auf die Funktionen.

Für einfache Zwecke reichen Analogmischer aus. Ein Beispiel für einen kleinen Trickmischer, der für die Bearbeitung von Signalen in Studios vorgesehen ist, die mit S-VHS-Geräten ausgerüstet sind, bzw. auf Y/C-Signalen beruhen, ist der Typ KM-D600 von JVC (Abb. 9.25). Dieses Gerät verfügt über vier Eingänge für Y/C oder FBAS-Signale. Es ermöglicht Mix, Wipe, Luminanz Key und Chroma Key und stellt darüber hinaus auch einige einfache Digitaleffekte minderer Qualität (z. B. Zoom) zur Verfügung. Eine Besonderheit ist hier der integrierte 2-kanalige TBC, der die Einbindung von Signalen ermöglicht, die nicht an den Studiotakt gekoppelt sind.

Abb. 9.25. Kleiner Y/C-Mischer [118]

Abb. 9.26. Analoger Komponentenmischer

Ein bekannter analoger Bildmischer für Composite- oder Komponenten-Signale war der Typ GVG 110 von der Firma Gras Valley Group, der in vielen Nachbearbeitungsstudios und kleinen Ü-Wagen zu finden war. Das Gerät hat 10 Eingänge, von denen acht mit einer »10 auf 3« Kreuzschiene frei belegbar sind, und bietet alle wesentlichen Effekt- und Keymöglichkeiten. Ein vergleichbares Gerät ist das Modell KM 3000 von JVC (Abb. 9.26) mit einer Benutzeroberfläche, die der oben beschriebenen entspricht.

9.1.6.1 Digitale Bildmischer

Kleine Bildmischer sind für einfache Aufgaben, z. B. für den Einsatz an linearen Schnittplätzen vorgesehen oder für Anwendungen außerhalb des Fernsehbereichs, wie für die Zuspielung bei Projektionen etc. Hier wird nur eine geringe Zahl von Videoeingängen gebraucht, dafür aber oft ein digitales Videoeffektsystem. Abbildung 9.27 zeigt als Beispiel die Benutzeroberfläche und das Anschlussfeld eines solchen Gerätes von der Firma GrassValley. Es kann analoge und digitale SD- und optional auch HD-Signale mit 1080i und 720p verarbeiten und von SD auf HD konvertieren. Hinzu kommt die Möglichkeit, auch Audiosignal zu mischen. Ein Videoeffektgerät ist integriert.

Abb. 9.27. Kleiner Digital-Mischer [219]

Mittelgroße und größere Digitalmischer sind häufig modular aufgebaut. Sie arbeiten digital, können aber mit seriell digitalen Anschlüssen ebenso ausgestattet werden wie mit analogen. Auch die Signalverarbeitungsstufen können modular eingesetzt werden, sodass z. B. nur die Key-Stufen erworben werden müssen, die tatsächlich gebraucht werden. Ein Beispiel für einen großen und verbreiteten Bildmischer nach diesem Konzept ist das in Abbildung 9.28 dargestellten Gerät Sony MVS 8000. Es kann für bis zu vier Ebenen mit 34 Eingängen und bis zu 24 Ausgängen modular konfiguriert werden. Die Abbildung zeigt die größte Ausbaustufe.

SD- und HD-Signale mit 1080i oder 720p können hier parallel verarbeitet und für bis zu 16 Kanäle auch intern konvertiert werden. Die Mischerelektronik, der Switcher-Processor, wird über ein Datenbussystem mit den Benutzeroberflächen verbunden, die auch als kompakte Control Panels mit nur einer Ebene vorliegen können. Im Multi Program Mode können gleichzeitig zwei Programme bereitgestellt werden, die Mix/Effekt- und Key-Stufen können der erforderlichen Signalverarbeitung angepasst und entsprechend flexibel konfiguriert werden.

Das System bietet die Integration eines zugehörigen digitalen Videoeffektgerätes (DVE) und mehrere Down Stream Keyer, sodass z. B. bei multinationalen Produktionen zeitgleich parallele oder kaskadierte Einblendungen vorgenommen werden können. Ein solcher Produktionsmischer arbeitet mit zwei redundanten Netzteilen und ist auch hinsichtlich der Computersteuerung auf Echtzeitsteuerung und Betriebssicherheit hin optimiert.

Die Digitaltechnik ermöglicht nicht nur die Herstellung aufwendiger Digitalmischer, sondern auf der anderen Seite auch eine erhebliche Verkleinerung der Geräte. Einfache Mischer können inzwischen so kompakt gebaut werden, dass sie zusammen mit einer Digitalisierungshardware auf einer Computersteckkarte integriert werden und damit den Kern computergesteuerter Schnittplätze darstellen können.

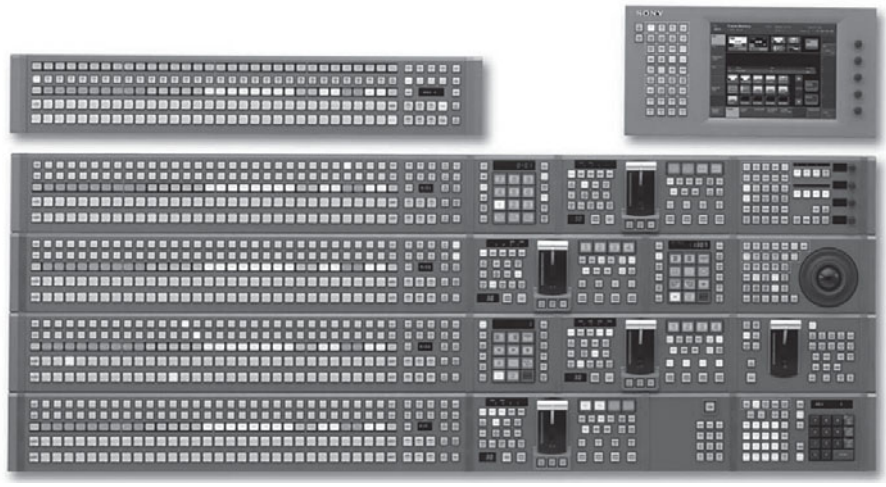


Abb. 9.28. Großer Produktionsmischer [208]

9.2 Videografik- und Effektgeräte

Grafik- und Effektgeräte arbeiten ggf. mit analogen Ein- und Ausgängen, intern aber ausschließlich auf digitaler Basis, da nur so die aufwendigen Funktionen realisiert werden können. Sie werden für den Live-Produktionsbereich als betriebssichere separate Geräte gebraucht, im Postproduktionsbereich aber zunehmend durch Lösungen auf Computerbasis ersetzt. Gleichzeitig verschwimmen zunehmend die Grenzen zwischen den Effektgerät-Gattungen, denn die Einzelgeräte werden in ihrem Funktionsumfang immer vielfältiger.

9.2.1 Schriftgeneratoren

Seit Beginn der Fernsehentwicklung spielt die Überlagerung von Videobildern mit Schrift und Grafik zur Darstellung von Titeln, Namenseinblendungen, Sportergebnissen etc. eine wichtige Rolle. Ihrer Bedeutung ist unabhängig von Modeerscheinungen, denen die Video-Trickeffekte oft unterworfen sind. Im einfachsten Fall kann die Grafik mit einer Insertkamera aufgenommen und in das Videobild eingestanzt werden, hier sind besonders kontrastreiche Bildvorlagen erforderlich. Heute werden die Schriften und Grafiken häufig elektronisch erzeugt, eine separate Kamera ist nicht erforderlich. Wenn der Schriftgenerator in der Lage ist, zusätzlich zum Video-Grafiksignal ein Key-Signal zu liefern, so kann dieses zum Extern Key-Eingang des Mischers geführt werden, sodass bei der Gestaltung der Grafik auch keine Rücksicht auf die Key-Fähigkeit genommen zu werden braucht (Abb. 9.29).

Bei der Gestaltung muss beachtet werden, dass das Videobild bei der Produktion oft anders wiedergegeben wird als bei den Konsumenten, die insbesondere bei SD-Wiedergabe auf einem Röhrenmonitor mit einem beschnittenen Bild versorgt werden. Damit die gesamte Schrift oder Grafik sichtbar ist und darüber hinaus auch noch von einem Rand umgeben ist, sollte das Titelfeld deutlich kleiner sein als das übertragene Bildfeld. Setzt man die Bildbreite des übertragenen SD-Bildfeldes zu 100%, so sollten sich die bildwichtigen Teile innerhalb von 90% der Bildbreite befinden (Safe Area), während für das Titelfeld nur 80% vorgesehen sind. Bei HD-Signalen soll der Titelbereich auf jeder Seite vertikal um 5% und horizontal um 10% verkleinert sein (Abb. 9.30).

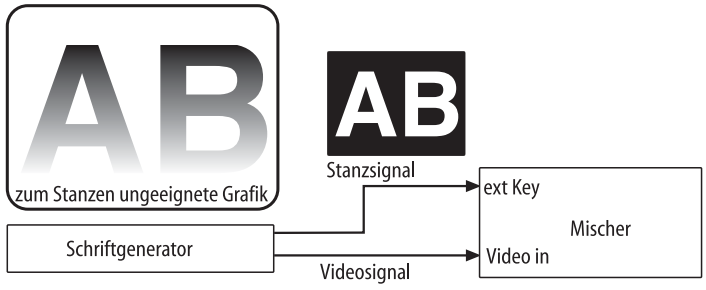


Abb. 9.29. Grafikeinblendung mit separatem Key-Signal

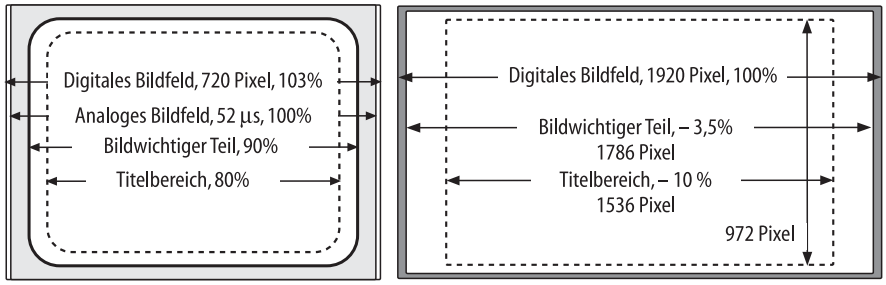


Abb. 9.30. Bildfeld und Titelbereich

Moderne Schriftgeneratoren sind wie Standard-Computer ausgestattet, meist mit Standard-Prozessoren und Festplatten zur Datenspeicherung (Abb. 9.31). Im Gegensatz zu Standard-PC ist eine aufwendige Videografikausgabeeinheit vorhanden, die oft mit doppeltem Video-RAM arbeitet und damit einen schnellen Austausch zweier Grafikseiten erlaubt. Eine wichtige Funktion ist dabei das Scrolling, d. h. einen Text als Lauftitel vertikal oder horizontal ablaufen zu lassen. Die Darstellungsqualität wird maßgeblich vom Antialiasing bestimmt. Die Geräte erlauben heute über Netzwerke eine Anbindung an PC-Systeme, sodass ein Direktimport aus PC-Textverarbeitungsprogrammen in ein vorgefertigtes Layout vorgenommen werden kann. Wichtige Schrift- und Grafikformate (Postscript, Tiff, Pict), die sich bei der Druckvorlagenerstellung etabliert haben, werden ebenso unterstützt wie Photoshop-Pixelgrafiken oder Vektor-Outlines aus dem Programm Illustrator. Die enge Anlehnung an Standard-PC ermöglicht den Zugriff auf die große Anzahl professioneller Schriften aus dem Druckbereich, die frei skaliert und mit allen möglichen Attributen versehen werden können. Der Schriftgenerator leistet vor allem den Transfer der Computerschriften und -grafiken ins Videoformat und bietet die Möglichkeit, die Schrift zu bewegen. Große Systeme sind diesbezüglich auch nicht mehr auf bestimmte Bewegungsrichtungen eingeschränkt, sondern bieten eine freie Animation.

Abbildung 9.31 zeigt den Schriftgenerator A 72 von Abekas. Bekannte Hersteller von Schriftgeneratoren sind die Firmen Aston und Chyron. Das Gerät

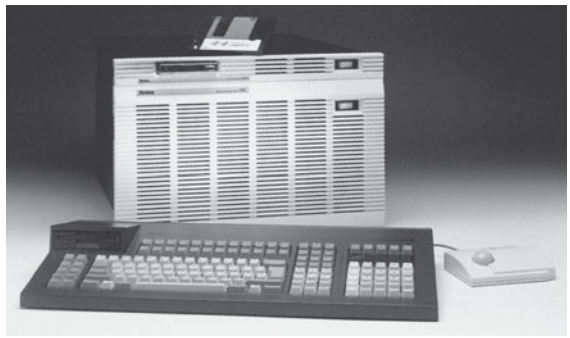


Abb. 9.31. Schriftgenerator [1]

Aston Ethos z. B. bietet Zweikanalbetrieb bei 32 Bit Farbtiefe sowie ein 256-stufiges Antialiasing für bis zu 4000 verschiedene Schrifttypen auf einer Seite. Die Schriftgröße kann zwischen 6 und 576 TV-Zeilen eingestellt werden. Es werden mehrzeilige Roll- und Kriechtitel unterstützt, die mit und ohne Beschleunigung ablaufen können, und jeder Buchstabe kann mit bis zu 20 Attributen wie Drop Shadow, Emboss, Outline, Doppelkante, Spiegelung etc. versehen werden.

Bei der Schriftwahl und -gestaltung sollte, insbesondere bei interlaced-Signalen, darauf geachtet werden, dass sehr feine Linien der Breite von nur einer Zeile nur in einem der beiden Halbbilder auftauchen und daher zu einer flimmernden Wiedergabe mit 25 Hz neigen. Sie sind zu meiden und es sind ggf. serifenlose Schriften zu bevorzugen.

9.2.2 Digitale Effektgeräte

Trickeffekte, die über das Schalten, Stanzen, Blenden und Mischen von Videosignalen hinausgehen, bleiben meist den digitalen Effektgeräten (DVE) überlassen. In diesen Geräten wird ein Bild komplett gespeichert und in sehr kurzer Zeit neu berechnet, um durch Veränderung der Lage und Art einzelner Pixel Bildverzerrungseffekte wie Bildverschiebung, elektronisches Zoom, Bildrotation, Farbänderung etc. zu erzielen. Der wesentliche Unterschied zur Bildbearbeitung in Standard-PC ist hier die Echtzeitfunktion für den Live-Einsatz und die Qualität. Man unterscheidet 2- und 3-dimensionale Effekte, je nachdem, ob

Abb. 9.32. 3D- und 2D-Digital-Videoeffekte (DVE)

Abb. 9.33. Sequenz von Digitaleffekten

sich bei der Positionsveränderung ein räumlicher Eindruck ergibt oder nicht. Ein zweidimensionaler Effekt ist z. B. eine einfache Bildverkleinerung (Abb. 9.32). Ein dreidimensionaler Effekt ist im Spiel, wenn ein Bild so erscheint, als sei es quasi um einen Körper herum gewickelt oder gewölbt. Ob 3D-Effekte möglich sind, hängt von der Leistungsfähigkeit des Gerätes ab, in jedem DVE kann aber zumindest das Bild vergrößert und verkleinert werden, wobei sich auch das Seitenverhältnis ändern lässt. Weiterhin ist die Position des Bildes auch meist so weit verschiebbar, dass es ganz aus dem sichtbaren Bereich verschwinden kann.

Unabhängig von Bildlagen- und Größenänderungen kann mit dem DVE die Bildauflösung oder Schärfe beeinflusst werden (Weichzeichnereffekt), außerdem sind oft auch Bewegungsunschärfen und Farbveränderungen einstellbar. Eine weitere Funktion ist die Programmierbarkeit ganzer Tricksequenzen. Hierbei wird ein Bild in verschiedener Weise in seiner Position verändert. Die jeweiligen Bildlagen und -größen werden gespeichert und in eine zeitliche Reihenfolge gebracht. Bei einem automatischen Ablauf der Sequenz werden zusätzlich die Bildpositionen zwischen den vorher gespeicherten sogenannten Keyframes interpoliert, sodass der Eindruck eines zusammenhängenden Bewegungsablaufs entsteht (Abb. 9.33).

Beim Einsatz eines DVE leidet die Bildqualität. Wenn z. B. ein Bild verkleinert wird, kann der Verkleinerungsfaktor so ungünstig liegen, dass Alias-Strukturen auftreten. Diese müssen durch aufwendige Filtertechniken und angepasste Steuerung der Bildauflösung verhindert werden. Dabei müssen die Tiefpassfilter zur Verhinderung der Alias-Effekte so gestaltet sein, dass die

Abb. 9.34. DVE-Gerät [1]

Grenzfrequenz an die jeweilige Bildkompression angepasst ist [185]. In der Praxis wird mit einem Satz umschaltbarer Filter gearbeitet, da eine kontinuierliche Änderung der Grenzfrequenz nicht möglich ist.

DVE arbeiten ein- oder mehrkanalig. Damit ist die gleichzeitige Manipulation eines oder mehrerer Videosignale möglich, sodass z. B. eine räumlich verkleinerte Bildsequenz über einer Sequenz im Hintergrund ablaufen kann. Häufig steht darüber hinaus noch ein Key-Channel zur Verarbeitung eines Stanzsignals zur Verfügung, über den parallel zu dem zu manipulierenden Bild auch das zugehörige Stanzsignal verändert wird. DVE's werden meist als separate Geräte ausgeführt, die an den Bildmischer angeschlossen werden und teilweise auch darüber bedient werden können. Beispiele für DVE-Geräte sind der ADO von Ampex, Kaleidoskop von GVG oder das in Abb. 9.34 dargestellte Gerät von Abekas.

Auch im Bereich der digitalen Effekte geht der Trend hin zu mehr Integration in ein Komplettsystem bei gleichzeitig vermehrtem Einsatz von Standardcomputern. Ein Beispiel dafür ist die PC-gesteuerte Hardwareeinheit Alladin von Pinnacle, die einen 4-Kanal-Bildmischer mit Chrominanz/Luminanz-Keying, einen Standbildspeicher und ein 3D-Effektgerät enthält. Das Gerät wird mit einer speziellen Steuersoftware bedient, die Animation und professionelle Untertitelung sowie die Effektprogrammierung mittels einer Timeline ermöglicht.

9.2.3 Grafik und Animation

Mit Grafik- und Animationssystemen können Videobilder und Bildfolgen nicht nur bearbeitet, sondern vor allem kreiert werden. Der Begriff Grafik ist im Videobereich noch immer mit dem Namen Paintbox verbunden, dem Paintsystem von der Firma Quantel (Abb. 9.35), das als erstes im professionellen Sektor eine weite Verbreitung gefunden hatte. Paintsysteme dienen der Erstellung oder Bearbeitung von Standbildern. Zur Bearbeitung fertiger Vorlagen werden diese mit einer Kamera oder einem Scanner aufgenommen und in digitaler Form gespeichert, soweit sie nicht bereits in einem digitalen Grafikformat wie Tiff, Pict, PCX etc. vorliegen.

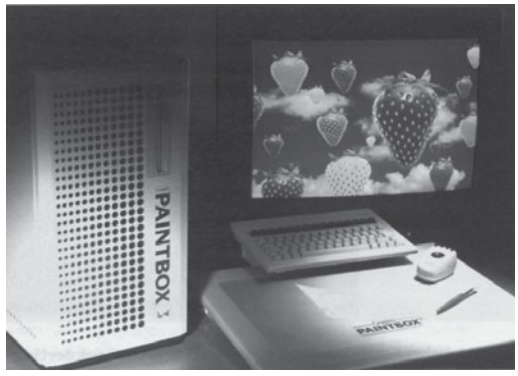


Abb. 9.35. Grafiksystem Paintbox [171]

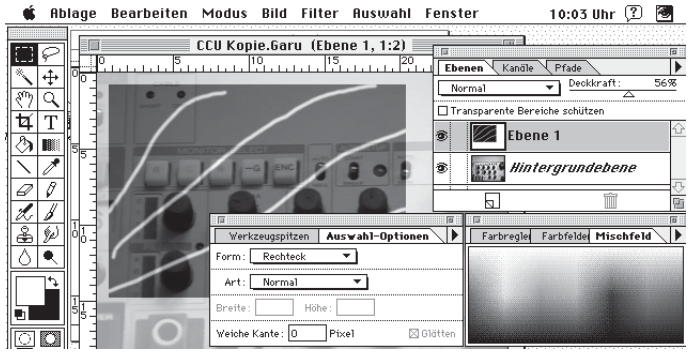


Abb. 9.36. Benutzeroberfläche eines Paint-Programms

Im Grafikprogramm stehen über Menüs Werkzeuge zur Erstellung vorgefertigter geometrischer Formen sowie für Freihandzeichnungen zur Verfügung. Um Freihandzeichnungen mit Werkzeugen wie Zeichenstift, Pinsel oder Airbrush (Sprühpistole) in möglichst gewohnter Form durchführen zu können, wird als elektronischer Pinsel oft ein druckempfindlicher Stift zusammen mit einem Grafiktablett zur Eingabe verwendet (Abb. 9.35). Die Bilder können farblich verändert werden, es können Teile ausgetauscht und retuschiert werden und es stehen Textwerkzeuge zur Verfügung. Mit Filtern kann auch der Schärfeeindruck verändert oder z. B. ein Effekt erreicht werden, der das Bild wie ein Aquarell aussehen lässt.

Viele der genannten Funktionen sind auch mit Programmen für Standardcomputer erreichbar (z. B. Photoshop, Abb. 9.36). Spezialisierte Video-Paintsysteme bieten aber eine Grafikausgabe als hochwertiges Videosignal und zusätzlich besondere, an den Videobereich angepasste Funktionen wie die Steuerung von MAZ-Maschinen.

9.2.3.1 Animation

In Animationssystemen werden nicht nur Einzelbilder erzeugt, sondern zusammenhängende Bildfolgen, die die Empfindung eines Bewegungsvorganges hervorrufen. Mit Hilfe leistungsfähiger Programme können fotorealistische 3D-Bilder erzeugt werden, die im Videobereich vor allem zur Erzeugung von Vorspännen und Logos, zur Verdeutlichung komplexer Zusammenhänge (Wissenschaftsfeatures) oder zur Erzeugung künstlicher Hintergründe eingesetzt werden.

Grafik und Animation sind Gebiete, in denen nicht immer ein Echtzeitbetrieb erforderlich ist. Es werden daher oft keine videospezifischen Geräte, sondern leistungsfähige Computer verwendet. Standard PC sind häufig flexibler und heute so weit entwickelt, dass sie viele Aufgaben in diesem Bereich erfüllen können und sie werden aufgrund des geringen Preises auch verwendet, obwohl ihre Anfälligkeit gegenüber Störungen für Sendeanstalten manchmal problematisch ist. Ein wesentlicher Aspekt für die Funktion ist der nichtlineare Zugriff auf die Bildsequenzen, der nicht an das Speichermedium Band gebunden ist.

9.3 Elektronische Schnittsteuersysteme

Die neue Zusammenstellung gespeicherter Bildsequenzen wird in Anlehnung an den Kinofilm, der zur Umstellung der Sequenzen tatsächlich zerschnitten und wieder zusammengeklebt wird, im deutschsprachigen Raum auch im Videobereich mit Schnitt bezeichnet, obwohl die französischen oder englischen Begriffe Montage oder Editing den Vorgang wesentlich treffender bezeichnen, da es ja im Wesentlichen um eine Neuanordnung der Sequenzen geht, die einen erheblichen Einfluss auf die Wirkung des Films als Ganzes hat.

Prinzipiell kann auch ein Videoband zerschnitten und wieder zusammengeklebt werden. Dieses Verfahren wurde aber nur bis Anfang der 60er Jahre bei den 2"-Quadruplexmaschinen angewendet. Auf speziellen Schneidlehren (Abb. 9.37) musste dabei mit Hilfe eines Oszilloskops ein elektronischer Impuls aufgesucht werden, der den Bildanfang markiert. Neben der problematischen Lokalisierung der Bilder und der Unzuverlässigkeit der Klebestellen entstanden hier Probleme mit dem Bild-Tonversatz. Allgemein lässt sich feststellen, dass die technische Schnittausführung entscheidend von der Art der Bildspeicherung abhängt. Film wird mechanisch geschnitten, digitale gespeicherte Bildfolgen werden durch Verweise umgeordnet und Bildsequenzen auf Videobändern werden kopiert, ein Vorgang der hier zunächst im Vordergrund steht.

9.3.1 Grundfunktionen

In Verbindung mit Schnittsystemen tauchen immer wieder die Begriffspaare linear/nonlinear Editing und Online/Offline-Schnitt auf. Der Begriff Linear Editing steht für das Kopieren von auf Magnetband gespeichertem Bildmaterial, wobei die gewählten Sequenzen in der gewünschten Reihenfolge linear hintereinander auf einem zweiten Band gespeichert werden (Abb. 9.38). Dieser Vorgang erfordert immer einen qualitätsmindernden Kopiervorgang, der den

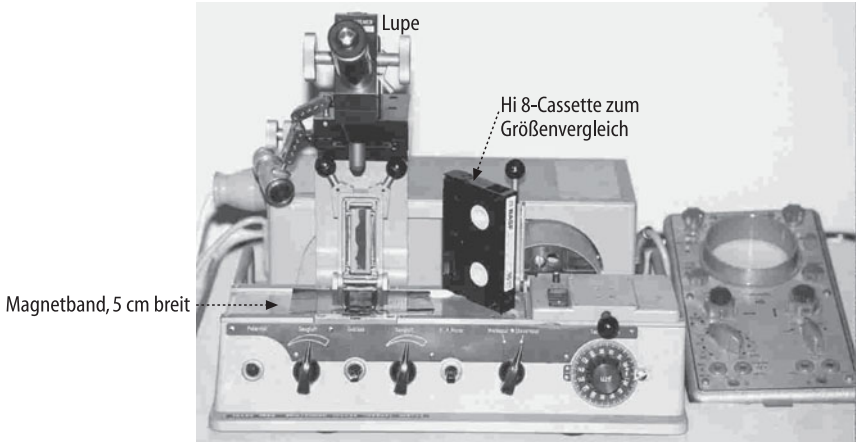


Abb. 9.37. Einrichtung für den mechanischen Schnitt eines 2"-Magnetbandes

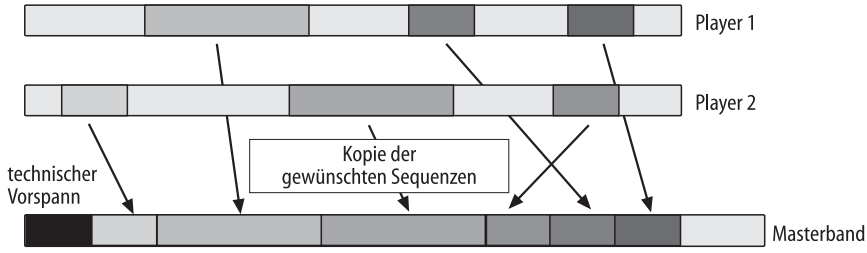


Abb. 9.38. Linearer Schnitt

Einsatz möglichst hochwertiger Aufzeichnungsmaschinen erfordert. Linearer Schnitt ist also auch HD-tauglich, die Auflösung hängt nur von den MAZ-Maschinen und ggf. dem Bildmischer ab.

Beim linearen Schnitt sind ein oder mehrere Zuspieler und ein Recorder beteiligt. Am Schnittpunkt wird der Recorder von der Wiedergabe- in die Aufnahmebetriebsart umgeschaltet. Die Maschinen müssen dabei fernsteuerbar sein, und die Bilder müssen einzeln lokalisiert werden können, was über eine Nummerierung mit dem Timecode erreicht wird. Dieser Vorgang ist aufgrund erforderlicher Vorlauf- und Umspuldauern mit relativ hohem Zeitaufwand verbunden, daher wird der lineare Schnitt fast nur noch dann verwendet, wenn mit bandgestützter Aufzeichnung gearbeitet wird und aus Zeitknappheit die Einspielung in ein nichtlineares System vermeiden werden soll.

Beim Nonlinear Editing wird die Wiedergabe der gewählten Bildfolgen in der gewünschten Reihenfolge einfach dadurch erreicht, dass eine Adressierung der Bildsequenzen in anderer Form als bei der Aufzeichnung vorgenommen wird (Abb. 9.39). Die Voraussetzung dafür ist ein System mit einer hohen Zugriffsgeschwindigkeit, sodass bei dem Übergang zu einer anderen Sequenz die Video-Bildfolgefrequenz aufrechterhalten werden kann. Dabei ist neben der Freiheit und der Geschwindigkeit beim Editing auch von Vorteil, dass der Editiervorgang nicht mit dem verlustbehafteten Kopiervorgang verknüpft ist.

Nonlineare Systeme arbeiten mit Digitalsignalen, die meist auf Magnetplatten gespeichert werden. Das Problem war daher früher die begrenzte Speicherkapazität, sodass oft mit datenreduzierten Signalen gearbeitet wurde.

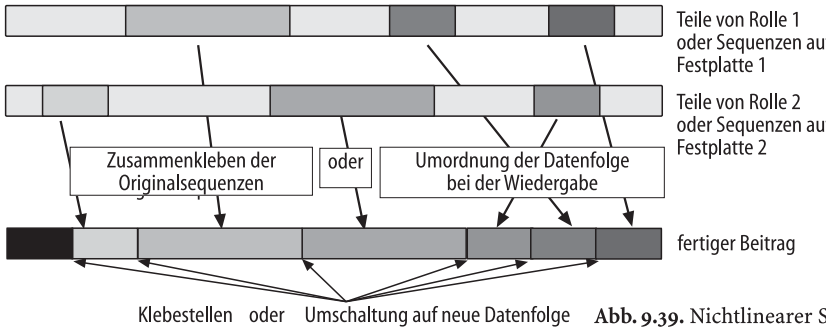


Abb. 9.39. Nichtlinearer Schnitt

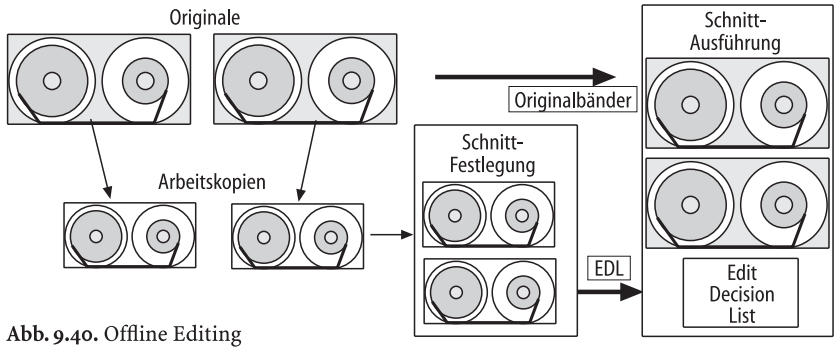


Abb. 9.40. Offline Editing

Beim Einsatz von Datenreduktionsverfahren ist es möglich, sogar mit stark qualitätsmindernder Reduktion zuarbeiten, nämlich dann, wenn beim Editiervorgang kein Endprodukt hoher Qualität (Online) entstehen, sondern zunächst nur die Schnittfolge bestimmt werden soll. Es wird hier eine Trennung zwischen Schnittbestimmung und Schnittausführung vorgenommen, die Schnittbestimmung geschieht offline (Abb. 9.40). Als Ergebnis des Offline-Schnitts entsteht eine Schnittliste (Edit Decision List, EDL), in der die gewünschten Sequenzen in der richtigen Reihenfolge aufgeführt sind. Diese Schnittliste wird dann mit dem nicht datenreduzierten Originalmaterial online abgearbeitet (Abb. 9.42). Einfache Formen des Offline-Schnitts sind auch ohne digitale Systeme möglich, indem z. B. eine einfache Kopie des Masterbandes erstellt wird, anhand derer unabhängig von teuren Schnittplätzen die Schnittfestlegung vorgenommen werden kann. Die gewünschten Sequenzen werden von Hand notiert oder in spezielle Computerprogramme eingegeben. Voraussetzung ist natürlich, dass eine Identifikationsmöglichkeit der Bandstellen gegeben ist, anhand derer die Sequenzen auch im Originalmaterial wiedergefunden werden können. Eine verbreitete Methode ist diesbezüglich, den Timecode des Originalbandes bei der Überspielung in das Bild einzublenden (Abb. 9.41).

Gegenüber diesem einfachen Vorgang bieten nonlineare Editingsysteme, bei denen die Bildfolgen in stark komprimierter Form mit sogenannter Offline-Qualität gespeichert werden, den wesentlichen Vorteil, dass lange Aufzeichnungsdauern zur Verfügung stehen. Generell bieten sie den Vorteil, dass die editierten Bildfolgen direkt abgespielt werden können, wodurch sofort ein Ein-

Abb. 9.41. Arbeitskopie mit Timecode im Bild

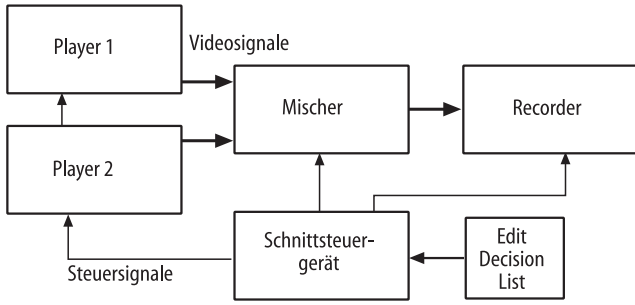


Abb. 9.42. Online Editing

druck des Endproduktes entsteht. Die Schnitte können hier bereits bildgenau festgelegt werden und das Ausprobieren verschiedener Versionen ist problemlos ohne Verbrauch von Speicherplatz möglich. Der Computer generiert schließlich selbsttätig eine Schnittliste, die in gängige Online-Systeme übertragen wird, wo der Editiervorgang fast selbsttätig ablaufen kann, sodass der teure Online-Schnittplatz nur kurze Zeit belegt ist. Der Editiervorgang kann allerdings nur bei einfachen Schnitten völlig ohne Personal ablaufen. Bei komplexen Bildübergängen ist ein Eingriff erforderlich, da die Ausführung dieser Übergänge von der speziellen Ausstattung des Online-Platzes abhängt.

Bei nichtlinearen Online-Systemen ist die Bildqualität so gut, dass direkt von der Festplatte »gesendet« werden kann. Diese Systeme arbeiten entweder mit Datenkompressionsverhältnissen, die keine sichtbare Bildverschlechterung bewirken (bis ca. 5:1) oder mit unkomprimiertem Material (s. Abschn. 9.4).

9.3.2 Timecode

Der Editiervorgang erfordert die genaue Lokalisierung jedes einzelnen Bildes. Diese wird bei linearen Editing-Systemen durch Bild- bzw. Zeitzählung vorgenommen. Man unterscheidet Tapetime und Timecode. Tapetime ist eine relative Zählung, die sich an einer vom Band bewegten Zählrolle oder an den Steuerspursignalen (CTL) orientiert. Der Nachteil ist hier die Zählung bezüglich eines willkürlichen Anfangspunktes, der nach Verstellung des Zählwerks oder bei Wiederverwendung des Bandes nicht exakt wiedergefunden werden kann. Daher wird die relative Tapetime-Zählung möglichst wenig eingesetzt und die absolute Zählung mittels Timecode bevorzugt.

Timecode (TC) bezeichnet eine eindeutige Nummerierung von Bildfolgen durch Zeitwerte, wobei die Nummerierungsdaten unabhängig vom Videosignal in separaten Spuren eines Bandes oder Dateien gespeichert werden. Die Nummern ergeben sich aus einer aufsteigenden Zählung anhand der Video-Synchronsignale, der Timecode ist damit fest mit dem Videosignal verkoppelt [158]. Die Bildnummern sind digital verschlüsselt, die Daten stellen acht Ziffern als Zeitwerte in Stunden, Minuten, Sekunden und Bildern (Frames) dar. In Europa werden die Vollbilder von 0 bis 24, die Sekunden und Minuten von 0 bis 59 und die Stunden von 0 bis 23 gezählt (Abb. 9.43).

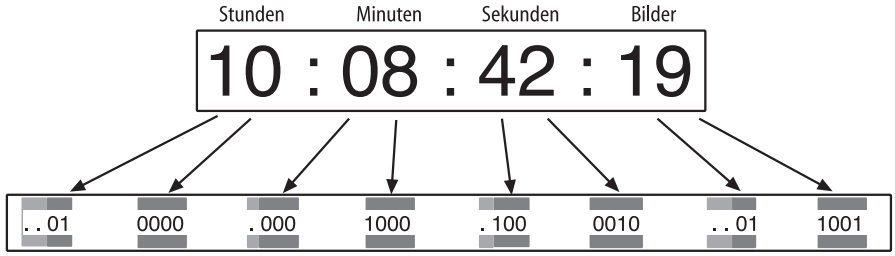


Abb. 9.43. Timecode-Anzeige und Binärcode

Die Nummer 00:00:00:01 ist dabei dem ersten Halbbild der PAL-8V-Sequenz zugeordnet, die nächste Sequenz beginnt bei 00:00:00:05 (Abb. 9.44). In den USA ist die Zählung gleich, nur die Bildzahlen laufen aufgrund der anderen Bildfrequenz von 0 bis 29. Hier ergibt sich das Problem, dass das NTSC-Farbfernsehsystem nicht mit exakt 30 sondern mit 29,97 frames per second (fps) arbeitet. Ein Timecode-Generator zählt nur ganzzahlig und läuft daher in einer Stunde um 3,6 Sekunden zu schnell. Zum Ausgleich kann mit dem Drop Frame-Modus in einem bestimmten Rhythmus ein Bild bei der Zählung ausgelassen werden. Man hat sich darauf geeinigt, pro Minute zwei Bilder auszulassen, außer bei jeder zehnten Minute. Die TC-Nummern werden beim Minutenwechsel weggelassen.

Für die insgesamt acht Timecode-Ziffern sind jeweils maximal 4 Bit vergeben. Für Zusatzinformationen steht ein gleich großer Datenbereich mit acht weiteren 4-Bit-Blöcken (User Bits) zur Verfügung. Die User Bits können frei belegt werden, z. B. zur Aufzeichnung einer kompletten zweiten Zeitinformation oder für ASCII-codierte alphanumerische Zeichen wie Kommentare oder Bandnummern. Die Gesamtinformation besteht aus $16 \cdot 4$ Bit = 64 Bit, die aber nicht alle zur Codierung der TC-Ziffern gebraucht werden. So läuft z. B. die Zehnerstelle der Bildnummer nur bis höchstens Zwei, eine Zahl, die mit nur zwei Bit binär codiert zu werden braucht (Abb. 9.45). Die freien Bits werden zur Kennzeichnung des Drop Frame-Modus und der Berücksichtigung der PAL-8V-Sequenz verwendet (Colour Frame Flag). Je nach TC-Typ werden die 64 Informationsbits um zusätzliche Synchronbits erweitert. Im professionellen Videobereich sind die TC-Arten LTC und VITC gebräuchlich. Beim Longitudinal-TC (LTC) werden die 64 Nutzbits um 16 Synchronbits auf 80 Bits erweitert, beim VITC steigt die Zahl auf insgesamt 90 Bits (Abb. 9.45).

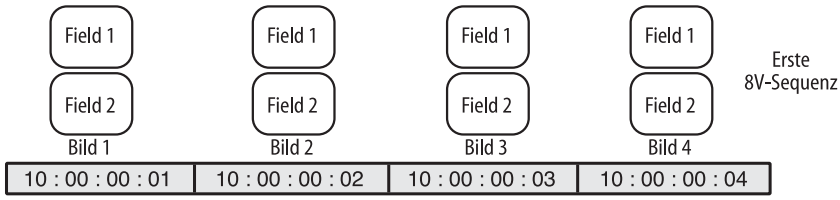


Abb. 9.44. Bildnummerierung und Zuordnung zur PAL-8V-Sequenz

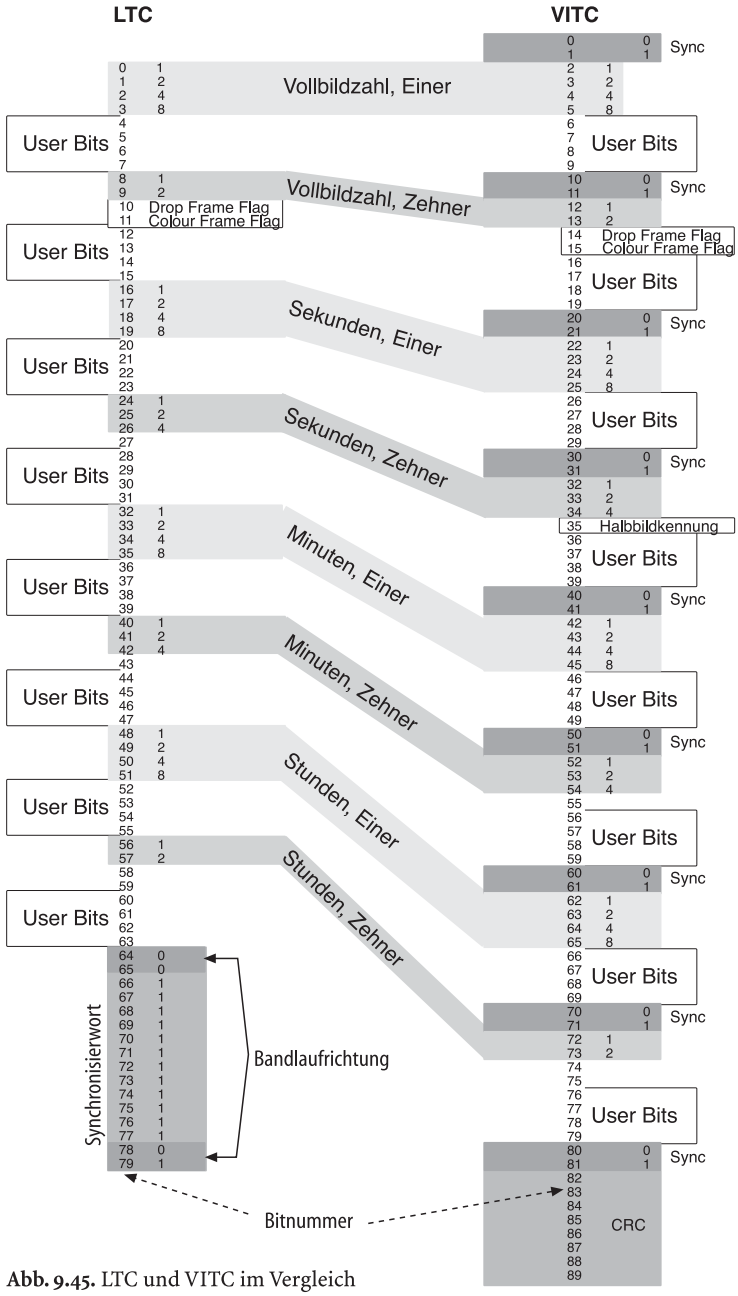


Abb. 9.45. LTC und VITC im Vergleich

9.3.2.1 Longitudinal Timecode

LTC war die erste Timecode-Form und wurde 1967 von der Society of Motion Pictures and Television Engineers (SMPTE) für die NTSC-Norm definiert. Von

der European Broadcast Union (EBU) wurde das Timecodesystem 1972 für PAL und SECAM übernommen und für 25 Bilder/s modifiziert. Das aufgezeichnete LTC-Signal umfasst 80 Bit für jedes Bild. Am Ende der 16 Datenblöcke folgen 16 Synchronbits, die im Kern aus zwölf aufeinander folgenden logischen Einsen bestehen, die an keiner weiteren Stelle im Datenstrom auftauchen können und damit das Ende des Datenwortes signalisieren. Die zwölf Bits sind von den Kombinationen 00 und 01 umgeben, anhand derer die Bandlaufrichtung erkannt werden kann (Abb. 9.45).

Die Daten liegen als serieller Bitstrom vor und werden mittels Bi Phase Mark Coding moduliert, damit aus der Datenfolge auch die Taktrate ableitbar ist (Abb. 9.46). Bei 80 Bit in 40 ms (Bilddauer in Europa) ergibt sich eine Zeiteinheit von 0,5 ms, in der ein ständiger Polaritätswechsel stattfindet. Liegt im Datenstrom eine »1« vor, so wird ein zusätzlicher Wechsel eingefügt, bei einer »0« wird er weggelassen. Die höchste mögliche Frequenz beträgt damit 2 kHz, die geringste 1 kHz. Der LTC belegt so den Frequenzbereich eines Audiosignals und lässt sich auf Magnetbandspuren aufzeichnen, die für Audiosignale vorgesehen sind. Wegen der geringen Maximalfrequenz bei Normallauf kann das Signal auch dann noch wiedergegeben werden, wenn das Band im schnellen Suchlauf abgespielt wird. Um Übersprechstörungen zu vermeiden, sollte der Aufzeichnungspegel um 3 dB bis 10 dB unter dem Vollaussteuerungspegel liegen. Bei der Aufzeichnung auf Audiospuren sollten eventuell vorhandene Rauschunterdrückungssysteme (z. B. Dolby) abgeschaltet werden. In professionellen Geräten steht für die LTC-Aufzeichnung eine separate Längsspur zur Verfügung. Die Verkopplung des LTC mit dem Videobild ist festgelegt. Der erste Nulldurchgang des TC-Wortes soll innerhalb der fünf Haupttrabanten des ersten Halbbildes stattfinden (Abb. 9.47). Die vollständige Information liegt erst nach Ablauf des Bildes vor, zum Ausgleich wird die ausgelesene Bildzahl automatisch um eins erhöht.

LTC ist die gebräuchlichste Timecode-Quelle im professionellen Videoschnitt und universell im Video- und im Audiobereich einsetzbar. Das Signal wird meist auf einer eigenen TC-Spur aufgezeichnet, die unabhängig vom Videosignal editierbar ist, die rechteckigen Flanken des Digitalsignals werden bei Aufnahme- und Wiedergabevorgang allerdings verschliffen. Nach mehreren Kopiervorgängen kann es vorkommen, dass das Signal nicht mehr ausgewertet

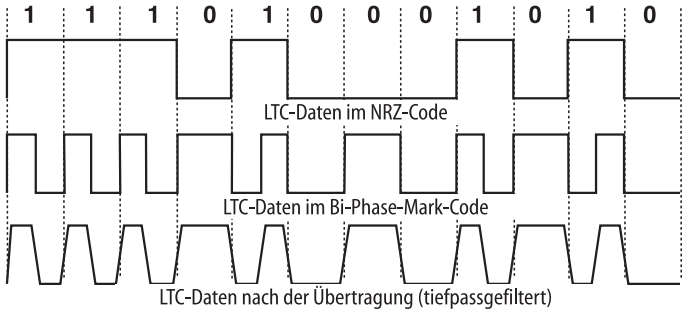


Abb. 9.46. LTC mit Bi-Phase-Mark-Code

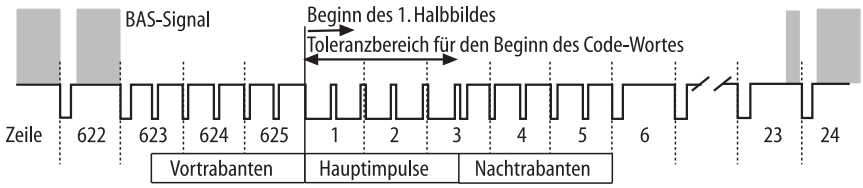


Abb. 9.47. LTC in Relation zum Bildbeginn

werden kann. Hieraus folgt, dass das LTC-Signal vor einer Kopie immer regeneriert, d. h. in seine ursprüngliche Form zurück gewandelt werden sollte. Der wesentliche Nachteil des LTC ist, dass er bei stehendem oder sehr langsam laufendem Band nicht ausgewertet werden kann.

9.3.2.2 VITC

Durch die Aufzeichnung des Vertical Interval Timecode (VITC) in der vertikalen Austastlücke des Videosignals wird es möglich, den TC auch bei Bandstillstand zu benutzen, da die Daten hier von den rotierenden Videoköpfen gelesen werden. Das VITC-Signal wird, ähnlich wie Videotext, digital aufgezeichnet. Daten mit 90 Bit pro Bild werden in eine Zeile (oder zwei) eines jeden Halbbildes eingetastet, dafür können bei analogen SD-Signalen die Zeilen 9 ... 22 des ersten und 322 ... 335 des zweiten Halbbildes gewählt werden.

Das Digitalsignal wird direkt in NRZ-Form aufgezeichnet (Abb. 9.48). Es enthält im Gegensatz zum LTC eine Halbbildkennung. Ein Nachteil des VITC ist, dass das TC-Signal nicht unabhängig vom Videosignal editiert und damit ggf. repariert werden kann. Aus Sicherheitsgründen wird der VITC daher doppelt, meist in den Zeilen 19/332 und 21/334, aufgezeichnet, außerdem werden 10 Bits als Fehlerkorrekturbits für den Cyclic Redundancy Check (CRC) eingesetzt, deren Verteilung in Abbildung 9.45 sichtbar wird. Das VITC-Signal umfasst daher 90 Bit bei gleicher TC- und Userbitkapazität wie beim LTC. Abbildung 9.49 zeigt einen LTC/VITC-Generator.

Beim Einsatz des VITC als D-VITC in Digitalsystemen werden die 90 Bits nicht einfach digital in den Datenstrom eingefügt, sondern der Verlauf des analogen VITC-Signals wird abgetastet. Die »1« des VITC-Signals wird als 192 (562 mV) und die »0« als 16 (0 mV) dargestellt. Jedem Bit sind dann 7,5 Luminanzabtastwerte zugeordnet, die Taktfrequenz des VITC beträgt dafür 1,8 MHz.

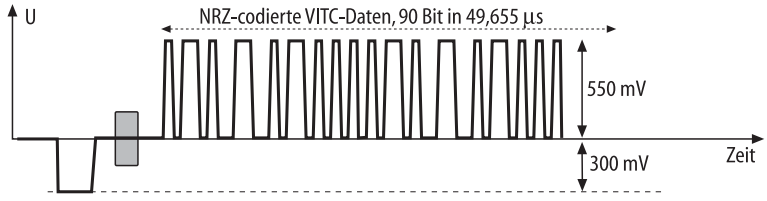


Abb. 9.48. VITC-Datenzeile

Im Prinzip reicht es aus, entweder mit LTC oder VITC zu arbeiten. Beide Arten haben aber ihre Vor- und Nachteile, sodass sie in der Praxis meist gleichzeitig benutzt werden, wobei LTC und VITC i. d. R. die gleichen Daten enthalten. Standard-Timecode ist der LTC, der auch im schnellen Suchlauf benutzt werden kann. Der Nachteil ist allerdings, dass die Daten bei sehr geringer Bandgeschwindigkeit oder bei Bandstillstand nicht gelesen werden können. Da ist es gut, wenn sich der TC-Reader auf die absoluten VITC-Daten umschalten kann. VITC kann auch bei Bandstillstand ausgewertet werden, eignet sich aber wiederum nicht für schnellen Suchlauf, da die Videospuren dann oft nicht störungsfrei abgetastet werden.

9.3.2.3 RCTC und GSE Rapid Timecode

Für die im Amateur-Camcorderbereich früher wichtigen Videoaufzeichnungsformate Video 8 und Hi 8 wurde ein eigener Timecode entwickelt, der nicht mit LTC oder VITC kompatibel ist. RCTC steht für Rewritable Consumer Time Code. Er ist unabhängig vom Bild editierbar bzw. nachträglich aufzeichenbar, da die Daten nicht in das Videosignal integriert sind, sondern in einem separaten Bereich der Schrägspur gespeichert werden, der zwischen dem Video- und dem PCM-Tonbereich liegt (s. Kap. 8.6). In professionellen Hi 8-Geräten werden die RCTC-Daten in das SMPTE/EBU-Timecode-Format umgerechnet.

Als zweites TC-Format aus dem Amateurbereich ist der Rapid-Timecode zu nennen, der als Standard einer TC-Variante für die Formate VHS und S-VHS festgeschrieben wurde. Da im VHS-Format kein separater Timecodebereich vorgesehen ist, behilft man sich mit der CTL-Spur. Die dort aufgezeichneten CTL-Impulse werden so moduliert, dass ein Zeitcode übertragen werden kann. Der CTL-Zeitcode umfasst 50 Bit. Pro Bild kann nur ein Bit übertragen werden, sodass erst nach Ablauf von jeweils zwei Sekunden ein vollständiger Timecode-Wert vorliegt. Die Lokalisation innerhalb dieses Zeitraums muss anhand relativer CTL-Impulszählung vorgenommen werden. GSE-Rapid ist die ungenaueste TC-Variante, jedoch die einzige Form, die es ermöglicht, VHS-Bänder ohne Verlust einer Audiospur mit Timecode zu versehen.

9.3.2.4 Timecode in der Praxis

Professionelle Videorecorder mit Timecode-Ausstattung beinhalten meist einen eigenen videoverkoppelten Timecode-Generator und -Reader. Dieser wird im Regelfall eingesetzt, es steht aber meist auch ein Anschluss für TC-Daten von einem externen Generator (Abb. 9.49) zur Verfügung, damit auf mehreren Maschinen ein identischer Timecode (z. B. Realzeit) aufgezeichnet werden kann.

Abb. 9.49. Timecode-Generator und -Reader

Der Timecode sollte ohne Sprünge gleichmäßig aufsteigen. Diese Forderung wird bei der Nachbearbeitung erfüllt, indem vorcodierte Bänder mit unterbrechungsfreiem Timecode eingesetzt werden. Bei der Aufnahme des Originalmaterials kommt das TC-Signal vom TC-Generator. Von dessen Betriebsart (Free Run oder Record Run) hängt es ab, ob die Forderung erfüllbar ist oder nicht.

Bezüglich der aufgezeichneten Zeitdaten unterscheidet man zunächst Realzeit, bei der die aktuelle Uhrzeit aufgezeichnet wird, Quasi-Realzeit mit durchlaufender Zählung, aber willkürlichem Anfangspunkt, und schließlich die Ablaufzeit (Elapsed Time), bei der die Uhr bei Bedarf angehalten werden kann. Entsprechend dieser Unterscheidung ist der TC-Generator auf verschiedene Modi umschaltbar. Zunächst wird die Startzeit gesetzt. Für die Aufzeichnung von Real- oder Quasi-Realzeit läuft der Generator dann im Free Run Modus. Realzeit- und Quasi-Realzeitaufzeichnung werden in der Praxis eher selten, meist nur zu Dokumentationszwecken und bei der Aufzeichnung eines Ereignisses auf mehreren Recordern eingesetzt. Der Nachteil ist, dass bei Aufnahmeunterbrechung ein TC-Sprung entsteht, der bei der Nachbearbeitung zu Problemen führen kann. In den meisten Fällen wird der Generator auf Record Run eingestellt, sodass er nur während der Aufnahme läuft und bei Unterbrechungen angehalten wird. In diesem Modus wird automatisch dafür gesorgt, dass die TC-Zählung bei der nächsten Aufnahme störungsfrei fortgesetzt wird, sodass immer ein lückenlos aufsteigender Timecode vorliegt.

Abbildung 9.50 zeigt die Bedienelemente des TC-Generators mit den Schaltern für Free und Rec Run an einer Betacam SP-MAZ. Hier kann der TC-Generator auch zwischen den Betriebsarten Preset und Regenerate umgeschaltet werden. Bei Preset startet der Generator mit einem vorher gesetzten Zeitwert, Regenerate steht dagegen für die Wiederherstellung des TC, bei der während des Schnitt-Vorlaufs (s. u.) der alte Timecode gelesen und der neue am Schnittpunkt lückenlos fortgeführt wird. Weiterhin kann an dem Bedienfeld auch zwischen LTC- und VITC-Auswertung gewählt werden. Im Modus »Auto« nimmt die Maschine je nach Gegebenheit automatisch eine Umschaltung vor. Ein weiterer Schalter (DF/NDF) dient zur Wahl der Betriebsarten Drop Frame und Non Drop Frame.

Für die Übertragung des Timecodes zu anderen Recordern oder einem Schnittsteuergerät können separate Leitungen mit BNC- oder XLR-Buchsen eingesetzt werden. Beim Anschluss an ein Schnittsteuergerät kann auf diese Verbindung verzichtet werden, da die TC-Daten in den seriellen Datenstrom zur Gerätefernsteuerung integriert werden (RS-422, s. Abschn. 9.3.5).

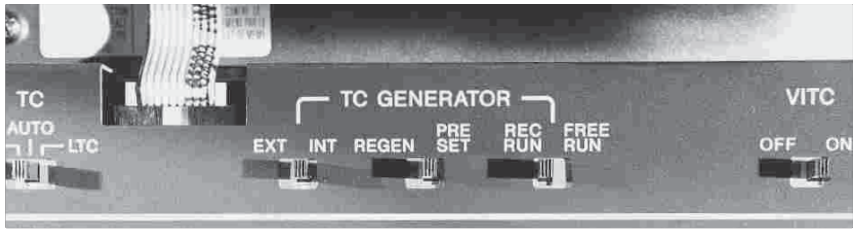


Abb. 9.50. Timecode-Einsteller an einem Betacam SP-Recorder

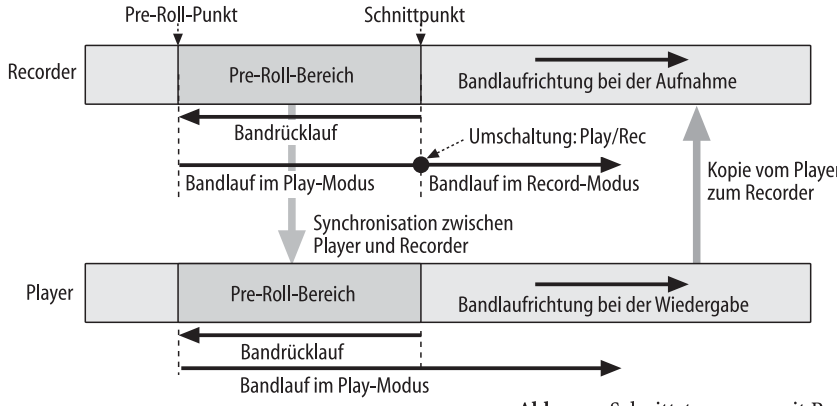


Abb. 9.51. Schnittsteuerung mit Pre-Roll

9.3.3 Linear Editing

Beim »linearen Schnitt« werden verschiedene Bildsequenzen auf einem Magnetband durch einen Kopiervorgang aneinander gereiht, dabei soll sich auch am Schnittpunkt eine unterbrechungsfreie Bildfolge ergeben. Es werden eine oder mehrere Zuspieler als Bildquellen eingesetzt. Das Erreichen einer störungsfreien Bildfolge ist aber mit großem Aufwand verbunden, denn ein Recorder wird aus dem Stand heraus gestartet. Der Vorgang ist zeitaufwendig, denn es ist eine gewisse Hochlaufzeit erforderlich, bis die Mechanik stabil läuft. Außerdem muss der Bandlauf auf das Video- oder CTL-Signal synchronisiert werden. Ein störungsfreier Schnitt erfordert also bei lineraren Systemen in jedem Fall eine Anlaufzeit der Maschinen (Pre Roll) in der Größenordnung von einigen Sekunden. Die Abbildungen 9.51 und 9.52 zeigen den Schnitt-Ablauf unter Einbeziehung eines Players, der auch eine Pre-Roll-Zeit braucht, und die Synchronisierung der beiden Maschinen während der Pre-Roll-Zeit [33].

Der störungsfreie lineare Schnitt wird in Assemble- und Insertschnitt unterschieden. Eine Aufnahme, bei der dagegen keine Rücksicht auf bereits auf dem Band befindliche Signale genommen wird, wird mit Crash Record bezeichnet. Diese Betriebsart ist im professionellen Bereich nur dazu zu verwenden, ein völlig unbespieltes Band mit einem kurzen Signal zu versehen, auf das sich die Maschine im Assemblemodus synchronisieren kann (s. u.). Crash Record wird bei vielen Heimrecordern angewandt. Die Auswirkungen zeigen sich daran, dass am Anfang und Ende der aufgenommenen Sequenz (oder des Filmbeitrags) immer eine Bildstörung oder ein Rauschen auftritt (Abb. 9.53).

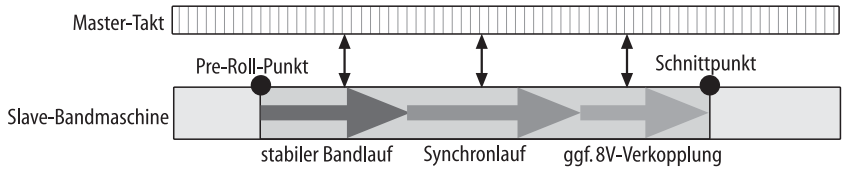


Abb. 9.52. Signalverkopplung während der Pre-Roll-Phase

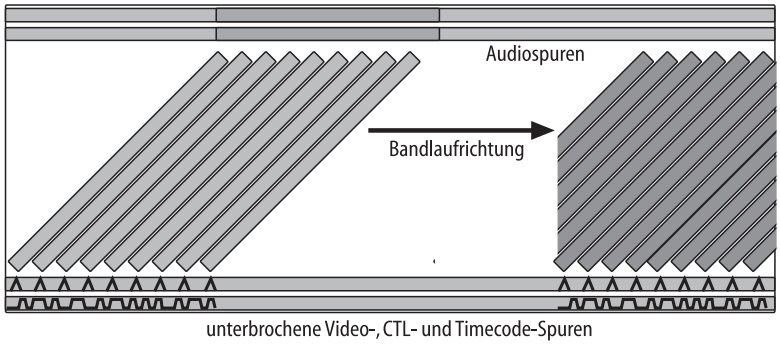


Abb. 9.53. Crash-Record

9.3.3.1 Assemble-Schnitt

Diese Schnitt-Betriebsart ermöglicht ein störungsfreies Ansetzen einer Bild/Ton-Sequenz bei Fortführung eines lückenlos aufsteigenden Timecodes. Eine Sequenz kann störungsfrei an die vorhergehende angesetzt, aber nicht in eine bestehende eingesetzt werden, denn nur der Schnitteinstieg, nicht aber der -ausstieg geschieht bildgenau. Diese Betriebsart eignet sich vor allem für den Produktionsbereich und wird z. B. im Camcorder realisiert, der zum Aneinanderreihen (Assembling) von Bildsequenzen dient. Für den bildgenauen Schnitteinstieg orientiert sich der Schnittrecorder an den bereits aufgezeichneten Video- und CTL-Signalen. Dazu wird jede Sequenz etwas zu lange aufgezeichnet. Die Aufnahme wird durch die Pausenfunktion unterbrochen und das Band wird automatisch um ca. 20 Bilder zurück gespult, wobei die Timecodedifferenz oder die CTL-Pulse registriert werden. Beim nächsten Wechsel vom Pausen- in den Record-Modus steht nun eine Vorlaufzeit (Pre Roll) zur Verfügung, während derer sich der Recorder zunächst im Play-Modus auf das auf dem Band befindliche Signal synchronisiert und sich der Timecodegenerator auf den aufgezeichneten Timecode einstellt (Abb. 9.54). Damit kann nach Ablauf der registrierten Pre Roll-Zeit synchron mit dem Kopfumschaltimpuls eine bildgenaue Umschaltung von Play auf Record vorgenommen werden.

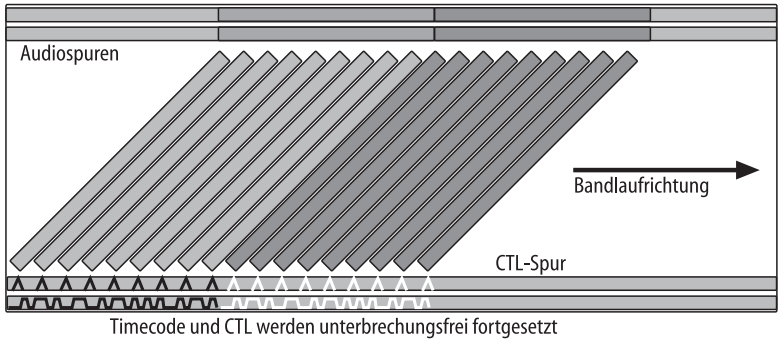


Abb. 9.54. Assemble-Schnitt

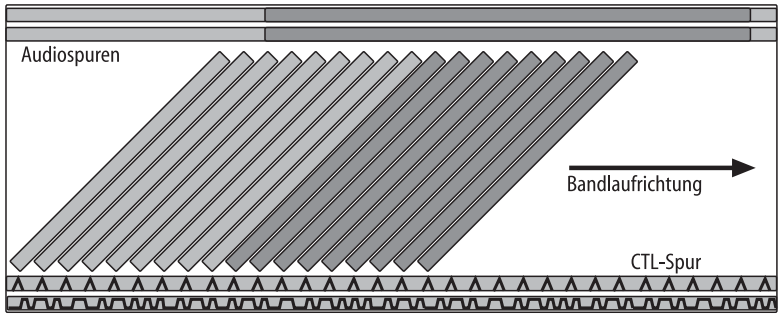


Abb. 9.55. Insert-Schnitt

Generell muss vor der Aufnahme das Band gelöscht werden. Beim Assemble-Schnitt kann dafür der stationäre Hauptlöschkopf eingesetzt werden, der alle Spuren erfasst und das Band auf voller Breite löscht. Damit wird klar, dass im Assemble-Modus Bild, Ton und Timecode nur gemeinsam, und nicht separat, editiert werden können. Eine einmal im Assemble-Modus begonnene Schnittsequenz kann nur in diesem und nicht im Insert-Modus fortgesetzt werden, da bei der Querspurlöschung auch die CTL-Spur gelöscht wird, die für den Insert-Schnitt unentbehrlich ist.

9.3.3.2 Insert-Schnitt

Insbesondere bei der Nachbearbeitung ist es erforderlich, dass in einer bestehenden Sequenz nachträglich Bilder durch andere ersetzt werden können. Eine weitere wichtige Forderung ist die separate Editierbarkeit von Video-, Audio- und Timecode-Signalen. Die Forderungen werden durch die Insert-Schnittbetriebsart erfüllt, die nicht nur einen bildgenauen Schnitteinstieg, sondern auch -ausstieg ermöglicht. Das Band darf dazu nicht mehr mit dem Hauptlöschkopf gelöscht werden. Die Videoschrägspuren müssen einzeln erfasst werden, was mit den auf der Kopftrommel montierten fliegenden Löschköpfen erreicht wird. Damit ist es möglich, in gleicher Weise wie beim Schnitteinstieg auch den Schnittausstieg durch Umschaltung vom Record- auf den Play-Modus und die Abschaltung der Löschköpfe vorzunehmen.

Der Verzicht auf die Löschung auf voller Spurbreite ermöglicht auch das selektive Löschen der Längsspuren, sodass im Insert-Modus insgesamt separate Bearbeitungsmöglichkeiten für Bild, Ton und Timecode gegeben sind. Das wichtigste ist, dass die CTL-Spur erhalten bleibt, denn mit dieser Orientierung ist immer ein kontrollierter Zugriff auf einzelne Videospuren möglich (Abb. 9.55).

Insert-Schnitt ist die Standardbetriebsart bei der Nachbearbeitung. Sie setzt allerdings voraus, dass das Band auf einer Länge, die mehr als der vorgesehenen Beitragslänge entspricht, vorcodiert, d. h. mit einer ununterbrochenen CTL-Spur und einem gleichmäßig aufsteigenden Timecode versehen ist. Als Videosignal wird dazu neben diesen Signalen ein Schwarzbild mit Farbburst aufgezeichnet (Preblacking the Tape).

9.3.4 Fernsteuersignale

Wenn bei der Produktion oder Nachbearbeitung mehrere MAZ-Geräte beteiligt sind, so müssen diese von einem Steuersystem über definierte Schnittstellen bildgenau gestartet und synchron gehalten werden können.

9.3.4.1 GPI und Parallele Fernsteuerung

Die einfachste Form der Gerätekommunikation besteht in der Übertragung eines simplen Ein/Aus-Befehls. Diese Fernsteuerung ist geräteunabhängig und kann für eine Vielzahl von einfachen Steueraufgaben eingesetzt werden. Sie wird als General Purpose Interface (GPI) bezeichnet. Der Ein- oder Auszustand wird meist über einen elektronischen Schließkontakt realisiert. Damit können einfache Aufgaben, wie z. B. der bildsynchrone Start eines Audiozuspielers oder der Aufruf eines Mischerzustands ausgeführt werden.

Wenn ein Gerät komplexer gesteuert werden soll, so sind viele spezifische Schaltsignale erforderlich, die über eine parallele Schnittstelle übertragen werden. Ein Beispiel ist ein Videorecorder, der bezüglich der Funktionen Start, Stop, Pause Rewind, etc. gesteuert werden muss. Wenn hier z. B. 30 Funktionen gefragt sind, werden auch 30 Leitungen zur Übertragung verwendet.

Die Realisierung einer Parallelschnittstelle ist sehr simpel, der große Nachteil ist aber, dass sie immer nur zu wenigen Geräten passt, außerdem sind die erforderlichen Verbindungsleitungen sehr unflexibel. Parallele Fernsteuereinrichtungen erlauben zudem meist keine Übertragung von Timecode-Daten.

9.3.4.2 Serielle Schnittstellen

Eine universellere Steuermöglichkeit wird durch serielle Schnittstellen und eine softwaremäßige Zuordnung von Daten und Funktionen erreicht. Die Daten werden in einem Protokoll definiert und seriell übertragen. Im Videobereich werden zur Fernsteuerung die gleichen seriellen Schnittstellen wie im Computerbereich verwendet. Die beiden verbreitetsten Typen sind unter den amerikanischen Bezeichnungen RS 232C und RS 422A bekannt, die europäischen Bezeichnungen lauten V.24 und V.11.

Die RS 232-Schnittstelle war bei Standard-PC weit verbreitet und diente oft zur Druckeranbindung. Die Bits werden in serieller Folge übertragen, für jedes Bit ist eine Zeiteinheit vorgesehen (Abb. 9.56). Acht Bits werden zu einer Informationseinheit zusammengefasst. Im Sender und im Empfänger muss festgelegt sein, mit welcher Datenrate die Übertragung stattfindet, übliche Werte sind 1200 Bit/s bis 19200 Bit/s. Die Daten werden mit Spannungen von ± 12 V übertragen. Die übliche Steckverbindung ist der 25-polige Sub D-Stecker. Die

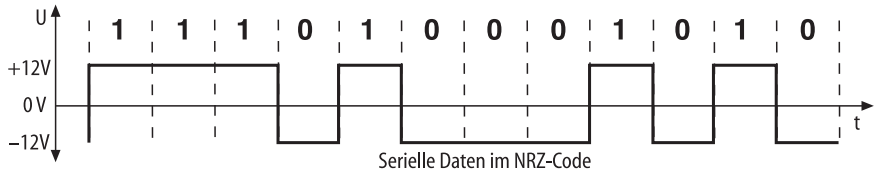


Abb. 9.56. Serielle Daten bei der RS 232-Schnittstelle

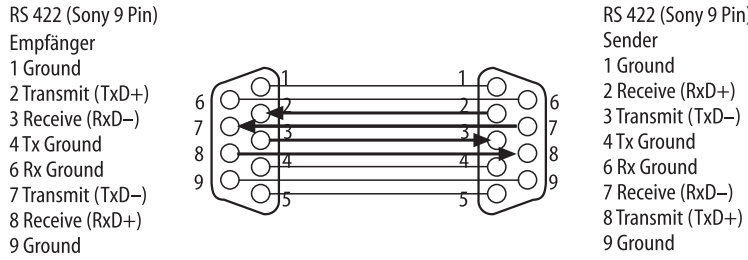


Abb. 9.57. RS 422-Verbindung zur seriellen Maschinensteuerung

RS 232-Schnittstelle erlaubt eine Datenübertragungsrate bis zu 20 Kbit/s, und eine Übertragungsdistanz von maximal 15 m.

Die RS 422-Schnittstelle arbeitet ähnlich wie die RS 232-Schnittstelle, allerdings wird durch eine symmetrische Datenübertragung eine hohe Störsicherheit erreicht, sodass auf Übertragungsstrecken von bis zu 15 m eine maximale Datenrate von 10 Mbit/s erreicht werden kann. Auf Strecken bis zu 1000 m sind bis zu 100 kbit/s möglich. Die Spannungsniveaus liegen bei ± 5 V.

Die RS 422-Schnittstelle ist die meist verwandte Schnittstelle zur Fernsteuerung professioneller Videorecorder. Im Videobereich hat sich eine Steckverbindung mit 9-poligen Sub D-Steckern als Quasi-Standard durchgesetzt: die sogenannte Sony 9 Pin-Schnittstelle, deren Steckerbelegung in Abb. 9.57 dargestellt ist. Die Datenrate beträgt 38,4 kbit/s, damit ist es möglich, neben den Fernsteuersignalen auch die Timecode-Daten zu übertragen.

Das Datenprotokoll ist geräteabhängig. Bezüglich der gängigen Befehle dominiert aber auch softwareseitig das Sony-Protokoll, sodass die Sony 9 Pin-Schnittstelle universell einsetzbar ist. Wenn z. B. eine Software auf einem Computer die entsprechenden Befehle sendet, kann ein professioneller Videorecorder auch vom Computer aus gesteuert werden. Für das 9-Pin-Protokoll ist immer ein Sender und ein Empfänger definiert. Gesendete Daten werden innerhalb 9 ms entweder einfach quittiert (ACK) oder es wird eine Information zurück gesandt. Wichtig ist die Geräteidentifikation und dann die Kommandostruktur für die eigentliche Steuerung. Tabelle 9.1 zeigt eine Übersicht für die Identifikation von Betacam SP-Recordern und Kommandobeispiele.

Tabelle 9.1. Beipieldaten des Sony-9Pin-Protokolls

Geräteidentifikation (Device Type)		Kommandos		Antwort	
BVW-65	2X 21	00 0C	Local Disable	10 01	Ack
BVW-95	2X 22	00 11	Device Type Request	12 11	Type
BVW-96	2X 23	00 1D	Local Enable	10 01	Ack
BVW-70	2X 24	20 00	Stop	10 01	Ack
BVW-75	2X 25	20 01	Play	10 01	Ack
BVW-D75	2X 46	20 02	Record	10 01	Ack
BVW-9000	2X 47	20 04	Standby Off	10 01	Ack
PVW-2600	2X 40	20 05	Standby On	10 01	Ack
PVW-2800	2X 41	20 0F	Eject	10 01	Ack

X = 0 bzw. 1 für 525- bzw. 625-Zeilensysteme

9.3.5 Schnittsteuersysteme

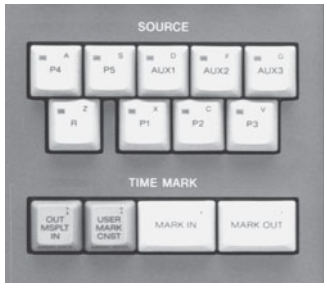
Bei der linearen Nachbearbeitung müssen ein Recorder und ein oder mehrere Zuspieler verkoppelt werden. Die zentrale Steuerung der MAZ-Maschinen wird mit Schnittsteuersystemen (engl. Editor) vorgenommen, die es oft auch ermöglichen, Ton- und Bildmischer mit in den Editiervorgang einzubeziehen.

9.3.5.1 Der Schnittsteuervorgang

Das Schnittsystem ist für die Steuerung eines Recorder (R) und mehrerer Player (P1, P2,..) vorgesehen, zur Anwahl der Quellen stehen entsprechende Tasten zu Verfügung (Abb. 9.58, links). Wenn ein Gerät angewählt ist, können seine Laufwerksfunktionen (Rec, Play, Stop etc.) vom Bedienpult gesteuert werden. Wenn Schnittmarkierungstasten gedrückt werden, beziehen sich auch diese auf die hier angewählten Geräte. Die Aux-Tasten dienen zur Anwahl von Signalen, die von einer Kamera oder vom Bildmischer kommen (Schwarz, Farbbalken).

Vor der Schnittausführung wird zunächst festgelegt, ob im Assemble- oder Insert-Modus geschnitten werden soll. Wegen des bei Insert erlaubten separaten Zugriffs auf Audio- und Videospuren stehen für Insert mehrere Tasten zur Verfügung (Abb. 9.58, links unten).

Für den Insert-Schnitt wird das Band vorcodiert und vor dem Programm wird ein technischer Vorspann aufgezeichnet. Die Art des Vorspanns ist in technischen Richtlinien der Sendeanstalten vorgegeben. Häufig werden eine Minute Farbbalken und ein 1 kHz-Pegelton bei −9 dB vorgeschrieben, die beim TC-Wert 09:59:29:24 enden, gefolgt von 30 Sekunden Schwarz und Stille, sodass das Programm mit Timecode 10:00:00:00 beginnt. Der Vorspann stellt eine Referenz dar, mit Hilfe derer Signalfehler aus nicht normgerechter Kalibrierung ausgeglichen werden können. Mit einem derart codierten Masterband kann der eigentliche Schnitt beginnen. Die Mastermaschine wird dazu



Tasten zur Quellenwahl und für IN/OUT-Punkte

Laufwerkskontrolltasten

Tasten zur Wahl der Schnittbetriebsart

Abb. 9.58. Bedientasten am Schnittsteuergerät [208]

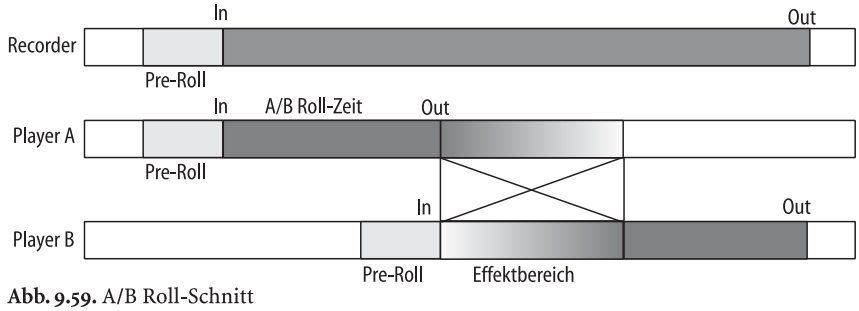


Abb. 9.59. A/B Roll-Schnitt

zum Beitragsbeginn gefahren und dieser Punkt wird als Einstieg mit der Taste IN markiert (Abb. 9.58). Die Schnittein- und Ausstiegspunkte (IN und OUT) können über eine Zahltastatur auch numerisch eingegeben werden. Die komfortable Suche eines Bildes wird durch ein Handrad ermöglicht, das im Shuttle-Modus das Gerät mit verschiedenen Suchlaufgeschwindigkeiten und im Jog-Modus mit verschieden schneller Einzelbildwiedergabe betreibt. Nachdem die gewünschten IN- und OUT-Punkte für den Player markiert sind, braucht nur noch die Taste REC (oder Edit) gedrückt zu werden, und der Schnittcomputer übernimmt den gesamten weiteren Steuerungsablauf. Wenn bei dem Editiervorgang die Schnittpunkte für Bild und Ton gegeneinander versetzt sind, spricht man von Split-Edit.

Die Verwendung von mehreren Zuspielern ermöglicht Bildüberblendungen und wird in Anlehnung an die in der Filmtechnik für ähnliche Effekte erforderlichen zwei Filmrollen mit AB Roll bezeichnet. Beim AB Roll-Schnitt (Abb. 9.59) muss das Schnittsystem mehr als zwei Geräte steuern können. Der Bildübergangseffekt wird zur Zeit des Synchronlaufs von zwei Zuspielmaschinen am Bildmischer ausgelöst. Komfortabel ist es, wenn vom Schnittsystem auch der Mischer gesteuert werden kann. Wenn auch der Tonmischer kontrolliert werden kann, ist es zudem möglich, parallel eine Tonblende auszuführen.

Größere Schnittsteuereinheiten erlauben auch die Anwahl einer variablen Wiedergabegeschwindigkeit vom Steuersystem aus. Sie bieten dabei mit einer Fit-Funktion die automatische Berechnung der erforderlichen Wiedergabegeschwindigkeit für Situationen, in denen Bildsequenzen in eine Lücke eines bestehenden Beitrags eingesetzt werden sollen.

Beim Schnitt von FBAS-Signalen ist das Colour Framing zu beachten. Der Begriff steht für die Einhaltung der PAL-8V-Sequenz. Das bedeutet, dass, wenn eine Bildfolge mit dem 6. Halbbild der 8V-Sequenz beendet wurde, dafür gesorgt werden muss, dass die nächste Bildfolge mit einem 7. Halbbild angeschnitten wird. Schnittsteuerungen und MAZ-Geräte können so eingestellt werden, dass bei der Schnittpunktfestlegung automatisch die Sequenz beachtet wird, dabei kann es jedoch vorkommen, dass der vom Anwender gewählte Schnittpunkt nicht in die 8V-Sequenz passt, und evtl. verschoben wird.

Wird die PAL-8V-Sequenz nicht beachtet, wird im TBC ein Regelvorgang ausgelöst, der den Farbträger phasenrichtig an das Referenzsignal angekoppelt. Dabei erscheint ein kleiner horizontaler Ruck, der aber nur erkannt wird, wenn dieselbe Szene nach einem Schnitt weitergeführt wird (Invisible Cut).

TITLE: TEST

EDT	REL	MODE	TYP	P	S	T	P-VTR IN	P-VTR OUT	R-VTR IN	R-VTR OUT
BLOCK 001										
001	002	V	C				00:00:00:00	00:00:01:05	10:00:00:00	10:00:01:05
002	001	VA1A2	C				00:00:00:06	00:00:07:01	10:00:01:05	10:00:09:10
	EFFECTS NAME IS CROSS DISSOLVE									
003	001	V	C				00:00:07:01	00:00:07:01	10:00:09:10	10:00:09:10
003	002	V	D		01:01		00:00:00:00	00:00:03:09	10:00:09:10	10:00:12:19
	EFFECTS NAME IS WIPE									
004	002	VA1A2	C				00:00:03:09	00:00:03:09	10:00:12:19	10:00:12:19
004	001	VA1A2	W000		00:22		00:00:00:22	00:00:05:09	10:00:12:19	10:00:18:06

Abb. 9.60. Beispiel einer Schnittliste (Edit Decision List, EDL)

9.3.5.2 Schnittdatenanzeige und -speicherung

Kleine Schnittsteuersysteme haben meist nur ein Display zur Anzeige der Timecode-Daten, größere Systeme sind mit einem Datenmonitor ausgestattet. Abbildung 9.61 zeigt ein typisches Display eines solchen Schnittsystems (Sony BVE 910). Oben links wird die Schnittnummer, die beteiligten Audio- und Videospuren sowie Bildmischer-Effektnummern angegeben. Oben rechts und in der Mitte befinden sich eine grafische Anzeige des Schnittvorgangs und eine Angabe über die Sequenzdauer und die Dauer des angewählten Effekts. Der umrandete Bereich darunter beinhaltet in Form einer Schnittseite alle relevanten Informationen über die Programmquellen und die Zeitwerte eines Schnitts. Jeder Quelle ist eine Zeile zugeordnet, ein Stern neben der Zeile kennzeichnet die angewählte Quelle.

Die nächste Spalte gibt den Timecode für die Position an, an der sich die Maschine gerade befindet, und daneben kann der Status der Maschine abgelesen werden (Play, Stop, Stand by off (SBOF)). Unter den Spaltenbezeichnungen In und Out werden die Timecode-Daten der Ein-und Ausstiegspunkte angegeben, und bei Duration wird die Länge des Schnitts angezeigt. Die Daten der Schnittseite werden nach der Schnittausführung automatisch in einer Schnittliste gespeichert ((Edit Decision List, EDL, Abb. 9.60).

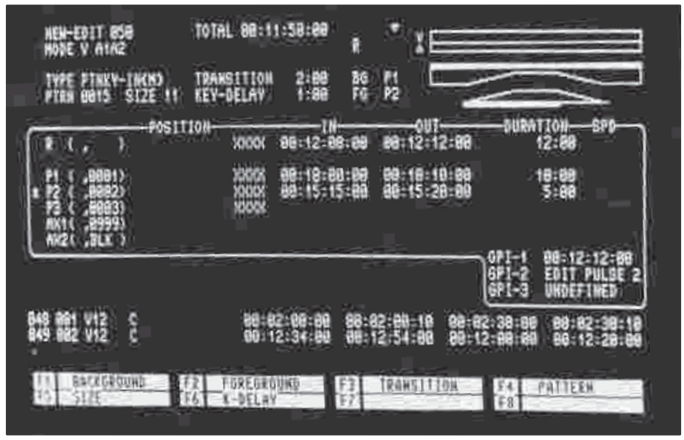


Abb. 9.61. Display eines Schnittsystems [32]

9.4 Nonlineare Editingsysteme

Cutterinnen und Cutter, die vom Filmschnitt her kommen, fühlen sich in ihrer Tätigkeit durch lineare Schnittsysteme oft stark eingeschränkt, da sie die Freiheit aufgeben müssen, bereits bearbeitete Sequenzen zu jeder Zeit verändern zu können. Diese beim Filmschnitt selbstverständliche Bearbeitungsmöglichkeit ist beim linearen Schnitt nicht vorhanden. Wenn hier in der Mitte einer bearbeiteten Sequenz eine Szene gelöscht werden soll, ist der lückenlose Fortlauf der Bildfolge nur mit einem Kopiervorgang erreichbar, der folgenden Szenen wieder neu ansetzt. Schon Mitte der 80er Jahre wurde daher versucht, die Freiheiten der nichtlinearen Bearbeitung auch mit Videosystemen zu erreichen. Dazu wurden damals eine Vielzahl von VHS- bzw. Betamax-Videorecordern über ein Computersystem gesteuert, wobei alle Maschinen mit identischem Zuspielmaterial ausgestattet waren, sodass ein schneller Zugriff gewährleistet war. Die gewünschten Szenen konnten markiert, in einer Edit Decision List (EDL) vermerkt und dem Onlinesystem übergeben werden. Die nächste Generation verwendete Laser Discs als Zuspielmedium, die Benutzeroberflächen verbesserten sich, das Signal war aber noch analoger Art. Erst die Einführung digitaler Speichersysteme mit ihrer hohen Zugriffsgeschwindigkeit brachte dann den Durchbruch der nichtlinearen Bearbeitung im Videobereich [104].

Nichtlineare Editingsysteme (NLE) werden heute auf Basis von Standard-PC mit entsprechender Software, wie Avid Media Composer, Adobe Premiere oder Apple Final Cut, realisiert. Die Videodaten liegen meist in komprimierter Form auf einer internen, externen oder einer Festplatte im Netzwerkverbund vor.

9.4.1 Grundlagen

Das große Problem beim linearen Videoschnitt ist, dass er auf einem Kopiervorgang beruht und eine festgelegte Reihenfolge der gewählten Bildsequenzen auf einem Masterband erfordert. Beim Filmschnitt wird dagegen nicht kopiert, sondern das Originalmaterial wird nur anders zusammengestellt. Das Gleiche gilt für nichtlineare Editingsysteme. Die Schnittfolge ergibt sich dadurch, dass das System den Lesekopf an jedem Schnittpunkt in sehr kurzer Zeit anders positioniert, womit die Wiedergabe an einer anderen Stelle fortgesetzt wird (Random Access). Da in Computersystemen die Daten immer mit Adressen versehen sind, ist der Steuerungsprozess sehr einfach. Bei den nichtlinearen Systemen wird einfach die gewünschte Adresse aufgesucht, zusätzlich ist nur ein Kurzzeit-RAM-Speicher erforderlich, der zu lange Zugriffszeiten auf neue Sequenzen überbrückt und damit eine konstante Bildrate von 25 oder 30 Frames per second gewährleistet. Das Ausprobieren verschiedener Versionen ist mühelos möglich, die Varianten belegen dabei keinen Speicherplatz für Videodaten, sondern nur für die Schnittinformationen mit ihrem sehr geringen Umfang. Da der Datendurchsatz herkömmlicher Computersysteme früher sehr beschränkt war, arbeiteten NLE anfangs mit hoher Datenreduktion und waren wegen der schlechten Bildqualität nur für den komfortablen Offline-Schnitt einsetzbar. Im Zuge der rasanten Entwicklung der Computertechnik konnten

immer mehr Funktionen in die Systeme integriert und die Datenreduktionsfaktoren erheblich gesenkt werden. Moderne Systeme sind sowohl für den Offline- als auch für den Online-Schnitt einsetzbar und größere Systeme arbeiten auch mit unkomprimierten Signalen. NLE-Systeme beinhalten meist Schriftgeneratoren, Audiobearbeitung und auch DVE, die eigentliche Schnittfestlegung ist dabei technisch der unaufwendigste Teil.

Gute Systeme arbeiten mit einer Vielzahl von Datenreduktionscodecs bei nativer Verarbeitung, d. h. dass sie ohne interne Signalkonvertierung, die ggf. die Bildqualität, beeinflussen, die codierten Signale ein- und ausgeben können. Fremde Codecs werden auch mittels der Umsetzung auf hochwertige Intermediate-Codecs verarbeitet. Weit verbreitet sind hier Avid DNxHD und Apple ProRes (3. Abschn. 3.10).

Zum Datenaustausch sollten NLE dabei offene Formate unterstützen. Ein Austausch-Standard ist das Open Media Framework (OMF) der Firma Avid, oder auch MXF und AAF, weitere Beispiele sind Quicktime (Apple) und Video for Windows/AVI (Microsoft) (s. Abschn. 10.3). NLE erlauben oft auch den Im- und Export von Grafiken in vielerlei Form (Pict, Tiff, BMP). Als Audioaustauschformate werden AIFF und WAV verwendet.

9.4.1.1 Hardware

Der Hauptanspruch an die Hardware besteht darin, die Videodatenströme schnell genug bereitzustellen. Weitere Ansprüche sind erst mit der Entwicklung schnellerer PC hinzugekommen, denn nichtlineare Editingsysteme waren früher nicht zur Bildbearbeitung konzipiert. Die dominante Plattform war zu Beginn das Macintosh-System, viele sind Editingsysteme aber heute auch für Windows-PC verfügbar sind. Eine Besonderheit stellen die Editingsysteme der Firma Quantel dar. Diese haben ihre Wurzeln in Videografik-Systemen und sind auf höchste Qualität abgestimmt, können also meist auch ohne Datenreduktion arbeiten. Quantelsysteme arbeiten nicht mit Standardcomputern, es sind geschlossene Systeme, die den Vorteil bieten, dass Soft- und Hardware optimal harmonieren. Sie verwenden zur Bedienung ein Grafiktablett, wobei die Funktionen mit einem Pen durch Zeigen auf Schaltflächen ausgelöst werden (Abb. 9.62). Als weitere Bedienungshilfen stehen auch von außen bedienbare Fadercontrol-Geräte für die Audiopegeleinstellung zur Verfügung.

Abb. 9.62. Externe Bediengeräte am nichtlinearen Schnittplatz

Für die Signaldarstellung sollte neben dem Computermonitor ein Videomonitor vorhanden sein, auf dem das Schnittbild ständig kontrolliert werden kann, da sich die Bildwirkungen aufgrund verschiedener Gamma-Einstellungen, Farbtemperaturen und auch wegen der Halbbilddarstellung erheblich unterscheiden können. Empfehlenswert ist auch die Verwendung von zwei Computermonitoren, sodass einer zur Darstellung der vorhandenen Materialien und der zweite als Schnittmonitor eingesetzt werden kann.

Nonlineare Editingsysteme sind meist so gut ausgestattet, dass außer einem PC, großen Festplatten und dem Betriebsprogramm nur noch ein Recorder oder ein Server als Massenspeicher vorhanden sein muss. Die erforderlichen Geräte finden alle auf einem Schreibtisch Platz, und man spricht in Anlehnung an den Begriff Desktop Publishing auch von Desktop Video. Für den Signalaustausch müssen die PC oft mit Video- und Audioschnittstellen ausgestattet werden. Dafür gibt es herstellerspezifische Lösungen wie z. B. die Hardware Avid Mojo oder Nitris. Mit der Verfügbarkeit schneller Schnittstellen an den PC, wie USB 3.0 oder Thunderbolt, ist es möglich geworden, die Verbindung auch über preisgünstige Geräte von Drittanbieter wie z. B. Blackmagic zu realisieren, die auch 10-Bit-Signale unterstützen (Abb. 9.63).

Anstatt über SDI oder andere In- und Outputs werden die Daten zunehmend jedoch per Netzwerk ein- und ausgespielt. Entsprechend wichtig wird die Möglichkeit der Netzwerkanbindung. Wenn das System in ein Netzwerk eingebunden wird, das mit hohen Datenraten arbeitet, ist es möglich, dass mehrere Anwender auf dieselben Daten zugreifen und im Verbund parallel am selben Projekt arbeiten können. Für diese Funktion kommt z. B. Fibre Channel oder Ethernet in Frage (s. Abschn. 10.3). Auf dieser Basis bietet z. B. die Firma Avid Storage Area Networks (SAN) an, die einen zentralen Großspeicher darstellen, an die per Netzwerk viele Video- und Audioeditingplätze anschlossen werden können, die, je nach Zuteilung und Rechten, alle auf das selbe Audio- und Videomaterial zugreifen können, sodass nur wenig Speicherplatz an den Workstations selbst zur Verfügung stehen muss [230].

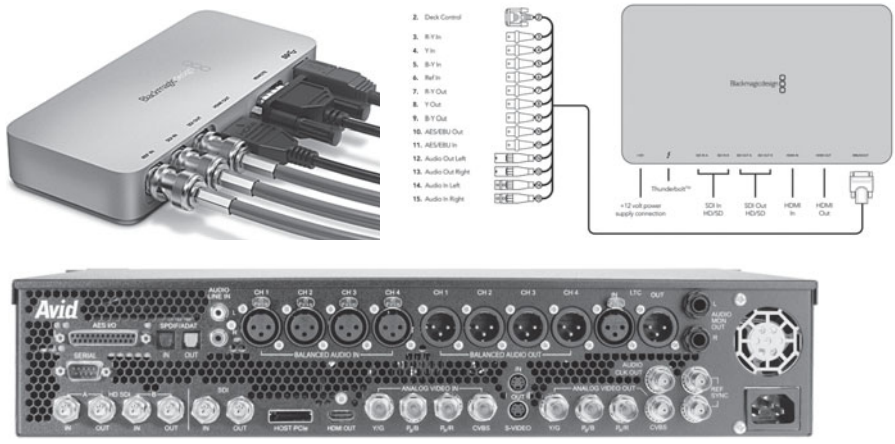


Abb. 9.63. Interfaces (Break Out Box) für nichtlineare Schnittsysteme [10]

9.4.1.2 Software

Ein nichtlineares Editingsystem zeichnet sich ein vor allem durch die Benutzeroberfläche und die einfache Bedienbarkeit umfangreicher Funktionen aus. Die meisten verbreiteten nonlinearen Schnittsysteme werden als Software allein oder als komplett konfigurierte Systeme verkauft, wobei sich die Anwender oft zwischen den Betriebssystemen MAC OS und Windows entscheiden können. Universelle und verbreitete Software für diesen Bereich sind die Programme Avid Media Composer und Adobe Premiere, die beide für MacOS und Windows erhältlich sind, sowie Final Cut Pro von Apple (Abb. 9.64).

Allen gemein ist, dass das Quellmaterial in einem Verwaltungsfenster angezeigt wird. Hier kann es mit Namen versehen und nach bestimmten Kriterien geordnet werden. Die Sequenzen werden dann in einem Player-Fenster platziert, betrachtet, grob beschnitten und anschließend in horizontaler Richtung in einer Timeline nacheinander angeordnet, wobei zur Ansicht der Schnittsequenz ein Recorder-Fenster dient. Die Benutzeroberflächen verschiedener Programme werden einander immer ähnlicher, hier ist das etablierte Konzept des Marktführers Avid dominant an das auch die genannten Begriffe und die folgenden Darstellungen angelehnt sind.

9.4.2 NLE-Betrieb

9.4.2.1 Eingabe des Materials

Nach dem Systemstart und dem Anlegen eines neuen Projekts beginnt die praktische Tätigkeit beim Einspielen der Daten in das Schnittsystem. Wenn kein Netzwerkbetrieb vorliegt werden die Daten dabei auf die Festplatte kopiert, wobei sie früher als analoge Daten zunächst den meist integrierten A/D-Wandler durchlaufen mussten, weshalb sich für diesen Vorgang der Begriff

Abb. 9.64. Typische Benutzeroberfläche eines Editing-Programms

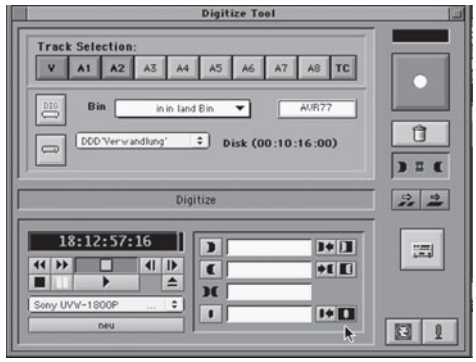


Abb. 9.65. Digitize Tool zum Einspielen des Materials

Digitalisieren erhalten hat. Auch Filmbilddaten werden über MAZ-Geräte eingespielt, sie werden mit Filmabtastern in ein Videosignal umgesetzt. Obwohl heute die meisten nichtlinearen Schnittsysteme in der Lage sind, Standardvideosignale in Online-Qualität zu verarbeiten, wird auch in diesem Bereich oft der Offline-Betrieb mit gering aufgelösten Bildern eingesetzt, um geringerem Speicherplatz auskommen zu können. Der gewünschte Modus wird bei der Einspielung über die Wahl der Bildqualität bzw. Video Resolution bestimmt. Beim Einspielen ist es meistens möglich, das Quellband einfach laufen zu lassen und dabei die gewünschten Szenen per Tastendruck »on the fly« zu übernehmen. Zum genaueren und komfortableren Aufsuchen der einzuspielenden Szenen erlauben gute Systeme die Fernsteuerung eines angeschlossenen MAZ-Gerätes über die RS 422- oder Firewire-Schnittstelle (Abb. 9.65). Dabei sollte auch der Timecode im System registriert werden. In diesem Fall kann zunächst geloggt werden, d. h. es werden nur grob die Ein- und Ausstiegspunkte der Szenen bestimmt und die eigentliche Einspielung erfolgt anschließend in dem automatisch gesteuerten Batch Digitize-Prozess. Mit dieser Funktion kann auch knapper Speicherplatz gut genutzt werden, indem zunächst in Offline-Qualität digitalisiert wird, sodass viel Material zur Auswahl zur Verfügung steht. Anhand der beim Offline-Schnitt entstehenden Schnittliste kann dann nachträglich nur das wirklich benötigte Material in hoher Qualität automatisch noch einmal digitalisiert und online verarbeitet werden.

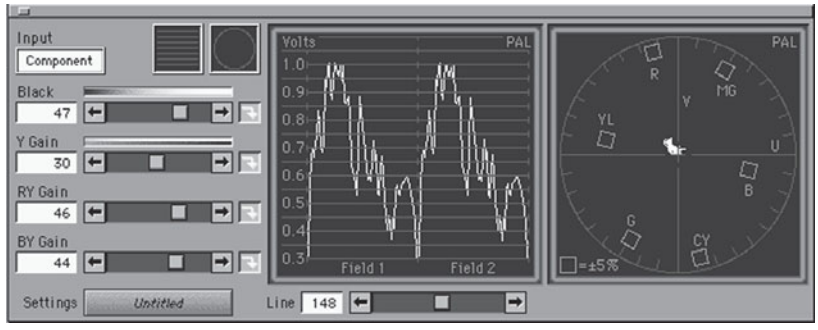


Abb. 9.66. Pegeleinstellung und -kontrolle

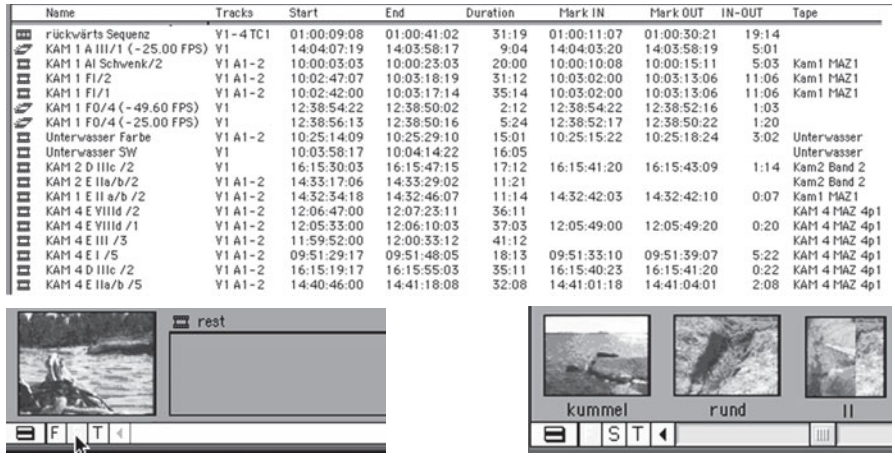


Abb. 9.67. Verschiedene Formen der Materialorganisation

Beim Einspielen analoger Daten ist es wichtig, dass das System Pegelkontrollgeräte für Video- und Audiosignale auf Softwarebasis zur Verfügung stellt und die Möglichkeit bietet, die Pegeleinstellungen vor der Digitalisierung auch verändern zu können (Abb. 9.66).

9.4.2.2 Materialorganisation

Wie beim Filmschnitt kommt es dann darauf an, den Überblick über das zur Verfügung stehende Material zu behalten. Die Qualität des NLE wird nicht zuletzt durch die Möglichkeiten der Materialorganisation bestimmt. Die Ablage und Verwaltung erfolgt in Bins oder Libraries, die thematisch oder nach dem Drehbuch organisiert sein können (Abb. 9.67). Dabei ist es möglich, eigene Verzeichnisse für die eingespielten Rushes, die daraus gebildeten Subclips und weitere Materialien wie Special Effects, Sounds, Sub Master etc. zu verwenden. Zu jedem Eintrag können Datum, Timecode, Dauer und Kommentare gespeichert werden. Wichtig ist dabei die Vergabe von Bandnummern, damit gleiche Timecodewerte von verschiedenen Bändern verarbeitet werden können.

Für den Filmschnitt gilt die Besonderheit, dass Bild und Ton getrennt aufgenommen und eingespielt werden. Hier bieten gute Schnittsysteme die Möglichkeit die als Klappen bezeichneten Bildsequenzen mit dem zugehörigen Tonmaterial komfortabel zu synchronisieren und als einen gemeinsamen Subclip im Bin abzulegen.

9.4.2.3 Editing

Nach der Organisation und Benennung der eingespielten Szenen wird der Rohschnitt erstellt. Dazu wird das Zuspielfenster geöffnet, die Szene betrachtet, mit Hilfe von In- und Outmarken beschnitten und in die Timeline übernommen (Abb 9.68). Oder die Szene wird direkt in die Timeline gezogen, wobei eine Einstellung gewählt werden kann, die bewirkt, dass sich die Szene auch bei ungenauer Positionierung mit der Maus lückenlos an die vorhergehende anfügt.

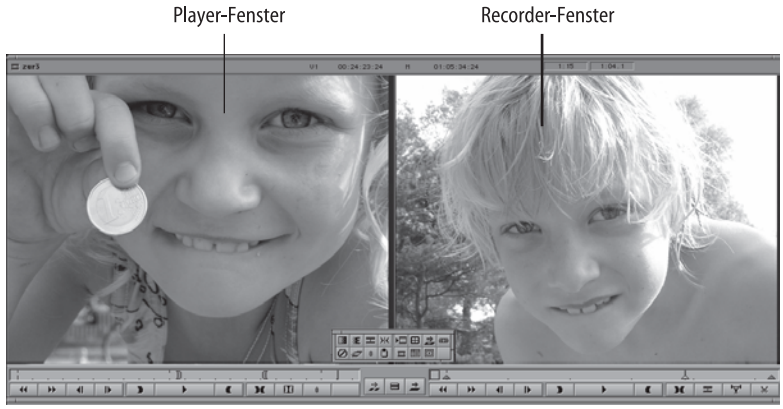


Abb. 9.68 Player- und Recorder-Fenster

Die Timeline symbolisiert in horizontaler Richtung die Zeitachse (Abb. 9.69), für die verschiedene Skalierungen eingestellt werden können, um wahlweise einzelne Bildübergänge oder den gesamten Film sichtbar zu machen (Abb. 9.69). Die auf der Timeline platzierten Sequenzen sind oft farblich gegeneinander abgehoben. Die eigentliche Schnittbearbeitung besteht in der Umordnung und Verlängerung oder Verkürzung der Sequenzen, was rechnerintern eine Umadressierung der auszuspielenden Szenen bewirkt. Bei ausschließlicher Verwendung von Hartschnitten ist eine Linie ausreichend. Für den Einsatz von Bildübergängen sollten mehrere Spuren verfügbar sein, die beteiligten Signale werden dazu meist auf Spuren in vertikaler Richtung angeordnet. Bei einfachen Systemen werden i. d. R. zwei Videosignale verarbeitet, die mit Clips aus einer dritten, der Grafikspur, überlagert werden können. Zwischen den beiden Videospuren befindet sich oft die Effects-Spur. Hier werden vordefinierte Bildübergänge von Videospur A zu Spur B platziert. So kann leicht auf die Effekte zugegriffen und ihre Dauer bestimmt werden. Unter den Video- sind die Audiotracks angeordnet, oft bis zu acht oder beliebig viele, die aber, wie auch die Videospuren, nur während der internen Verarbeitung und nicht separat zum Ausspielen in dieser Anzahl zur Verfügung stehen. Die Sequenzen können mit Effekten versehen werden, die auch auf der Timeline erscheinen. Hier wird die Benutzerfläche übersichtlicher, wenn diese Informationen nicht ständig, sondern nur wahlweise erscheinen (Abb. 9.69).

Beim Editing können die platzierten Sequenzen beliebig verändert oder ersetzt werden. Dazu muss zunächst bestimmt werden, welche Spuren an dem Vorgang beteiligt sein sollen. Bei der Platzierung einer neuen Sequenz zwischen zwei bereits platzierten, kann dann bei der Übernahme in die Timeline entschieden werden, ob die folgenden Sequenzen dabei überschrieben (Overwrite) oder nach hinten verschoben werden (Insert) (Abb. 9.70). Es ist auch möglich, das neue Material so zu platzieren, dass exakt eine Sequenz überschrieben wird (Fit to Fill). Komfortable Systeme erlauben darüber hinaus die anschließende Verschiebung des eingefügten Materials innerhalb des durch die Sequenzlänge gegebenen Zeitfensters (Slip Mode, Abb. 9.70). Bei all die-

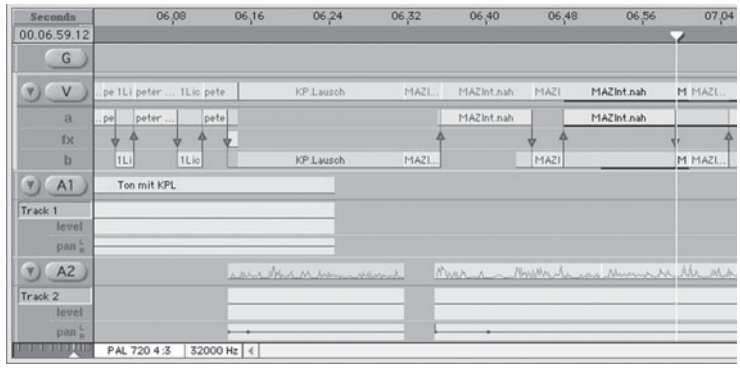


Abb. 9.69. Ein- und ausblendbare Details der einzelnen Spuren auf der Timeline

sen Vorgängen ist zu bedenken, dass die Verlängerungen und Verschiebungen von Sequenzen, sowie Überblendungen immer nur soweit möglich sind, wie es durch die Länge des digitalisierten Materials vorgegeben ist. Aus Gründen der Flexibilität und um auch Bildübergänge zu ermöglichen, wird beim Einspielen darauf geachtet, dass für jede Szene genügend sogenanntes Blendfleisch (Handles) zur Verfügung steht.

Nachdem die Sequenzen in der richtigen Folge angeordnet sind, kann mit dem Feinschnitt, dem Trimming, begonnen werden. Nach der Anwahl der zu bearbeitenden Spuren erscheinen im Trim-Mode zwei Fenster, die die Ein- und Ausstiegspunkte der beiden angrenzenden Szenen zeigen. Durch Anwahl eines oder beider Fenster wird bestimmt, ob die endende, die beginnende oder beide Szenen verändert werden sollen. Nur in letzterem Fall bleibt die Gesamtlänge des Films unverändert. Die Bearbeitung geschieht durch Verschiebung des

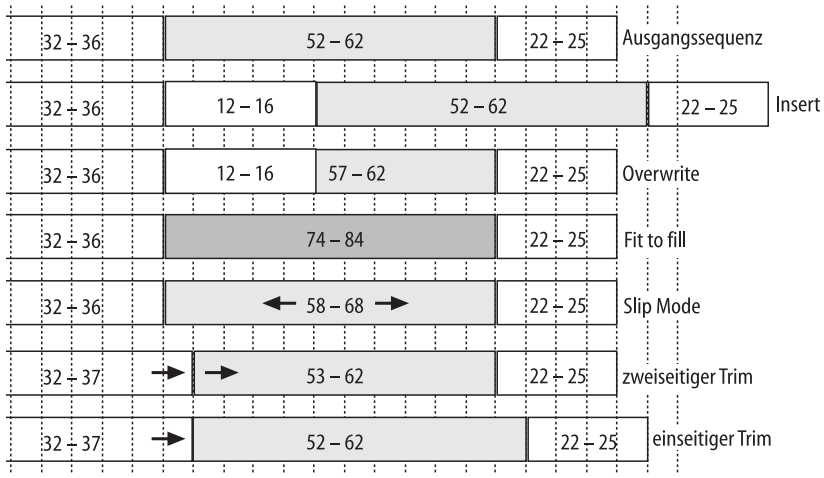


Abb. 9.70. Verschiedene Modi beim Einfügen neuer Sequenzen

Abb. 9.71. Bildausschnitt des Trim-Mode-Fensters

Schnittpunktes in Einer- oder Zehnerbildschritten. Um Synchronität zu gewährleisten, können Spuren vor dem Trimvorgang verkoppelt werden (Abb. 9.71).

9.4.2.4 Effekte und Compositing

Mit der steigenden Leistungsfähigkeit der Computersysteme wurden immer mehr Videobearbeitungsfunktionen in die Schnittsysteme integriert. Schon länger existieren Schnittstellen, die es erlauben computergenerierte Grafiken einzubinden. Die meisten Systeme enthalten auch Schriftgeneratoren, die sich allerdings bezüglich der Qualität von Antialiasing und Titelbewegung erheblich unterscheiden können.

Aufwendigere Bildbearbeitung erfordert die Möglichkeit, das Endbild aus vielen verschiedenen Video- und Standbildquellen zusammenzusetzen. Wenn diese Funktion gegenüber der Schnittfunktion überwiegt, spricht man von Compositing-Systemen (s. Abschn. 9.5). Viele Effekte werden als Software-Plug-Ins von Fremdherstellern angeboten. Effekte innerhalb eines Layers werden dabei als Filter bezeichnet, z. B. die Farbkorrektur für eine Sequenz.

9.4.2.5 Audiosignalverarbeitung

Die meisten NLE enthalten je nach Ausstattung eine digitale Audiobearbeitungsmöglichkeit für meist unbeschränkt viele Kanäle. Die Bearbeitung ist aber nur innerhalb des Systems mit sogenannten virtuellen Kanälen in dieser Anzahl möglich. Die Zahl der physikalischen Ein- und Ausgabekanäle ist auch bei größeren Systemen meist auf acht beschränkt, denn falls mehr erforderlich sind, wird man die Bearbeitung des Projekts an ein Audiobearbeitungssystem übergeben. Es werden vorwiegend die professionellen Parameter, also 16 oder 24 Bit Amplitudenauflösung und 44,1 kHz bzw. 48 kHz Abtastfrequenz, verwendet, die eine hohe Audioqualität ermöglichen. Auch in Offline-Systemen kann damit bereits ein großer Teil der Audiobearbeitung vorgenommen werden. Audiodaten können komprimiert gespeichert werden, in den meisten Fällen wird jedoch darauf verzichtet und mit höchster Qualität gearbeitet, da die Datenmenge im Vergleich zu hochqualitativ aufgezeichnetem Videosignal vernachlässigbar gering ist. Bei der Abschätzung des erforderlichen Speicherplatzes für den Offline-Schnitt mit starker Datenreduktion sollten die Audiodaten

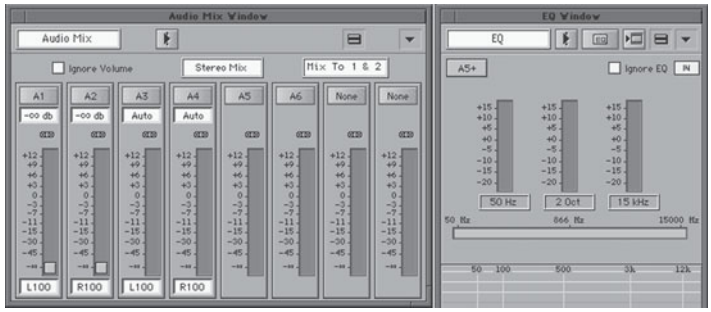


Abb. 9.72. Tools zur Audiobearbeitung

allerdings berücksichtigt werden, denn die ca. 100 kB/s Audiodaten pro Kanal sind natürlich relativ zu ca. 500 kB/s Videodaten (Reduktionsfaktor 40:1) nicht mehr einfach vernachlässigbar.

Die Bearbeitungsmöglichkeiten für die Audiosignale beschränken sich meistens auf Pegel- und Panoramaeinstellungen (Abb. 9.72). Diese können bei einigen NLE neben der Audiowellenform auch in der Timeline eingeblendet werden, sodass z. B. eine Pegelsenkung sehr einfach einer Videosequenz zugeordnet werden kann, denn der Pegelverlauf kann dann meist auch in der Timeline durch einfaches Setzen und Verschieben von Bearbeitungspunkten verändert werden (Rubberband, Abb. 9.73). Auch einfache Klangbeeinflussungsmöglichkeiten (Equalizer) sowie weitere Filterfunktionen sind oft standardmäßig verfügbar, wobei jedoch die Qualität recht unterschiedlich sein kann. Gute Ergebnisse werden dagegen erzielt, wenn auch hier Plug-Ins von Drittanbietern eingebunden werden können, die auf hochqualitative EQ- und Dynamikbearbeitung oder Halleffekte spezialisiert sind.

Für den Audioimport und -export werden üblicherweise die gewöhnlichen Audiodatenformate unterstützt. Dabei akzeptieren die Systeme oft auch Abtastraten, die ihre eigene Audiohardware nicht unterstützt. Die Einbeziehung derartiger Files erfordert dann eine Konversion der Abtastraten, die manchmal zu einer Minderung der Audioqualität führt. Gute Systeme unterstützen auch eine Voice Over-Funktion mit der ein Kommentarton in das System eingespielt werden kann, während gleichzeitig eine Videosequenz vom System ausgegeben wird.

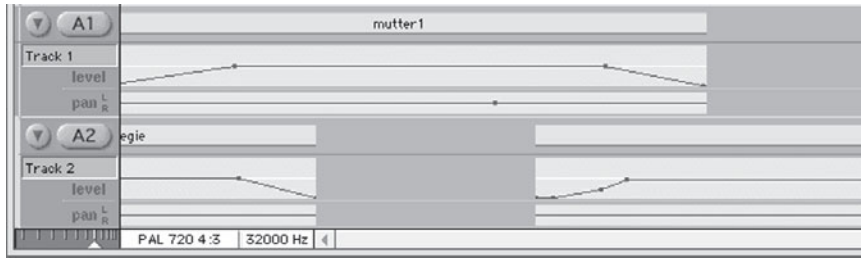


Abb. 9.73. Audiopegelbeeinflussung in der Timeline

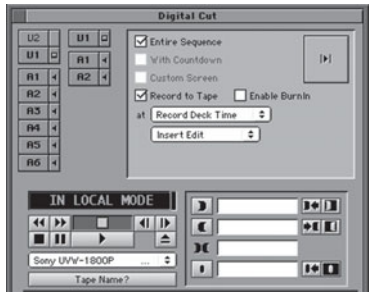


Abb. 9.74. Ausgabefenster für Digital Cut

9.4.2.6 Ausgabe

Zur Ausgabe des Films kann eine Schnittliste (EDL) erzeugt werden, die an einem linearen Online-Platz oder in einem teuren nichtlinearen System abgearbeitet wird, das durch den Vorschnitt zeitlich nicht belastet sein soll. Das System sollte möglichst den EDL-Export in verschiedenen Formaten beherrschen, die sich durch die Verbreitung verschiedener linearer Schnittsteuersysteme etabliert haben (CMX, Sony, Abb. 9.75). Außerdem kann der fertige Film als sogenannter Digital Cut auch auf Magnetband aufgezeichnet (Abb. 9.74) oder als digitaler Videofilm z. B. im Quicktime-Format exportiert werden. Viele Hersteller bemühen sich inzwischen auch um eine Ausgabemöglichkeit im MPEG-2- und MPEG-4-Format, damit das Schnittsystem im Zusammenhang mit dem DVD- und BD-Authoring verwendet werden kann.

Vom Digital Cut auf dem Masterband wird eine Sendekopie erstellt, die den üblichen technischen Vorspann enthält, der oft aus 30 Sekunden Schwarz, 60 Sekunden 75%-Farbbalken und 1 kHz-Sinuston bei –9 dB besteht, auf den dann wieder 20 Sekunden Schwarz und acht Sekunden ein Countdown mit Textangaben über den Beitrag folgen, bis schließlich noch zwei Sekunden Schwarzbild bis zum eigentlichen Beitragsbeginn zur Verfügung stehen.

Bei den Ausgabemöglichkeiten sollte auch auf eine komfortable Back up-Möglichkeit geachtet werden. Damit können die Projekt- zusammen mit den Mediadaten in einer Form vorgehalten werden, die jederzeit eine sofortige Weiter- oder Neubearbeitung des Materials zulassen. Als Back-Up-System wird z. B. die Speicherung auf Magnetbändern mit dem DLT-Verfahren (bandgestütztes Massenspeichersystem) eingesetzt.

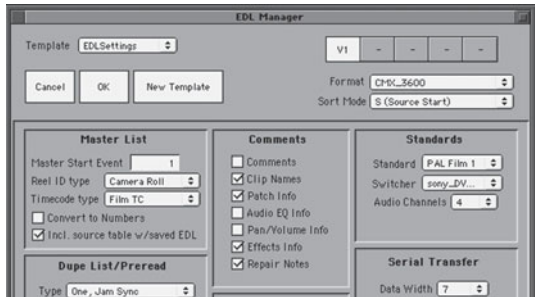


Abb. 9.75. Ausgabefenster für den EDL-Export

9.4.3 Filmschnitt

Auch beim Filmschnitt wird meistens mit nichtlinearen Videoschnittsystemen gearbeitet. Dabei sind sehr viele filmspezifische Umstände zu beachten, die eine genauere Darstellung und die Kenntnis der Filmparameter (s. Kap. 5), der Filmproduktion und des klassischen Filmschnitts erfordern.

9.4.3.1 Klassischer Filmschnitt

Für die Filmproduktion gibt es wesentliche Unterscheidungskriterien, die erheblichen Einfluss auf den technischen Ablauf haben. Dies ist zunächst die Frage, ob dokumentarisch gearbeitet wird oder szenisch, also eine Spielhandlung inszeniert wird. Die zweite wichtige Frage ist die Auswertungsform, d. h., ob für Kino oder Fernsehverwendung produziert wird. Beide Fragen hängen eng mit dem finanziellen Aufwand zusammen, der die technische Realisierung natürlich erheblich beeinflusst. Diese Aspekte bestimmen z. B., ob auf 16-mm oder 35-mm gedreht wird, die verfügbaren Kameras und Objektive, den Lichtaufwand, das Personal etc.

Bei Kinofilmproduktionen soll i. d. R. ein 35-mm-Positivfilm entstehen. Aufnahmeseitig wird dafür 35-mm-Material verwendet, aus Kostengründen aber teilweise auch auf (Super) 16-mm gedreht und das Format beim Kopierprozess angepasst. Der Ton wird separat aufgezeichnet, wobei mit Hilfe der Filmklappe und ggf. von Timecode die Synchronisation sichergestellt wird. Der belichtete Film wird am Ende eines jeden Drehtages entwickelt und als Muster (rushes) in Form eines Filmpositivs oder als Videoabtastung zur Beurteilung bereitgestellt. Dazu werden die gewünschten Takes von den unerwünschten (Nichtkopierern) getrennt, im Negativ aneinander geklebt und auf Positivfilm kopiert bzw. durch Filmabtastung (FAT) auf Videoband übertragen. Zur Überprüfung wird das Tonmaterial kontrolliert, und wenn alles Material für eine Szene vorliegt, kann bereits mit dem Rohschnitt begonnen werden.

Für den Schnitt im klassischen, analogen Filmbereich wird vom Negativmaterial eine Arbeitskopie als Positiv angefertigt. Der Ton wurde dabei früher vom Aufnahmemedium, z. B. 1/4“-Tonband (Senkel) auf perforiertes Magnetband überspielt. Die Synchronisation war dann unproblematisch, sie erfolgte über die mechanische Verkopplung der Antriebe für die Perfobänder.

Vor der Schnittfestlegung sind einige technische Vorarbeiten zu erledigen. Zunächst sollte geprüft werden, ob in den Mustern die Fußnummern vorhanden sind und ob der Ton so überspielt wurde, dass er zu der verwendeten Bildfrequenz passt, d. h. ob einer Sekunde Audiomaterial 24 oder 25 Bilder entsprechen. Als Nächstes gilt es, die zusammengehörigen Bild- und Tonteile am Schnittplatz zu synchronisieren. Dieser Vorgang wird Anlegen genannt. Als wesentliche Merkmale werden dafür die Klappen benutzt, die im Bild- und Tonmaterial zum selben Zeitpunkt sicht- und hörbar sind, wenn das Material richtig angelegt ist. Um sich in der Flut des Materials orientieren zu können, müssen zunächst die Einstellungen beschriftet werden. Die Anfänge und Enden werden mit A und E bezeichnet und die Einstellungsnummer notiert, weiterhin wird vermerkt, ob eine Schlussklappe vorliegt und ob ohne Ton oder ohne Klappe gedreht wurde.

Die Arbeit geschieht am Schneidetisch. Er gewährleistet die Synchronisation und bietet Sicht- und Abhörmöglichkeiten für Bild und Ton. Weit verbreitet waren die Schneidetische der Firma Steenbeck. Abbildung 9.76 zeigt einen solchen Tisch mit vier Tellern. Der Filmtransport wird mit dem sogenannten Steenbeck-Hebel gesteuert, der unten rechts zu sehen ist.

Der Rohschnitt des gesamten Films, d. h. die aneinandergefügten Szenen nach den Vorgaben im Drehbuch, liegt bereits kurz nach Beendigung der Dreharbeiten vor. Dann beginnt die Phase der Postproduktion. Dabei wird der Feinschnitt durchgeführt, bei dem anhand des zur Verfügung stehenden Materials die künstlerische Gesamtwirkung festgelegt und der Erzählrhythmus herausgearbeitet wird.

Nach dem Anlegen werden die Einstellungen in Klappenreihenfolge hintereinander angeordnet und möglichst durchgehend nummeriert. Dann kann eine Schnittliste angefertigt werden, in der das Material hinsichtlich seiner Verwendung klassifiziert und ausgemustert wird. Nun beginnt der eigentliche Schnitt, nach dessen Beendigung die fertig bearbeitete Arbeitskopie vorliegt. Für den Negativschnitt werden die Fußnummern in einer Negativschnittliste notiert, die im Kopierwerk abgearbeitet werden kann. Als Orientierung steht dort auch die Arbeitskopie zur Verfügung.

Mit dem Bild wird auch der Ton geschnitten, im Audiobereich kommen weiterhin noch Geräusche, Nachsynchronisationen und die Filmmusik hinzu, die in einem Mischatelier unter kinoähnlichen Abhörbedingungen abgemischt werden. In der Tonmischung werden die zu den Schnittsequenzen synchronen Originaltöne mit Geräuschen, Musik und Audioeffekten zu einem homogenen Klanggebilde gestaltet, das die Gesamtwirkung des Films in erheblichem Maße bestimmt. Als Orientierung steht auch hier die fertig geschnittene Bildarbeitskopie zur Verfügung. Die Schnittinformationen gelangen dann in Form der Schnittliste zum Kopierwerk, wo das Originalnegativ montiert, ggf. Trickteile eingefügt und die fertige Tonmischung auf ein Tonnegativ belichtet wird, sodass schließlich Bild- und Tonnegativ gemeinsam auf die Vorführpositive kopiert werden können. Für das Bild gibt es dann als wesentlichen Schritt noch die Lichtbestimmung, bei der die Helligkeit und Farbstimmungen der einzelnen Szenen einander angeglichen und für das Positiv optimiert werden.

Abb. 9.76. Filmschneidetisch

9.4.3.2 Filmschnitt mit NLE-Systemen

Damit beim Filmschnitt mit nichtlinearen Videoschnittsystemen gearbeitet werden kann wird das Filmmaterial zunächst in die digitale Form umgesetzt. Das Tonmaterial wird entweder direkt digital oder auf analogen Magnetbandmaschinen aufgezeichnet und anschließend digitalisiert bzw. eingespielt. Bei der Aufnahme wird dafür gesorgt, dass Ton- und Bildaufnahme synchron verlaufen. Zur Markierung auf der Audioseite dient dabei der Timecode. Die Filmbilder sind mit einem maschinenlesbaren Keycode (Fußnummern) versehen, der in Timecode umgesetzt werden kann. Nach Beendigung des Offline-Schnitts wird eine Schnittliste erzeugt und zusammen mit einem Playout auf Videoband an das Kopierwerk übergeben, wo anhand dieser Informationen der Film, wie oben beschrieben, physikalisch geschnitten wird [2].

Das wesentliche Problem beim digitalen Schnitt ist die unterschiedliche Zeitbasis mit 24 Frames per Second (fps) beim Film und 25 fps bzw. 29,97 fps bei Videosystemen. Die Anpassung der Filmbildfrequenz an die EU-Videofrequenz kann einerseits dadurch geschehen, dass der Filmabtaster um 4,16% zu hoch eingestellt wird und damit 25 statt 24 fps erzeugt. Dabei werden Bild und Ton geringfügig zu schnell wiedergegeben, was mit einer Tonhöhenänderung einher geht. Die zweite Möglichkeit besteht darin, in bestimmten Abständen einzelne Bilder zu wiederholen, um auf diese Weise ohne Tonhöhenänderung 25 fps zu erhalten. Hierzu kann mit dem sogenannten PAL-Pulldown jedes 12. Filmbild auf 3 Videohalbbilder ausgedehnt werden, während die vorhergehenden 11 Bilder wie üblich jeweils in nur zwei Halbbilder zerlegt werden. Für NTSC existiert ein eigenes 3-2-Pulldown-Verfahren (s. Kap. 5).

Die Filmbildfrequenz von 24 fps stellt kein Problem dar wenn alle Systeme der Bearbeitungskette diese Bildrate beherrschen. Das würde bedeuten, dass z. B. das Material über eine MAZ eingespielt wird, die in der Lage ist, mit 24 Bildern zu arbeiten und auch die Verarbeitung beim Schnitt tatsächlich auf 24 fps beruht. In diesem Falle sind die Bearbeitungsschritte direkt mit denen im Videobereich vergleichbar.

Problematisch wird es dagegen, wenn die 24-fps- und 25-fps-Bereiche gemischt werden. Die Schnittsysteme stammen durchweg aus dem Videobereich und basieren damit auf 25 fps. Damit sie für den Filmbereich nutzbar werden, müssen sie mit einer sogenannten Filmoption ausgestattet sein. Zur Erklärung muss der Arbeitsablauf betrachtet werden, der hier anhand des Avid-Systems dargestellt wird.

Zunächst wird das Filmmaterial dem Videosignal angepasst, indem es bei der Filmabtastung um 4,1% zu schnell, d. h. statt mit 24 fps mit 25 fps wiedergegeben wird. Bei der Abtastung werden auch die Fußnummern des Negativmaterials erfasst und in fester Zuordnung zu den Timecode-Werten des Videobandes in einer log-Liste gespeichert. Das Bildmaterial wird als 25-fps-Videosignal auf ein Magnetband überspielt und von dort aus wiederum mit 25 fps in das Schnittsystem eingelesen. Nach der Eingabe des Audiomaterials werden dann im nächsten Arbeitsschritt die zusammengehörigen Bild- und Tonelemente synchronisiert. Dazu können die Bildclips in eine neue Sequenz digitalisiert und der Ton anhand der Klappe angelegt werden. Mit der Funktion Autosync entstehen daraus dann synchrone, zusammenhängende Subclips.

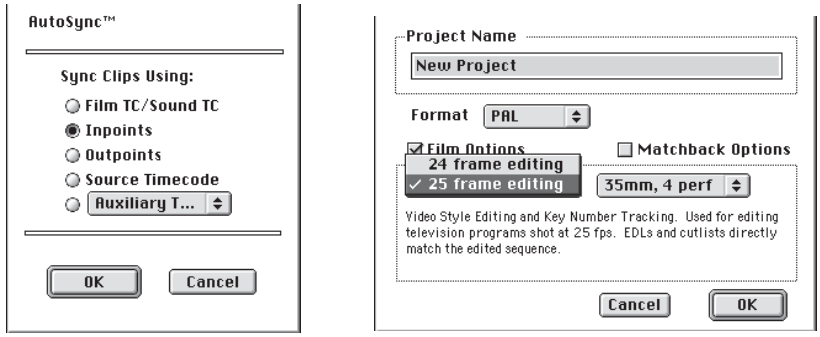


Abb. 9.77. AutoSync-Funktion und Filmprojektfenster

Für die Feinsynchronisation erlaubt das Avid-System mit Hilfe von Slip-Perf die Verschiebung der Video- gegenüber den Audioteilen mit einer Genauigkeit von 1/4 der Bilddauer (Abb. 9.77).

Danach kann mit dem Schnitt begonnen werden, wobei nun aber die Bildgeschwindigkeit der Originalszene dargestellt werden soll. Um diese in dem Videosystem und dem angeschlossenen Videomonitor realisieren zu können, wird vom Schnittsystem automatisch nach jedem 12. Bild ein Halbbild verdoppelt, sodass neben den 24 Bildern pro Sekunde zwei zusätzliche Halbbilder vorliegen, die zu einer Videofrequenz von 25 fps führen. Durch die Halbbildverdopplung tritt eine Bildruckelstörung auf, die aber nur bei schnell bewegten Objekten oder schnellen Kameraschwenks sichtbar wird.

Die zusätzlichen Halbbilder werden nur mit der Aktivierung der Filmoption erreicht. Beim Avid-System muss das Projekt dafür bereits vor der Einspielung der Daten für die Filmbearbeitung eingerichtet werden (Abb. 9.77). In den Film Settings (Abb. 9.78) wird dann festgelegt, dass der Transfer vom Film zum Videosystem mit der Video Rate geschehen soll, wenn die Abtastung mit 25 fps vorgenommen wurde. Für den Audio Transfer wird Film Rate gewählt, damit die Einspielung ohne Geschwindigkeitsänderung stattfindet.

Um die Orientierung im 24-fps-Projekt zu erleichtern, kann die Timecode-Anzeige im Recorderfenster zwischen der Zählung von 24 oder 25 Bildern/s

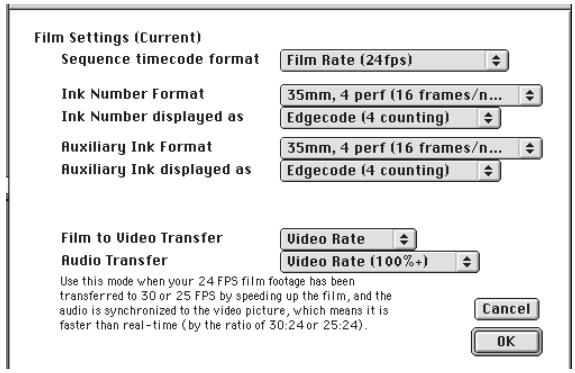


Abb. 9.78. Filmsettings

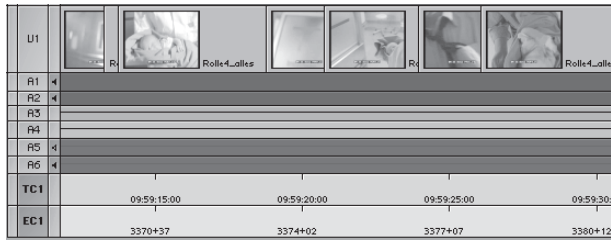


Abb. 9.79. Timecode und Edgecode in der Timeline

umgeschaltet werden. Beide Zähleinheiten können auch in der Timeline erscheinen, zusätzlich zu einer EC-Spur für die Anzeige der Fußnummern des Filmnegativs (Edgecode) (Abb. 9.79). Die Fußnummernverwaltung hat natürlich auch bei 25-fps-Filmprojekten, wie sie bei Filmproduktionen für den TV-Bereich verwendet werden, ihre Bedeutung.

Nach der Schnittfestlegung muss eine Ausgabe erfolgen, die es ermöglicht, dass im Kopierwerk das Originalnegativmaterial fehlerfrei physikalisch geschnitten und geklebt werden kann. Dazu gibt es mehrere Möglichkeiten. Erstens: Das Kopierwerk erhält vom Editor eine Schnittliste (Edit Decision List, EDL), die direkt die Zuordnung von Timecode- und Fußnummern beinhaltet. Das setzt voraus, dass beim Abtastvorgang eine log-Liste erstellt und an den Schnittplatz übergeben wurde, die den Bezug zwischen den Fußnummern und dem Timecode des Videobandes enthält. Zur Sicherheit werden die Fußnummern dabei oft in das Bildmaterial eingeblendet (Abb. 9.80). Die zweite Möglichkeit besteht darin, dass die Fußnummern bei der Abtastung in den VITC-Bereich (Bereich für die Übertragung von Timecode-Daten in der V-Austastlücke des Videosignals) des Videobandes geschrieben wurden, sodass sie beim Digitalisieren mit Hilfe eines Zusatzgerätes direkt in das Schnittsystem eingelesen werden können. Die dritte Möglichkeit ist, dass das Kopierwerk die bei der Abtastung erstellte Zuordnungsliste zwischen Fußnummer und Timecode des Videobandes behält. Dann braucht mit dem Schnittsystem nur eine Ausspielung auf Magnetband zu erfolgen (Digital Cut), das die Sequenzfolge mit den Timecodes des Einspielbandes enthält. Bei der Ausgabe auf Band muss in diesem Zusammenhang darauf geachtet werden, dass die Ausspielung bildidentisch ist, also keine Halbbildverdopplungen enthält. Zu diesem Zweck wird als Digital-Cut-Option Video Rate (100%+) angewählt. Die Abbildungen 9.81 und 9.82 zeigen die Einstellmöglichkeiten für die Ausgabe.

Die bildrichtige Ausspielung ist für den Schnitt erforderlich, sie ist jedoch zwangsläufig nicht zeit- bzw. längenrichtig. Eine längenrichtige Ausspielung wird aber auch benötigt, nämlich, wenn das geschnittene Bildmaterial als Referenz für die Tonmischung zur Verfügung gestellt wird. Für diesen Modus ist bei den Ausspieloptionen die Film Rate anzuwählen (Abb. 9.81).

Abb. 9.80. Eingeblendeter Keycode

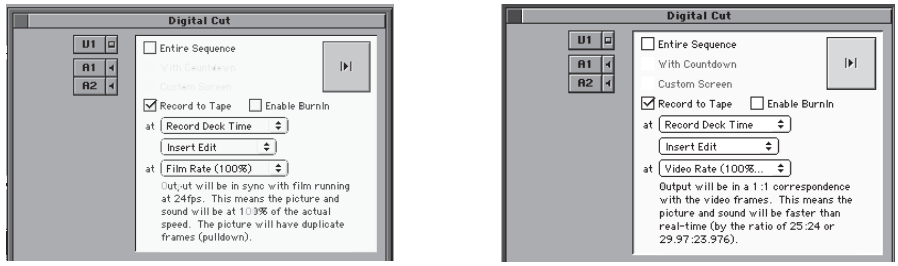


Abb. 9.81. Optionen für den Digital Cut

Das geschnittene Tonmaterial selbst wird dem Mischstudio möglichst in digitaler Form übergeben, wobei es hilfreich ist, wenn dabei auch direkt die Zuordnung zu verschiedenen Tonspuren erhalten bleibt.

9.4.3.3 NLE für 3D

Editingsysteme, die für die Verarbeitung der Halbbildsquenzen stereoskopischer Aufnahmen geeignet sind, müssen im Prinzip doppelt soviel leisten wie gewöhnlich. Ihre Hauptaufgabe ist, zwei Datenströme synchron und in Echtzeit wiederzugeben. Die Datenströme müssen so verkoppelbar sein, dass sich alle Editing- und Farbkorrekturschritte gleichzeitig auf beide Teile auswirken.

Daneben muss es aber auch auf möglichst einfache Weise möglich sein, die Bilder horizontal und vertikal gegeneinander zu verschieben, um Parallaxenfehler auszugleichen und vor allem, um die Stereobildwirkung durch Verschiebung des Konvergenzpunktes verändern zu können. Editing bedeutet hier also auch, Einfluss auf die Lage der Objekte in der Bildtiefe zu nehmen. Damit dies gut gelingt, muss die 3D-Wirkung in Echtzeit am Schnittplatz beurteilt werden können, d. h. es müssen 3D-taugliche Wiedergabesysteme vorhanden sein.

Stereoskopische Wiedergabeverfahren sind in Kapitel 7 dargestellt. Aufgrund der starken Kopfpositionsabhängigkeit kommen autostereoskopische Systeme beim Editing kaum in Frage. In den meisten Fällen wird mit übereinander projizierten Bildern gearbeitet oder mit zwei LCD, die im rechten Winkel zueinander angeordnet und über einen halbdurchlässigen Spiegel betrachtet werden. In beiden Fällen erfolgt die Bildtrennung mit Polarisations- oder Shutterbrillen.

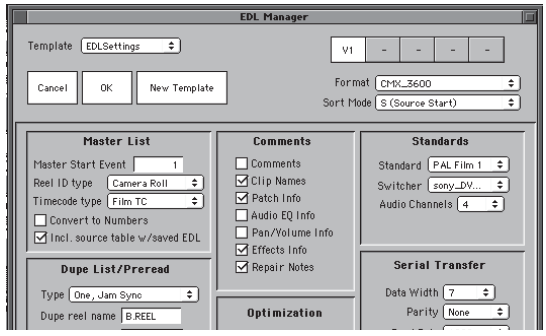


Abb. 9.82. EDL für den Filmschnitt

9.5 Compositingsysteme

Compositingsysteme dienen nicht dem Schnitt, sondern der Manipulation der Bilder selbst und ihrer Kombination. Diese Manipulationsmöglichkeiten standen oft im Zusammenhang mit Schnittsystemen, da in diese im Laufe der Zeit immer mehr Bildbearbeitungsfunktionen integriert wurden. Dann ging aber der Trend dahin, auf der gleichen technischen Basis wie bei den Editingsystemen eigenständige Bildbearbeitungssysteme einzusetzen, u. a. weil für die Arbeitsgebiete Schnitt und Bildbearbeitung sehr unterschiedliche Qualifikationen der Mitarbeiter erforderlich sind. Die Systeme werden dann als Compositingsysteme bezeichnet, weil die zu produzierende Bildsequenz meist aus vielen verschiedenen Video- und Standbildquellen zusammengesetzt wird und dazu oft Bildteile, z. B. durch Stanz- oder Maskenverfahren, kombiniert werden. Bei der grafischen Darstellung auf den Benutzeroberflächen werden die Bildkombinationselemente meist untereinander angezeigt, während die zeitliche Abfolge der Sequenzen in der Horizontalen, der Timeline, erscheint. Damit haben sich für die Differenzierung zwischen Bildschnitt und Compositing auch die Begriffe horizontaler und vertikaler Schnitt etabliert.

9.5.1 Grundfunktionen

Compositingsysteme sind mit Grafikprogrammen wie z. B. Adobe Photoshop vergleichbar, mit dem Unterschied, dass sie für die Verarbeitung von zusammenhängenden Bildfolgen ausgelegt sind. Zu den Grundfunktionen von Compositingsystemen gehören daher die Funktionen, die auch die Grafikprogramme bieten, wie Farbveränderungen, Mischungen von Bildern und Bildverzerrungen (Transformation, Abb. 9.83). Weiterhin die Möglichkeit der Freistellung einzelner Bildelemente durch Maskierung und Keying. Es existieren Paint- und Textmodule zur Erstellung von Farbflächen und Schriftzügen.

Eine zentrale Funktion hat die Möglichkeit der Verwendung separater Ebenen für Bildteile, die dann in verschiedener Weise gemischt und überblendet werden können. Hinzu kommen dann die Bearbeitungswerkzeuge, die über den Grafikbereich hinausgehen und sich auf die Bildfolge und den Zeitbereich beziehen. Dazu gehören einfache Schnittfunktionen, Animation und Tracking zur Verfolgung von Bildobjekten sowie die Möglichkeit der Einbeziehung von Audiosignalen inklusive ihrer Wellenformdarstellung.

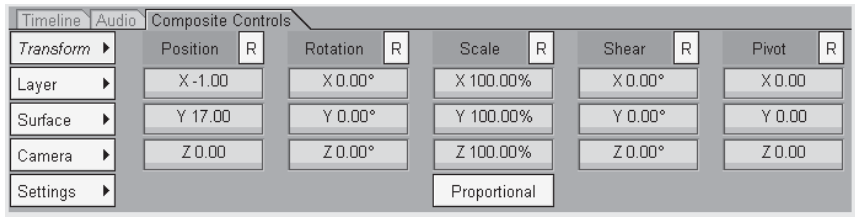


Abb. 9.83. Bildtransformationsparameter

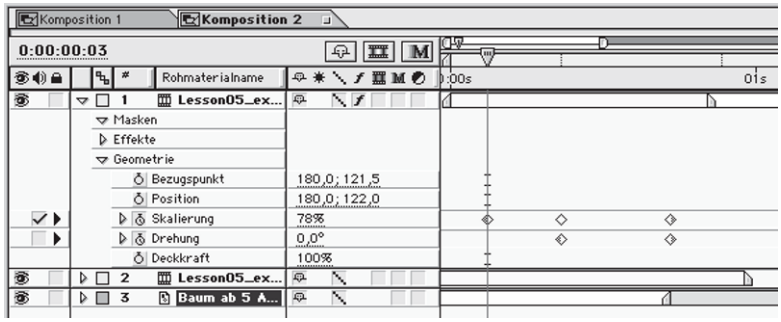


Abb. 9.84. Darstellung der Keyframes innerhalb der Layer

Wie auch im Grafikbereich ist der Einsatzzweck von Compositingsystemen nicht nur die Erschaffung vordergründiger Effekte, sondern oft die Bildmanipulation, die nicht auffallen soll. In diesem Zusammenhang ist es ratsam, die Einsatzmöglichkeiten und Beschränkungen der durch das System bereitgestellten Werkzeuge bereits vor den Dreharbeiten der später zu manipulierenden Bildsequenzen zu berücksichtigen und Effektspezialisten hinzu zu ziehen.

9.5.1.1 Layer

Die zentrale Orientierung, die bei Schnittsystemen die Timeline bietet, wird bei Compositingsystemen über das Layering erreicht, d. h. über das Übereinanderlegen der verschiedenen Bildelemente, die kombiniert werden sollen. Die Layer oder Ebenen haben eine Hierarchie, wenn unten angeordnete sichtbar werden sollen, müssen darüber liegende Layer wenigstens teilweise transparent sein. Die Layer können ein- und ausgeschaltet und animiert werden, d. h. die enthaltenen Bildteile werden mit der Zeit in ihrer Position verändert. Die Festlegung der damit verbundenen Bewegung erfolgt wie auch bei anderen animierten Effekten mit Hilfe von Keyframes anhand der Timeline, indem zu den per Keyframe definierten Zeitpunkten bestimmte Positionen definiert werden, zwischen denen das System dann interpoliert (Abb. 9.84). Die Veränderungen zwischen den Keyframes können oft separat und komfortabel editiert werden (Abb. 9.85).

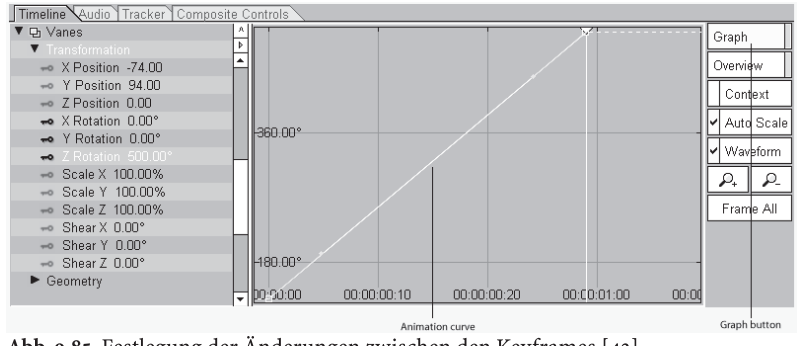


Abb. 9.85. Festlegung der Änderungen zwischen den Keyframes [42]

Abb. 9.86. Layeranordnung und Darstellung der Kamera in perspektivischer Darstellung

Neben der Definition der Layerpriorität über die Anordnung, bei der gewöhnlich immer die höchste Ebene Vorrang hat, bieten einige Systeme auch die Möglichkeit, die Ebenen in einem 3-D-Raum darzustellen (Abb. 9.86). Die Layer müssen dabei nicht mehr zwangsläufig parallel zueinander liegen, sondern können im Raum gekippt werden, und sich gegenseitig durchdringen. Die Priorität ist hier durch die Entfernung des Objekts vom Betrachtungspunkt definiert. Der 3-D-Raum ist zudem hilfreich bei der Positionierung des Betrachterstandpunktes, der durch eine Kamera symbolisiert wird, und von zusätzlichen virtuellen Lichtquellen, die z. B. eingesetzt werden, um ein farbiges Spotlight auf eine Bildfolge zu werfen.

9.5.1.2 Masken

Eine der zentralen Funktionen eines Compositingsystems ist die Möglichkeit, Einzelteile eines Bildes zu isolieren, denn nur wenn sie sauber freigestellt sind, können sie so mit anderen Bildern verbunden werden, dass die ungestörte Illusion eines neuen Bildes entsteht. Die Freistellung geschieht mit Hilfe von Masken, die um die Umrisse der freizustellenden Objekte gelegt werden. Die Masken können unabhängig vom Bild animiert und auch separat dargestellt werden. Dabei erscheinen völlig abgedeckte bzw. transparente Bereiche schwarz bzw. weiß (Abb. 9.87), während die Grauwerte teiltransparente Gebie-

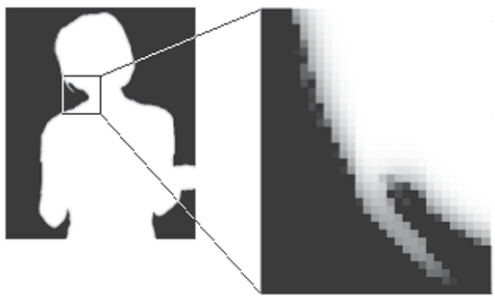


Abb. 9.87. Maske mit teiltransparenter Grenze [46]

te darstellen. Abbildung 9.87 zeigt eine Maske mit teiltransparentem Bereich am Maskenrand, die es ermöglichen, dass sich das Bild ohne harte Übergänge in den Hintergrund einpassen kann. Wenn die Graustufeninformationen des Maskenbilds als zusätzlicher Kanal zu den Farbwerten übertragen werden, so bezeichnet man diesen als Alpha-Kanal, ein RGB-Bild mit Alpha-Kanal trägt daher die Bezeichnung RGBA.

9.5.1.3 Keying

Die Erstellung der Masken geschieht entweder manuell oder automatisch, indem die freizustellenden Gebiete anhand charakteristischer Merkmale in jedem Bild der Sequenz erkannt werden. Diese Funktion wird dann als Keying oder Stanze bezeichnet. In den meisten Fällen werden als Merkmale bestimmte Helligkeits- oder Farbwerte gewählt, entsprechend wird mit Luminanz- oder Chrominanzkey gearbeitet. Wenn z. B. Personen in einen fremden Hintergrund gesetzt werden sollen, werden sie oft freigestellt, indem sie vor einer gleichmäßig ausgeleuchteten Blauwand aufgenommen werden und ggf. von blauen Elementen abgestützt werden (Abb. 9.88). Da das Blau kaum in der menschlichen Gesichtsfarbe vorkommt, lässt sich die Person relativ leicht vom Hintergrund trennen und vor ein neues Bild platzieren. Diese Funktion ist aus der täglichen Fernsehproduktion bekannt und ein Standardbestandteil bei Bildmischern (s. Abschn. 9.1.3). Die Keyqualität gewöhnlicher Bildmischer kann aufgrund der dort geforderten Echtzeitfähigkeit i. d. R. jedoch nicht mit der von Compositingsystemen konkurrieren. Im Filmbereich wird neben dem Dreh vor Blau oft auch vor Grün produziert, was den Vorteil hat, dass der Grünauszug des Films weniger Rauschanteile enthält.

Die Erstellung eines sauberen Keys ist keine einfache Aufgabe. Die Compositingsysteme unterscheiden sich in der Qualität der Keyer erheblich. Hersteller, die selbst nur Keyer in Standardqualität in ihre Systeme integrieren, bieten aber oft die Möglichkeit, hochwertige Keyer, wie z. B. die der Firma Ultimatte, per Software-Plug-In einzubinden. Keyeinstellungen sind dadurch problematisch, dass das gewählte Stanzkriterium wie z. B. die Blauwand eine ungleichmäßige Blauwertverteilung und Rauschen aufweisen kann. Weiterhin sind natürlich Farben innerhalb der auszustanzenden Gebiete kritisch, die

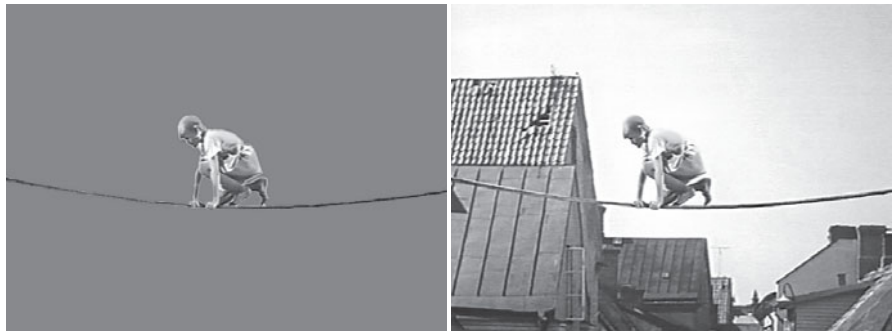


Abb. 9.88. Vordergrundbild vor Blau und Gesamtbild nach dem Key

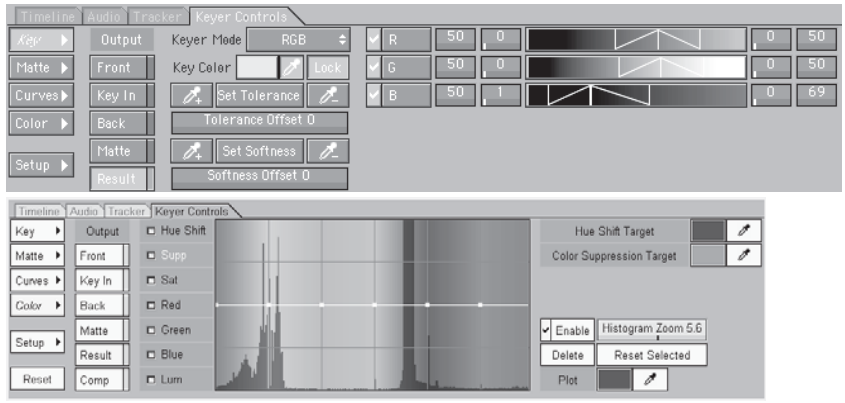


Abb. 9.89. Einstellmöglichkeiten bei hochwertigen Keyern [46]

ähnliche Blauwerte wie die Stanzfarbe aufweisen, da hier „Löcher“ im Objekt entstehen können, durch die der Hintergrund sichtbar wird. Durch die Farbgebung des Objekts lässt sich dieses Problem allein nicht lösen, denn in den meisten Fällen lässt sich nicht verhindern, dass Licht von der Blauwand auf das Objekt reflektiert wird. Diese Reflektionen (Color Spill) sind besonders an den Objektkanten sichtbar. Schließlich sind fein strukturierte und teiltransparente Gebiete besonders kritisch, wie z. B. Glas, durch die die Blauwand hindurchschimmert. Hochwertige Keyer erlauben die Farbseparation auf Basis verschiedener Farbräume (RGB, YUV, HLS) und die Einschränkung des auszustanzenden Bereichs mittels Masken. Sie bieten durch die separate Behandlung von Keymask und Keyfill eine Möglichkeit zum Umgang mit teiltransparenten Bereichen sowie mit der Color Spill Suppression eine Hilfe zur Unterdrückung der störenden Reflexionen. Abbildung 9.89 zeigt oben das Keyer-Bedienfeld und darunter ein Feld, in dem die Farbtöne, die den Color Spill hervorrufen, angezeigt und reduziert werden können. Die Keyparameter können auch zeitlich der Sequenz angepasst werden um z. B. einer über die Sequenzdauer stattfindenden Farbwertänderung gerecht zu werden. Um das separierte Objekt dann gut in den Hintergrund einpassen zu können stehen schließlich Bearbeitungswerkzeuge zur Verfügung, die die Kanten eingeschränkt (shrink) oder unscharf (blur) oder nach innen auswaschend (erode) darstellen (Abb. 9.90).

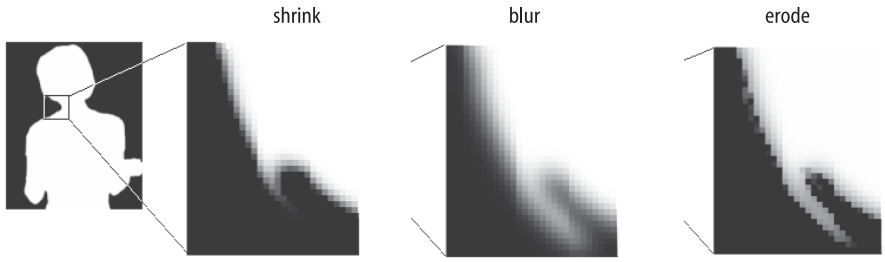


Abb. 9.90. Möglichkeiten zur Behandlung der Kanten ausgestanzter Objekte [46]

9.5.1.4 Tracking

Mit diesem Begriff wird ein Werkzeug zur Bearbeitung der zeitlichen Dimension einer Bildsequenz bezeichnet. Tracking dient der Verfolgung von Bildteilen über der Zeit und der Erstellung eines Bewegungspfades, der wiederum auf andere Bildelemente angewandt werden kann. Die Bewegung wird anhand markanter Punkte, wie z. B. deutlicher Helligkeits- oder Farbunterschiede, analysiert (Abb. 9.91, oben links). Die Punkte werden vom Nutzer ausgesucht und dann automatisch verfolgt, wobei die Größe des Suchbereichs abhängig von der Bewegungsgeschwindigkeit variiert werden kann. Der ermittelte Bewegungsverlauf wird meist als zeitliche Folge von Keyframes dargestellt und ist veränderbar. Der Bewegungsverlauf kann dann beliebigen Bildteilen, Masken oder Layern zugeordnet werden, die sich dann entsprechend mitbewegen.

Abbildung 9.91 zeigt als Beispiel ein Segelschiff, dessen Mastspitze vor einer Kaimauer erscheint. Um die Illusion zu erreichen, dass das Schiff auf offenem Meer fährt, kann die Mauer ausmaskiert und durch eine Bildsequenz ersetzt werden, die Himmel mit bewegten Wolken zeigt. Da die Rechteckmaske hier die Segelspitze abschneidet, wird diese separat maskiert. Die Tracking-Funktion dient dann dazu, die Segelspitze zu verfolgen und ihre Bewegung entsprechend der Schiffsbewegung vor dem neuen Hintergrund fortzusetzen [46].

Zur Erfassung der Bewegung von Flächen, die sich mit der Zeit perspektivisch ändern, werden mehrere Tracking-Punkte verwendet. So kann mit Hilfe des 4-Point-Tracking z. B. eine Werbetafel auf der Tür eines Autos angebracht werden, das in die Tiefe des Bildes hineinfährt, und die Tafel ändert sich entsprechend. Die mit der Tracking-Funktion erzielte Bewegungsanalyse kann darüber hinaus auch zur Stabilisierung der Bildsequenz verwendet werden.

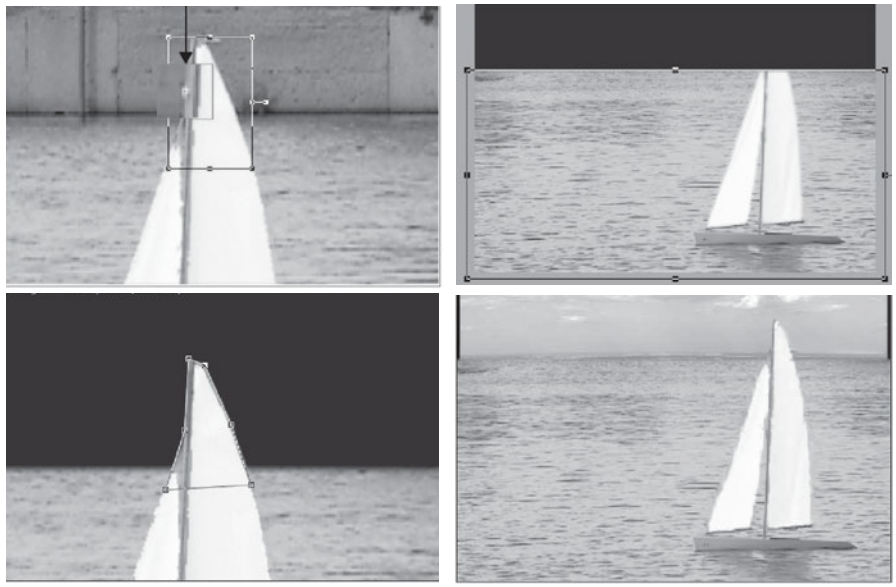


Abb. 9.91. Beispiel für die Anwendung einer Tracking-Funktion [46]

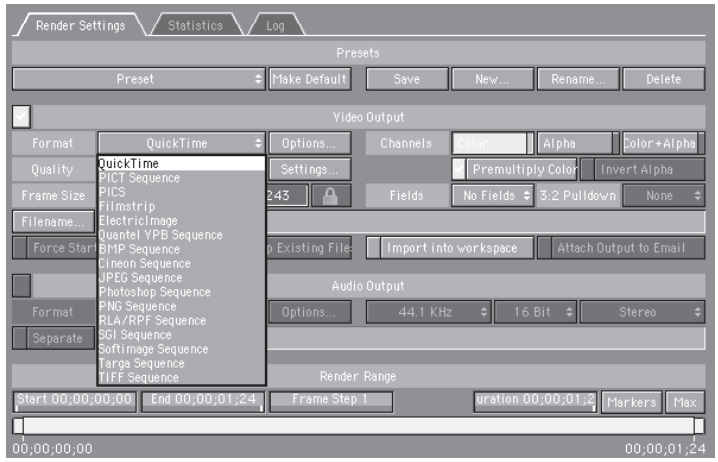


Abb. 9.92. Renderingtool

9.5.1.5 Rendering und Ausgabe

Wenn die zu bearbeitende Bildfolge aus den gewünschten Teilen zusammengesetzt ist, wird das Composit mit dem sogenannten Rendering endgültig berechnet und als Bildsequenz abgelegt. Im Rendering-Tool können dazu alle notwendigen Parameter bestimmt werden (Abb. 9.92). Bei aufwendigen und langen Sequenzen kann das Rendering auch mit modernen Rechnern eine erhebliche Zeit in Anspruch nehmen. Daher bieten gute Compositingsysteme die Möglichkeit, die Aufgabe auf mehrere Rechner in einem Verbund zu verteilen. Um sich während der Arbeit auch ohne Rendering einen Überblick über die Bewegungsabläufe verschaffen zu können, werden kurze Sequenzen über den RAM-Speicher in Echtzeit ausgespielt. Die Dauer dieser Bildfolgen hängt vom Composit-Aufwand und natürlich von der verfügbaren RAM-Größe ab, die auch in einfachen Fällen mehrere GB betragen sollte.

Für die Ausgabe ist neben dem Speicherort und dem Filenamen auch das gewünschte Formt anzugeben. Hier stehen meist die gängigen Bildformate wie Tiff, BMP zur Verfügung, weiterhin QuickTime mit seinen verschiedenen Kompressionsalgorithmen (s. Kap. 10) sowie die bekannten Fileformate aus dem Filmbereich (Cineon, DPX, s. Kap. 5). Vor der Ausgabe können die Sequenzen bezüglich der Bildauflösung mit unterschiedlichen Interpolationsverfahren skaliert werden. Schließlich ist es auch möglich, bei der Ein- und Ausgabe die Halbbilder des Zeilensprungverfahrens in verschiedener Art zu berücksichtigen.

9.5.2 Ausführung von Compositingsystemen

Compositingsysteme existieren in großer Vielfalt. Beim Vergleich der technischen Parameter ist zum Ersten die Auflösungsunabhängigkeit zu beachten, zweitens die Möglichkeit mit verschiedenen Farbtiefen zu arbeiten. Die aufwendigen Berechnungen erfordern eine hohe Rechnerleistung, auch diese

kann von verschiedenen Systemen unterschiedlich effektiv genutzt werden. Oft wurden früher Maschinen von Silicon Graphics (SGI) verwendet, mit dem Betriebssystem IRIX, einem UNIX-Derivat. Mit steigender Leistungsfähigkeit der Standard-PC kamen dann immer mehr Compositingprogramme für Rechner unter Windows und MacOS auf den Markt.

Ein verbreitetes und preiswertes Programm aus dem unteren Preisbereich ist After Effects von Adobe, das für Windows und MacOS verfügbar ist. Trotz des günstigen Preises bietet es viele hochwertige Funktionen. Es lässt sich zudem sehr einfach und effektiv in einer Kombination mit dem Bildbearbeitungsprogramm Adobe Photoshop und dem Filmschnitt-Programm Adobe Premiere betreiben.

Einen guten Namen im Bereich Compositing hat sich die Firma Autodesk (früher Discreet) gemacht, die eine breite Palette von Software aus allen Leistungsklassen anbietet, mit dem großen Vorteil, dass die Benutzeroberflächen sehr gut an die individuellen Bedürfnisse der Nutzer anpassbar und einander sehr ähnlich sind. Auf der unteren Leistungsstufe steht hier das Programm Combustion, das unter Windows und MacOS läuft und bereits eine Vielzahl von Werkzeugen beinhaltet, die z. T. den großen Systemen entlehnt sind, wie z. B. die hochwertige Farbkorrektur und das sehr gute Keymodul (Abb. 6.93). Die leistungsfähigeren Systeme werden mit Flint, Flame und Inferno bezeichnet [70]. Sie unterscheiden sich neben der Geschwindigkeit im Wesentlichen durch unterschiedliche Möglichkeiten bei der Filmbearbeitung.

Abb. 9.93. Benutzeroberfläche eines Compositing-Programms

9.6 Computeranimation

Ebenso wie die Bereiche Editing und Compositing haben auch die Gebiete Computeranimation und Compositing Ähnlichkeiten. Als Abgrenzung lässt sich anführen, dass Compositingsysteme vor allem der Manipulation existierender Bildfolgen dienen, während es bei der Computeranimation um die Erschaffung neuer Bildfolgen geht, die auf Objekten in einem dreidimensionalen Raum basieren, die bildweise so verändert werden, dass ein Bewegungsablauf erscheint (Animation). Auf diese Weise können irreale Kreaturen und Fantasiewelten im Film, für Computerspiele oder zur Architekturvisualisierung entstehen.

Computeranimationssysteme erfordern ähnliche Hardware wie Compositingsysteme. Die Software ist in den meisten Fällen aus einzelnen abgegrenzten Modulen aufgebaut. Sie dienen der Erstellung und Modellierung der Objekte, der Bestimmung der Oberflächenstruktur und der Definition des Verhaltens bei Beleuchtung und bei Bewegungen, weiterhin der Animation und schließlich der Berechnung der Bildfolgen [31].

Die Erstellung von 3D-Animationen ist in die Phasen Modelling und Rendering unterteilt. Beim Modelling wird zunächst ein Drahtgitter erzeugt dessen Umrisslinien Volumen und Form des Objekts andeuten (Abb. 9.94). Dabei wird das Objekt oft aus einfachen Elementen wie Kugeln, Rechtecken und Zylindern zusammengesetzt. Komplexere Oberflächen bestehen aus einer Vielzahl von Polygonen, deren Anzahl die Genauigkeit der modellierten Oberfläche bestimmt (Abb. 9.95). Anstatt das Modell von Hand zu entwerfen können auch fertige Modelle gekauft und importiert werden oder die Daten werden gewonnen, indem ein reales Objekt mit einem 3-D-Scanner in allen Raumdimensionen abgetastet wird.

Nach der Modellierung wird die Oberflächenbeschaffenheit bestimmt. Dazu wird im einfachen Fall eine 2-D-Textur auf Teile des Objekts gelegt, die eine Wirkung wie Stein, Holz, Stoff etc. hervorruft. Die Texturen können wieder selbst erstellt werden oder sie kommen aus einer zugehörigen Bibliothek bzw. werden von Drittanbietern zugekauft.

Um einen Bildeindruck zu gewinnen, der natürlichen Objekten entspricht, muss zusätzlich bestimmt werden, wie sich das Material bei Lichteinfall verhält, d. h. ob es das Licht eher absorbiert oder reflektiert, ob Glanzlichter ent-

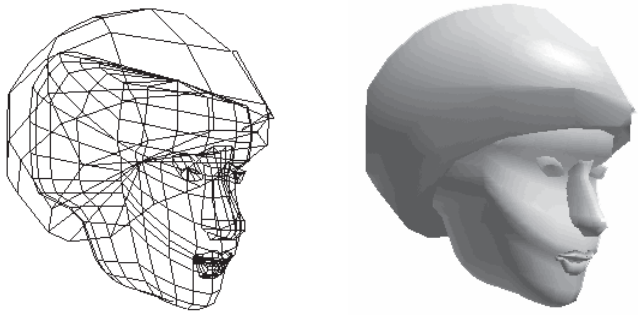


Abb. 9.94. 3D-Bild als Wireframe und nach dem Rendering

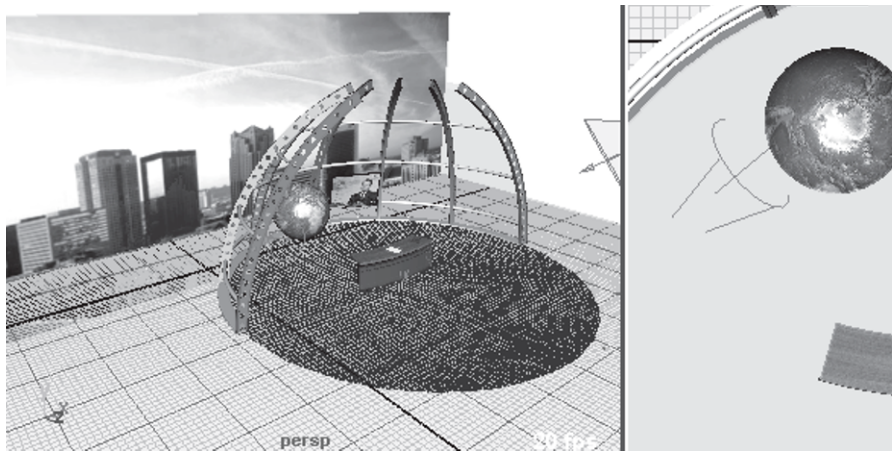


Abb. 9.95. Komplexes Wireframe-Modell vor Hintergrundbild in 3D-Ansicht und Aufsicht

stehen etc. (Abb. 9.96). Darüber hinaus kann das Reflexions- oder Transparenzverhalten mit den Grauwerten der Textur verkoppelt werden. So entsteht bei der Beeinflussung der Reflexion mit dem sogenannten Bump Mapping die Illusion von Unebenheiten auf der Oberfläche. Die Beeinflussung der Transparenz wird als Transparency Mapping bezeichnet, während das Displacement Mapping eine Verformung der Oberfläche durch die Grauwerte hervorruft.

Nachdem die Objekte definiert sind, können sie zueinander in Beziehung gesetzt und animiert werden. Interessant ist dabei die Möglichkeit, die Freiheitsgrade der Bewegung einzelner Elemente einzuschränken, damit sie mit anderen verkoppelt werden können, sodass z. B. der Eindruck erzeugt wird, dass sich ein Unterarm in typischer Abhängigkeit vom Oberarm bewegt.

Neben der Bewegung im Raum werden hier auch die Lichtquellen und der Kamerastandpunkt bestimmt. Die Lichtquellen können farblich und bezüglich der Lichtverteilung (Spot, diffus) verändert werden. Für die Objekte wird dagegen bestimmt, ob sie Schatten werfen sollen und welcher Art diese sind.

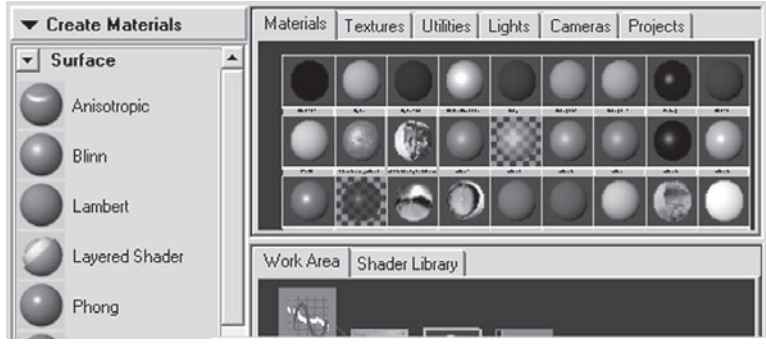


Abb. 9.96. Festlegung der Oberflächenbeschaffenheit

Nach der Festlegung aller Parameter im 3D-Raum müssen schließlich die zweidimensionalen Bilder des zu erstellenden Films errechnet werden, was mit Rendering bezeichnet wird. Dabei können unterschiedliche Filme schon dadurch entstehen, dass bei gleicher Objektbewegung eine unterschiedliche Bewegung der in der Software definierten Kamera zugrunde liegt. Das Rendering kann auch bei leistungsfähigen Rechnern eine erhebliche Zeit in Anspruch nehmen. Einen großen Einfluss nimmt hier das Beleuchtungsmodell, das wiederum die Realitätsnähe der Bilder bestimmt. Die einfachste Form ist die Darstellung des Schattenwurfs (Shading), komplexer ist das Raytracing, die Berechnung von Reflexionen und Schatten durch Strahlverfolgung von der Quelle bis zum Beobachter. Ähnlich aufwendig ist das Radiosity-Verfahren, dem zugrunde liegt, dass sich die Objekte unter Einfluss der gesetzten Lichtquellen gegenseitig mit diffusem Licht bestrahlen.

Computeranimationssysteme gibt es in großer Zahl. Zunächst wurden die Software-Pakete für die High-End-Rechner von SGI entwickelt, sie stehen aber heute meist auch für Standard-PC unter Windows und MacOS zur Verfügung. Eines der bekanntesten war lange Zeit Softimage von der gleichnamigen Firma. Zu Beginn des neuen Jahrhunderts hat sich vor allem das Software-Paket Maya als führend durchgesetzt, das inzwischen für alle drei genannten Rechnerplattformen verfügbar ist (Abb. 9.97).

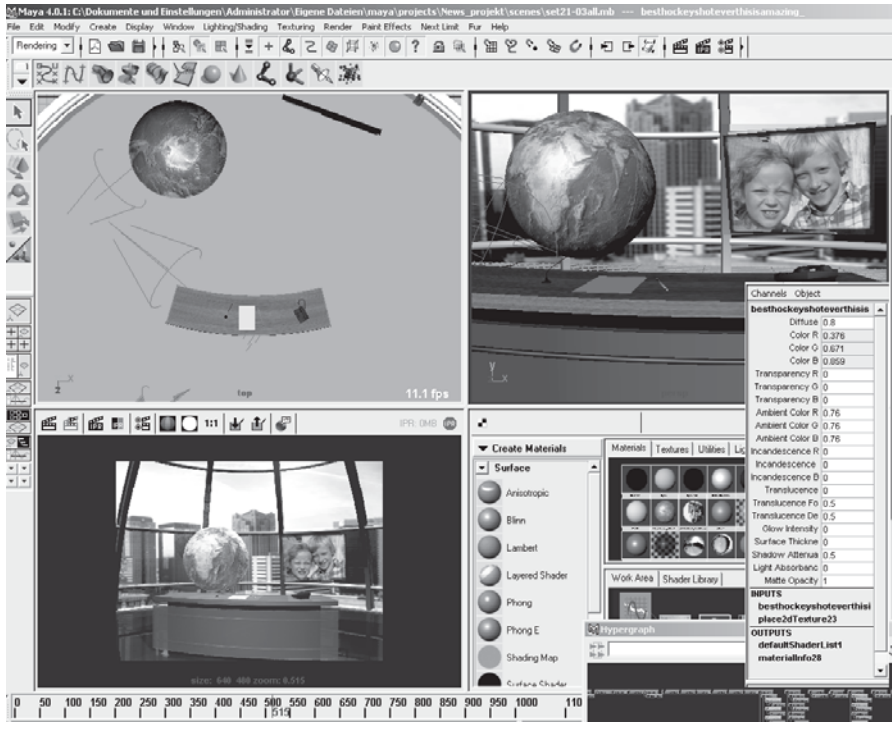


Abb. 9.97. Benutzeroberfläche eines Computeranimationsystems

10 Videostudiosysteme

Im Schlusskapitel dieses Buches geht es um die Verknüpfung der in den vorangegangenen Kapiteln betrachteten Grundlagen und Geräte zu Produktions-, Postproduktions- und Studiosystemen. Dabei reicht es nicht mehr aus, nur die Fernsehstudios zu betrachten, denn die Anwendung von Videosystemen zieht immer weitere Kreise. Zwar bleibt der Fernsehbereich nach wie vor das dominante Einsatzgebiet von Videosystemen, doch kommen immer mehr Gebiete wie DVD-Produktion, Videoübertragung im Internet und Videoanwendungen im Filmbereich hinzu.

Die Orientierung in dem von „Multimedia" beeinflussten Broadcastbereich ist nicht leicht. Eine relativ klare Unterteilung ergibt sich noch hinsichtlich der Bildauflösung. In vielen Fällen wird noch die Standardauflösung (Standard Definition, SD) mit 720 x 576 aktiven Bildpunkten verwendet. High Definition (HD) mit 1280 x 720 oder 1920 x 1080 Pixeln und dem Bildseitenverhältnis 16:9 findet aber immer stärker Eingang in den Fernsehbereich. Gegenwärtig werden diese und höhere Auflösungen, wie 2k und 4k, auch für Filmanwendungen verwendet. Am anderen Ende beschreibt Low Definition mit 352 x 288 Bildpunkten eine Auflösung von einem Viertel des Standardformates. Diese oder noch geringere Auflösungen werden in Kombination mit erheblicher Datenreduktion für die Videoübertragung bei kleiner Datenrate, z. B. im Internet eingesetzt. Um die Orientierung zu erleichtern, lässt sich Tabelle 10.1 nutzen, in der die Auflösungen verschiedenen Gebieten aus den Bereichen Produktion und Postproduktion zugeordnet werden.

Historisch junge Bereiche sind die Computeranimation und die Generierung virtueller Welten. Man spricht von Virtual Reality und meint Umgebungen, die komplett im Rechner entstehen. Als Vorstufe existiert die augmented Reality, d. h. die Abbildung der realen Welt, die durch virtuelle Elemente unterstützt wird. Diese Bereiche stehen zwischen den Gebieten Produktion und Postproduktion, ebenso wie die interaktiven Systeme. Interaktive Systeme gewähren dem Nutzer einen erheblichen Einfluss auf den Umgang mit dem Medienprodukt. Ein einfaches Beispiel ist die DVD (s. Kap. 8), bereits hier können die Nutzer – natürlich nur bei entsprechender Vorproduktion – den Fortgang der Geschichte selbst bestimmen. Noch weit komplexere Möglichkeiten sind durch die Manipulierbarkeit von Objekten bei MPEG-4 gegeben (interactive storytelling). Um deren Vielzahl gerecht zu werden, ist ein großer (Post-) Produktionsaufwand in einem eigenen Bereich, dem sogenannten Authoring erforderlich, der bei guter inhaltlicher Qualität mit immensen Kosten verbunden sein kann.

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4_10, © Springer-Verlag Berlin Heidelberg 2013

Tabelle 10.1. Übersicht über Produktions- und Postproduktionsbereiche

	Produktion	Postproduktion	
	Aquisition	Authoring	Editing
HD, 2k und 4k	Filmproduktion	Animation	Schnitt, Compositing, Grading
SD und HD	Studio, EB, EFP	DVD, MPEG-4	Schnitt, Effekte, Animation
Low Definition	Studio	Webdesign	Schnitt, SD/LD-Anpassung

10.1 Signale im Produktionsbereich

Im Videostudio stehen die in den vorangegangenen Kapiteln beschriebenen Geräte zur Verfügung, mit denen die Videosignale aufgenommen, gespeichert, gemischt, geschnitten und bearbeitet werden. Dabei entstehen besondere Anforderungen an das Videosignal, vor allem bei der Verwendung datenreduzierter Signale. Daher werden in diesem Kapitel zunächst noch einmal die Videosignale hinsichtlich der Verwendung im Produktionsbereich betrachtet.

10.1.1 Analoge Signalformen

Das in Kap. 2 beschriebene PAL-FBAS-Signal ist nicht nur als ein analoges Sendeformat geeignet, sondern kann auch in allen Stufen der Produktion und Postproduktion eingesetzt werden. Das geschieht aber heute kaum noch, zumindest als Produktionssignal. Aufgrund der Tatsache, dass es den Vorteil hat, alle Signalbestandteile zu integrieren und damit die einkanalige Signalführung zu erlauben, findet man es aber doch noch relativ oft in der Peripherie des Produktionsbereichs, auch im HD-Umfeld, wie z. B. bei der Signalverteilung zu Vorschauzwecken an die Ton- oder Lichtregie. Von Vorteil ist, dass das Signal nicht mit merklichen Laufzeitproblemen behaftet ist, die Latenz ist bei Digitalsignalen meist deutlich größer. In Ländern, in denen mit SECAM-Signalen gearbeitet wird, kann das FBAS-Signal nicht zur Produktion verwendet werden, der Produktionsbereich ist in Komponenten- oder PAL-Technik ausgestattet und das SECAM-Signale wird erst zum Schluss durch Umsetzung oder Transcodierung gewonnen [195].

Auch das Y/C-Signal kann PAL-codiert sein, es wurde fast ausschließlich im semiprofessionellen und im Heimbereich verwendet. Dagegen hatte sich im professionellen Bereich das Komponentenformat durchgesetzt. Es bietet eine hohe Chromabandbreite und ist mit keinem der oben genannten Probleme behaftet. Das Komponentensignal war bis zum Einzug der Digitaltechnik der Standard bei hochwertiger analoger Produktion und Postproduktion. Mit Betacam SP stand hier ein hochwertiges MAZ-System zur Verfügung, das die Qualität der FBAS-Direktaufzeichnungsverfahren erreicht, dabei aber preisgünstiger und wegen der Verwendung von Kassetten leichter handhabbar ist. Der Nachteil des analogen Komponentenformats ist die dreifache Leitungsführung.

10.1.2 Digitale Signalformen

Analoge FBAS- und Komponentensignale lassen sich direkt in die digitale Form überführen. Das Digital-Composite-Signal hatte aber nur dort Bedeutung, wo digitale Systeme, insbesondere MAZ-Geräte, in eine vorhandene FBAS-Infrastruktur eingebunden werden sollten. Heute ist in der digitalen Studiotechnik allein das digitale Komponentensignal mit serieller Übertragung (DSK) von Bedeutung (s. Abschn. 3.3), soweit konventionelle, unidirektionale Geräteverbindungen ohne Netze und Server verwendet werden.

Schon früh waren das DSK-Signal als CCIR-601 (heute ITU-R BT 601) und die zugehörige serielle Schnittstelle (Serial Digital Interface, SDI) nach ITU-R 656 definiert. Da auch der Verbindungsaufwand wieder auf eine einzige Leitung beschränkt ist, konnte sich das DSK-Signal auch aus praktischen Gründen auf breiter Front durchsetzen – erst im SD- und dann auch im HD-Bereich (ITU-R BT 709). Das Format ist, ggf. in unterabgetasteter Form, auch das Ausgangssignal für alle Formen datenreduzierter Signale. Als hochwertige Variante findet sich im Produktionsbereich manchmal auch ein digitales RGB-Format mit zusätzlichem, den Transparenzgrad definierenden Alphakanal mit dem Abtastratenverhältnis 4:4:4:4.

Auch wenn die Leitungslängen bei der Verwendung von HD-SDI-Schnittstellen geringer sind (< 100 m) als bei SDI (< 300 m), so stellen sie und die damit verbundene DSK-Signalform einen robusten, bewährten und weit verbreiteten Übertragungsweg dar, der quasi latenzfrei ist.

Im HD-Bereich sind drei Signalformen (720p/50, 1080i/25 und 1080p/25) definiert, für die die HD-SDI-Schnittstelle mit ihren 1,485 Gbit/s ausreicht (s. Kap. 3). Bei höheren Ansprüchen (1080p/50) wird eine 3G-Schnittstelle mit verdoppelter Datenrate benutzt, oder es werden per Dual Link zwei Leitungen parallel betrieben. In der Festlegung SMPTE 425 wird der erste Fall als Level A, direct mapping bezeichnet, während der zweite Fall, Level B, in DL (Dual Link mapping nach SMPTE 372) oder DS (Dual stream mapping, 2 x HD-SDI) unterschieden wird. Die 3G-Infrastruktur wird heute schon relativ oft verbaut, sie hat sich jedoch noch nicht auf breiter Front als durchgängiger Produktionsstandard etabliert.

Bezüglich der Signalaussteuerung sind bei SDI Head- und Feetroom vorgesehen, die Übersteuerungen auffangen sollen. Die technischen Produktionsrichtlinien der deutschen Rundfunkanstalten legen diesbezüglich fest, dass dieser Bereich nur für sehr kurzfristige Übersteuerungen reserviert ist und nicht für gestalterische Zwecke benutzt werden darf, die z. B. das Ziel haben, den Dynamikbereich im Signal zu erweitern. Für die Überwachung und Kontrolle der Gültigkeit der Signale wird eine RGB-Darstellung auf einem Waveformmonitor als ausreichend angesehen [112].

Bezüglich der Farbwiedergabe muss beachtet werden, dass das digitale SD- einen anderen Farbraum aufweist als das digitale HD-Signal. Für die Bildung des Y-Signals gilt bei SD-Signalen: $E'_Y = 0{,}299\ E'_R + 0{,}587\ E'_G + 0{,}114\ E'_B$ und bei HD: $E'_Y = 0{,}212\ E'_R + 0{,}715\ E'_G + 0{,}072\ E'_B$. Bei Konvertierungen zwischen und SD und HD muss entsprechend auch auf die Umcodierung der Farbräume und die Auswahl der richtigen Zielmatrix geachtet werden.

Im Gegensatz zu analogen Verfahren ermöglicht die digitale Signalübertragung die Entscheidung darüber, ob die Daten in Echtzeit übertragen werden sollen oder nicht. Die Echtzeitübertragung ist vor allem bei der Produktion erforderlich und durch HD-SDI-Schnittstellen gewährleistet. Andere Vorgänge, wie z. B. die Überspielung von Beiträgen oder der Datenaustausch mit Grafiksystemen, können dagegen in »Nicht-Echtzeit« vorgenommen werden, d. h. sowohl schneller als auch langsamer, sodass z. B. für die Übertragung auch ein Computernetzwerk benutzt werden kann und zudem datenreduzierte Signale im Produktionsbereich Verwendung finden können.

10.1.2.1 Datenreduzierte Signale im Studio

Aufgrund der Nachfrage nach möglichst kostengünstigen Produktionsmitteln wird angestrebt, vermehrt mit datenreduzierten Signalen und Netzwerken zu arbeiten (s. Abschn. 10.3). Datenreduzierte Signale verringern den Aufwand bei der Zuspielung und Archivierung, gleichzeitig sind sie aber mit Qualitätsminderung behaftet. Auch wenn diese zunächst unsichtbar ist, sollte sie möglichst gering sein, um nicht zu hohe Generationsverluste zu erzeugen. Die Anforderungen sind abhängig von den Produktionsbereichen. So steht in einigen Fällen die Forderung nach Schnittfähigkeit im Vordergrund, in anderen die Bildqualität oder eine möglichst geringe Datenrate. Daraus resultieren verschiedene Grade der verwendeten Datenreduktion. Während im Bereich Grafik und Produktion oft ganz darauf verzichtet wird, kann sie bei Zuspielung und vor allem bei der Ausstrahlung erhebliche Ausmaße annehmen. Im Produktionsbereich finden sich Systeme mit so geringer Datenreduktion, dass fast nur Redundanzinformation entfernt wird (z. B. das MAZ-Format Digital Betacam mit der Reduktion 2:1). Etwas stärker reduzierend arbeiten Server und Editingsysteme, mit Reduktionsfaktoren zwischen 3 und 5, z. B. mit IMX oder DV-basiert. Im Bereich der HD-Signale ging der Trend hin zu einer Datenrate, die auf dem Niveau von SD liegt, wie vor allem bei XDCAM HD422, das mit 50 Mbit/s und zeitabhängigen Bildern (long GOP) arbeitet. Neuerdings werden aber auch bei HD wieder rein intraframe-basierte Codecs, wie AVC-I mit mindestens 100 Mbit/s, favorisiert. Abbildung 10.1 zeigt eine Übersicht. [104].

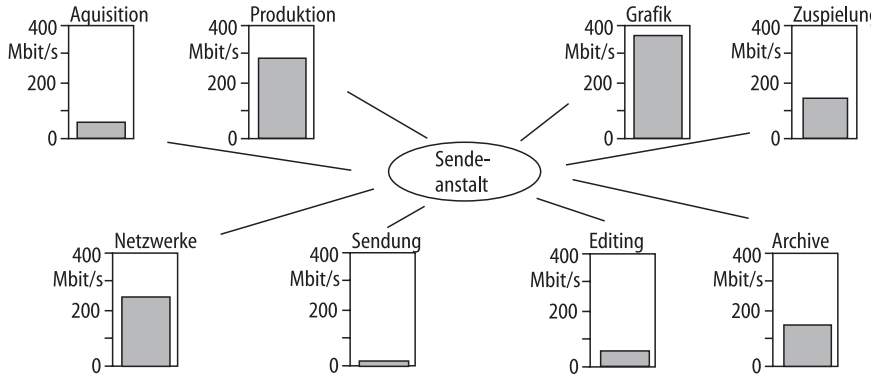


Abb. 10.1. Datenraten für verschiedene Anwendungen

Tabelle 10.2. Professionell genutzte Signalformate im Videobereich

Verfahren	Nutzung	Format	Datenrate	Merkmal
SDI	Studio/Live	4:2:2	270 Mbit/s	unreduziert
HD-SDI	Studio/Live	4:2:2	1,485 Gbit/s	unreduziert
Digital Betacam	Bandaufzeichnung	4:2:2	108 Mbit/s	SD, intraframe
DVCPro/50	Banda./Übertrag.	4:1:1/4:2:2	25/50 Mbit/s	SD, intraframe
MPEG-2-422@ML	Aufz./Editing	4:2:2	bis 50 Mbit/s	SD, I-frame only
DVCProHD	Aufz./Editing	3:1,5:1,5	100 Mbit/s	HD, intraframe
HDCAM	Aufz./Editing	3:1:1	118 Mbit/s	HD, I-frame only
MPEG-2-422@HL	Aufz./Ed.	4:2:2	bis 50 Mbit/s	HD, I-, B-, P-Sequ.
AVC-I 100	Aufz./Editing	4:2:2	112 Mbit/s	HD, I-frame only
MPEG-4 SStP	HDCAM SR	4:2:2/4:4:4	440/880 Mbit/s	HD, I-frame only
DNxHD	Intermedaite	4:2:2	185 Mbit/s	HD, I-frame only
ProRes	Intermediate	4:2:2/4:4:4	bis 600 Mbit/s	HD, bis 4k
XAVC	Aufz./Editing	4:2:2/4:4:4	bis 600 Mbit/s	HD, bis 4k

Die Verwendung von Datenreduktion erlaubt für die Fernsehübertragung die Wiedergabe mit sehr guter Qualität und wenig Artefakten, der Einsatz im Produktionsbereich muss jedoch differenziert betrachtet werden. Für alle Anwendungsbereiche stehen hier viele Codecs zur Verfügung. Tabelle 10.2 zeigt eine Übersicht, darin dominieren die Codecs, die I-frame-only arbeiten. Tabelle 10.3 zeigt die Gegenüberstellung der Eigenschaften der Intraframe-Codierung (wie bei IMX oder AVC-I) im Vergleich zu einer Hybrid-Codierung mit Bewegungskompensation (wie bei bestimmten MPEG-2-Varianten).

Die Auswahl und die generelle Festlegung des Signalformats im Produktionsbereich hat große Bedeutung, denn die gewählte Signalform soll möglichst weitgehend beibehalten werden. Eine Kaskadierung verschiedener Reduktionsverfahren vermindert ebenso die Qualität wie ein Übergang zum DSK-Signal oder zurück, da jeweils erneut decodiert und codiert werden muss.

Im Bereich der SD-Produktion wurden die Festlegungen zu Zeiten getroffen, als die Produktion noch weitgehend bandgestützt arbeitete. Hier etablierten sich bei MAZ-Systemen, die mit Datenreduktion arbeiten, die Verfahren auf DV- und MPEG-2-Basis. Sie sind weitgehend standardisiert und es wurde für den SD-Bereich empfohlen, auch MAZ-unabhängig nur DV und MPEG-2, möglichst mit 50 Mbit/s zu verwenden. Entsprechend werden von den Rundfunkanstalten für SD-Produktionen die Formate DVCPro50 und IMX (MPEG-2, I-Frame only, 4:2:2, 50 Mbit/s) akzeptiert.

Für den Bereich der HD-Produktion legten die deutschen Rundfunkanstalten Ende 2011 fest, dass nur noch Signalformen erlaubt sind, die HD-Bilder mit 1920 x 1080 Pixels bei einer Abtastung von 4:2:2 unterstützen. Für Mainstream-Produktionen des Fernsehtagesgeschäfts sind hier die Codecs XDCAM-HD422 und AVC-I 100 geeignet. Ersterer arbeitet auf MPEG-2-Basis bei 4:2:2 mit zeitabhängigen Bildern (long GOP) und 50 Mbit/s, Letzterer auf H.-264-Basis bei 4:2:2 und 112 Mbit/s, wobei keine zeitliche Abhängigkeit zwischen den Bildern besteht. Geräte, die mit weiterem Subsampling und weniger als 50 Mbit/s arbeiten (z. B. 4:2:0 bei 35 Mbit/s) werden nur für besondere Ausnahmefälle vorgesehen.

Tabelle 10.3. Eigenschaften der Datenreduktionstypen

Codierung	Intraframe	Interframe mit Bewegungskompensation
Effizienz	gering	hoch
Laufzeitverzögerung	gering	groß
Bildqualität	kaum –	stark – bewegungsabhängig
Einzelbildzugriff	uneingeschränkt	eingeschränkt
Aufwand im Coder	gering	hoch
Aufwand im Decoder	gering	mittel
Codec-Symmetrie	symmetrisch	unsymmetrisch

Diese Codecs sollen sowohl bei der Akquisition als auch bei der Zuspielung zum Einsatz kommen und sollen weiterhin auch im Studio und bei der Verteilung über Netzwerke beibehalten werden.

Vor dem Hintergrund, dass sowohl SD-Produktionsmittel als auch noch sehr viele SD-Beiträge zur Verfügung stehen, die in HD-Programme integriert werden, wird angestrebt, dass ein HD-Programm über die gesamte Produktionskette hinweg in HD produziert und SD-Material nur zu einem geringen Teil enthalten sein soll. Für alle Konvertierungen zwischen SD und HD ist der Einsatz von hochqualitativen Konvertern dringend zu empfehlen, denn die Qualität von Schnittprogrammen oder gewöhnlicher Software-Applikationen reicht für diesen Zweck in der Regel nicht aus.

Als Datenträger für den Austausch sind die Professional Disc mit XDCAM HD422-Codec, oder P2-Karten für AVC-I 100 verfügbar und geeignet. Das gilt vor allem für Mainstream-Produktionen. Davon unterschieden werden TV-Produktion wie große Sportevents oder szenische Produktionen. Hier bestehen höhere Ansprüche an die Bildqualität, entsprechend wird vorausgesetzt, dass die Produktion durchgängig hochqualitativ durchgeführt wird. Das gilt für den Einsatz hochwertiger Kameras (z. B. Arri Alexa) ebenso wie für hochwertige Codecs und Datenträger. Hier kommen die Intermediate-Codecs ProRes oder DNxHD in Frage. Als Bandformat ist hier HDCAM SR geeignet. Künftig werden auch Premiumproduktionen in einem hochbitratigen Codec produziert und auf bandlosen Speichern ausgetauscht werden.

10.1.2.2 DV-Signale im Studio

Generell sind die Vor- und Nachteile des Einsatzes von DV-basierten Signalen viel schneller dargestellt als bei MPEG-2, da DV viel stärker festgelegt ist als das sehr offen konzipierte MPEG-2-System. Bei DV müssen weniger Eckpunkte beachtet werden. Das DV-Signal ist in Abschn. 3.6 genauer beschrieben. Es basiert auf einer Irrelevanzreduktion mit DCT. Der große Unterschied zu MPEG-2 ist, dass die zeitliche Dimension nicht zur Datenreduktion verwendet wird. Da das DV-Signal ursprünglich für den Einsatz in MAZ-Geräten konzipiert wurde, zeichnet es sich dadurch aus, dass auf aufwendige Weise eine konstante Datenrate erzielt wird [55]. DV-Recorder wurden für den Heimanwender-Bereich entwickelt und bei der Adaption für den Produktionsbereich von verschiedenen Herstellern unterschiedlich modifiziert. Hier existieren zwei DVCPro-Varianten und das DVCAM-System.

DVCAM ist stärker an das Ursprungsformat angelehnt. Das Signal wird mit einer 4:2:0-Unterabtastung aus dem DSK-Signal gewonnen und auf 25 Mbit/s reduziert und nur im semiprofessionellen Bereich benutzt.

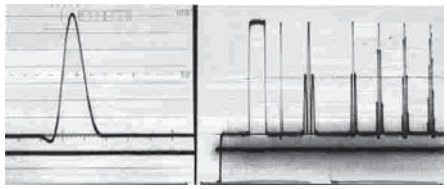
Beim DVCPro-System liegt das Signal liegt in einer Variante mit 25 Mbit/s (Reduktion 5:1) und mit 50 Mbit/s (Reduktion 3,3:1) vor. Letztere wird als DVCPro50 bezeichnet und wird von den Rundfunkanstalten akzeptiert. Die Signalverarbeitung basiert auf der 4:2:2-Abtastung mit einer Datenreduktion von 3,3:1. Für DVCPro HD wird die Rate noch einmal verdoppelt.

Das DV-Verfahren bietet eine hohe Signalqualität. Verluste können mit klassischen messtechnischen Mitteln kaum nachgewiesen werden. So zeigt sich auch die mehrfache Codierung und Decodierung, wie sie beim Kopiervorgang über die SDI-Schnittstelle auftritt, anhand eines 2T-Impulses in Abbildung 10.2 als messtechnisch unkritisch [231]. Das gilt ebenso bei der Kaskadierung von DVCPro und DVCAM über mehrere Generationen. Die bei den Kopien verwendeten Geräte waren Digital Betacam DVW 500, DVCPro AJ D 750 und DVCam DSR 85. Nur der Vergleich von DVCPro und Digital Betacam bezüglich mehrerer Kopiergenerationen über die analoge Schnittstelle zeigt ein schlechteres Verhalten von DVCPro, was aber eher auf die Wandler als auf das Reduktionsverfahren zurückzuführen ist.

10.1.2.3 MPEG-2-Signale im Studio

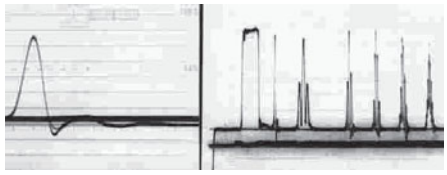
Auch die MPEG-Codierung ist in Kap. 3 vorgestellt worden. MPEG-2 bietet viele Varianten. Abbildung 10.3 zeigt noch einmal das System der Profiles und Levels. Darin zeigt sich, dass MPEG-2 die Codierung verschiedener Bildauflösungen von LDTV mit 352 H x 288 V über SDTV mit 720 H x 576 V bis hin zu HDTV mit 1920 H x 1152 V Bildpunkten unterstützt.

Die Abbildung zeigt auch die MPEG-2-Variante 422P@ML (Professional Profile), die speziell für den SD-Produktionsbereich wurde. Die Verwendung dieses Codecs, der ja eigentlich dezidiert für die Auslieferung an die Zuschauer konzipiert wurde, hat bei der Einführung im Produktionsbereich umfangreiche

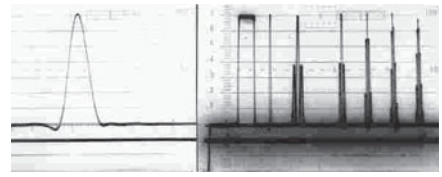


Digital Betacam 10. Kopie über analoge Komponenten

DCVPro 150. Kopie über digitale Komponenten, SDI



DVCPro 10. Kopie über analoge Komponenten



DCVPro + DVCam 50. Kopie über digitale Komponenten

Abb. 10.2. Vergleich verschiedener Formate bei Kopien über mehrere Generationen [231]

Profile Level	Simple Profile 4:2:0 (no B-Frames)	Main Profile 4:2:0	Professional Profile 4:2:2	Scalable Profile 4:2:0	High Profile 4:2:0 or 4:2:2
High Level < 60 fps		1920 x 1152 < 80 Mbit/s			1920 x 1152 < 100(80) Mbit/s
High 1440 < 60 fps		1440 x 1152 < 60 Mbit/s		1440 x 1152 (Spat.) < 60 (40) Mbit/s	1440 x 1152 < 80 (60) Mbit/s
Main Level < 30 fps	720 x 576 < 15 Mbit/s	720 x 576 < 15 Mbit/s	720 x 576 < 50 Mbit/s	720 x 576 (SNR) < 15 (10) Mbit/s	720 x 576 < 20 (15) Mbit/s
Low Level < 30 fps		352 x 288 < 4 Mbit/s		352 x 288 (SNR) < 4 (3) Mbit/s	

Werte in Klammern geben die Typen (Spatially oder SNR) und Alternativen bezgl. der Skalierung an

Abb. 10.3. Profiles und Levels bei MPEG-2

Diskussionen und Tests hervorgerufen, die hier erläutert werden. MPEG-2 ist im Professional Profile ist für Standardproduktionen mit 4:2:2-Abtastung und hohen Datenraten konzipiert, denn wie bei DV hat sich gezeigt, dass die Chromaunterabtastung (4:2:0 oder 4:1:1) zu hohe Qualitätsverluste mit sich bringt. Für den Akquisitionsbereich standen mit Betacam SX und dem IMX-Recoder zwei MAZ-Systeme für diese Variante zur Verfügung. Daneben auch die optical Disc (XDCAM) sowie News- und Archivsysteme, sodass die Signalform durchgängig bis zur Ausstrahlung beibehalten werden kann.

Beim Professional Profile 422P@ML ist die Bildauflösung festgelegt (720 H x 576 V), jedoch noch keine Entscheidung über die verwendete Datenrate und vor allem über die verwendete Struktur der Group of Pictures gefallen. MPEG-2 erzielt ja seine hohe Codiereffizienz aufgrund der Ausnutzung zeitlicher Abhängigkeiten zwischen den Nachbarbildern. So gibt es unabhängige, intraframe-codierte Bilder (I), sowie ein- und zweidimensional prädizierte Bilder (P, B).

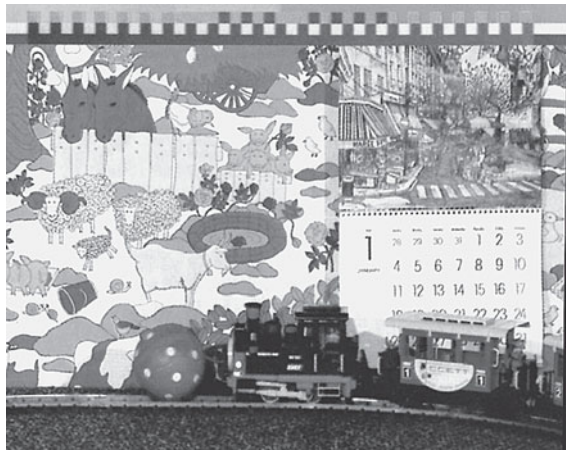


Abb. 10.4. Testbild aus der Bildsequenz Mobile + Calendar

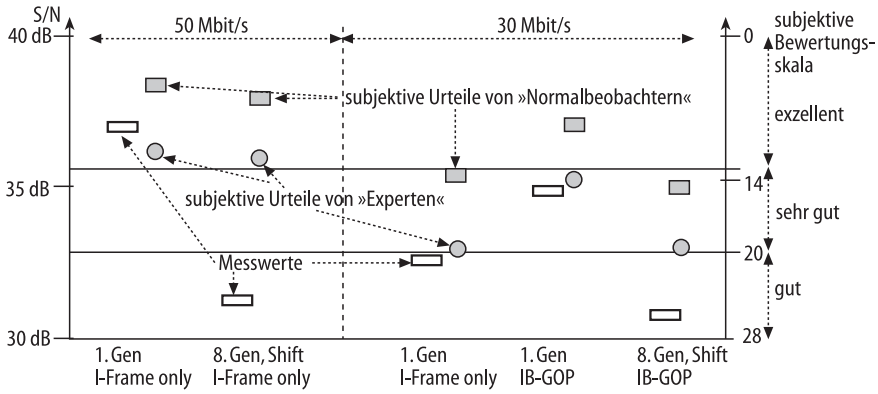


Abb. 10.5. Qualitätsbeurteilung unterschiedlicher Codierverfahren beim 422P von MPEG-2

Bei einer Datenrate zwischen 30 Mbit/s und 50 Mbit/s ist bei sehr guter SD-Bildqualität eine reine Intraframe-Codierung möglich. Abbildung 10.5 zeigt, dass dann bei subjektiver Beurteilung die Bildqualität das MPEG-2-Signal im Professional Profile als exzellent eingestuft wird, d. h. dass Unterschiede zum Original nicht sichtbar sind [83]. Dabei wurde die schwierig zu codierende Testsequenz Mobile + Calendar (M + C, Abb. 10.4) zugrunde gelegt und zudem eine achtfache Codierung mit zwei zeitlichen Verschiebungen durchgeführt, sodass sich der Codiertyp des Einzelbildes zeitlich ändert. Der Wert 12,5 auf der Qualitätsskala entspricht der Grenze der Sichtbarkeit von Fehlern. Kleinere Werte bedeuten exzellente Qualität, bis hin zu 28 kann die Bildqualität als gut bezeichnet werden. Wenn die Fehler nicht sichtbar sind, spricht man auch von visueller Transparenz der Codierung.

Die Bilder sind bei I-frame-only-Codierung auch nicht mehr zeitlich abhängig, der Zugriff auf jedes Bild wird möglich und der größte Teil weiterer Diskussionen um Probleme beim Einsatz von MPEG-2 im Produktionsbereich wird hinfällig.

Die Diskussion muss nur geführt werden, wenn stärkere Datenreduktion erwünscht ist. Diese geht mit der Einführung zeitlich abhängiger Bilder einher und wird z. B. bei XDCAM HD422 genutzt. Das Problem ist dann, dass im Zuge der Signalverarbeitung ein Bild über mehrere Generationen hinweg in unterschiedlicher Form (I, B oder P) codiert sein kann, worunter die Qualität leidet. Außerdem kann nicht ohne weiteres zwischen den Bildern einer GOP geschnitten werden, da dann die Abhängigkeiten der Bewegungsvektoren, die sich auf dann veränderte Nachbarbilder beziehen, nicht mehr stimmen.

Eine einfache Qualitätsbeurteilung datenreduzierter Signale wird durch die Bestimmung des »Signal/Rauschverhältnisses« ermöglicht, das über

$$S/N = 10 \log Q^2/MSE$$

aus der maximalen Anzahl der Quantisierungsstufen Q und dem mittleren quadratischen Fehler MSE gebildet wird [83]. Fehler bedeutet hier die Abweichung aller Bildpunkte vom Original. Dieser Wert ist einfach zu ermitteln, kor-

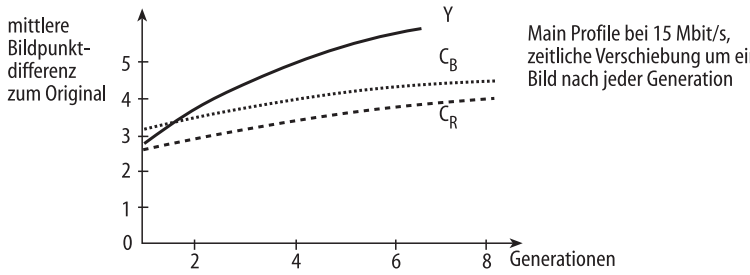
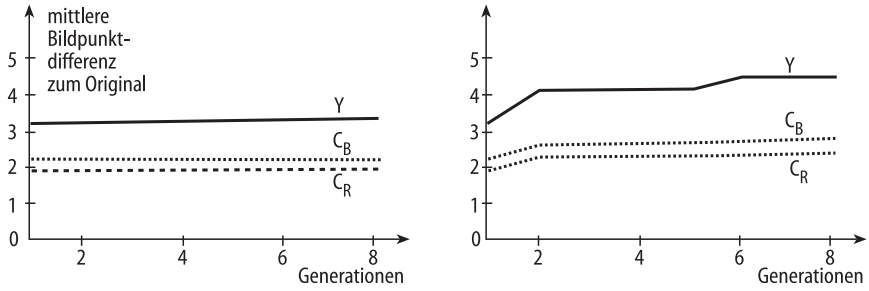


Abb. 10.6. Multigenerationsverhalten beim Main Profile von MPEG-2

reliert aber nur zu ca. 80% mit dem visuellen Eindruck. Abbildung 10.5 zeigt den Zusammenhang von S/N-Werten verschiedener MPEG-SD-Codierungen mit subjektiven Urteilen. Die Übereinstimmung zwischen MSE und visuellem Eindruck ist in der ersten Generation gut, in der achten Generation ist der visuelle Eindruck wesentlich besser als es der S/N-Wert suggeriert. Trotzdem kann der S/N-Wert zum Vergleich verschiedener Datenreduktionsverfahren benutzt werden, wenn immer dieselbe Bildsequenz verwendet wird.

Das dargestellte Generationsverhalten beruht auf einer fortwährenden Decodierung und Codierung des Datenstroms. Diese ist z. B. erforderlich, wenn mit Hilfe der Umsetzung in DSK-Signale eine Bildbearbeitung durch Videomischer, DVE etc. vorgenommen werden soll oder eine MAZ-Maschine mit SDI eingesetzt wird. Zum Vergleich mit dem Professional Profile zeigt Abbildung 10.6 am Beispiel der Testsequenz M + C anhand der entstehenden Bildpunktabweichungen, zunächst das schlechte Multigenerationsverhalten des Main Profile (4:2:0) mit IBBP-Bildfolgen bei der dort maximal zulässigen Datenrate von 15 Mbit/s. Dabei wurde nach jeder Generation eine zeitliche Verschiebung um ein Bild vorgenommen. Es wird deutlich, dass die Anzahl der Bildpunktabweichungen in Abhängigkeit von den Generationen der Codierung und Decodierung stetig steigt. In Abbildung 10.7 wird am gleichen Beispiel deutlich, dass das Multigenerationsverhalten beim Einsatz des SD-422-Profile mit z. B. 30 Mbit/s viel unkritischer ist, selbst wenn nach jeder Generation die Aufzeichnung auf



422-Profile mit I-Frame-only bei 30 Mbit/s, kaskadiert mit einem Videorecorder mit 2:1-Datenreduktion, Testsequenz Mobile+Calendar

422-Profile mit I-Frame-only bei 30 Mbit/s, und Einfluss eines DVE bei örtlicher Verschiebung um 2 Pixel horizontal und vertikal nach der 1. und 5. Generation

Abb. 10.7. Multigenerationsverhalten beim 422-Profile von MPEG-2

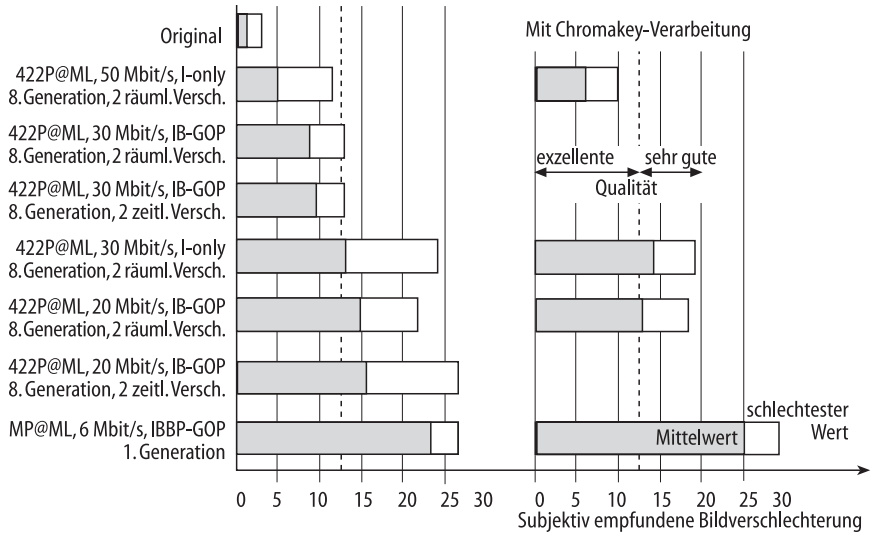


Abb. 10.8. Subjektive Bewertung verschiedener Codierungen

eine digitale MAZ-Maschine zwischengeschaltet wird, die ihrerseits mit 2:1-Datenreduktion arbeitet. Außerdem wird deutlich, dass I-only-Bildfolgen mit 30 Mbit/s auch dann noch eine sehr hohe SD-Bildqualität bieten, wenn eine Bildbearbeitung mittels DVE vorgenommen wird [179].

Die subjektive Beurteilung des MPEG-SD-Signals mit 422-Profile entsprechend der in Abbildung 10.5 dargestellten Skala ist in Abbildung 10.8 dargestellt. Sie zeigt den Vergleich verschiedener Codiertypen nach der 8. Generation der Codierung und Decodierung, wieder anhand der kritischen Sequenz Mobile + Calendar. Es liegen die in der Abbildung genannten Codierungspa-

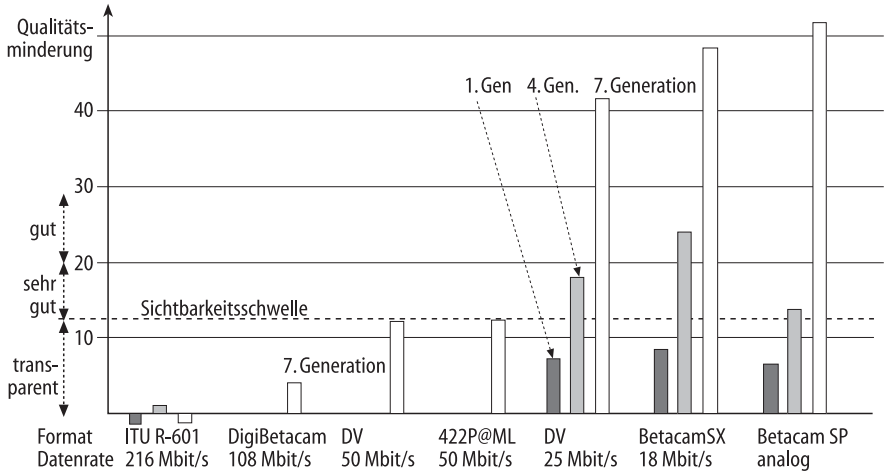


Abb. 10.9. Vergleich der subjektiven Bildqualität verschiedener Aufzeichnungsverfahren

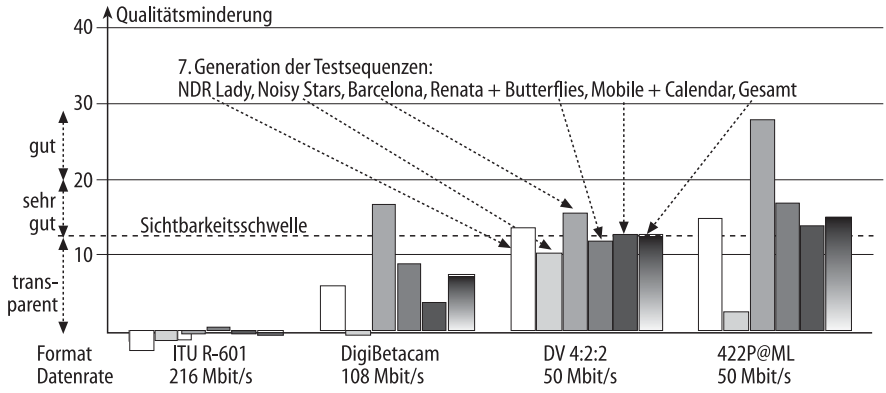


Abb. 10.10. Vergleich der Qualität der Kompressionsverfahren bei diversen Testsequenzen

rameter vor, inklusive zweier Bildmanipulationen, nämlich einer horizontalen und vertikalen Verschiebung um zwei Bildpunkte nach rechts unten zwischen der 1. und 2. Generation und nach links oben zwischen der fünften und der sechsten. Exzellente Qualität wird auch in diesem Fall noch mit I-Frame only bei 50 Mbit/s oder mit I/B-Frame-Codierung bei 20 ... 30 Mbit/s erreicht, obwohl die Bildverschiebung dazu führt, dass die für die Diskrete Cosinus Transformation gebildeten Blöcke innerhalb der Bilder nach der Bildmanipulation unterschiedliche Inhalte haben. Abbildung 10.8 zeigt unter den gleichen Umständen schließlich noch die subjektive Beurteilung einer Chroma-Key-Nachbearbeitung des MPEG-2-Signals im 4:2:2-Professional Profile [179].

Interessant ist auch der Vergleich der Qualität des Professional Profile mit anderen Datenreduktionsverfahren (Abb. 10.9), wie z. B. den bei DV oder Digital Betacam eingesetzten Algorithmen. Dieser Vergleich ist anhand der Bandformate DVCpro, DVCPro50, Digital Betacam und Betacam SX dargestellt [85]. Auch hier gilt, dass die subjektive Bewertung entscheidend vom Bildinhalt abhängt. So liegen zwar die Ergebnisse der 4. Generation sowohl für DV25 als auch für Betacam SX über der Sichtbarkeitsschwelle, es muss jedoch bedacht werden, dass sich die dargestellten Werte nur auf die Sequenz Mobile + Calendar beziehen. Beim Vergleich mit fünf weiteren Bildvorlagen zeigt sich, dass sowohl bei DV25 als auch bei Betacam SX nur zwei von sechs Werten oberhalb der Sichtbarkeitsschwelle liegen. Bei der Betrachtung unterschiedlicher Bildsequenzen bei 50 Mbit/s (Abb. 10.10) lässt sich feststellen, dass die Bildqualität bei Verwendung des DV-Algorithmus weniger vom Bildinhalt abhängt, als bei der Codierung nach MPEG-2 und auch insgesamt etwas besser abschneidet.

10.1.2.4 MPEG-2-Long GOP

Die hier dargestellten Tests rechtfertigen bei Verwendung des Professional Profile insgesamt den Einsatz von MPEG-2 im SD-Produktionsbereich. Entsprechend ist der darauf basierende IMX-Codec von den Rundfunkanstalten für den Produktionsbereich empfohlen. IMX arbeitet mit SD-Signalen ohne zeitlich abhängige Bilder bei 50 Mbit/s. Im Zuge der Verbreitung von HD-Signa-

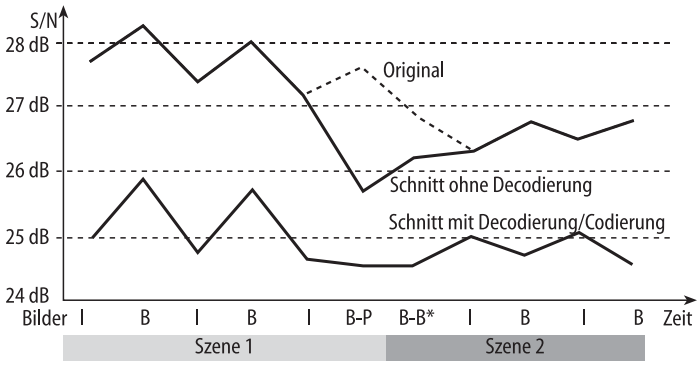


Abb. 10.11. Signal-Rauschverhältnis beim Schnitt mit und ohne Decodierung

len etablierte Sony dann das XDCAM-HD 422-Verfahren als Codec für 8-Bit-HD-Signale mit 1920 x 1080 Bildpunkten und 4:2:2-Abtastung. Trotz des ca. 5-fachen Informationsgehalts gegenüber SD wird hier weiter die Datenrate von 50 Mbit/s verwendet, was den Vorteil hat, dass ein Großteil der Speicherinfrastruktur, die sich bei IMX entwickelte, unverändert beibehalten werden kann. Der Nachteil ist, dass die Datenmenge entsprechend stärker reduziert werden muss. Das gelingt nur durch die Verwendung zeitabhängiger Bilder und umfangreicher Group of Pictures (Long GOP), was wiederum spezielle Probleme, vor allem beim Editing der Signale und der Zeitverzögerung bei der Signalverarbeitung mit sich bringt.

Betrachtet man beim Editing zunächst nur den eigentlichen Schnitt, das direkte Aneinanderfügen von Sequenzen, so spricht man auch von Splicing, das nicht nur hier, sondern generell beim Umschalten zwischen Videosignalen oder der Zusammenstellung vorproduzierter Beiträge bei Signalen mit 4:2:2- (Professional Profile) oder auch 4:2:0-Abtastung (DVB) auftritt. Hier kann eine hohe Signalqualität beibehalten werden, indem der Datenstrom auf der Ebene der datenreduzierten Signale geschaltet wird. Da dies nicht ohne Störungen an jeder Stelle erlaubt ist, ist es vorteilhaft, die erlaubten Punkte, durch Splicepoints zu kennzeichnen. Am Splicepoint bestehen keine Abhängigkeiten zu den Nachbarbildern, d. h. ein Beitrag darf nur mit einem I-Frame beginnen und nur mit einem I- oder P-Frame enden. Dass dabei der gewünschte Ein- oder Ausstiegspunkt des Beitrags ggf. um einige Bilder versetzt werden muss, ist bei abgeschlossenen Beiträgen unkritisch. Kritisch ist es hingegen, den Füllgrad des Puffers unter Kontrolle zu halten [89]. Hierzu kann mit dem sogenannten Seamless Splicing (nahtlose Verbindung) der Füllstand am Splicepoint auf einen festen Wert angeglichen und dafür ggf. das Bild gröber quantisiert werden. Beim non-seamless Splicing kann es dagegen zu einem Unter- oder Überlauf des Puffers kommen.

Ein Schnitt, der nur an Splice-Points erlaubt ist, ist generell nicht akzeptabel. Um dies bei einer IB-Codierung zu umgehen, kann durch Umwandlung einzelner Bilder auch recht einfach ein bildgenauer Schnitt erreicht werden. Der Szeneneinstieg am B-Frame erfordert dabei eine Umcodierung in ein nur

rückwärts prädiziertes sogenanntes B*-Bild. Bei diesem Schnittverfahren bleibt das Signal fast ausschließlich in der codierten Ebene und behält eine bessere Qualität als bei einem Schnitt in der decodierten Ebene. Abbildung 10.11 zeigt den Vergleich zu einem Schnittvorgang, für den eine Decodierung und erneute Codierung vorgenommen wurde [83]. Im ungünstigsten Fall tritt beim Schnitt ohne Decodierung eine Verschlechterung nur über zwei Bilder auf Ein ähnliches Verfahren lässt sich auch bei längeren GOP anwenden, wobei die Sequenzen, die qualitativ schlechter sind, etwas länger werden. Zum Ausgleich des Verlusts erlaubt MPEG-2 jedoch die Anhebung der Datenrate.

Sollen MPEG-2-Signale allgemein editiert werden, ist die Decodierung und Recodierung unumgänglich. Hier lässt sich ein Qualitätsverlust dadurch minimieren, dass man die Datenströme synchronisiert, d. h. dafür sorgt, dass die Codierreihenfolge erhalten bleibt und die einzelnen Bilder ihren Codiertyp (I, B, P) nicht ändern. Durch die Decodierung werden MPEG-2-Signale in ein DSK-Signal gewandelt, damit sie mit Studiogeräten bearbeitet werden können, die für unkomprimierte Verarbeitung ausgelegt sind. Wenn bei der Neucodierung nach der Bearbeitung die ursprüngliche Codierreihenfolge wieder erscheinen soll, so müssen Zusatzinformationen über die GOP-Struktur, Bewegungsvektoren, Makroblockstrukturen und Codiertabellen auch im DSK-Datenstrom übertragen werden.

Als weiteres Problem bei Long-GOP-Verfahren gibt es die Zeitverzögerung, die umso stärker auftritt, je komplexer die Signalstruktur wird. Auch beim 4:2:2-Profil gilt, dass bei der Verwendung zeitlich abhängiger Bilder die ursprüngliche Reihenfolge nicht gleich der Codierreihenfolge ist, denn die bidirektionalen B-Frames hängen von den vorhergehenden und nachfolgenden Bildern ab, sodass diese Bezugs-I- und P-Frames vor den B-Frames dem Decoder bekannt sein müssen und entsprechend übertragen werden [83]. Der Decoder muss die I- und P-Bilder zunächst in einen Pufferspeicher übernehmen und kann anschließend ein dazwischen liegendes B-Frame berechnen. Dieser Vorgang erfordert eine gewisse Zeit, sodass die MPEG-2-Decodierung mit einer relativ großen Zeitverzögerung einhergeht, die mit der Länge der GOP steigt (Abb. 10.12).

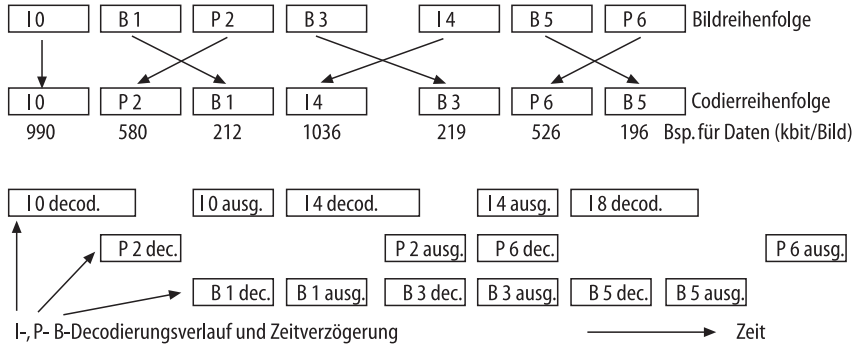


Abb. 10.12. Zeitliche Bezüge der I-, P- und B-Frames bei MPEG-2

Beim Einsatz von MPEG-Codern können Signalverzögerungen bis in den Sekundenbereich hinein auftreten, was insbesondere bei Live-Interviews über größere Distanzen sehr störend wirken kann. So ist z. B. manchmal bei Übertragungen mit Hilfe von SNG-Fahrzeugen (s. Abschn. 10.5) ein zugeschalteter Moderator zu sehen, der noch auf sein per Satellit und MPEG-Codierung übertragenes Rückbild wartet. Gründe für die Verzögerungen liegen einerseits in der GOP-Struktur selbst, anderseits in Maßnahmen zur Verbesserung der Bildqualität, wie Rauschfilterung vor der Codierung. Ein weiterer Grund ist der Ausgleich der Unterschiede zwischen den Quantisierungsstufen der verschiedenen Makroblöcke bzw. der I- und zugehörigen P-Bilder. Um Laufzeiten gering zu halten, muss ggf. weniger komplex und damit weniger optimal, also mit höheren Datenraten codiert werden.

10.1.2.5 H.264-Formate im Produktionsbereich

Angesichts der dargestellten Probleme bei Codecs mit Long GOP werden moderne Intraframe-Codecs interessant, die effektiver arbeiten als DV oder MPEG-2 im I-frame-only-Modus. Intraframe-Codecs bieten besonders dann Vorteile, wenn geringe Latenzzeiten von einem zum anderen Ende der Übertragungsstrecke (End to End, E2E) gefordert sind und wenn Programme nachbearbeitet werden sollen. Ihr einziger Nachteil gegenüber Long GOP-Verfahren ist die höhere Datenrate, doch sinkt dessen Bedeutung mit der stetig steigenden Leistungsfähigkeit der Netze und Speicher. Generell sind intracodierte Datenreduktionsverfahren sehr gut für den professionellen Bereich geeignet.

Neben MPEG-2-Intra kommen als Intraframe-Codecs JPEG2000, VC-2, H.264/AVC-I und XAVC in Frage, die in Kapitel 3 dargestellt sind. JPEG2000 und VC2 arbeiten effektiv auf Basis der Wavelet-Transformation, VC2 wird jedoch nur sehr selten implementiert und für JPEG2000 ist keine konsistente und interoperable Definition für den Transportstrom definiert. Daher setzten sich unter den Bezeichnungen AVC-I und XAVC vor allem Intraframe-Codecs durch, die auf H.264/AVC basieren. Unter der Bezeichnung AVCHD gibt es auch HD-taugliche Long-GOP-Varianten, aber im Gegensatz zur Situation bei MPEG-2 werden sie nicht im professionellen Bereich verwendet.

AVC-I arbeitet effektiver als MPEG-2-Intra und ähnlich leistungsfähig wie JPEG2000 ist jedoch als Untermenge des allgemeinen H.264/AVC-Standards in der Lage, mit allen gängigen Systemen zu interagieren, einschließlich der verbreiterten Containerformate MPEG-2-Transportstrom oder MXF-Files (s. nächster Abschnitt). Tests haben ergeben, dass AVC-I kritisch auf verrauschte Signale, z. B. das Filmkorn, reagiert, dass es bei genügend hoher Datenrate

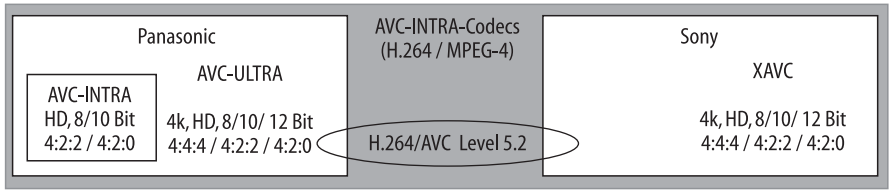


Abb. 10.13. Professionelle Intraframe-Codecs auf Basis von H.264/AVC

Tabelle 10.4. Videoqualität der AVC-I-Codierung von HD-Material mit 1080i/25 und 720p/50

AVC-I-Bitrate	Bildqualität
< 50 Mbit/s	zu gering für Produktionsanwendungen
50 – 75 Mbit/s	akzeptabel bei Quellen mit geringem Rauschen
75 – 90 Mbit/s	akzeptabel
90 – 110 Mbit/s	gut
110 – 150 Mbit/s	ausgezeichnet
> 150 Mbit/s	visuell verlustlos

aber gut für den HD-Produktionsbereich geeignet ist (Tabelle 10.4) [137]. Es hat eine geringe E2E-Verzögerung. Für den HD-Produktionsbereich bietet sich vor allem das High-422-Intra-Profil mit 10 Bit an, das bei der Datenrate 110 Mbit/s auch von den deutschen Rundfunkanstalten akzeptiert ist. Für höhere Anforderungen unterstützt AVC-I auch Daten bis 4k-Auflösung und eine 4:4:4-Abtastung mit Abtastwerten, die mit bis zu 12 Bit codiert sind.

AVC-I ist eine Bezeichnung und eine Codec-Familie, die bereits früh von der Firma Panasonic propagiert und als AVC-Ultra für hohe Ansprüche weiterentwickelt wurde. Später entwickelte die Firma Sony unter der Bezeichnung XAVC einen ähnlichen, H.264-basierenden Codec, der von vorn herein für höchste professionelle Ansprüche und 4k-Daten ausgelegt war. XAVC arbeitet bei 4k maximal mit RGB-Daten (4:4:4) bei bis zu 12 Bit pro Abtastwert, unterstützt aber auch 4:2:2-Daten mit 10 oder 8 Bit und 2K- oder HD-Auflösung. Als maximale Datenrate wird bei der Einführung des Codecs 600 Mbit/s angegeben. Da XAVC und AVC-I beide den Level 5.2 von H.264 unterstützen, soll es möglich sein, dass ein XAVC-Decoder auch AVC-I-Daten wiedergeben kann (Abb. 10.13).

10.1.3 Fileformate

Im vorigen Abschnitt wurde bereits deutlich wie wichtig die Beschränkung auf wenige Datenreduktionsverfahren ist. Trotzdem ist in der Praxis heute eine Vielzahl von Codecs zu finden. Diese können jedoch in ein einheitliches Fileformat gebracht werden das als Codec-unabhängiger Datencontainer oder als sogenannten Wrapper dient. Bereits existierende Fileformate wie Apple QuickTime sind im PC-Umfeld entstanden und eher für den Heimanwenderbereich relevant. Sie konnten sich im professionellen Bereich nicht fest profilieren. Daher wurden einige Fileformate speziell für den professionellen Videodatenaustausch entwickelt, die die Arbeitsabläufe im Produktionsbereich wesentlich erleichtern. Als eine Voraussetzung sind zunächst universelle Möglichkeiten zur eindeutigen Identifizierung des Materials erforderlich.

10.1.3.1 Unique Material Identifier

In einer großen Netzwerkumgebung ist es unverzichtbar, dass Datenbestände, also Audio-Files, Videoströme, Bilder und Metadaten, unverwechselbar gekennzeichnet sind, um einen eindeutigen Zugriff zu ermöglichen, besonders dann,

Universelles Label	12 Byte
Länge	1 Byte
Instanz-Nummer	3 Byte
Material-Nummer	16 Byte

Wann: Datum, Zeit	8 Byte
Wo: Geogr. Lage	12 Byte
Wer: User-Info	12 Byte

Abb. 10.14. Basic UMID und Source Pack

wenn zu einer Einheit gehörende Daten auf verschiedenen Servern oder anderen Medienträgern verteilt gespeichert sind. Daher wurde von der Society of Motion Picture Engineers (SMPTE) der Unique Material Identifier, kurz UMID, definiert. Aus dieser Nummer lassen sich zahlreiche Informationen gewinnen, u. a. über den Autor und das Erstellungsdatum der Datei.

Der UMID ist ein 32 oder 64 Byte langes Codewort und kann für jede Art audiovisueller Daten Verwendung finden. Die ersten 32 Byte bezeichnet man als Basic UMID, sie reichen prinzipiell zur Identifizierung aus. Der Basic UMID (Abb. 10.14) teilt sich in ein universelles 12-Byte-Label, wobei die ersten 10 Byte ein registrierter ISO-Eintrag sind, welcher von der SMPTE verwaltet wird, es folgen 2 Byte, die einen Rückschluss auf die Art des Inhalts erlauben. Das 13. Byte gibt die Länge der restlichen Bytekette an, deren Wert im Basic UMID auf 19 festgelegt ist. Am Ende des Eintrags befindet sich ein 16-Byte-Wert, für dessen Zusammensetzung verschiedene Methoden möglich sind, die gemein haben, dass sich eine global eindeutige Zahl ergeben muss. Ein Beispiel wäre, diese Material-Nummer aus der MAC-Adresse (Media Access Control) des Gerätes zu kreieren, die für jede an einem Netzwerk teilnehmende Maschine eine eigene ist, zusammen mit einem Zeitstempel der Datei-Erstellung. Damit wäre der Dateninhalt bereits eindeutig bestimmt. Der UMID bietet aber noch mehr Funktionalität, weswegen zusätzlich eine 3 Byte lange Instanz-Nummer erzeugt wird. Anhand dieser lässt sich ersehen, ob das File neu angelegt wurde oder ob es sich um eine weiterbearbeitete oder exportierte Datei handelt.

Zusammen mit dem ebenfalls 32 Byte langen Source Pack (Abb. 10.13) ergibt sich aus dem Basic UMID der Extended (erweiterte) UMID. Er bietet die Möglichkeit, dem Code die Information hinzuzufügen, wann und an welchem Ort eine Videosequenz aufgezeichnet worden ist, und die Angabe über den dazugehörigen Urheber.

Die ersten 8 Byte des Source Packs geben das Datum und die Uhrzeit der Datenerzeugung an. Dies gilt auch für einzelne Samples oder Frames. Die 12 Byte lange Wo-Komponente liefert geografische Daten über den Entstehungsort. Die letzten 12 Byte teilen sich gleichmäßig in eine Identifikation des Landes, der Firma und des einzelnen Nutzers auf. Der Ländercode wird nach ISO 3166-1 bestimmt und eine alphanumerische Identifikation für eine Organisation ist bei der SMPTE zu erwerben.

Bei der Neubearbeitung eines Files gibt es zwei Möglichkeiten zur Erzeugung des UMID. Entweder wird der neu entstehenden Datei ein ganz neuer UMID zugewiesen oder aber es wird nur die Instanz-Nummer geändert. In letzterem Fall bleibt ein eindeutiger Bezug zu den ursprünglichen Daten erhalten. Dies ist besonders dann geeignet, wenn für verschiedene Anwendungen

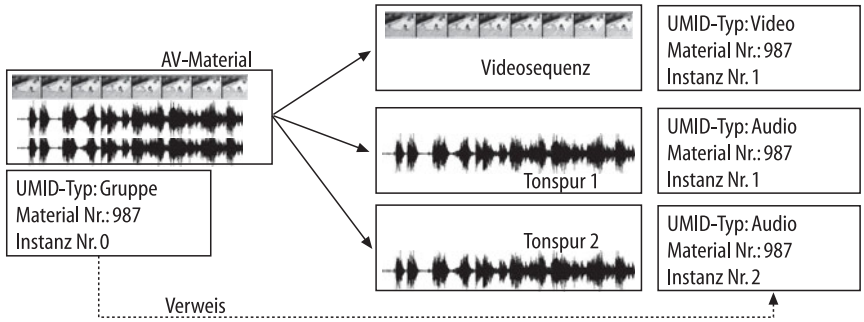


Abb. 10.15. Verweis auf die ursprüngliche Quelle bei UMID

unterschiedliche Formate erstellt werden müssen, z. B. einmal für eine DVD und einmal für die Distribution via DVB, der Verweis auf das Original besteht weiterhin. Ebenso können die Informationen über den Ursprung von Daten erhalten bleiben, wenn zwar der Basic UMID geändert wird, das Source Pack allerdings in seinem Zustand verbleibt. Jeder ursprüngliche Clip behält dann einen Verweis auf seine Quelle. Eine weitere denkbare Variante ist, allen Daten, die aus einem audiovisuellen Material extrahiert werden, die gleiche Material-Nummer zuzuordnen, sie aber durch den jeweiligen Typ (Audio oder Video) zu unterscheiden und unterschiedliche Daten gleichen Typs, wie beispielsweise mehrere Audiokanäle, durch differierende Instanznummern (Abb. 10.15).

10.1.3.2 Advanced-Authoring-Format

Vor über dreißig Jahren entwickelte die Firma CMX für ihre Schnittsysteme die Edit Decision List (EDL). Aus dieser Textdatei lässt sich erkennen, welche Dateien für einen Schnitt verwendet wurden, wo die In- und Out-Punkte gesetzt sind. Ebenso wird ersichtlich, welche speziellen Effekte Verwendung finden sollen. Die Vielfalt der aus einer EDL zu entnehmenden Daten ist beschränkt, aber der Grundgedanke, anwendungsunabhängig eine Schnittliste als Referenz erstellen zu lassen, ist bis heute geblieben.

Um die Standardisierung auch in diesem Bereich voran zu treiben wurde das Advanced-Authoring-Format entwickelt. Basierend auf dem Open Media Framework (OMF) der Firma Avid wurde damit ein offener Standard entwickelt, der die Möglichkeit bietet, digitale Daten eines nichtlinearen Editings, Multimedia Authorings oder Compositings plattformübergreifend und unabhängig von den Dateiformaten auszutauschen. Das AAF beschreibt alle seine verwendeten Inhalte zusammen mit deren Zusammenstellung und darüber hinaus Bild- und Tonmanipulationen, Spezialeffekte etc. Damit ist die Möglichkeit gegeben, mit unterschiedlichen Anwendungen gleichzeitig an demselben Inhalt zu arbeiten.

Die zentrale Funktion ist der Erhalt der Metadaten, die sonst häufig bei einem Datenaustausch auf der Strecke bleiben. Sie werden bei AAF besonders intensiv genutzt. Bei der Verwendung des AAF bleiben die ursprünglich verwendeten Medien-Files erhalten, es werden aber die Anweisungen über alle Schritte aufgezeichnet. Im Nachhinein ist also ein Zugriff auf die Rohdaten,

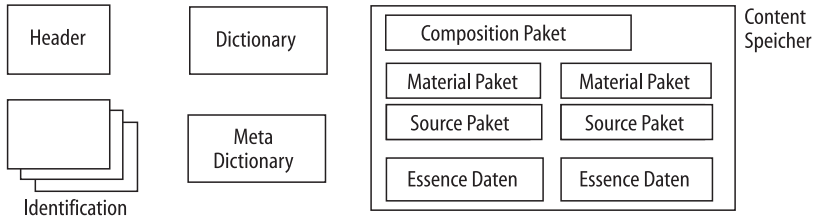


Abb. 10.16. Objekte eines AAF-Files

auch mit einer anderen, AAF unterstützenden Anwendung möglich. Ebenso können sämtliche Zwischenschritte nachvollzogen werden.

Sämtliche Bild- und Tondaten, die in irgendeinem Format vorliegen, werden in der AAF-Spezifikation als Essence bezeichnet. Die grundlegenden Objekte eines AAF-Files sind in Abbildung 10.16 aufgezeigt. Im Datei-Header finden sich für die gesamte Datei benötigte Metadaten, u. a. wird hier ein Stempel mit dem Zeitpunkt aufgeprägt, an dem die Datei erstellt bzw. zum letzten Mal modifiziert wurde [96]. Hier wird auch die AAF-Versionsnummer angegeben. Dem Header untergeordnet sind die Identifikationsobjekte, welche Verweise auf die Anwendungen enthalten, mit denen das File bereits bearbeitet worden ist. Die Dictionaries (Wörterbücher) (Abb. 10.16) erlauben die Erweiterbarkeit des AAF, hier sind sämtliche Definitionen abgelegt, die für die Datei benötigt werden, u. a. über Effekte und Überblendungen. Dem Content-Speicher sind sämtliche Objekte untergeordnet, die Essence beschreibende Daten oder die Essence selbst beinhalten.

Die beschreibenden Informationen (Metadaten) sind in diverse Pakete aufgeteilt, von denen jedes einzelne durch einen UMID identifiziert werden kann. Die Pakete geben Aufschluss über deren Art und beinhalten Befehle, wie sie kombiniert und verändert werden sollen: Das Physical-Source-Paket beschreibt das analoge Medium, welches zur Digitalisierung der jeweiligen Daten genutzt wurde. Unterschieden wird dabei je nach Quelle zwischen dem Film-Source-Paket und dem Tape-Source-Paket (Abb. 10.17).

Das hierarchisch am höchsten liegende Paket im Content-Speicher ist das Composition-Paket. Es vereint sämtliche audiovisuellen Informationen über den Schnitt und das Compositing der Daten in sich. Das Material-Paket legt schließlich die Verbindung zwischen der Composition und dem File-Source-Paket fest. Hier wird die Beschreibung über eventuell unterschiedliche Versionen von zu verwendender Essence hinterlegt sowie Effekte, die sich auf den gesamten Clip auswirken, wie Filter oder eine Farbkorrektur. Außerdem synchronisiert das Material-Paket Bild und Ton.

Die Composition wird aus verschiedenen Elementen, den Slots, zusammengesetzt. Es gibt drei Arten von Slots, den TimelineSlot, den StaticSlot und den EventSlot. Letzterer wird verwendet, um im Ablauf des Compositings Werte zu festgelegten Zeitpunkten zu steuern, um mit der Peripherie wie z. B. einer MIDI-Schnittstelle zu kommunizieren, um Interaktivität zu ermöglichen und zu einem bestimmten Zeitpunkt Kommentare bereitzuhalten. Ein StaticSlot beinhaltet Einzelbilder, während ein TimelineSlot eine Bild- oder Tonsequenz enthält, die sich kontinuierlich über die Zeit verändert. Die Dauer eines sol-

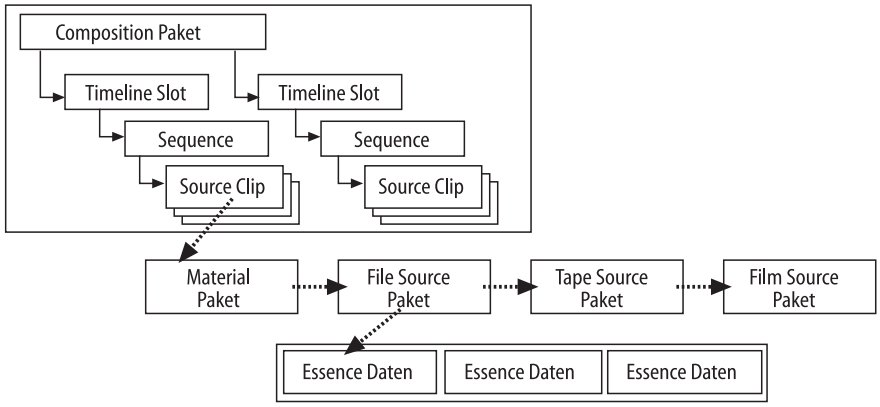


Abb. 10.17. Paketbildung bei AAF

chen Slots wird in Edit Units angegeben, die intern als Zeitreferenz dienen.

Die Source Clips oder Segmente werden in der angegebenen Länge und in der Reihenfolge angeordnet, wie sie in der Sequenz abgelegt sind. Einzufügende Überblendungen werden zwischen die Clips gesetzt. Die Enden der Clips werden dabei übereinander gelegt, sodass sie sich überlappen. Zur Veranschaulichung ist dieser Vorgang in Abbildung 10.18 dargestellt.

Damit die Sequenz bei einem Löschen der Überblendung nicht ungewollt länger wird, sollte die an der Datei arbeitende Anwendung die Clips entsprechend der Blendenlänge trimmen. Informationen über Clips, die nicht nacheinander, sondern mit Bild-in-Bild-Effekten oder einem Layering gleichzeitig abgespielt werden, sind in Unterobjekten der Sequenz abgelegt. Standardüberblendungen und häufig genutzte Bildeffekte wie ein Chroma Key oder Filter sind bereits im AAF implementiert, ebenso standardgemäße Audio-Manipulationen. Damit AAF alle ihm von einer Anwendung übergebenen Effekte speichern kann, besteht die Möglichkeit, weitere Effekte und auch Codecs über ein Plug-In-Modell zur Verfügung zu stellen. Der neue Effekt bekommt einen Namen und eine Beschreibung des erforderlichen Plug Ins [96].

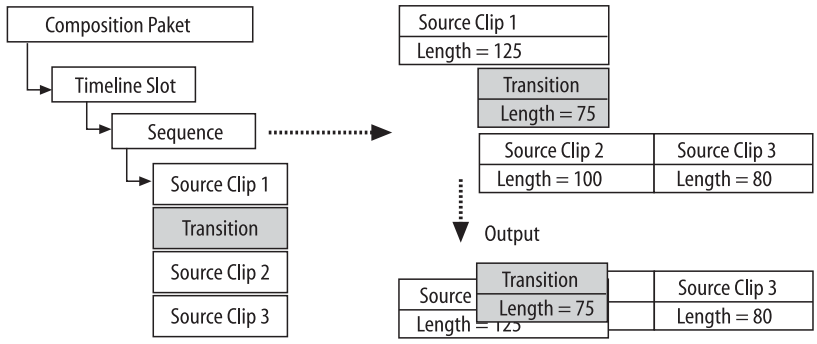


Abb. 10.18. Beispiel einer Clipanordnung unter AAF

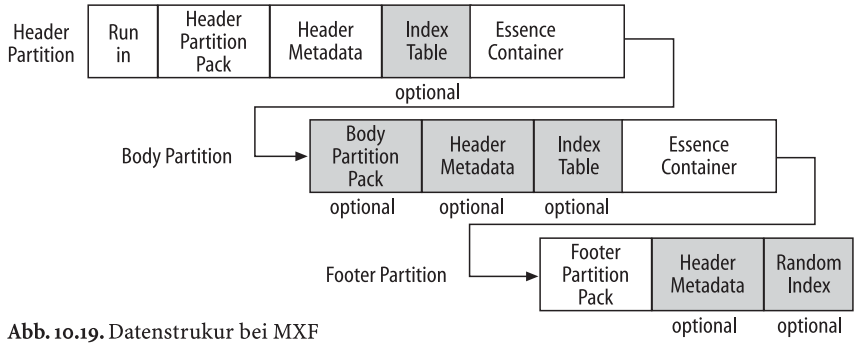


Abb. 10.19. Datenstrukur bei MXF

10.1.3.3 Material-eXchange-Format

Um Programmmaterial in professionellen Netzwerken möglichst einfach als Bitstreams austauschen zu können, wurde ein Datenaustauschformat entwikkelt, das Material-eXchange-Format (MXF), seltener auch Media-eXchange-Format genannt wird. Es bietet eine einfache Möglichkeit, format-, kompressions- und plattformunabhängig audiovisuelle Inhalte inklusive Metadaten mittels IT-Technologie zu übertragen. Der Ansatz ist ähnlich wie bei AAF, der einzige Unterschied besteht darin, dass MXF eine begrenzte Funktionalität aufweist, die speziell auf einen schlanken Datenaustausch angepasst wurde.

MXF basiert auf dem Data Encoding Protocol SMPTE 336M, KLV. Die Abkürzung steht für Key, Length, Value und bedeutet, dass jedes Dateipaket mit einem Key, einer einzigartigen, 16 Byte langen, registrierten SMPTE-Nummer beginnt, die Aufschluss über seinen Inhalt gibt. Die Length gibt die Länge des anschließenden Values an, also des Content-Datenpaketes. Sollte ein Interpreter ein Paket nicht entschlüsseln können oder nicht benötigen, kann er die Bytes des Pakets, dessen Länge ja bekannt ist, komplett überspringen.

Die MXF-Dateistruktur ist in Abbildung 10.19 gezeigt. Im Header, der am Anfang jeder Datei steht, werden sämtliche für das File benötigten Metadaten angegeben. Anschließend folgt der Datei-Body, der einen oder mehrere Essence-Container enthält. Sollte der Body eines MXF-Files nicht definiert sein, enthält die Datei nur Metadaten. Der Footer schließt das File ab.

Die Header Partition kann optional mit einem sogenannten Run-In beginnen. Dieser ist nicht nach KLV codiert und nur für spezielle Anwendungen erforderlich, denen er eine Synchronisation ermöglicht. Das nun folgende Header Partition Pack enthält u. a. den Verweis, ob die Partition offen oder geschlossen ist. Eine geschlossene Partition enthält die beschreibenden Metadaten für die gesamte restliche Datei, sollte sie offen sein, so ist nur der Anfang der Datei beschrieben. Eine eingefügte Index-Tabelle gibt den Standort aller Mediendaten in der Datei wieder. In den Essence-Containern befindet sich schließlich der audiovisuelle Inhalt.

Wie bereits erwähnt, handelt es sich bei den Metadaten im Header um Daten ohne die eine Dateiübertragung nicht möglich wäre. Das hier untergebrachte MXF-Material-Paket (Abb. 10.20) repräsentiert die auszugebende Timeline des Files und legt die Edit Unit Rate fest, mit der intern als Zeitreferenz

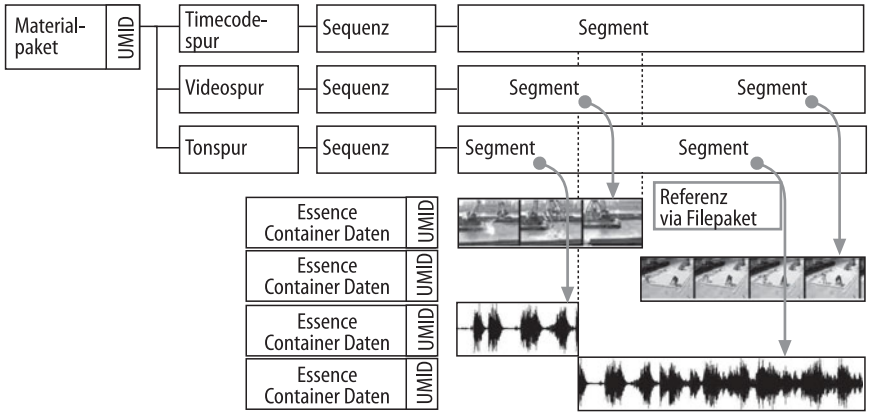


Abb. 10.20. Zuweisung von Spuren, Sequenzen und Essence bei MXF

gearbeitet wird. Identifiziert wird das Paket über einen UMID und es enthält z. B. einen Timecode-, einen Bild- und einen Ton-Track. Deren einzelne Segmente werden über das File-Paket einer Essence innerhalb eines Containers zugewiesen. Das Source-Paket, welches ebenfalls auf die Essence verweist, gibt Auskunft über den Ursprung des Inhaltes, wie es auch bei AAF der Fall ist. Während der Timecode-Track kontinuierlich ist und direkt den Output widerspiegelt, können die Segmente einer Timeline im Bild- und Ton-Track durch harte Schnitte miteinander verbunden werden. Andere Übergänge werden von MXF nicht unterstützt, denn es kommt hier lediglich auf die Reihenfolge und auf ein streamingfähiges Format an.

Um die Komplexität skalieren zu können, wurden für MXF verschiedene Operational Pattern als 3 x 3 Kombinationen definiert (Abb. 10.21). Weist ein Material-Paket (MP) nur ein einziges Item auf, spricht man von einem Sin-

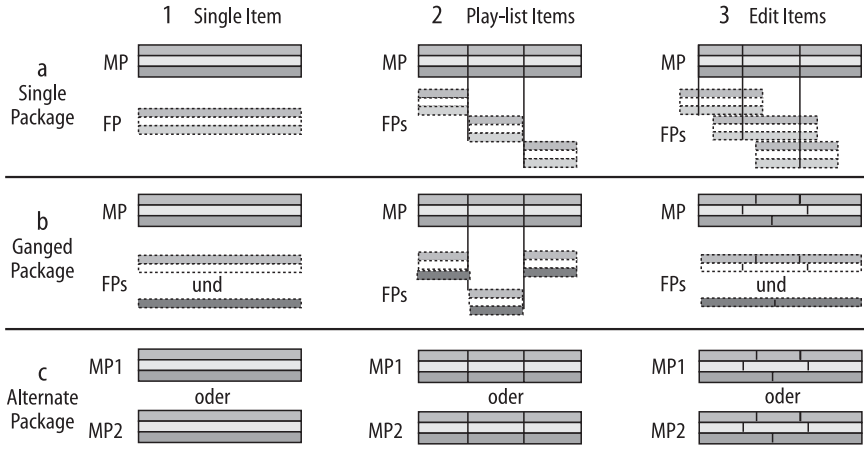


Abb. 10.21. Operational Patternse bei MXF

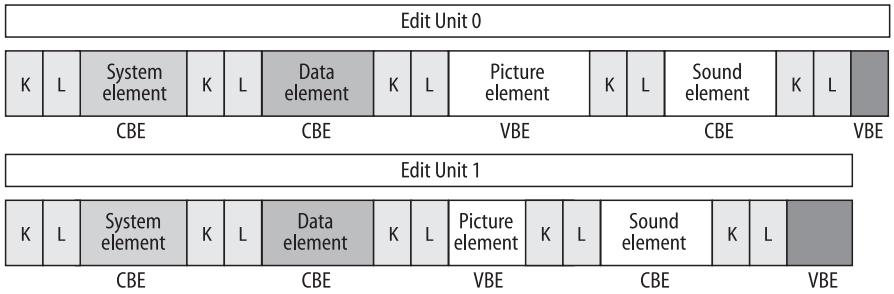


Abb. 10.22. Beispiel für eine KLV-codierte MXF-Datei

gle Item, welches über die gesamte File-Länge abgespielt wird. Playlist Items werden in einer gegebenen Reihenfolge hintereinander gehängt, während für Edit Items zum Schnitt noch In- und Out-Punkte gesetzt werden, ähnlich einer sehr einfachen EDL.

Wenn das Material-Paket nur auf ein einziges File-Paket (FP) zur Zeit zugreifen muss, so spricht man in der Paket-Komplexitätshierarchie in einfachster Form von einem Single Paket. Bei Zugriff auf mehrere File-Pakete spricht man von gruppierten Paketen (Ganged Packages). Im dritten Fall hat man verschiedene alternative File-Pakete zur Auswahl, die z. B. eine unterschiedliche Sprache oder Bildauflösung beinhalten können.

Gegenwärtig wird meist die einfachste der neun möglichen Kombinationen, nämlich das MXF Opertional Pattern (OP) 1a, also Single Item/Single Package benutzt. Auch die Rundfunkanstalten in Deutschland haben für ihren Filetransfer die Variante OP 1a (SMPTE 378M) festgelegt. Es ist sehr gut geeignet für den Programmaustausch mit Bild und Ton. Für einfachste Editing-Anwendungen erscheint es in manchen Fällen noch zu komplex, sodass nachträglich das OP-Atom (Fa. Avid) definiert wurde, das nur eine einzige Art von Essence enthält, also z. B. entweder Ton- oder Bild-Dateien, doch nicht beide zusammen.

In dem MXF-Container werden die Essence-Elemente in Edit Units aufgeteilt und in einem Datenstrom KLV-codiert abgelegt. Als praktisches Beispiel seien hier in der Abbildung 10.22 drei Elemente mit fester Länge angegeben, das System-Element, das Daten- und Ton-Element (Constant Bytes per Element, CBE). Das Bildelement wird gesondert behandelt, ihm wird eine variable Länge zugewiesen (Variable Bytes per Element, VBE).

MXF ist flexibel und erweiterbar angelegt. AV-Daten können komprimiert oder unkomprimiert, übertragen werden. Bei der Verwendung datenreduzierter AV-Daten ist dabei aber zu bedenken, dass MXF z. B. MPEG-2-Daten ebenso aufnimmt wie DV-basierte, allerdings nicht für eine Umsetzung sorgt. Das bedeutet, dass ein Endgerät, das nur den DV-Algorithmus beherrscht, die ggf. in MXF enthaltenen MPEG-2-Daten nicht entschlüsseln kann. Die Decodierbarkeit hängt vom Vermögen des empfangenden Systems ab.

10.1.3.4 General-eXchange-Format

Das General-eXchange-Format (GXF) ist ein von der Grass Valley Group (GVG) entwickeltes Austauschformat mit Echtzeitfähigkeit. Das bedeutet z. B.,

dass ein von einer Kamera geliefertes Live-Signal sofort über GXF weitergesendet werden kann, sodass ein GXF-File nicht auf einem Datenträger abgelegt werden muss.

Über die Jahre wurde das Format weiterentwickelt um auch AV-Datenströme inklusive einiger Metadaten zwischen Nutzern, die den relativ simplen, als SMPTE 360M festgelegten Standard vor allem im Sport- und Nachrichtenbereich nutzten, austauschen zu können. Während die erste GXF-Implementation nur JPEG-Sequenzen und unkomprimierte Videosignale zuließ, wurden später auch MPEG-ES, DVCPro und M-JPEG als Videokompression unterstützt. Die Übertragung der Audio-Daten erfolgt direkt als PCM-Strom, im AC-3- oder Dolby E-Format. Da GXF schon vor der Existenz des SMPTE-KLV-Schemas entwickelt wurde, ist es nicht KLV-kompatibel. Dennoch können KLV-kompatible Daten als User-Metadaten übertragen werden [96].

Die Zielrichtung und Leistungsfähigkeit von GXF ist ähnlich wie bei MXF, doch während sich Letzteres auf breiter Front als Fileformat im professionellen Sektor durchsetzte hat GXF heute kaum noch praktische Bedeutung.

10.1.3.5 OMF, DPX und PSD

Das Open Media Framework (OMF), das auch als Open Media Framework Interchange (OMFI) bezeichnet wird, ist das Format, auf dem AAF aufbaut. Es wurde lange unter den Bemühungen von Avid Technology in seiner Entwicklung vorangetrieben. Ebenso wie sein Nachfolger sollte es Daten für den Transport und den Gebrauch in Kompatibilität mit anderen Anwendungen bereitstellen. Die Tatsache, dass OMF anscheinend nicht endgültig plattformübergreifend funktioniert hat, ist zum Teil darin begründet, dass OMF kein von der SMPTE oder ITU verabschiedeter offener Standard, sondern Eigentum der Firma Apple ist.

Das unter SMPTE 268M seit 1994 standardisierte Digital Moving Picture Exchange Format, kurz DPX, wurde vor allem entwickelt, um den Transfer unkomprimierter Bilder von Telecine-Maschinen (Filmabtastern) zu unterstützen. Es trägt Metadaten zur detaillierten Beschreibung der Bilder und wird auch bis heute zur Übertragung unkomprimierter Bilder genutzt. Es ist in Kapitel 5 genauer beschrieben.

Oft wird auch das Bildformat des Bildbearbeitungsprogramms Adobe Photoshop, PSD, als Austauschformat für Einzelbilder genutzt. Es unterstützt Multi-Layering mit einer Vielzahl von Effekten und Metadaten. Alternativ können Bildfolgen auch als TIFF-Sequenzen ausgetauscht werden.

10.1.3.6 QuickTime

Die professionellen Fileformate für den Datenaustausch, also AAF und MXF sind relativ jungen Datums. Im Heimanwenderbereich gibt es ähnliche Formate schon seit längerem. Sie sind vor allem im PC-Bereich dominant und integraler Bestandteil der Betriebssysteme. Die beiden am häufigsten anzutreffenden Formate sind hier Apple QuickTime für MacOS und Microsoft Windows Media für Windows. Die Wiedergabe-Software für die beiden Formate ist auf jedem PC vorinstalliert und plattformübergreifend, z. B. auch unter Linux nutzbar.

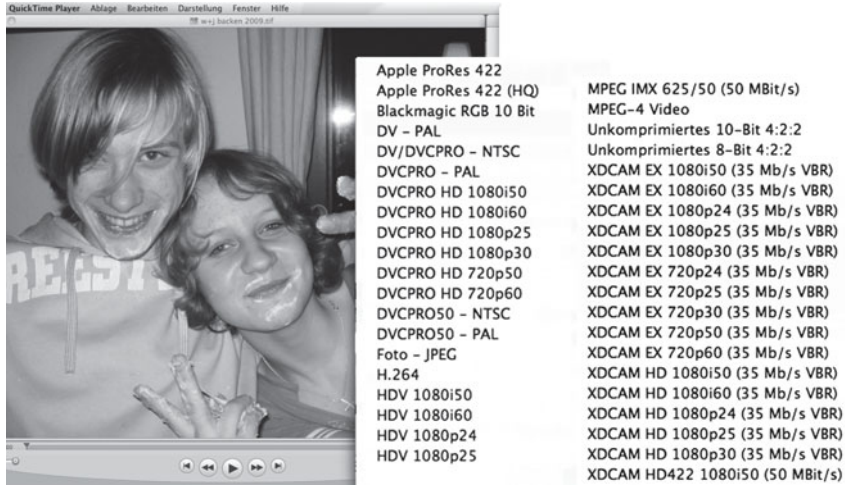


Abb. 10.23. QuickTime-Player und Auswahl von unterstützten Codecs für den Dateiexport

QuickTime (QT) ist ein von Apple Computer Inc. entwickeltes und bereits 1991 auf den Markt gebrachtes Format für multimediale Inhalte. Neben einem wachsenden Funktionsumfang hat Apple stets die Abwärtskompatibilität bewahrt, ohne dass das Format dabei rückständig oder ausgedient wirkt. Im Gegenteil bietet QuickTime ausgefeilte Funktionen als Dateiformat und wird sehr aktuell gehalten. Unter QuickTime können verschiedene Codecs laufen, sodass z. B. die 4:1:1-Variante von DV ebenso unterstützt werden kann wie HDV, XDCAM HD422, Apple ProRes oder unkomprimierte Daten (Abb. 10.23).

QuickTime bietet als eine der Besonderheiten die Möglichkeit, eine Referenz auf verwendete Daten zu erstellen, ohne die Notwendigkeit, diese als neuen Medien-Datenstrom innerhalb des erstellten Files abzuspeichern. So können beispielsweise für einen Videoschnitt oder ein Layering, also einer schichtweisen Überlagerung mehrerer Videoströme, lediglich Metadaten über das Zusammenspiel der verwendeten Bilder angelegt werden. Beim Abspielen muss dann auf deren externe Referenzquellen zugegriffen werden. Diese Vorgehensweise entspricht der des neu entwickelten Advanced-Authoring-Formats. Der Vorteil dieser Art des Umganges mit Content-Daten liegt darin, dass eine mit QuickTime erstellte Datei u. U. sehr klein ausfallen kann, obwohl sie mehrere Gigabyte an Mediendaten verwaltet. Selbstverständlich ist weiterhin auch die herkömmliche, komplette Neu-Speicherung aus dem Resultat des Editings möglich.

QuickTime weist eine klar nachzuvollziehende Entwicklung auf. Im einfachsten Fall stellt das Format einen Container für verschiedene AV-Inhalte dar, die in Spuren organisiert sind und unter dem System synchron gehalten werden. Eine Spur könnte z. B. eine DV-codierte Filmsequenz enthalten, eine zweite eine zusätzliche Tonspur und eine dritte dazu synchrone MIDI-Daten, also reine Steuersignale, die dazu dienen, Töne an elektronischen Musikinstrumenten auszulösen. Der Funktionsumfang wird stetig erweitert, z. B. ist ab QT

4 der Import von Shockwave-Daten (Flash) möglich. Neben dem ebenfalls von Apple entwickelten QuickTime Virtual Reality bietet QT viele aufbauende Applikationen, die es selbst im Datenaustausch bei professionellen Produktionen zur beliebten Anwendung haben reifen lassen.

Hinzu kommen die umfangreichen Möglichkeiten zur Verknüpfung von Objekten, die über sogenannte Atome definiert werden. Aufgrund dieser Vielseitigkeit begann sich auch die ISO (International Organization for Standardization) bei der Suche nach einem Grundstein für MPEG-4, vor allem mit Blick auf die Möglichkeiten des interaktiven Umgangs mit Medienobjekten, für das Format von Apple zu interessieren. Sie übernahm es dann 1998 dafür als Schlüssel-Komponente, weshalb QuickTime 6 als erster wirklicher MPEG-4-Player genannt wird.

10.1.3.7 AVI und Windows Media

Ab 1991 stattete Microsoft das Betriebssystem Windows mit dem AVI-File-Format aus. AVI steht für Audio-Video-Interleave und beinhaltet einen Audio- und Videostrom. Nach einer teilweise unübersichtlichen Entwicklung unterstützte Windows Media 9 seit Anfang 2003 High-Definition-Video und 5.1-Kanal-Surround-Sound. Die AV-Formate tragen ab dann die Bezeichnung Windows Media Video und Audio mit den Dateiendungen .wmv und .wma [96].

Bis heute ist AVI weit verbreitet. Das wesentliche Charakteristikum ist seine Einfachheit. Es ist leicht zu adaptieren, aber die Aufnahme von Datenmaterial beschränkt sich lediglich auf Audio und Video, und zwar jeweils auf genau einen Strom. Ein weiteres Manko ist die Tatsache, dass sich die beiden Spuren kaum synchronisieren lassen. Sollte die Möglichkeit des Interleavings nicht genutzt werden, besteht die Gefahr, dass die Audio- und Videospuren nach einiger Zeit auseinander laufen.

Ein weiterer Nachteil, den AVI mit sich bringt, ist die Tatsache, dass ein aus Rohdaten komponiertes Video komplett neu in einem AVI-File gespeichert wird, wodurch eine immense Redundanz verursacht wird, da die Rohdaten neben dem fertigen File auf der Festplatte vorliegen. Außerdem ist AVI sehr anfällig für Fehler. Sollte dem File nur der am Ende der Datei liegende Index-Teil fehlen oder sollte der defekt sein, versagen viele Player das Abspielen der Datei oder verhindern einen wahlfreien Zugriff.

Der Aufbau von AVI entspricht einem typischen Container-Format mit der Bezeichnung RIFF (Resource Interchange File Format). Hier gibt es auch eine eigenständige Audio-Variante die als WAVE-Format bezeichnet aber von Microsoft offiziell durch Windows Media Audio abgelöst wird.

Aufbauend auf dem WAVE-Format hat die European Broadcasting Union (EBU) das Broadcast-Wave-Format (BWF) als unabhängiges Austauschformat für Audio-Daten herausgegeben. Ein BWF-File behält die gesamte Struktur des WAVE-Formats bei, es wird lediglich eine sogenannte Broadcast Audio Extension hinzugefügt.

Im Kontext der Bemühungen um ein Medienformat für Windows ist auch die DivX-Codierung entstanden. Der Kompressionsalgorithmus sollte ursprünglich den MPEG-4-Codec Version 3 für Windows bereitstellen, aber Hacker patchten (engl. flickten) den Code und gaben ihn als DivX;-) 3.11 für AVI heraus.

10.2 Signalverteilung

Videosignale müssen in Studios und Produktionsbereichen vielfältig und flexibel verteilt werden können. Dabei sind zum Teil auch große Distanzen zu überbrücken z. B. wenn Beiträge verschiedener Landesfunkhäuser zu einem Gesamtprogramm zusammengestellt werden müssen. Man spricht dann von Programmzuspielung oder auch Contribution, während die Verteilung an die Endgeräte der Zuschauer als Distribution bezeichnet wird.

Die Signalverteilung geschieht im lokalen Bereich leitungsgebunden. Dabei ist es bei analogen Signalen so, dass sie jeweils eine einzelne Leitung beanspruchen während Digitalsignale auch in dieser Form, aber auch sehr effektiv in Netzwerken übertragen werden können. Für Weitverkehrsverbindungen kommen ebenfalls Netzwerke in Frage oder auch eine Signalverteilung via Satellit.

10.2.1 Verteilung von Analogsignalen

Zur Signalführung werden für analoge Signale Koaxial-Leitungen mit einer Impedanz von 75 Ω eingesetzt. Diese verursachen eine Signalverzögerung und eine frequenzabhängige Dämpfung (Abb. 2.38), die besonders bei analogen Signalen ggf. durch Leitungsentzerrer und Laufzeitanpasser ausgeglichen werden müssen.

10.2.1.1 Synchronisation

Bei der Signalverteilung und Bildmischung sind mehrere Videosignale beteiligt. Da diese der Periodizität des Synchronsignals unterliegen, ist die wesentliche Voraussetzung für eine fehlerfreie Funktion die absolute Synchronität der Bildgeneratoren. Der Gleichlauf wird im Studiobetrieb erreicht, indem alle angeschlossenen Bildquellen an einen zentralen Studiotakt angepasst werden (Abb. 10.24). Aufwendige Systeme nutzen dazu als Taktgeber hochstabile Generatoren mit temperaturstabilisierten Quarzen. In kleinen Nachbearbeitungsstudios reicht es meistens aus, das im Bildmischer generierte Black-Burst-Signal, ein FBAS-Videosignal mit Schwarz als Bildinhalt, als Studiotakt zu verwenden. Da nicht alle Geräte die gleiche Signalverarbeitung aufweisen, und die zur Ver-

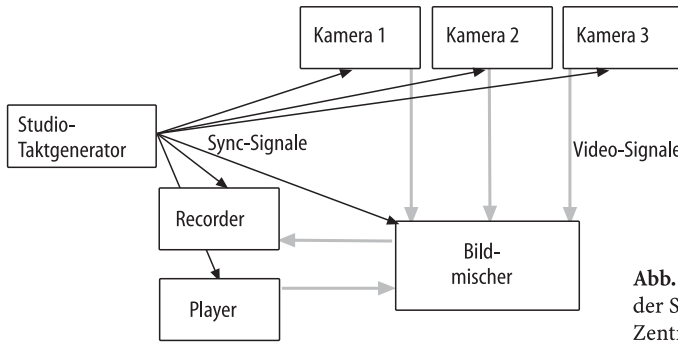


Abb. 10.24. Synchronisation der Studiogeräte mit einem Zentraltakt

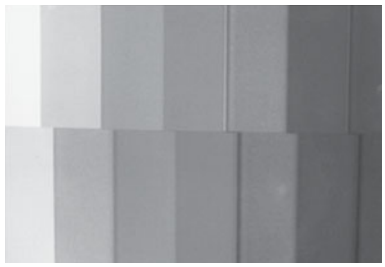


Abb. 10.25. Trickmischbild (Wipe) zweier Farbbalkensignale mit H-Phasenfehler

bindung dienenden Leitungen unterschiedliche Längen haben ist der Gleichlauf durch den zentralen Takt allein noch nicht gewährleistet. An allen professionellen analogen Geräten ist deshalb eine Möglichkeit zum Feinabgleich der Phasenlage zwischen dem Videosignal und dem Studiotakt gegeben.

Es reicht aus, die horizontalen Bildlagen (H-Phase) und bei FBAS-Signalen auch die Farbträgerphasen (Subcarrier-Phase) in Übereinstimmung zu bringen, da die Zeitabweichungen maximal im µs-Bereich liegen. Ein H-Phasenfehler äußert sich in einem seitlich versetzen Bild. Er ist leicht zu erkennen, wenn z. B. zwei Bandmaschinen das gleiche Bild wiedergeben. Dann erscheinen die gemischten Signale auf dem Monitor gegeneinander verschoben (Abb. 10.25). Der H-Phasenabgleich kann gut mit Hilfe eines 2-Kanal-Oszilloskops vorgenommen werden. Der Studiotakt und das anzugleichende Signal werden dazu auf die beiden Kanäle gegeben, und die H-Phase des Signals wird so eingestellt, dass die Flanken der H-Synchronimpulse exakt übereinander liegen (Abb. 10.26). Wenn ein Waveformmonitor zum H-Phasenabgleich herangezogen werden soll, der nur ein Signal darstellen kann, so muss dafür gesorgt werden, dass er extern auf das Taktsignal getriggert wird, da sonst bei der Kanalumschaltung auch die Triggerung auf das jeweilige Signal reagiert, und damit eine Verschiebung nicht sichtbar wird. Analoge Taktsignale werden auch in digitaler Infrastruktur eingesetzt, manchmal wird hier aber auch per SDI-Signal getaktet.

Während bei analogen SD-Signalen das Black-Burst-Signal (BB) mit dem bipolaren Synchronimpuls als Taktsignal verwendet wird, kommt im HD-Studio ein Black-Signal mit Tri-Level-Sync (TLS) zum Einsatz (Abb. 10.27). Da dieses aufgrund der höheren Zeilenzahl ein eigenes Timing aufweist, brauchen SD- und HD-Studios unterschiedliche Taktsignale. Wenn SD- und HD-Programme parallel verarbeitet werden sollen, müssen entsprechend zwei Taktebenen vor-

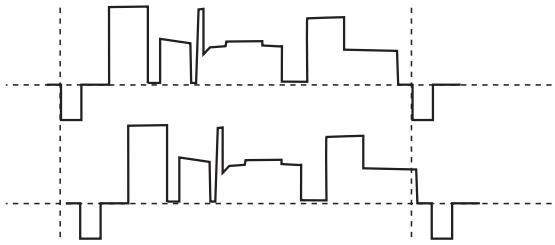


Abb. 10.26. Oszilloskopdarstellung zweier Signale mit H-Phasenfehler

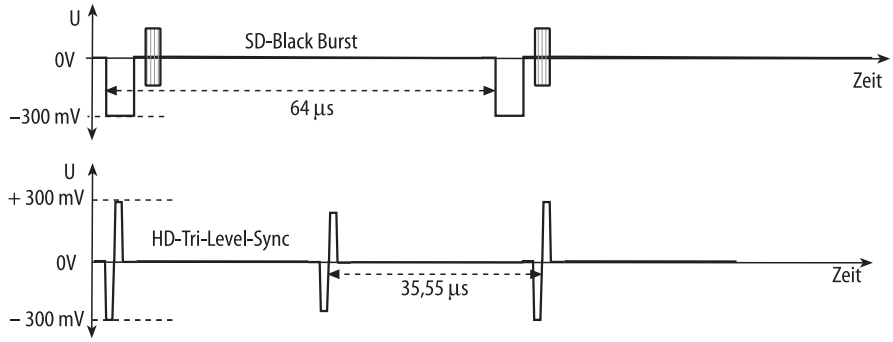


Abb. 10.27 Bi- und Tri-Level-Sync bei SD- und HD-Signalen

handen sein (Abb. 10.27). Moderne Synchron Pulse Generatoren (SPG) stellen Black Burst und Tri Level Sync parallel bereit [186].

10.2.1.2 Synchronizer

Falls Signale in die Mischung eingebunden werden sollen, die von Geräten stammen, die nicht an den Studiotakt anpassbar sind (Amateurgeräte), oder die soweit vom Studio entfernt sind, dass der Takt nicht übertragbar ist, werden Framestore-Synchronizer eingesetzt, mit denen beliebige asynchrone Videosignale an den Studiotakt angepasst werden können.

Die Synchronizer arbeiten wie ein digitaler TBC (s. Kap. 8), sind aber in der Lage, mehrere Bilder zu speichern. Das asynchrone Signal wird dazu in den Speicher eingelesen, das Auslesen richtet sich nach dem Studiotakt (Abb. 10.28). Die Fähigkeit mindestens ein Bild im Synchronizer zu speichern kann auch dazu genutzt werden, einzelne Bilder für Einspielungen vorzuhalten.

Generell ist im professionellen Bereich dafür zu sorgen, dass alle an der Mischung beteiligten Signale bezüglich H-und SC-Phase abgeglichen sind, die Einstellung ist regelmäßig zu kontrollieren. Es gibt aber auch Bildmischer, die standardmäßig mit Synchronizern ausgerüstet sind, sodass hier der Phasenabgleich automatisch geschieht.

Auch bei Digitalmischern müssen die an der Mischung beteiligten Signale im Gleichtakt laufen, hier ist aber keine Anpassung und kein Abgleich erforderlich, denn die erwähnten Synchronizer sind dort standardmäßig enthalten. Sie weisen meist einem Korrekturbereich von ca. 50 µs auf. Beim Einsatz von

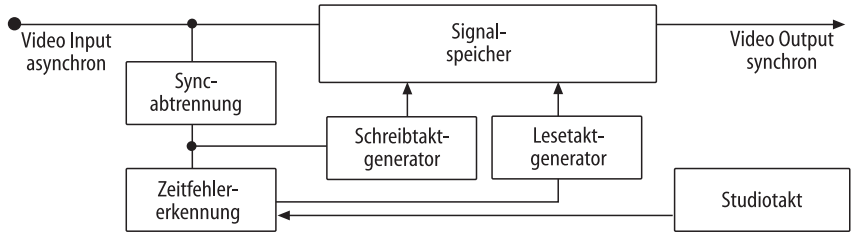


Abb. 10.28. Funktionsprinzip des Synchronizers

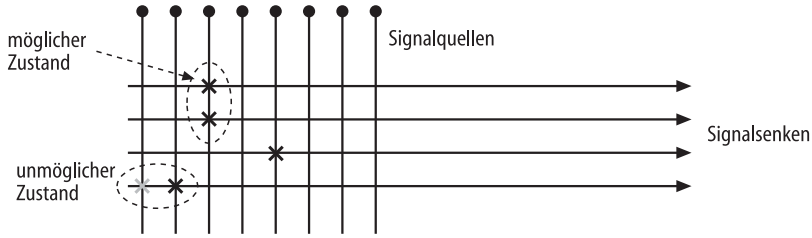


Abb. 10.29. Kreuzschiene 8 x 4

Synchronizern ist zu bedenken, dass sie eine Verlängerung der Signallaufzeit bewirken. Das kann bei der digitalen Signalverarbeitung mit ihrer ohnehin schon relativ langen Prozessdauer dazu führen, dass das Bild gegenüber einem parallel geführten Tonsignal sichtbar verzögert ist. Eventuell müssen zur Abhilfe Tonverzögerungsgeräte eingesetzt werden. Moderne Synchronizer sind oft mit integriertem Audiodelay ausgestattet.

10.2.1.3 Steckfelder und Kreuzschienen

Die einfachste Art der Kombination von Signalquellen und -senken ist die Nutzung eines mit Koaxialbuchsen bestückten Steckfeldes (Abb. 10.30), das auch Bildwähler genannt wird. Die auf der Rückseite fest angebrachten Signalleitungen können auf der Vorderseite über Brückenstecker oder kurze Koaxialleitungen verbunden werden. Um den richtigen Leitungsabschluss zu gewährleisten, sind in die Buchsen oft 75 Ω-Widerstände integriert, die bei der Entfernung eines Steckers automatisch auf das Leitungsende geschaltet werden.

Wie der Bildwähler dienen auch Kreuzschienen einer flexiblen Verteilung mehrerer Signalquellen an mehrere Videogeräte, die Auswahl wird aber nicht durch Stecker, sondern auf Tastendruck vorgenommen. Sehr häufig werden sie zur Aufschaltung von Signalquellen an Regiegeräte eingesetzt und sind wichtiger Bestandteil von Bildmischern. Die Ein- und Ausgänge bilden eine Matrix (Abb. 10.29). Jedem Kreuzungspunkt ist ein elektronischer kontaktloser Schalter zugeordnet. Auf Tastendruck kann ein Eingang auf ein oder mehrere Ausgänge geschaltet werden, es ist aber nicht möglich, mehrere Eingänge auf einen Ausgang zu legen. Abbildung 9.7 zeigt mögliche Schaltzustände.

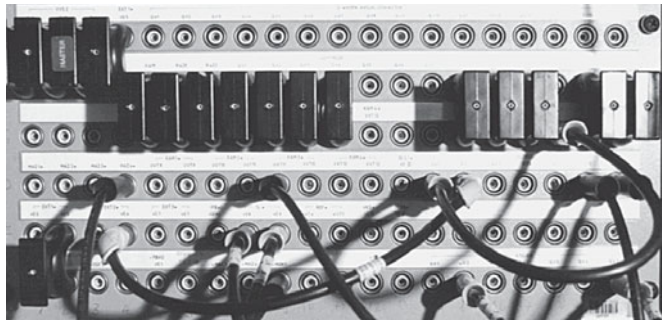


Abb. 10.30. Video-Steckfeld

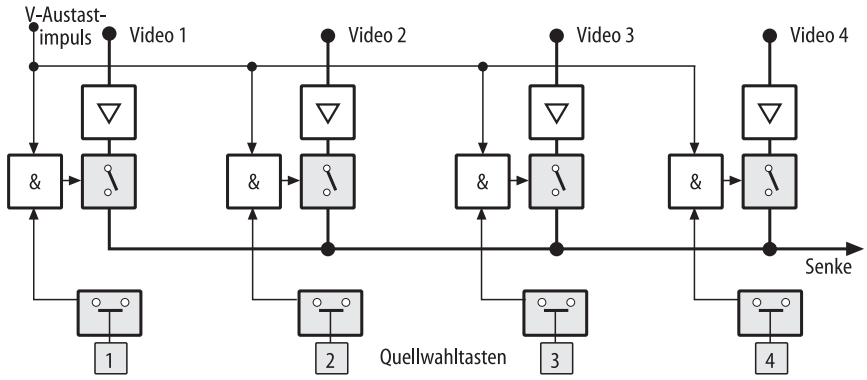


Abb. 10.31. Funktionsprinzip der Videokreuzschiene

Eine einfache Umschaltung zwischen zwei Videosignalen führt i. d. R zu Bildstörungen, da die Synchronsignalstruktur gestört wird, wenn mitten im Halbbild auf ein anderes Signal gewechselt wird. Aus diesem Grund wird dafür gesorgt, dass mit Hilfe elektronisch verriegelter Schalter die Umschaltung nur in der vertikalen Austastlücke, also jeweils nach einem Bild, geschieht. Die Steuerlogik für die Schalter entspricht der eines UND-Gatters. Mit dem Tastendruck wird der Wechsel nur vorbereitet; erst wenn auch der V-Synchronimpuls vorliegt, wird der Schaltvorgang tatsächlich ausgelöst (Abb. 10.31). Die Umschaltung soll in der 6. bzw. 319. Zeile des Videosignals vorgenommen werden.

Videokreuzschienen werden auch als Havariekreuzschienen ausgeführt und so geschaltet, dass der Betrieb bei Ausfall der Hauptkreuzschiene störungsfrei fortgeführt werden kann. Daneben gibt es Kreuzschienen für andere Studiosignale, die z. T. mit der Hauptkreuzschiene verkoppelt werden.

10.2.2 Digitale Signalverteilung

Auch im digitalen Umfeld werden Steckfelder eingesetzt, noch mehr sind aber Kreuzschienen verbreitet. Moderne Kreuzschienen sind sowohl für SD- als auch HD-SDI-Signale geeignet. Wie analoge Kreuzschienen bewirken auch digitale Typen eine Umschaltung in der V-Austastlücke, orientiert am Timing Reference Signal. Wie bei Analogsignalen soll in der Zeile 6 bzw. 319 umgeschaltet werden und zwar bei SD-Signalen in einem Zeitfenster zwischen 25 µs und 35 µs nach der negativen Flanke des Synchronisationsimpulses.

Das Signal wird zunächst entzerrt und einem Komparator zugeführt, der die verzerrten Signalzustände eindeutig rekonstruiert. Bei der Digitalübertragung tritt aber nicht nur eine Verzerrung der Amplitude auf, sondern auch Zeitbasisschwankungen (Jitter). Wenn diese unverändert weitergegeben werden, so spricht man von transparenten Kreuzschienen, die für alle möglichen Datenraten bis zu einem Maximalwert von z. B. 2 Gbit/s geeignet sind. Kreuzschienen, die mit einer Reclocking-Stufe zur Jitter-Verringerung ausgestattet sind (Abb. 10.32), erfordern dagegen eine Festlegung der Datenrate. Es ist nicht generell

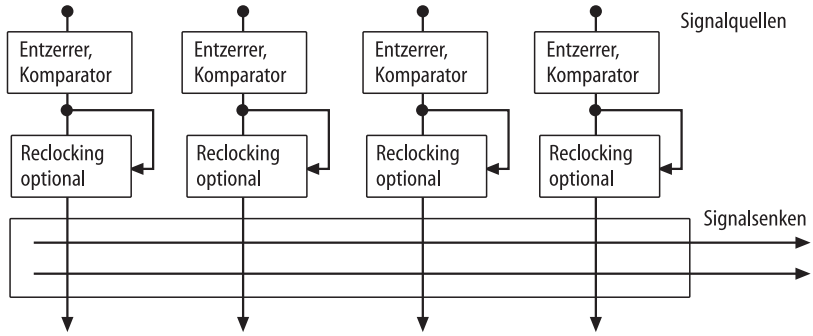


Abb. 10.32. Elemente einer Digitalkreuzschiene

von Vorteil, Reclocking-Stufen einzusetzen, da sie eventuell zu zusätzlichem Jitter führen können [180].

Abbildung 10.33 zeigt die Anschlüsse auf der Rückseite eines modularen Kreuzschienensystems. Über Verstärker werden die Signale hier Matrixkarten zugeführt, sodass insgesamt bis zu 128 Kanäle geschaltet werden können. Die Matrix ist in der Lage, 128 Eingänge auf 32 Ausgänge zu legen und diese wieder so zu verteilen, dass auch bis zu 128 Ausgänge zur Verfügung stehen. Bevor die Ausgangssignale auf die Ausgangstreiber geführt werden findet ein Reclocking statt. Viele Kreuzschienen sind so ausgelegt, dass sie parallel ein zum Videosignal synchrones Umschalten zugehöriger Audiosignale gestatten [180].

Die Steuerung der Kreuzschiene erfolgt über Hardware-Controller oder mittels eines externen PC, oft via RS 485-, RS 422- oder RS 232-Schnittstellen. Die PC-Software bietet meistens eine graphische Benutzeroberfläche mit Hilfe derer das System übersichtlich bedient werden kann, wobei den Quellen und Senken Namen gegeben werden können und Standardkonfigurationen abspeicherbar sind. Alternativ kann die KS-Steuerung auch über Computernetzwerke und damit aus Newsroomsystemen erfolgen. Die komplexen Kontrollmöglichkeiten erlauben dann auch statt der großen zentralen KS viele kleine, dezentrale einzusetzen. Da die Kreuzschiene das Herzstück eines Produktionskomplexes ist, muss besondere Rücksicht auf den Havariefall genommen werden. So können sich Hard- und Softwaresteuerung ergänzen und wenn beide ausfallen, wird der letzte Zustand beibehalten. Aufgrund möglicher Havarie wird auch mit zwei redundanten Netzteilen gearbeitet.

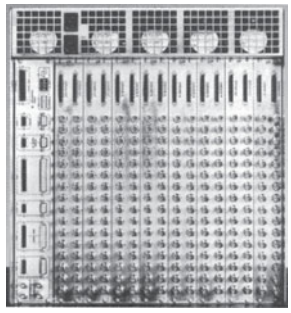


Abb. 10.33. Anschlussfeld einer Digitalkreuzschiene [180]

10.2.3 Videosignalschnittstellen

Die Übertragungssysteme aus diesem Bereich sind noch eng mit dem Videosignal verknüpft und aus diesem Grund als echtzeitfähige Streaming-Systeme einzuordnen. Man unterscheidet hier die synchrone Übertragung bei der die Daten über einen separaten Takt synchronisiert werden und den weniger anspruchsvollen isochronen Datentransfer, bei dem es ausreicht, die Datenrate weniger konstant zu halten, z. B. soweit, dass sich eine stabile Bildwechselfrequenz einstellt. Streaming wird mit komprimierten und unkomprimierten Signalen durchgeführt und kann auch in Netzwerken aus dem Computer- und Telekombereich realisiert werden.

10.2.3.1 SDI

Für digitale Videosignale wurde als Schnittstelle eigens das Serial Digital Interface (SDI) entwickelt, das in Abschnitt 3.3 vollständig dargestellt ist. Es bietet eine Datenrate von 270 Mbit/s. Für hochaufgelöste Videoströme steht das technisch sehr ähnliches HD-SDI mit ca. 1,5 Gbit/s zur Verfügung über das durch Erweiterung zum Dual-HD-SDI oder 3G-SDI auch 10 Bit-RGB-Daten bei hohen Qualitätsansprüchen übertragen werden. SDI ist für die unkomprimierte Datenübertragung ausgelegt. Es ist frei schalt- und steckbar, quasi latenzfrei, robust und stellt die Standardschnittstelle im Bereich der Echtzeitanwendungen dar.

10.2.3.2 SDTI

Nach der Einführung standardisierter Datenreduktionsverfahren wie DV und MPEG-2 entstand der Wunsch, dieses echtzeitfähige Interface zu nutzen, um Signale in der datenreduzierten Ebene auszutauschen und damit Qualitätsverluste durch die Umsetzung in die DSK-Ebene und ggf. anschließende erneute Codierung zu vermeiden. Da mit SDI ein verbreitetes System mit einer großen Gerätepalette (Kreuzschienen, Router) zur Verfügung steht, lag der Gedanke

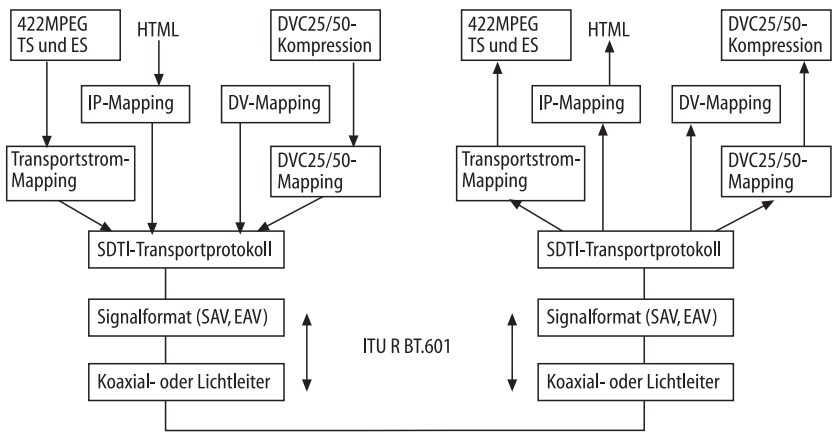


Abb. 10.34. Layer-Modell von SDTI

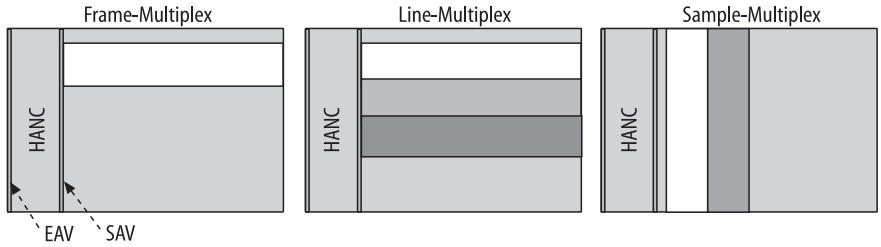


Abb. 10.35. Multiplex-Varianten bei SDTI

nahe, dieses auch für die Übertragung der Datenpakete des datenreduzierten Signalstroms einzusetzen. Da ein einzelnes datenreduziertes Signal von z. B. 50 Mbit/s die SDI-Kapazität von 270 Mbit/s nicht ausnutzt, besteht dabei die Möglichkeit das reduzierte Signal schneller als in Echtzeit zu übertragen oder den SDI-Kanal für mehrere Datenströme zu nutzen.

Nach dieser Idee entstanden zunächst die herstellereigenen Schnittstellenstandards SDDI (Serial Digital Data Interface) und CSDI (Compressed serial Data Interface). Die Standardisierungsgremien der EBU und SMPTE entwickelten dann aus beiden Vorschlägen auf Basis des SDI die Kompromisslösung SDTI (Serial Data Transport Interface).

Bei der Entwicklung von SDTI wurde darauf geachtet, die SDI-Grundstruktur unangetastet zu lassen. Für die Einbettung der Datenpakete kommt nur der für die aktive Videozeile definierte und der Ancillary-Bereich in Frage. Die EAV- und SAV-Synchrondaten, die durch die Bitfolgen FF3 000 000 besonders hervorgehoben sind, bleiben unangetastet, sie sind für die Gesamtstruktur von größter Bedeutung. Das Serial Digital Transport Interface setzt auf die SDI-Definitionen auf, wobei ein eigenes SDTI-Transportprotokoll eingeführt wird. Abbildung 10.34 zeigt dies anhand des ISO-Schichtenmodells.

Für die anwendernahen Ebenen war es erforderlich, dass für jedes Datenformat eine eigene Mapping-Spezifikation entwickelt wurde. Für den SD-Bereich sind dies: DV-based, d. h. DV25 und DV50-Datenströme, die Applikation Consumer-DV, die Übertragung des MPEG-2-Transportstroms, die SDTI Content Package (SDTI-CP) zum Multiplex vom MPEG-2-Elementarstrom mit Audio- und Metadaten und schließlich SDTI-IP zur Übertragung nach dem Internet-Protokoll.

Zur Signalisierung der paketierten Daten wird ein SDTI-Header eingeführt, der 53 Bytes umfasst und direkt auf die EAV-Kennung folgen muss [102]. Ob im Ancillary-Bereich der H-Austastung (HANC) weitere Daten übertragen werden, wird im Header besonders gekennzeichnet (Data Extension Flag). In den meisten Fällen werden die paketierten Daten in den Bereich des aktiven Bildes eingesetzt, wobei oft das sogenannte Line-Multiplexing verwendet wird, was insbesondere die Übertragung schneller als Echtzeit bei variabler Blocklänge ermöglicht. Alternativ kann das Sample Multiplexing verwendet werden, bei dem feste Datenblöcke innerhalb einer Videozeile eingesetzt werden, oder das Frame-Multiplexing, was große Datenblöcke und aufgrund der wenigen Headerinformationen eine hohe Nettoausnutzung ermöglicht, dafür aber mit langen Verarbeitungsdauern verbunden ist (Abb 10.35).

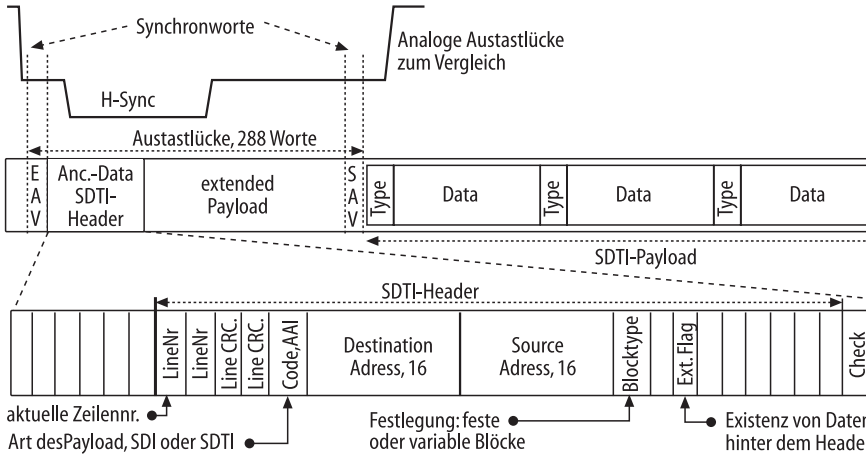


Abb. 10.36. Zeilenstruktur und Header bei SDTI

Abbildung 10.36 zeigt die Zeilenstruktur des SDI-Signals und die Struktur des Headers. Der Header enthält die Zeilennummer und die Kennzeichnung der verwendeten Adressierung, sodass mit Hilfe der Angabe von Zieladressen ein Point to Multipoint-Verkehr und der Datentransfer mittels IP-Adressen möglich wird. Der Block Typ definiert, ob eine variable Blocklänge eingesetzt wird, wie es bei der Übertragung des MPEG-2-Transportstroms über SDTI der Fall ist, oder ob eine der vordefinierten fixed Block Strukturen verwendet wird, wie es bei DV-Anwendungen mit ihren festen Datenraten geschieht. Jedem Datenblock ist eine Type-Identifikation vorangestellt, an dem der verwendete Anwendungstyp erkannt wird. Die Art der Nutzdatenlast (Payload) ist also nicht durch den Header bestimmt [102].

Bei der Verwendung der variablen Blockstruktur muss die Länge der Datenpakete durch einen Wordcounter angegeben werden, wobei aufeinander folgende Pakete durch Endmarker getrennt werden. Der große Vorteil von SDTI ist, dass es auf Basis einer weit verbreiteten Infrastruktur die Möglichkeit eröffnet, mehrere Videodatenströme parallel oder einen Strom schneller als in Echtzeit zu übertragen. Beim Transfer der Daten treten keine Transcodierungsverluste auf, dafür müssen die Sende- und Empfangsgeräte aber natürlich den richtigen Datentyp verarbeiten können. Der Nachteil ist die Unidirektionalität und die Point to Point bzw. Multipoint-Topologie, sodass die Flexibilität von echten Netzwerken nicht erreicht werden kann.

10.2.3.3 ASI

Diese Abkürzung steht für Asynchronous Serial Interface. Die Grundfunktion ist ähnlich wie bei SDTI: Die etablierte SDI-Infrastruktur wird für die Übertragung datenreduzierter Signale, hier insbesondere MPEG-2-Transportströme (TS), genutzt. Die Taktinformation des Transportstroms ist dabei unabhängig von der SDI-Datenrate. Mehrere TS können im Multiplex übertragen werden. ASI wird z. B. bei der Zuführung von MPEG-2-Daten im DVB-Umfeld genutzt.

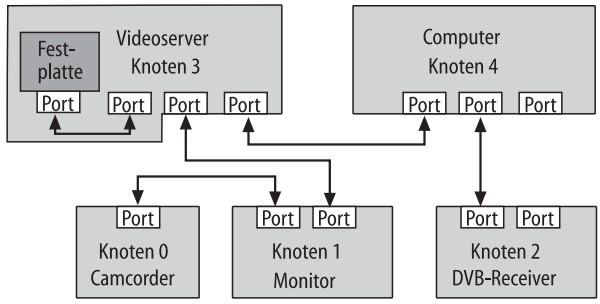


Abb. 10.37. Beispiel für Verbindungen bei IEEE 1394

10.2.3.4 IEEE 1394

Diese Bezeichnung steht für ein Bussystem das als Firewire vom Computerproduzenten Apple entwickelt, als IEEE 1394 standardisiert und heute unter der Bezeichnung i-Link vor allem im Consumer-Videobereich verbreitet ist. Es ist videogeeignet, aber auch als allgemeines Dateninterface konzipiert. Die Variante IEEE 1394a ermöglicht den Echtzeitaustausch von Video-, Audio- und Kontrolldaten mit Datenraten von 100, 200 und 400 Mbit/s und bietet als IEEE 1394b bis zu 1,6 Gbit/s. i-Link ist ein bidirektionaler Datenbus, der am ehesten mit einem SCSI-Bus vergleichbar ist. Es erfordert wie SCSI intelligente Endgeräte, im Gegensatz zu SCSI können aber Geräte während des Betriebs verbunden werden (hot plugging), erfordern keine Abschlusswiderstände und keine ID-Zuweisung der Nutzer. Es lassen sich ohne die Verwendung von Zusatzsystemen bis zu 63 Geräte in eine Kette integrieren.

Abbildung 10.37 zeigt ein Beispiel eines IEEE 1394-Netzwerkes. Ein Gerät wird darin als Modul bezeichnet und besitzt einen oder mehrere Knoten. Jedem Knoten wird eine Adresse zur Identifizierung zugeordnet, die zwischen 0 und 62 liegt. Neben den 6 Bit für die Adressierung des Knotens deklarieren die ersten 10 der gesamten 64 Bit langen Adresse einen von 1023 Busbereichen. Die letzten 48 Bit dienen zur Adressierung interner Registerbereiche des angesprochenen Knotens. Die Point-to-Point-Verbindung der Knoten erfolgt über Ports, von denen jedes Gerät 1 bis 27 Stück besitzt. Geräte mit mehreren Ports dienen als Repeater, d. h., dass Daten von einem Port zum nächsten durchgereicht werden. Dies funktioniert auch bei ausgeschalteten Geräten, da diese über die Schnittstelle eine Betriebsspannung beziehen können. Bei der Verkabelung von Ports gibt es keine festgelegte Struktur, die einzelnen Knoten können also beliebig in Stern- oder Baum-Topologie verbunden sein, nur Schleifen sind nicht erlaubt.

Sobald ein Knoten aus dem System entfernt oder ein neuer hinzugefügt wird, wird eine Selbstkonfiguration des Netzwerkes eingeleitet. Diese beginnt mit einem Reset, durch den alle Knoten ihre ID und ihre Kenntnis über das Netzwerk aufgeben. Daraufhin signalisieren alle Knoten erneut ihre Beteiligung an dem Netzwerk. Sobald das Netz als Ganzes erkannt ist (Tree Identification), leiten alle Knoten einen Selbstidentifizierungsprozess ein und teilen den anderen Teilnehmern ihre Adresse und maximale Geschwindigkeit mit.

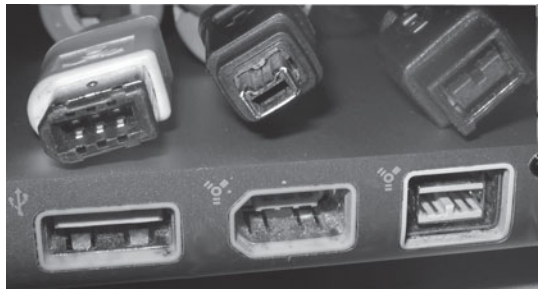


Abb. 10.38. Steckverbindungen bei IEEE 1394a und IEEE 1394b (rechts)

Die Datenübertragung erfolgt mit dem NRZ-Code seriell über einfache Leitungen mit 4- oder 6-Pin Steckern, wobei über Letztere auch die Stromversorgung (8 V bis 40 V bei max. 60W) bereitgestellt werden kann (Abb. 10.38 und 10.39). Die Drahtpaare übertragen Signale nach 1394a immer jeweils in die gleiche Richtung, also im Halbduplex. Die Übertragung über IEEE-1394b-Ports, den sogenannten Beta-Ports, erfolgt über neunpolige Stecker (Abb. 10.38, rechts). Neben den Leitungen der sechspoligen Variante enthält diese jeweils einen zusätzlichen Massedraht pro Twisted Pair und eine Leitung für zukünftige Erweiterungen. Die Signalübertragung erfolgt im Vollduplex, also über beide Twisted Pairs gleichzeitig in entgegengesetzte Richtungen.

Eine Verbindung nach IEEE 1394 ermöglicht sowohl den Filetransfer-Betrieb als auch die Realzeitübertragung in Streaming-Technik mit definiertem Delay. Das Protokoll ist im Kern paketorientiert, die für den Videoeinsatz interessante Realzeitübertragung wird erreicht, indem den Paketen der isochronen Daten eine besonders hohe Priorität verliehen wird. Dabei wird die Datenübertragung in Zyklen aufgeteilt, von denen jeder eine Länge von 125 µs hat. So wird garantiert, dass 66% der Bruttobitrate für den Echtzeittransfer unter Einhaltung einer gegebenen Quality of Service zur Verfügung stehen.

Die Schnittstelle ist universell nutzbar. Die Übertragung spezieller Formate, wie DV und MPEG-2-Transportstrom, wird mit dem Standard IEC 61883 erreicht, der auf der IEEE 1394-Spezifikation aufsetzt, dabei wird mit dem Functional Channel Protocol (FCP) zusätzlich eine Möglichkeit des Austauschs von Daten zur Gerätesteuerung definiert.

Eine IEEE 1394-Verbindung ist sehr anwenderfreundlich im Plug and Play-Verfahren verwendbar. Als Anschlussbezeichnung ist sie auch unter DV IN/OUT bekannt und wird häufig eingesetzt, um digitale Camcorder und Videorecorder mit einem PC oder entsprechenden Videoschnittgeräten zu verbinden. Das System ist als Low Cost-Lösung also vor allem für den Heimanwender-Bereich konzipiert, im professionellen Bereich wird es eher selten verwendet. Es ist nicht sehr störsicher, aus diesem Grund und wegen der integrierten Stromversorgung sollten die Verbindungslängen 4 m nicht überschreiten.

Die Weiterentwicklung IEEE 1394b unterstützt optische Interfaces und lässt bei Distanzen bis zu 100 m Bitraten im Gbit-Bereich zu. Die Zukunft der Schnittstelle ist aber fraglich, da angesichts der Verfügbarkeit sehr schneller Schnittstellen wie USB 3 oder Thunderbolt die Firma Apple die Entwicklung nicht mehr voran zu treiben scheint.

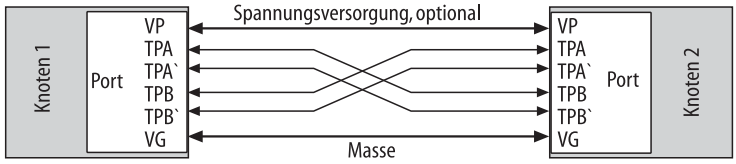


Abb. 10.39. Steckerbelegung bei IEEE 1394

10.2.3.5 USB

Der Universal Serial Bus (USB) wurde entwickelt, um Peripheriegeräte verschiedener Art an Computer anzuschließen. USB 1 war für Drucker, Tastaturen etc. mit einer Maximaldatenrate von 12 Mbit/s vorgesehen. USB 2 unterstützt bis zu 480 Mbit/s und USB 3 erlaubt Datenraten bis 4,8 Gbit/s. Die bis zu 127 möglichen Geräte am Bus werden nach dem Anschluss automatisch erkannt.

USB 2 ist weit verbreitet und so leistungsfähig, dass es auch für die Übertragung datenreduzierter Videosignale z. B. zwischen PC und externen Speichern eingesetzt wird. Im Gegensatz zu Firewire ist USB aber nicht für hohe Videoperformance und Echtzeitbetrieb optimiert. Es ist keine bestimmte Datenrate garantierbar, denn USB basiert auf einer Master/Slave-Architektur bei der der PC den Datenverkehr regelt. Abbildung 10.38 zeigt links unten eine USB 2-Buchse. USB 3 scheint eine ähnliche Verbreitung zu finden. Es ist abwärtskompatibel zu USB 2 und 10-mal so schnell. Im Gegensatz zu USB 2 unterstützt es aber eine Vollduplexverbindung mit Leitungen, die 8 Drähte enthalten von denen zwei die Kompatibilität zu USB 2 sicherstellen. Über die Schnittstelle ist eine Spannungsversorgung für externe Geräte möglich dabei stellt USB 2 maximal 500 mA und USB 3 bis zu 900 mA zur Verfügung.

10.2.3.6 Thunderbolt

Thunderbolt ist eine Datenschnittstelle die von den Firmen Intel und Apple entwickelt wurde und mit 10 GBit/s in der ersten Ausbaustufe doppelt so schnell arbeitet wie USB 3. Sie ist als universelle Schnittstelle konzipiert und ist im Kern eine aus dem PC-Bereich stammende PCI-Express-Verbindung die nach außen geführt ist. Sie unterstützt entsprechend zwei Protokolle, nämlich PCI-Express (PCIe) und Display-Port (DP) und erlaubt den Anschluss von Speichergeräten und Displays gleichzeitig, wobei die Geräte in Reihe geschaltet werden. Die Datenübertragung erlaubt gleichzeitiges Senden und Empfangen in beiden Richtungen (Vollduplex) sowie die Stromversorgung von externen Geräten mit Leistungen bis zu 10 W.

Thunderbolt-Schnittstellen wurden als erstes bei Apple-Computern verbaut und verbreiten sich bei anderen Herstellern eher langsam. Viele Herstellet setzen eher auf USB 3, da es erheblich kostengünstigere Kabel erlaubt.

Für dem Einsatz im Videobereich sind sowohl für USB 3 als auch für Thunderbolt externe Geräte verfügbar, die die Videosignale an SDI oder HDMI-Schnittstellen umsetzen und zu den Computern führen, sodass selbst für Laptops HD-SDI-Editing und -Monitoring verfügbar wird. Bei Thunderbolt gilt dies auch für die Bearbeitung von 2K- oder 3D-Datenströmen. Abbildung 9.63 im vorigen Kapitel zeigt entsprechende Geräte.

10.3 Videodaten in Netzwerken

Wie in vielen Bereichen der Technik, bietet die Vernetzung von Arbeitsplätzen, Produktions- und Postproduktionskomplexen auch in der Videotechnik erhebliche Zeit- und Kostenvorteile. Die digitalen Daten werden im Netzwerk über kürzere oder längere Strecken (LAN, WAN) verteilt, wobei die Netze ring- oder sternförmig, oder als Point to Point Verbindungen aufgebaut sein können (Abb. 10.40). Ein Netzwerk verbindet viele Geräte durch ein gemeinsames Übertragungsmedium, wobei die Daten an ein oder mehrere Geräte adressiert sind, sodass die Geräte überprüfen müssen, ob die Daten für sie bestimmt sind, was bei der direkten Point to Point Verbindungen natürlich nicht erforderlich ist. Als ein Beispiel für eine Point to Point Verbindung kann bereits die digitale Signalübertragung über SDI betrachtet werden, womit hier eher eine Interfacedefinition als ein echtes Netzwerk vorliegt. Doch wird an diesem Beispiel bereits deutlich, dass Netze für Videoanwendungen besonderen Anforderungen genügen müssen. Erstens ist die Datenrate auch beim Einsatz datenreduzierter Signale vergleichsweise hoch. Zweitens kann das Erfordernis bestehen, mit einer festen Datenrate einen synchronen Betrieb zu gewährleisten. Diese Forderung gilt aber nicht generell. Aus Gründen der Flexibilität wird möglichst auf sie verzichtet, denn dann kann man mit paketierten Datenströmen arbeiten, die inhaltlich unabhängig nicht nur für Videosignale, sondern in gleicher Weise für Audio-, Zusatz- und hinhaltsbeschreibende Metadaten verwendbar sind. Diese können asynchron und auch schneller oder langsamer als in Echtzeit übertragen werden.

Die Art des Austausches der Programminhaltsdaten in Form des Filetransfers oder als sogenanntes Streaming ist ein wichtiges Unterscheidungsmerkmal. Beim Filetransfer steht die fehlerfreie Übertragung im Vordergrund. Die Inhaltsdaten (Content) werden mit Headerdaten versehen und als Datenpakete übertragen. Bei fehlerhafter Übertragung können Pakete wiederholt gesendet werden, bis sie den Empfänger fehlerfrei erreicht haben. Es gibt keine Synchronisation zur zeitlichen Struktur der Inhaltsdaten, dafür kann die Übertragung auf Anforderung des Empfängers ausgelöst werden (Pull-Betrieb). Als Filetransferprotokoll wird für diese Anwendungen Universal FTP empfohlen, das auf TCP/IP, dem Internetprotokoll beruht [103].

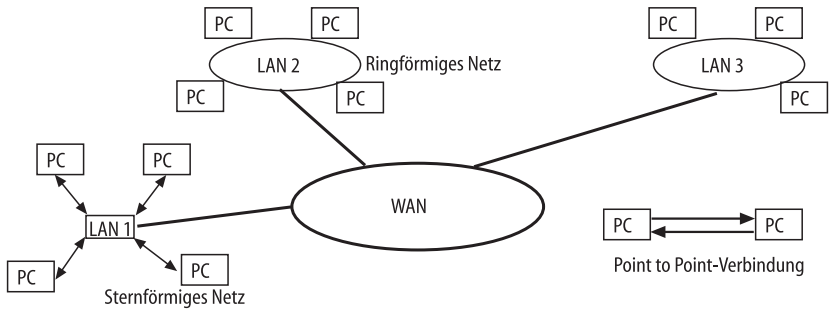


Abb. 10.40. Netzwerk-Topologien

Dagegen ist das Streaming eine realzeitbezogene Übertragung. Wie beim typischen Broadcastfall müssen sich die Empfänger auf den Sender einstellen und aus den ggf. fehlerhaften Daten das Beste herausholen (Best Effort Transfer). Die Sendung wird ohne Rücksicht auf den Empfänger oder eine Rückmeldung ausgestrahlt (Push-Betrieb), dabei muss sich der Empfänger jederzeit zum Datenstrom synchronisieren können.

Ein weiterer genereller Gesichtspunkt ist der Quality of Service (QoS). In modernen Produktionssystemen, wie z. B. beim Newsroom, wird mit Videodaten verschiedener Bildqualität gearbeitet, sodass es hilfreich ist, wenn die Netzwerkprotokolle es erlauben, die Qualität der Beiträge vom Anwender bestimmen zu lassen. Wenn eine gegebene Bandbreite von mehreren Anwendern genutzt wird, muss dafür gesorgt werden, dass eine Änderung des QoS für einen Nutzer keine Auswirkung auf die Signalqualität anderer Anwender hat. Neben der Bandbreite sind die Bitfehlerrate, Jitter und die Zeitverzögerungen beim Datenzugriff wichtige QoS-Parameter.

10.3.1 Systembetrachtung Signale und Netze

Zur Strukturierung der Netzfunktionen lässt sich gut das nach den internationalen Standardisierungsgremien bezeichnete ISO/OSI-Schichtenmodell verwenden, auf das in mehr oder weniger reiner Form alle Kommunikationssysteme abgebildet werden können. Die ISO/OSI-Betrachtung zeigt mit Hilfe von sieben aufeinander aufbauenden Schichten wichtige Aspekte der Systeme.

Die Informationen durchlaufen die Schichten der Reihe nach. Jede Schicht dient der darüberliegenden als Transportmedium (Abb. 10.41). Nicht bei jedem Verfahren sind alle Schichten zu finden [84]. Die unterste Schicht ist die physikalische Schicht (physical Layer). Hier sind die Art des Mediums und die elektrischen oder mechanischen Eigenschaften festgelegt (z. B. Brief, Telefon und elektrische Leitungen). Darauf aufbauend folgt die Verbindungsschicht (data link). Hier werden Punkt-zu-Punkt-Verbindungen und Netzwerke unterschieden, Zugriffsverfahren festgelegt und ggf. der Fehlerschutz definiert. (z. B.

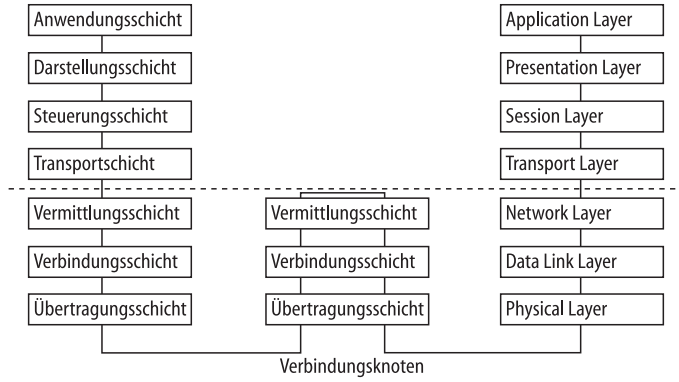


Abb. 10.41. ISO/OSI-Referenzmodell

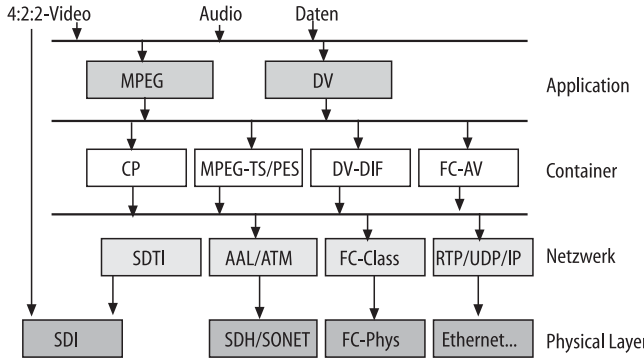


Abb. 10.42. Beispiel für zusammengefasste Netzwerkschichten

Zugriff auf das Telefon per Impuls oder Tonwahl). Die darüberliegende dritte Schicht ist die Vermittlungsschicht (Network Layer). Sie dient der Adressierung, der Befehlssatzdefinition, der Blockbildung und dem Verbindungsaufbau (z. B. Adressierung über die Telefonteilnehmernummer). Schicht vier ist die Transportschicht (Transport), in der der eigentliche Datentransport stattfindet und die Verbindung bestätigt wird.

Diese vier untersten Schichten bestimmen die Eigenschaften des Netzes direkt, die folgenden drei darüberliegenden bestimmen, wie mit der hergestellten Verbindung umgegangen wird. Schicht 5, der Session Layer (Kommunikationssteuerung), definiert die Netz- und Datenorganisation während der Verbindung (z. B: die Ausgabe von Impulsen für einen Gebührenzähler). Die sechste Ebene, die Darstellungsschicht (Presentation Layer), betrifft vor allem die verwendete Sprache, die Syntax, Kompression und Verschlüsselung (z. B. Art der Sprache am Telefon), während schließlich in der siebten, der Anwendungsschicht (Application Layer) der eigentlich Inhalt steht (z. B. Nachrichteninhalt am Telefon), zusammen mit Zugangsberechtigungen und anwenderspezifischen Protokollen.

Oft werden verschiedene Ebenen zusammengefasst und z. B. ein Daten· interface aus Elementen von Schicht 6 und 7 gebildet, das das Datenformat und die Kompressionsart festlegt. Dieses Interface ist dann für die Anwender dominant. Hier haben sie direkten Zugriff und können z. B. feststellen, dass ein Datenformat nicht zu einem zweiten kompatibel ist. Der Wechsel des Datenreduktionsverfahrens von z. B. DV auf MPEG-2 betrifft diese Schichten (Abb 10.42). Weiterhin kann beispielsweise aus den Schichten 4 und 5 ein Netzwerkinterface gebildet werden. Diese Art der Strukturierung ermöglicht, dass die anwendernahen oberen Ebenen nicht verändert zu werden brauchen, wenn die Netzinfrastruktur in den Ebene 1-3 z. B. von Kupfer- auf Glasfaserleitungen umgestellt wird.

Im Bereich Telekommunikation und Datentechnik sind Netzwerke seit längerem etabliert. Für die Videoanwendung wird versucht, an diese verbreiteten und damit kostengünstigen Systeme anzuknüpfen. Das Angebot bestehender Technologie ist sehr vielfältig und es ist nicht leicht festzustellen, welches Sys-

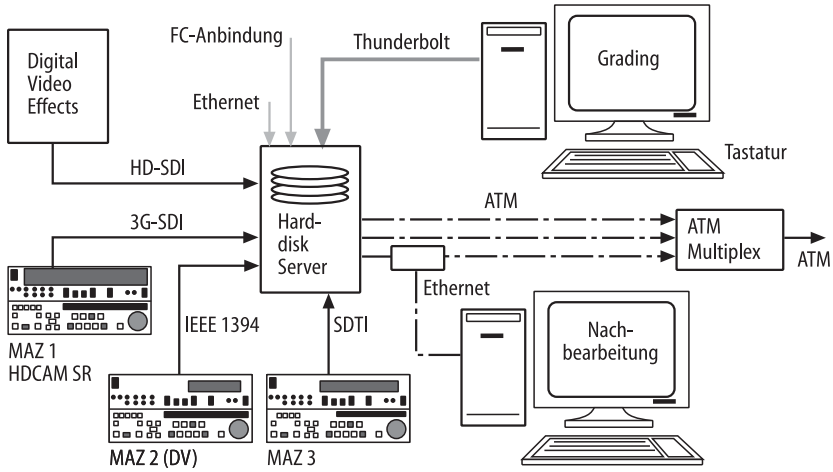


Abb. 10.43. Verschiedene Signale und Netze bei der Videoproduktion

tem für den Videobereich das geeignetste ist. Folgende Anforderungen werden an Netze für Videoanwendungen gestellt:

- Hohe Bandbreite. Auch wenn mit datenreduzierten Signalen gearbeitet wird, muss die Datenrate hoch sein, um mehrkanalig arbeiten zu können. Die 270 Mbit/s des SDI stellen daher eher die untere Grenze dar,
- Nutzung verschiedener Signalformate,
- Unterstützung unterschiedlicher und auch variabler Datenraten,
- Isosynchrone Übertragung für einen Teil der Infrastruktur. In diesem Bereich ist Echtzeitübertragung gefordert und die Daten müssen sich mit dem Studiotakt synchronisieren lassen,
- Parallele Nutzung von realtime- und nonrealtime Übertragung,
- Einsatz verbreiterter Netzprotokolle aufgrund der immer enger werdenden Verbindung zu Telekommunikations- und Datensystemen,
- Vorsorge für Havariefälle,
- synchrone Audioanbindung,
- Schnittstellen zu HD-SDI, SDTI und analogen Videosignalformen,
- Routing der Daten möglichst in einer Form, wie sie im Videostudio gehandhabt wird,
- keine unnötigen Ketten aus Codierung und Decodierung,
- Möglichkeit zur festen Verbindung zwischen Quelle und Senke.

Insgesamt lässt sich feststellen, dass die Anforderungen aus den verschiedenen Produktionsbereichen sehr unterschiedlich sind, sodass verschiedene Netzwerke zum Einsatz kommen. SDI und SDTI für die synchrone Übertragung konventioneller Art, Ethernet multifunktionell auch für Zusatz- und Steuerdaten, Fibre Channel und ATM für Anwendungen mit hoher Datenrate im LAN und WAN. Abbildung 10.43 zeigt eine Mischung möglicher Signalführungen in einem Postproduktionsstudio. Die Entwicklung geht dabei eindeutig hin zu Netzwerken, wie Ethernet, die kaum noch videospezifisch sind.

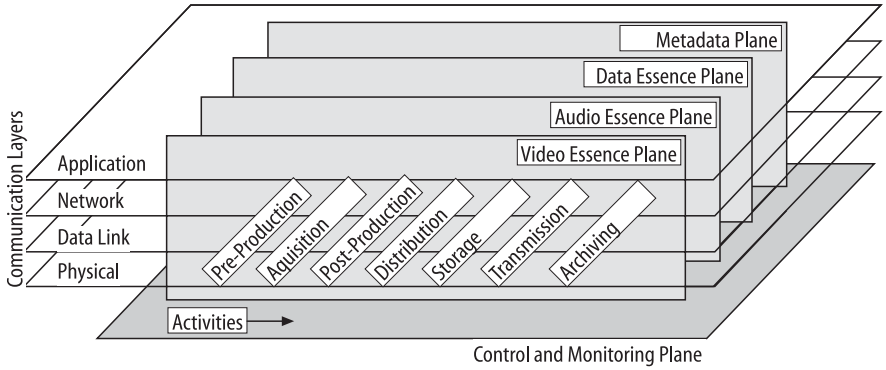


Abb. 10.44. Systemmodell mit Layers, Activities und Planes

Die Einführung gewöhnlicher IT-Technologie erlaubt eine große Flexibilität und auch die Verbindung von verschiedenen Medienproduktionsbereichen, sodass z. B. Rundfunkanstalten trimedial arbeiten, d. h. TV- mit Hörfunk- und Internetprogrammen sehr effektiv verbinden können [36]. Damit ist allerdings das Problem verbunden, dass die IT-Technologie sehr viele Varianten anbietet, die sich zudem in sehr kurzer Zeit ändern, sodass langfristige Entscheidungen über Investitionen in medienspezifische Infrastrukturen oft nicht leicht getroffen werden können.

Die Vielfalt und Komplexität ermöglicht die Optimierung für den Einzelfall, birgt aber auch die Gefahr der zu starken Diversifikation, die dazu führen kann, dass Anwender verunsichert werden, neue Technologien verspätet eingeführt werden und damit die Produktivität sinkt. Vor diesem Hintergrund haben die Standardisierungsgremien EBU und SMPTE eine Task Force of Harmonized Standards (TFHS) gegründet, die die Entwicklung von Einzellösungen so kanalisieren soll, dass sie in einem Gesamtsystem harmonisieren. Das Potenzial der Anwendung datenreduzierter und paketierter Signale im Videoproduktionsbereich soll damit weitgehend erschlossen werden.

Wichtige Themenbereiche sind:

- Art der Datenreduktion,
- Alle Speicherkonzepte (RAM-, Disk-, Bandgestützt),
- Nichtlinearer Zugriff auf Daten,
- Gleichzeitiger Mehrfachzugriff,
- Datentransport in realtime oder schneller bzw. langsamer als in Echtzeit,
- Transport und Bearbeitung von Zusatzdaten und Zusatzdiensten.

Die Task Force of Harmonized Standards empfiehlt bezüglich der Datenreduktion die Beschränkung auf wenige Reduktionsfamilien. Im SD-Bereich kommen hier z. B. DV und MPEG-2 bei 50 Mbit/s in Frage, für HD neben MPEG-2 die professionelle H.264-Variante AVC-I mit 112 Mbit/s.

Bezüglich des Datenaustauschs konnte eine Referenzarchitektur für File- und Realzeittransfer über SDI/SDTI, Fibre Channel und ATM entwickelt werden. Die Verpackung der Inhaltsdaten (Content) und Begleitdaten (Metadaten) in Datencontainern wurde mit Hilfe Wrappern definiert. Wrapper sind so

strukturiert, dass sie den Filetransfer oder die Archivierung von Content und Metadaten optimieren. In diesem Bereich hat sich das MXF-Format durchgesetzt und zwar meist mit dem einfachsten operational pattern.

Besonders wichtig bei der TFHS ist die Betrachtung der Systemebene. Hier ist das Systemmodell von großem Nutzen, das alle relevanten Merkmale übersichtlich erfasst und in Abbildung 10.44 dargestellt ist [82]. Auf der Horizontalachse sind die Aktivitäten, bzw. Produktionsbereiche von der Akquisition über die Postproduktion bis zur Archivierung aufgetragen. In der Raumtiefe sind die verschiedenen Datentypen dargestellt. Die eigentlichen Video- und Audiodaten (Essence) werden durch allgemeine Daten, wie z. B. Textdaten für die Teleprompterzuspielung ergänzt. Hinzu kommt die Ebene der Metadaten, in der alle Daten untergebracht sind, die, wie z. B. Timecode, inhaltsunabhängig eine eigene Relevanz haben. Die Metadaten sind besonders bei der Archivierung wichtig.

Die zwei Dimensionen aus Activity und Dataplanes sind durchzogen von den Schichten der Kommunikation, wobei die sieben Schichten des OSI-Modells hier als vier zusammengefasst sind: Applikation, Netzwerk, Verbindung und physikalische Schicht. Dem Ganzen liegt eine Control-Ebene zugrunde, von der aus das Gesamtsystem gesteuert und überwacht wird. Die Systemsteuerung sorgt u. a. dafür, dass nur sinnvolle Verknüpfungen auftreten und ein DSK-Datenstrom nicht einem MPEG-Decoder zugeführt wird. Sie beherrscht das Data-Management und entscheidet ob Daten im Strom (streaming) oder per Filetransfer übertragen werden, also ob eher Echtzeitanwendung oder Fehlerfreiheit und Datensicherheit im Vordergrund steht. Die Steuerung sorgt auch dafür, dass die Zuordnung von Inhalts- und Zusatzdaten in allen Bearbeitungsstufen erhalten bleibt. So kann hier die MPEG-Synchronisation implementiert werden, um Qualitätsverluste bei der Kaskadierung dieser Signale zu vermeiden, oder ein Audiobegleitsignal zusätzlich verzögert werden, wenn sich aufgrund einer komplexeren Codierung des Videosignals die Signalbearbeitungsdauer verlängert.

Zur Beherrschung der Steuerungskomplexität wird die Object modelling technique angewandt [82]. Die Systemkomponenten werden dabei vereinfacht und weitgehend unabhängig von ihrer Funktion dargestellt, sodass eine möglichst einfache Kontrolle und Weiterentwicklung, z. B. durch neue Software-Module, möglich wird. Dabei spielt auch die Steuerungsebene eine immer größere Rolle, da medienspezifische Geräte in immer umfangreicherer Form für besondere Aufgaben konfiguriert werden können. Dienste und Geräte werden als Objekte behandelt und für den Kontrollzweck angeordnet. Die Systemarchitektur ist so insgesamt dem immer komplexer werdenden Produktionsbereich angepasst.

Ein universelles Vernetzungssystem für den Videobereich hat sich noch nicht etabliert. Es sind viele Verfahren im Gebrauch. Im Folgenden werden Verfahren aus dem Telekom-Bereich, insbesondere die Weitverkehrstechniken, und aus dem Computerbereich die LAN-Systeme Ethernet und Fibrechannel näher erläutert, dabei wird in diesem Kapitel nur die Vernetzung im professionellen Bereich betrachtet, auf sogenannte In-House-Netze für den Heimbereich wird nicht eingegangen.

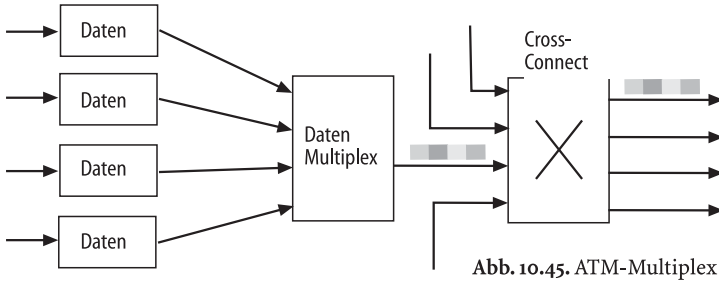


Abb. 10.45. ATM-Multiplex und Vermittlung

10.3.2 Netzwerke aus dem Telekom-Bereich

Dieser Bereich ist vor allem für Weitverkehrsverbindungen interessant, wie z. B. für den Programmaustausch zwischen voneinander entfernten Produktionsstätten oder Rundfunkanstalten. Vor der Digitalisierung wurde hier oft mit analogen Richtfunkstrecken gearbeitet, danach etablierten sich digitale Daten- und Netzstrukturen, die zunächst für allgemeine Telekommunikationsaufgaben konzipiert waren. Dazu gehört das ältere SDH-Netz, das verbindungsorientiert arbeitet, wobei die Verbindungen durch den Netzbetreiber aufgesetzt werden. Das SDH-Netz benutzt den seriellen Transfer Mode STM, der an der plesiochronen digitalen Hierarchie (PDH) orientiert ist und bietet eine Datenrate von 155,52 Mbit/s die für den Videobereich über sogenannte TV-Container angepasst wurde wobei Zugangsdatenraten von 140 Mbit/s und 34 Mbit/s spezifiziert waren. Höhere Datenraten können durch Multiplexing vieler Einzelverbindungen erreicht werden. SDH wurde im Zuspielnetz der ARD eingesetzt, hatte aber als universelles Netz kaum Chancen, weil es keine Variante für lokale Netzwerke gibt, die Datenraten schlecht zu den videotypischen Datenraten passen und der synchrone Übertragungsablauf bei Weitverkehrverbindungen Probleme bereiten kann. Besser ist hier der modernere Asynchronous Transfer Mode (ATM) geeignet. Das Verfahren ist als universelles Formate für Weitverkehrsverbindungen im Bereich der Telekommunikation konzipiert und dort weitgehend etabliert.

10.3.2.1 ATM

ATM ist prinzipiell besser als universelles Netz für Weit- und Nahverkehrseinsätze geeignet, da es asynchron arbeitet. ATM sollte auch das früher von den Rundfunkanstalten für Überspielungen genutzte VBN (vermittelnde Breitbandnetz) ersetzen. Es ermöglicht selbstwählfreie Verbindungen. ATM bietet eine Standard-Übertragungsrate von 155,52 Mbit/s bei einer Nutzbitrate von 100 Mbit/s, die bis zu 2 Gbit/s gesteigert werden kann. Es ist gut an bestehende LAN-Strukturen und an das Integrated Services Digital Network (ISDN) der Telekom angepasst. ATM basiert auf Datenpaketen und erlaubt die Realisierung mehrerer Verbindungen über nur eine Leitung. Die Übertragung ist sehr effizient, da eine (virtuelle) Verbindung nur aufgebaut wird, wenn sie auch gebraucht wird. Dies geschieht entweder vorkonfiguriert oder auf Anforderung als Permanent Virtual Circuit (PVC) bzw. Switched Virtual Circuit (SVC).

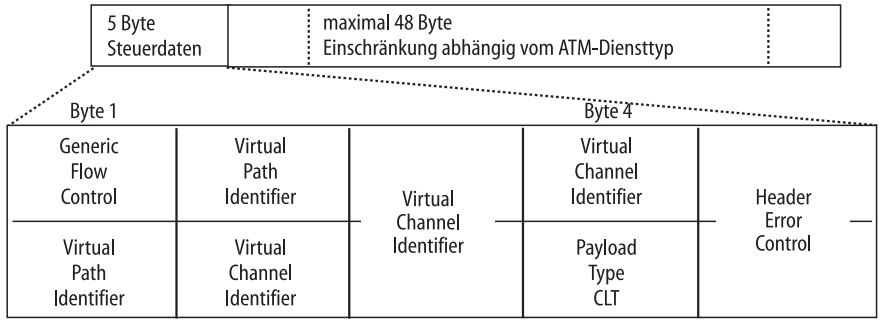


Abb. 10.46. ATM-Zellformat

Der Weg durch das Netz wird vor dem Aufbau der Verbindung festgelegt. Das ATM-Datenpaket umfasst 53 Bytes, von denen 5 für Steuerinformationen benutzt werden, die übrigen 48 Bytes stehen für Nutzdaten zur Verfügung. Eine Quelle mit konstanter Bitrate erzeugt eine konstante Rate von Datenpaketen (ATM-Zellen), deren Abstände in Abhängigkeit von der Bitrate variieren. Die Zellen haben keine feste Position im Datenstrom (asynchron). Signale verschiedener Quellen werden zusammen auf eine Leitung geschaltet, dabei müssen bei gleichzeitigem Zugriff einige Zellen warten und es entsteht eine variable Zellstromdichte. Die verschiedenen Daten werden in einem ATM-Multiplex zusammengefasst und über Crossconnects verteilt (Abb. 10.45). Die einzelne Verbindung wird über Steuerdaten im Kopf des Datenpakets festgelegt.

Abbildung 10.46 zeigt die Aufteilung der Steuerdaten, die u. a. den Virtual Channel Identifier (VCI) und den Virtual Path Identifier (VPI) enthalten, mit denen Kanäle und Pfade für den Verbindungsaufbau definiert werden. Mit der Header Error Control (HEC) ist auch ein Fehlerschutzmechanismus vorhanden. Einzelne Verbindungen werden mit Schalteinheiten realisiert, die die VPI- und VCI-Informationen auswerten. Aufgrund der festen Paketgröße können hier kleine, sehr schnelle Hardware-Elemente eingesetzt werden.

Virtuelle Kanäle können zu virtuellen Pfaden zusammengefasst werden. Bei einer Nachrichtenübertragung zwischen zwei Stationen A und B lassen die Crossconnects den Channel-Wert unverändert, während die Verbindungsstelle die Pfadinformation auswertet und die Einträge entsprechend der vorher festgelegten Vermittlungstabelle ändert. Neben Crossconnects und ATM-Verbindungsstellen gibt es ATM-Multiplexer, in denen mehrere Verkehrsströme zusammengefasst werden. Mit einer ATM-Inter Working Unit (IWU) können darüber hinaus Anpassungen an andere Netze vorgenommen werden.

Die bei ATM integrierte Netzwerkdefinition unterliegt einer dem ISO-Modell ähnlichen Schichtenstruktur. Der unterste Layer ist die physikalische Schicht, die von der darüberliegenden ATM-Schicht völlig unabhängig ist. In der ATM-Schicht werden die VCI- und VPI-Daten ausgewertet und das Multiplexing der Nutzdaten vorgenommen. Der ATM-Adaption Layer (AAL) dient als Verbindungsschicht zwischen der Datenform des Anwenders und der paketorientierten Datenstruktur. Mit Hilfe des Adaption Layers können unterschiedliche Diensteklassen gebildet werden, sodass ATM verschiedensten Anwendungen

Abb. 10.47. ATM-Schichtenstruktur

wie Telefon-, Fax- oder Videodatenübertragung gerecht werden kann (Abb. 10.47). AAL 1 findet bei ISDN Verwendung, AAL 3/4 wurde für Datenanwendungen konzipiert, bietet Fehlerschutz und die Möglichkeit der Datenwiederholung bei fehlerhafter Übertragung. AAL 5 beinhaltet keinen eigenen Fehlerschutz, es eignet sich aber für Videoanwendungen aufgrund der vollen 48 Byte Payload pro Zelle und der einfachen Implementierung.

ATM ist sehr flexibel und ermöglicht neben Filetransfer- auch den Streaming-Betrieb. Für den Filetransfer wird das Internetprotokoll empfohlen (IP over ATM). Um den Streaming-Betrieb zu realisieren, werden Synchroninformationen als Time-Stamps mit übertragen, gegenwärtig sind aber noch nicht alle Anforderungen aus dem Videobereich, hinsichtlich Delay, Jitter und Erhalt des Quality of Service ausreichend erfüllt bzw. standardisiert.

10.3.3 Netzwerke aus dem Computerbereich

Netzwerke aus dem Computerbereich dienen meist auf dem Datenaustausch mit dem File-Transfer-Verfahren. Bei dieser Konzeption hat die Datensicherheit höhere Priorität als die Garantie einer bestimmten Datenrate. Falls Daten fehlerhaft übertragen sind, werden sie wiederholt. Die Daten werden in einer Vielzahl verschiedener Formate gespeichert. Diese sind mit einem Header versehen, der die Identifizierung ermöglicht. Die Relation zwischen Headergröße und Länge der Nutzdaten ist ein Kompromiss. Kleine Nutzlast (Pay Load) vermindert die Effektivität, große Nutzlast ist mit großen Zeitverzögerungen verbunden, denn die Daten müssen erst gesammelt werden. Diese Aspekte machten es zunächst schwierig, die Netze für Videoanwendungen einzusetzen

10.3.3.1 Ethernet

Das im Computerbereich am weitesten verbreitete lokale Netzwerk wird Ethernet genannt und trägt die offizielle Bezeichnung IEEE 802.3x. Es entwickelte sich von Standard- und Fast-Ethernet mit Datenraten von 10 bzw. 100 Mbit/s zu den aktuellen Standards Gigabit-Ethernet (GbE) und 10 Gigabit-Ethernet (10 GbE), wobei die Namen die maximale Datenrate beinhalten.

Ethernet wurde zunächst eher zögerlich zur Vernetzung im Videobereich eingesetzt, da es für Echtzeitanwendungen, wie die Übertragung der Video- und Audiodaten, strukturell schlecht geeignet war. Das Problem ist, dass es die Filetransferprotokolle bei Ethernet erlauben, dass alle an das Netz angeschlos-

Tabelle 10.5. Parameter von Ethernet-Verbindungen auf Kupferkabelbasis

Datenrate	Cat.	Adernpaare	Codierung	Bandbreite	erforderl. SNR
10 Mbit/s	3	2	Manchester	16 MHz	15 dB
100 Mbit/s	5	2	MLT3	100 MHz	18 dB
1000 Mbit/s	5e	4	PAM 5	100 MHz	21 dB
10 Gbit/s	6a	4	PAM 16/DSQ128	400 MHz	26 dB
40 Gbit/s		4	PAM 16/DSQ128	1600 MHz	26 dB
40 Gbit/s		4	PAM 32/DSQ512	1200 MHz	32 dB

senen Clients jederzeit Kapazitäten beanspruchen können. Sie benutzen ein Zugriffsverfahren namens CSMA/CD (Carrier Sense Multiple Access/Collision Detection) bei der ein Transferverfahren neu begonnen werden muss, wenn ein beteiligter Teilnehmer mittels Abhörens (Carrier Sense) feststellt, dass das Netz belegt ist (Collision Detection). Das Verfahren sieht vor, dass der anfragende Teilnehmer dann eine unbestimmte Zeit wartet und die Anfrage erneut startet. Damit entsteht ein großer System-Overhead und es kann keine Dienstgüte garantiert werden.

Mit der Einführung von GbE und 10 GbE konnte dieser Nachteil aufgehoben werden, indem Switches eingesetzt werden, die die Kollisionen verhindern. Switches zerlegen das Netz in Segmente für jeweils ein Endgerät, das dann im Vollduplexbetrieb gleichzeitig senden und empfangen kann. Das CSMA/CD-Zugriffsverfahren wird dabei vor allem aus Kompatibilitätsgründen weiter beibehalten.

Die Übertragung selbst kann drahtlos oder drahtgebunden sein oder sie erfolgt auf Basis von Glasfasern. In LAN-Umgebungen werden sehr häufig verdrillte Kupferkabel mit Adernpaare (Twisted Pairs) und RJ45-Stecker benutzt. Die Wicklung der Adern ist dabei ist so ausgeführt, dass das Übersprechen und die Störausstrahlung minimiert werden. Die Leitungen sind in verschiedene Kategorien eingeteilt. Für GbE und 10 GbE sind Kabel nach CAT5 und CAT6 spezifiziert, die Grenzfrequenzen von 100 MHz bzw. 250 MHz aufweisen. Für das GbE werden z. B. vier Adernpaare verwendet. Die Daten werden dabei pulsamplitudenmoduliert mit 5 Zuständen übertragen (PAM5), sodass bei einem Grundtakt von 125 MHz pro Zustand ein Byte übertragen werden kann.

Obwohl noch weiter gesteigerte Bitraten vor allem mittels Glasfaserverbindungen realisiert werden, kann abgeschätzt werden, dass mit Kupferkabeln bei Längen bis zu 50 m mit modernen Modulationsverfahren Datenraten bis zu 40 Gbit/s erreicht werden können. Tabelle 10.5 zeigt in der Übersicht auch die dabei entstehenden Bandbreiten und die dafür erforderlichen Störsignalabstände [53].

Wegen der großen Marktdurchdringung und den damit verbundenen günstigen Preisen ist Ethernet einer schnellen Weiterentwicklung unterworfen und es werden immer neue Anpassungen entwickelt, die insbesondere den Broadcastanforderungen, nämlich der Integration von Echtzeitanwendungen mit geringer Latenz und der Betriebssicherheit, immer mehr gerecht werden. Entsprechend ist zu beobachten, wie GbE und 10 GbE in immer größerem Maße für viele Videovernetzungen eingesetzt werden. Man kann sogar behaupten, dass

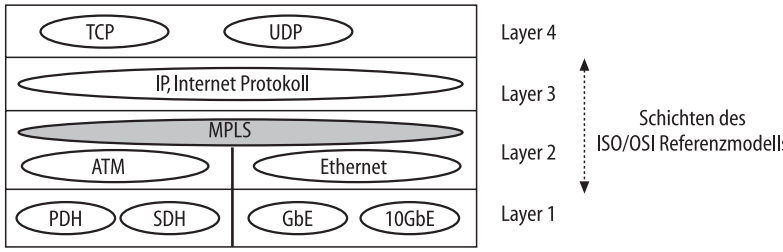


Abb. 10.48. MPLS in der Schichtenstruktur des Refernzmodells

mit dem Erreichen der hohen Bitraten ein genereller Wandel im Videoproduktionsbereich eintritt, der Produktionen erheblich kostengünstiger und daneben auch flexibler werden lässt. Der Wandel begann mit der Loslösung der Datenformate von Datenträgern, verbunden mit der Tatsache, dass Videodaten nur noch wenig spezifische Unterschiede zu anderen Digitaldaten aufweisen. Als Speicher kommen schon seit längerem viele videounabhängige Systeme der IT-Technik zum Einsatz und beim Datenaustausch wird nach Möglichkeit darauf geachtet, die Daten nicht zu recodieren, d. h. also ein Kompressionsformat wie MPEG-2 oder AVC-I über möglichst weite Strecken beizubehalten. Die Verteilung von MPEG-2-Daten mittels ASI oder SDTI über die SDI-Infrastruktur ist dafür ein Beispiel.

Auf der anderen Seite entstand im Bereich der Distribution mit der Videodatenverteilung im Internet auf Basis des Internetprotokolls die Möglichkeit, Programme mit Bitraten zwischen 2 und 20 Mbit/s mittels IPTV o. ä. als sogenanntes Streaming sicher zu übertragen. Dafür wurden neben dem etablierten File Transfer Protokolle wie UDP und RTP eingesetzt werden, die Multicastverbindungen ermöglichen (s. Kap.4 und Abb. 10.48).

Aufgrund der Verfügbarkeit der schnellen Netze können nun die genannten Aspekte zusammengebracht werden, sodass auch die weniger stark datenreduzierten Signale des Produktionsbereichs übertragen werden können. Selbst wenn die Datenrate von 10 GbE nicht völlig ausgeschöpft werden kann, so sind in einer solchen Verbindung z. B. mehrere 3G-Datenströme für unkomprimierte 1080p/50-Signale ebenso übertragbar wie bis zu 80 HD-Signale, die in sehr hoher Qualität z. B. mit AVC-I 100 codiert sind. [35]

Zu den videospezifischen Anforderungen gehört die Betriebssicherheit. Ein Ausfall eines Steuercomputers ist im Broadcastfall nicht akzeptabel. Weiterhin die Echtzeitfähigkeit und eine geringe Latenz. Die Echtzeitfähigkeit konnte mit schnellen Netzen wie GbE für datenreduzierte Signale schon recht früh erreicht werden. Daher entwickelte sich die IP-basierte Netzwerkübertragung zuerst im Postproduktionsbereich, wo an die beiden anderen Forderungen nicht allzu strenge Maßstäbe gelegt werden müssen. Die Weiterentwicklung erlaubt es heute, den Einsatz der Netzwerktechnik in immer zeitkritischere Bereiche auszuweiten. Unter sehr kontrollierten Netzwerkbedingungen können Latenzen von weniger als der Dauer eines Bildes erreicht werden.

Die Erfüllung der videospezifischen Anforderungen gelingt durch spezielle Verfahren, die meist auf der Protokollebene realisiert werden. Zunächst ist es

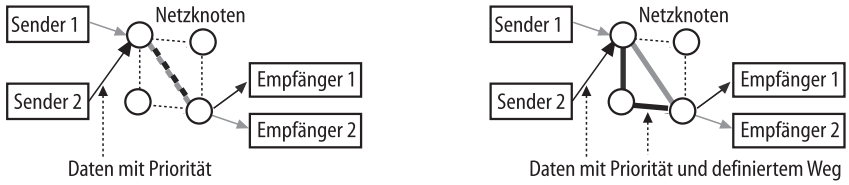


Abb. 10.49. Priorisierung von Daten und definierte Wegführung im Netz

wichtig, die bei Ethernet möglichen Störungen zu unterbinden, die aufgrund von Wartezeiten durch Anfragen neuer Netzteilnehmer entstehen. Dafür kann das Netz abgeschottet und spezifisch nur für bestimmte Dienste genutzt werden, was die Flexibilität jedoch manchmal eingeschränkt.

Daher wird in universell genutzten Netzen vor allem mit einer Priorisierung von Daten gearbeitet, d. h. dass den kritischen Datenströmen, die unterbrechungsfrei übertragen werden müssen, Vorrechte gegenüber anderen Datenpaketen eingeräumt werden. Diese Maßnahme verhindert jedoch nicht, dass das System insgesamt dadurch an seine Grenze kommen kann, dass die Signalpakete eines Videofiles über verschiedene Pfade des Netzwerkes laufen, von denen einer überbelegt sein kann [107]. Abhilfe schaffen hier Verfahren, die die Wege, in denen das Signal innerhalb des Netzes läuft, kontrollieren können. Ein Beispiel dafür ist das MPLS (Multiprotocol Label Switching), das dem Layer 2 im ISO/OSI-Referenzmodell zugeordnet werden kann (Abb. 10.48). Bei diesem System erhalten die Pakete eines Signals jeweils ein Label, das im Netzwerk entlang des vorgesehenen Pfades (Label Switched Path, LSP) über alle Knotenpunkte (Router) weitergereicht wird und auf diese Weise einen eindeutigen und kontrollierten Signalpfad für die priorisierten Daten etabliert. Durch sogenannte Policer kann dann zusätzlich sichergestellt werden, dass das Signal die Kapazität des angelegten Pfades nicht übersteigt (Abb. 10.49).

Mit der Nutzung von IP-basierten, vernetzten Produktionsumgebungen findet insgesamt ein Paradigmenwechsel statt, der von den synchronen Übertragungsverfahren wie SDI oder SDH aus dem Telekommunikationsbereich, die eine Erbschaft der dort vermittelten Telefonverbindungen sind, wegführt, hin zu paketvermittelten Netzen mit speziellen Mechanismen, die den Anforderungen aus dem Broadcastbereich gerecht werden. Dabei ergeben sich sehr viele Vorteile: Echtzeitdaten und andere Daten können im gleichen Netz verteilt werden. Abbildung 10.50 zeigt, wie dabei z. B. die Kapazität einer 10 GbE-Verbindung dynamisch zugeteilt wird, indem allgemeinen Daten, die z. B. auf File

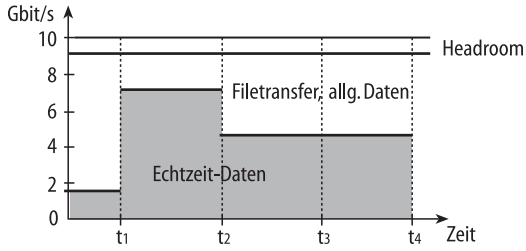


Abb. 10.50. Dynamische Datenraten

Transfer Protokollen (FTP) beruhen, in Abhängigkeit von den priorisierten Echtzeitdaten mehr oder weniger Datenrate eingeräumt wird.

Die Qualität der übertragenen eigentlichen Videodaten hängt nicht vom Netz ab. Relativ stark komprimierte AVC-Daten sind ebenso transportierbar wie hochbitratige 4k-Daten aus einer digitalen Filmumgebung. Die Videosignalqualität ist nur vom Codec bestimmt. Bereits auf IP basierende Daten können sehr einfach in die Netzstruktur integriert werden, wie z. B. IP-basierte Daten aus dem Videoproduktionsbereich, die dort seit langem eingesetzt werden, um Broadcastgeräte zu konfigurieren, zu steuern und zu warten.

Das IP-Netz erlaubt auch eine einfache Verteilung von Videoinhalten an sehr viele Bearbeitungsstationen in einem Produktionskomplex. Es ist einfach, ein ganzes Programmbouquet in das Netz zu senden, ohne dass es einer Station mittels Kreuzschienen separat zugeordnet werden muss. Ein Nutzer kann sich einfach aufschalten und mit Protokollen wie IGMP (Internet Group Management Protokolle) kann mittels intelligenter Switches dafür gesorgt werden, dass Multicastströme nur an Empfänger gesandt werden, die diese auch anfordern. Sie müssen dann nicht den gesamten Datenstrom verarbeiten und so kann z. B. ein 100 Mbit/s-Interface für den Empfang ausreichen. Das Verfahren ermöglicht z. B. auch die einfache Integration von MPEG-Transportströmen, die mittels ASI-Interfaces sonst über SDI-Infrastrukturen geschickt werden. In diesem Zusammenhang steht auch der Vorteil von IP-basierten Netzwerken für die Bereitstellung sehr große Verteilstrukturen. Z. B. ist eine Verteilungsmatrix mit 512 x 512 Koppelpunkten auf IP-Basis erheblich einfacher und kostengünstiger zu realisieren als mit einer SDI-Kreuzschiene die mit 512 x 512 Koppelpunkten an ihre Größengrenze stößt. Vor allem wenn man bedenkt, dass dabei mit der Kreuzschiene oft SDI-Signale geschaltet werden, die vorher in der datenreduzierten Ebene vorgelegen haben und auch wieder in diese konvertiert werden.

Angesichts der Vorteile und der Flexibilität scheint sich IP over Ethernet auch im Weitverkehrsbereich gegenüber ATM durchzusetzen. Dabei ist nicht zuletzt auch der Kostendruck ein erhebliches Argument für die Einführung IP-basierter Netzwerke, denn Geräte aus dem IT-Bereich sind in der Regel preisgünstiger als videospezifische Geräte aus dem Broadcastsektor.

Den vielen Vorteilen steht gegenüber, dass es im IT-Bereich ein unübersichtliche Marktangebot gibt und geeignete, betriebssichere Switches und Protokolle nicht einfach zu finden sind. Die Geräte und die Konfigurationen müssen sorgfältig gewählt und gewartet werden. IT-Abteilungen gewinnen somit in den Rundfunksanstalten eine immer größere Bedeutung.

Channels			Networks			
IPI	SCSI	HIPPI	802.2	IP	ATM	FC4
Common Service						FC3
Framing Protocol / Flow Control						FC2
Encode / Decode						FC1
133 Mbit/s		266 Mbit/s		531 Mbit/s	1062 Mbit/s	FC0

Abb. 10.51. Schichtenstruktur bei Fibre Channel

10.3.3.2 Fibre Channel

Fibre Channel (FC) ist vor allem durch seine hohe Flexibilität gekennzeichnet. Es erlaubt die Verwendung verschiedenster Protokolle und physikalischer Verbindungen in unterschiedlichen Topologien und ist somit auch nicht etwa nur auf die namengebende Glasfaserverbindungen (Fiber) beschränkt. FC ist sowohl als Ein- und Ausgangskanal als auch als Netzwerk verwendbar.

Der Begriff Kanal beschreibt einen Datentransfer mit direkter Verbindung. Die Unterscheidung von Daten, die eindeutig nur für die beiden Endgeräte bestimmt sind, von Netzwerkdaten, die für alle Teilnehmer bestimmt sein können, geschieht mit Hilfe von Protokollen, die die Kommunikationssteuerung festlegen. Bei Fibre Channel werden die Geräte im Netzwerk als Knoten (node) bezeichnet, die über Schnittstellen (Ports) mit dem Netz verbunden sind. Jeder Port hat einen Ein- und einen Ausgang. Eine Verbindung (link) bezieht sich auf zwei Knoten, während eine Topologie die Verknüpfung von Links darstellt.

Für Fibre Channel sind drei Topologien definiert: Als erstes Point to Point, als einfachste Verbindung zwischen zwei Teilnehmern. Zweitens die sogenannte Arbitrated Loop, die ähnlich wie Token Ring eine Netzverbindung für bis zu 126 Teilnehmer bildet. Bei Einsatz von Kupferleitungen darf dabei die Entfernung zum FC-Verteiler (Hub) bis zu 30 m betragen, mit Glasfaserleitungen sind Entfernung bis 10 km möglich. Arbitrated Loop ist für Videoverbindungen, wie z. B. den Datenaustausch zwischen zwei Editing-Stationen, gut geeignet. FC arbeitet dabei bidirektional mit voller Bandbreite in beiden Richtungen. Die dritte Topologie wird als Fabric bezeichnet, sie bietet höchste Flexibilität. Ein Umschalter (Router) arbeitet hier als eine Art Kreuzschiene und kann Point to Point und Arbitrated-Loop-Verbindungen herstellen, sodass insgesamt eine sehr große Teilnehmerzahl über große Strecken verbunden werden kann [72]. Auch bei diesem Netzwerk sind wieder funktionale Schichten definiert. In der obersten, anwendernahen Schicht können verschiedene Datenprotokolle (SCSI, ATM, IP...) verwendet werden (Abb. 10.51).

Fibre Channel stellte schon früh Datenraten von mehr als 4 Gbit/s zur Verfügung. Die Datenübertragung beruht auf Zellen, oder Frames, die mit einem Header versehen sind. Abbildung 10.52 zeigt das Frame-Format. Die Nutzlast in einem Frame kann bis zu 2 kByte betragen, damit ist FC gut für den schnellen Transport großer Datenfiles geeignet. Auch hier kann für den Filetransfer das Internetprotokoll zum Einsatz kommen. Das Streaming wird wie bei ATM mit Hilfe von Timing-Referenzsignalen erreicht, diesbezüglich ist ein FC-A/V-Protokoll entwickelt worden. Trotz seiner Flexibilität und guter Performance verliert Fibre Channel gegenüber Ethernet an Bedeutung.

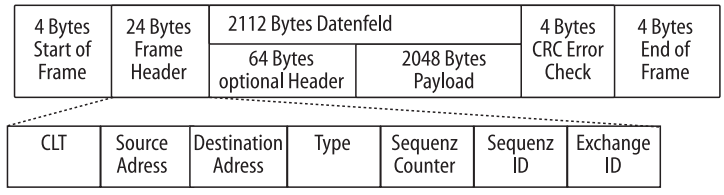


Abb. 10.52. Frame Format von Fibre Channel

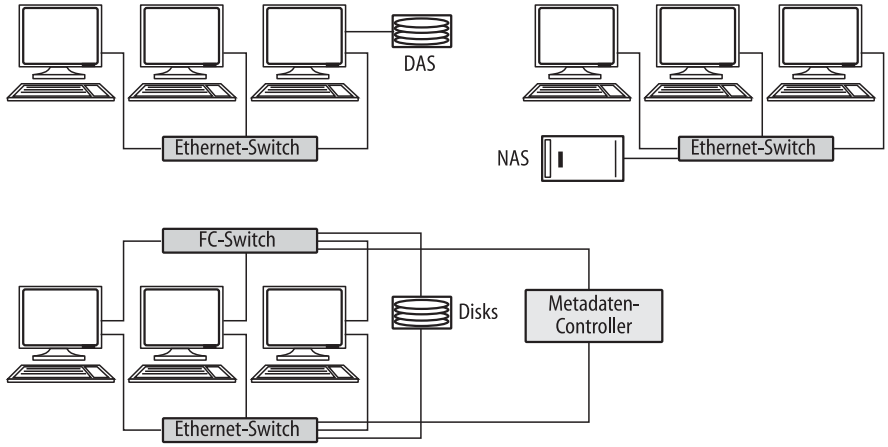


Abb. 10.53. Grundaufbau der Netzwerk-Speichersysteme DAS, NAS und SAN

10.3.3.3 Netzwerk-Speicher

Für Computeranwendungen mit hohen Datenraten, wie sie für die digitale Videoübertragung typisch sind, müssen die Computer mit großen und schnellen Speichern ausgestattet werden. Um dabei auch die Flexibilität zu steigern und die Datenzuordnung zu den Arbeitsplätzen zu vereinfachen sowie einen gemeinsamen Zugriff auf die Daten durch mehrere Nutzer zu ermöglichen, werden die Speicher oft auf verschiedene Weise mit Netzwerken verbunden. Grundsätzlich werden dabei die Verfahren DAS, NAS und SAN unterschieden.

DAS steht für Direct Attached Storage und bezeichnet die einfachste Art der Speicher-Computer-Verbindung (Abb. 10.53 links oben). Der Speicher wird hier ohne Netzwerk direkt angeschlossen und zwar über SCSI/SAS, ATA/SATA, Firewire etc. Der Nachteil eines DAS ist, dass die Daten auf dem Speichervolumen nur für den direkt angeschlossenen Computer (Host) zur Verfügung stehen. Wenn sie nach außen zugänglich gemacht werden sollen, muss der Host als Server im Netz agieren.

NAS bedeutet Network Attached Storage, also ein an ein Netzwerk angeschlossener Speicher (Abb. 10.53 rechts oben). Dieser ist meist als RAID vorkonfiguriert und agiert mit eigenem Betriebssystem als Server. Die Verbindung der Geräte erfolgt oft über Switches im Ethernet. Für den Datenzugriff werden Netzwerkprotokolle genutzt, um ganze Dateien anstelle der bei DAS verwendeten Blöcke zu übertragen. NAS-Systeme sind separate Geräte und einfach zu konfigurieren. Nachteilig ist, dass das Netzwerk die Nutzdaten gleichzeitig mit den Steuerdaten, zur Regelung der Kommunikation für Anfragen an den Server etc. transportiert.

Ein SAN (Storage Area Network) bildet ein echtes Netzwerk von Speicher und Computer (Abb. 10.53 unten). Mittels Switches können alle im Netzwerk eingebundenen Rechner auf eines oder mehrere Speichersysteme zugreifen. Aufgrund der umfangreichen Konfigurationsmöglichkeiten ist eine aufwendige Kontrolle erforderlich, die über eine Management-Software realisiert wird [130].

Physikalisch werden oft zwei Netzwerke genutzt, z. B. Fibre Channel per Glasfaser, für den schnellen Datentransfer und Ethernet per Twisted Pair für die Abwicklung der Datenzugriffe, die hier, wie beim DAS, wieder blockorientiert erfolgen. Insgesamt ist ein SAN die leistungsfähigste Variante der Netzwerk-Speichersysteme, es erfordert aber auch den höchsten Administrationsaufwand.

Ein bekanntes Beispiel für ein SAN ist das Avid Unity Media Network [245]. Es basiert in einer bestimmten Ausbaustufe z. B. auf der Kombination von 4-Gigabit Fibre Channel und Gigabit Ethernet, womit in Echtzeit bis zu 46 Clients bei 25 Mbit/s unterstützt werden oder bis zu acht Dual-Stream Clients bei dem Einsatz von unkomprimierten 10-Bit-HD-Daten. Das System zeichnet sich durch ein gutes Management der Speicherzuteilung aus, die es unproblematisch erlaubt, dass viele Nutzer gleichzeitig mit demselben Material arbeiten [82]. Als bekannte Alternative zu Avid Unity gibt es das System EditShare.

10.3.3.4 Zuspielung per Netzwerk

Der Begriff Zuspielung bezeichnet den Programmaustausch zwischen zusammengehörigen, aber doch weit entfernten Produktionseinheiten. Die Zuspielung erfolgt heute leitungsgebunden oder unter der Bezeichnung Satellite News Gathering (SNG) mittels Satellitenübertragung. Die Zuspielung ist insbesondere für die Landesfunkhäuser der ARD (Arbeitsgemeinschaft der Rundfunkanstalten Deutschlands) von großer Bedeutung, da deren Beiträge zu einem Gesamtprogramm (Das Erste) zusammengefasst werden.

Die Zusammenführung geschieht im Sternpunkt Frankfurt (STP) und wurde zu Zeiten der Analogtechnik über Richtfunkstrecken abgewickelt. Nach der Digitalisierung der Videotechnik wurde der Austausch dann durch das von der Telekom (DTAG) betrieben hybride Breitbandnetz der ARD (Hybnet) realisiert. Das Hybnet ist glasfaserbasiert und erlaubte zunächst eine Datentransferrate von 2,5 Gbit/s. Es ist als ein Dauerleitungsnetz für alle Bedarfe der Anstalten konzipiert, d. h. für die Zuführung von Hörfunk- und Fernsehprogrammen der ARD via ATM ebenso wie für generelle Kommunikationsbedarfe innerhalb des Verbundes. Dabei wurde auch die allgemeine Datenübertragung in das Netz integriert. Abbildung 10.54 zeigt die Netztopologie. Zu den dort dargestellten

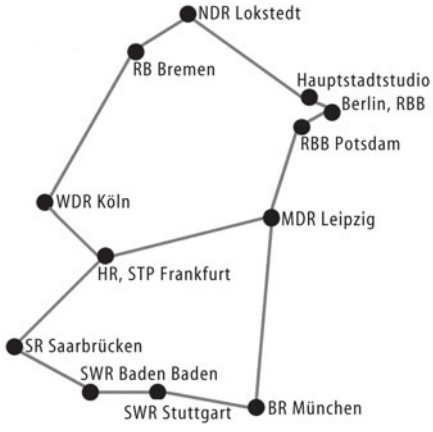


Abb. 10.54. Netztopologie des Hybnet zur Verbindung der ARD-Rundfunkanstalten

Tabelle 10.6. Signalformen im Hybnet-Neu

Diensttyp	Anwendung	Schnittstelle	Datenrate	Anzahl	Laufzeit
F1, F2	Fernsehen, SDTV	SD-SDI	270 Mbit/s	165	17 ms
F5, F6	Fernsehen, HDTV	HD-SDI	1, 485 Gbit/s	85	17 ms
H1 H2	Hörfunk-Echtzeit	AES/EBU	3 Mbit/s	218	17 ms
T2.4	Ethernet-Tunnel	GbE	1 Gbit/s	22	8,9 ms
T1.2	Ethernet-Tunnel	10 GbE	10 Gbit/s	7	8,5 ms
T2.1	STM-Tunnel	STM	155 Mbit/s	2	15,5 ms

Fernsehanstalten kommt noch die Anbindung der Deutschen Welle und des Deutschlandradio Berlin hinzu. Der Programmaustausch für die Fernsehprogramme erfolgte beim Hybnet vorwiegend via DVB-Codecs, d. h. mit MPEG-2 bei einer Datenrate von 45 Mbit/s im 422p@ML-Format. [150]

Ab 2010 wurde das Hybnet zum Hybnet-Neu weiterentwickelt, wobei es möglichst unabhängig von speziellen Codecs werden sollte. Es basiert auf einem Glasfasernetz von Vodafone und ist geeignet, die in Tabelle 10.6 dargestellten Dienste zu realisieren, und damit Audio-, Video- und Datendienste mit Datenraten zwischen 2 Mbit/s und 10 Gbit/s zu integrieren. Das Netz ist damit nun auch in der Lage, u. a. unkomprimierte HD-Videosignale zu übertragen.

Das Hybnet-Neu sollte flexibel uni- und bidirektionale Übertragungen mittels Uni- und Multicast ermöglichen und dabei weitestgehend ausfallsicher sein. Aufgrund der hohen Gesamtdatenrate kommt als Hauptleitungsnetz (Backbone) die Glasfasertechnik zum Einsatz. Die Fasern werden dabei mittels Wavelength Division Multiplex (WDM), also durch Verwendung verschiedener Lichtwellenlängen, effektiv genutzt, sodass insgesamt 98 sogenannte Tunnel realisiert werden konnten, die jeweils eine Kapazität von 10 Gbit/s aufweisen.

In den über dem physical Layer liegenden Netzebenen kommen im Layer 2 Ethernet und im Layer 3 IP-Protokolle zum Einsatz. Im Adaptionslayer werden die Daten an die gewünschten Schnittstellen, wie HD-SDI oder ASI, umgesetzt. Die geringe Latenz des Netzes konnte mittels des im vorigen Abschnitt angeführten MPLS im Layer 2 erreicht werden, da hier die Daten anhand der Labels geleitet werden und die IP-Header nicht ausgelesen werden müssen. Mit Hilfe der MPLS-Technik kann auch die Qualität des Service sichergestellt werden. Entsprechend werden beim Hybnet 2 drei Klassen gebildet; für anspruchsvolle Audio/Video-Echtzeitanwendungen, für allgemeine Datentransferaufgaben und für simple Managementaufgaben. Auch für die größte Entfernung, die in diesem Netz bei ca. 1500 km liegt, können Laufzeiten von nur 8,5 ms erreicht werden. Diese Werte erhöhen sich bei der Ausgabe der konvertierten Daten ggf. durch verschiedene Signaladaptionen (Tabelle 10.6) [35].

Die hohe Verfügbarkeit des Netzes wird durch mehrere Redundanzstufen erreicht. Zunächst wird das Netz physikalisch doppelt aufgebaut wobei die Leitungen über räumlich entfernte Trassen laufen (A- und B-Netz). Die Übergabepunkte an den Funkhäusern wurden von zwölf auf sieben verringert und jeder wird sowohl mit seinen zwei Nachbarn als auch mit dem Sternpunkt Frankfurt direkt verbunden. Die Signale werden stets doppelt verteilt. Eine Havarieumschaltung auf das Netz B erfordert weniger als 20 ms.

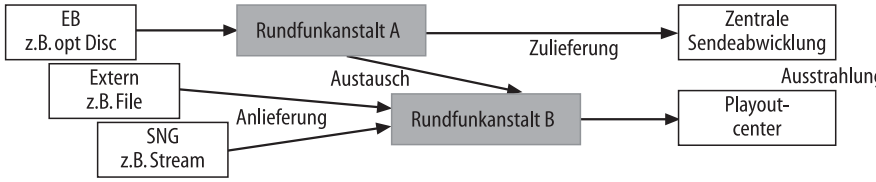


Abb. 10.55. Programmaustausch zwischen Produktionseinheiten

10.4 Produktionsanforderungen

In diesem Abschnitt werden die Anforderungen zur Gewährleistung einer hohen Bild- und Tonqualität über die gesamte Signalkette in einem Videoproduktionsbereich erörtert: von der Entstehung (Akquisition) über die Anlieferung und Zuspielung (Contribution), die Bearbeitung (Produktion) bis hin zur Verteilung der Signale an den Endkunden oder TV-Zuschauer (Ausstrahlung, Distribution) (Abb. 10.55). Dabei werden technische Parameter dargestellt, die zur Erreichung der konsistenten Qualität insbesondere für den Programmaustausch zwischen Produktionsstätten und Funkhäusern festgelegt werden.

Solche Vereinbarungen sind nicht universell bestimmbar. An eine digitale Filmproduktion werden andere Ansprüche gestellt als an eine TV-Produktion, die Anforderungen ändern sich zudem ständig. Ein gutes Beispiel dafür ist der Übergang zu hoch aufgelösten Bildern beim Umstieg von SDTV zu HDTV. Für die Praxis ist bei den Festlegungen außerdem zu beachten, dass die erforderlichen Geräte für die zu definierenden Formate erhältlich sind, dass die geprüft und verlässlich sind und auch längerfristig verfügbar. Entsprechend müssen die Anforderungen ständig neuen Gegebenheiten angepasst werden.

10.4.1 Produktionsempfehlungen der EBU

Für den Bereich der Fernsehproduktion gibt es die Empfehlung R132 der EBU (European Broadcasting Union) aus dem Jahre 2011. Als übergreifende Empfehlungen ist hier angegeben, dass neue Produktionsstätten für HD ausgelegt und möglichst auch 3G-tauglich sein sollen. Die subjektive Qualität, sollte als end-to-end-Qualität für die gesamte Strecke geprüft werden. Als gesamte Qualitätsminderung für das Bild wird höchstens ein Viertel einer Stufe aus der fünfstufigen Qualitätsskala (s. Abschnitt 10.2) zugelassen [52].

Bezüglich der Bildqualität ist gefordert, dass die Produktionen handwerklich gut gemacht sein sollen, vor allem hinsichtlich der Belichtung und Schärfe. Dem letzten Punkt ist nach der Etablierung von HDTV viel mehr Beachtung zu schenken, als zu Zeiten in denen SDTV produziert wurde und Unschärfen viel weniger auffielen. Neben der eigentlichen Bildqualität sollte ein Bearbeitungsspielraum beachtet und gewährleistet werden (processing headroom). Als Konsequenz daraus wird empfohlen, keine Signalformen zu benutzen, deren Abstaststrukturen geringere Verhältnisse als 4:2:2 aufweisen. Bezüglich der Quantisierung sollten durchgängig 10 Bit pro Sample verwendet werden.

Tabelle 10.7. Empfehlungen der EBU R132

Bereich	Anforderung	Parameter
Allgemein	Prüfung der End-to-end-Qualität	max. Verlust: 1/4 Stufe von 5 Stufen
Bildformat	1080/25, 1080p/50,	unkomprimiert, 4:2:2, 10 Bit
Konversion	mit höchster Qualität	Up, Down bewegungskompensiert
Datenreduktion	I-Frame-only / Long-GOP	mind. 100 Mbit/s / 50 Mbit/s
Tonproduktion	unkomprimiert oder Dolby E	4- oder 8-kanalig

Die Verwendung von interlaced-Signalen sollte nach EBU R132 in Zukunft vermieden werden, da sie problembehaftet sind und ausschließlich den Vorteil einer Datenreduktion um den Faktor 2 bieten, der mit modernen Algorithmen erheblich effektiver erreicht werden kann.

Datenreduzierte Signalformen sind für das Bild im Produktionsbereich akzeptiert. Um hier den Qualitätsanforderungen gerecht zu werden, sollten I-Frame-only-Codecs mit mindestens 100 Mbit/s oder Long-GOP-Codecs mit Datenraten von mindestens 50 Mbit/s verwendet werden. Weniger strenge Vorgaben gibt es für den Bereich des schnellen News-Akquisition bei dem auch Geräte mit 4:2:0-Abtastung bei 50 Mbit/s bzw. 35 Mbit/s für I-Frame-only bzw. Long-GOP-Codecs verwendet werden dürfen. Im Bereich der Nachbearbeitung sind die Anforderungen dagegen strenger, hier werden I-Frame-only-Codecs mit mehr als 160 Mbit/s oder unkomprimierte Videodaten gefordert. Generell wird auch hier hervorgehoben, dass ein Datenformat durchgängig beibehalten werden sollte.

Bezüglich der Konversion zwischen verschiedenen Videostandards sowie Geräten, die als Up-Converter von SD zu HD oder Cross-Converter von 1080i nach 720p arbeiten, wird empfohlen, dass diese unbedingt mit modernen hochwertigen Algorithmen ausgestattet sein sollten, die bewegungskompensiert arbeiten (Tabelle 10.7). Konverter, die in Aufzeichnungssystemen oder Editingstationen enthalten sind, werden als qualitativ zu schlecht eingestuft, die Konversion mit diesen Geräten sollte vermieden werden [52].

Neben den Empfehlungen für den Bildbereich gibt es noch viele weitere. So wird auf den Komfort bei Lesen von Buchstaben auf dem TV-Bildschirm geachtet und empfohlen, dass diese mindestens eine Höhe von 40 Zeilen des HD-Bildschirms haben sollen und dass ein Text den Bildschirm nicht bis zum Rand füllen sollte (safe area).

Für den Audiobereich wird empfohlen, ohne Datenreduktion zu arbeiten. Die Verwendung von 5.1-Surround für HDTV wird nahegelegt. Dabei wird die milde Datenreduktion der Dolby E-Verfahren erlaubt, allerdings sollten Codierungs- Decodierungsketten sorgfältig implementiert werden. Weitere Empfehlungen dazu gibt es im White Paper WHP 175 der BBC. Für die Tonspurbelegung auf Recordern und für den Filetransfer gibt es die Empfehlungen EBU R48 und R123. Bezüglich des Tons gilt es in Zeiten der Digitalverarbeitung, auf die Laufzeiten und insbesondere auf die Synchronisation zu achten. Die Empfehlung EBU R37 legt hier eine maximale Abweichung fest, die für die ganze Übertragungskette ≤ 40ms für den Fall vorsieht, dass der Ton eher als das Bild (early) wiedergegeben wird und ≤ 60ms für den Fall Bild vor Ton (late).

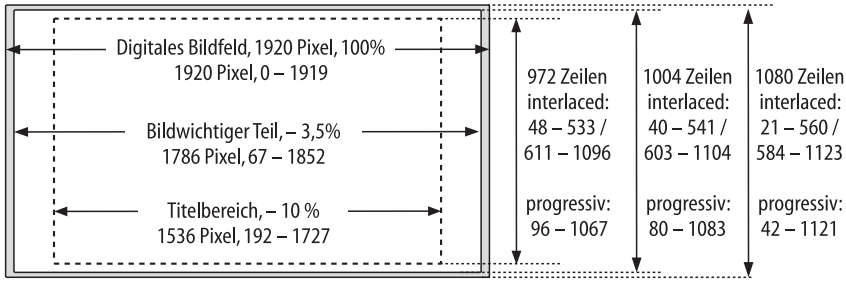


Abb. 10.56. Bildfeldfestlegungen für die Abtastformate 1080i/25 und 1080p/25 sowie 1080 psf/25

10.4.2 Technische Richtlinien - HDTV für ARD, ZDF und ORF

In der Festlegung „Technische-Produktions-Richtlinien-Fernsehen" (TPRF) werden die Vorgaben zur Herstellung von Fernsehproduktionen für ARD, ZDF und ORF definiert. Sie basieren auf den EBU-Empfehlungen und spezifizieren sie unter Anpassung an spezielle Gegebenheiten und unter Berücksichtigung am Markt erhältlicher und bewährter Geräte und Codecs.

Der folgende Abschnitt bezieht sich auf die TPRF von 2011, die Anforderungen für HD-Produktionen definieren. Auch wenn sich die Festlegungen im Einzeln zukünftig ändern, geben sie doch einen Einblick und zeigen, welche Aspekte bei derartigen Richtlinien generell bedacht werden müssen [112].

10.4.2.1 TPRF - Bild

Als Produktionssignal und als Basissignal für eine Datenreduktion ist das digitale HD-Komponentensignal bestimmt. Als Abtastformate sind 1080i/25, 1080p/25 und 720p/50 zugelassen. Der Einsatz des Letzteren scheint sich immer mehr auf den Distributionsbereich zu beschränken. Solange 1080p/50 noch nicht etabliert ist, wird vorwiegend mit 1920 x 1080 Pixels bei 25 Bildern/s produziert. Dabei wird bei typischen TV-Produktionen meist die relativ hohe Bewegungsauflösung mit 50 Halbbildern in der Form 1080i/25 gewählt. Bilder dieses Formats können mit hochwertigen Konvertern in sehr hoher Qualität in das Format 720p/50 gewandelt werden. Die Geräte sollten die Umsetzung dabei bewegungskompensiert vornehmen. Bei TV-Beiträgen aus dem szenischen Bereich wird, in Anlehnung an die Kinobildanmutung, aus ästhetischen Gründen meist mit 1080p/25 gearbeitet. Bilder in diesem Abtastformat können auf zwei Halbbilder verteilt und als 1080psf/25-Signal übertragen werden. Für Produktionen im internationalen Kontext ist festgelegt, dass eine Bildrate von 24 fps/1,001 gegenüber 30 fps/1,001 bevorzugt werden sollte.

Für das Bildfeld ist definiert, dass es keinen Overscan wie bei SD-Signalen gibt, sondern dass der aktive Bildbereich komplett gefüllt ist. Für Titel und Grafiken ist zu beachten, dass ein unbeschriebener Randbereich bestehen bleibt, sodass der bildwichtige Teil um 3,5% und der Titelbereich auf beiden Seiten um 10% kleiner ist als das Bildfeld. Abbildung 10.56 zeigt, welche Pixel für die Abtastformate 1080i/25 und 1080p/25 dabei im Einzelnen definiert sind.

Tabelle 10.8. Produktionsplattformen nach TPRF

Plattform	Mainstream	Premium
Beispiele	Tagesschau, Talkshow	Tatort, Olympiade
Bildformat	1080i/25, 720p/50	1080p/25, 1080p/24, 1080i/25
Kameraparameter	2/3"-Sensor, 1920 x 1080 Pixel	35-mm-Filmsensor, 1920 x 1080 Pixel
Kamerabeispiele	Sony HDC 2500, PMW 200	Arri Alexa, Sony F5/55
Datenreduktion	AVC-I-100, XDCAM HD422	wenig/un-komprimiert, 4:2:2, 10 Bit
Speichermedium	XDCAM-Disc, P2-Karte	HDCAM-SR

Bei einer Textdarstellung sollen große und klare Schrifttypen, möglichst ohne Serifen, verwendet werden. Während deren Höhe bei SDTV mindestens 20 aktive Zeilen umfasste, wird für den HD-Bereich die doppelte Anzahl, also mehr als 40 aktive Zeilen empfohlen. Zur Sicherstellung einer guten Lesbarkeit von horizontal oder vertikal bewegten Schriften (Rolltitel) wird auf die Testpattern in der Recommendation EBU R3325 verwiesen. Bei Bildverarbeitungen, die eine Bildverschiebung hervorrufen, müssen fehlende Bildteile ergänzt werden, damit das Bildfeld immer gänzlich gefüllt ist.

HD-Programme sollten durchgängig in einem der drei genannten Formate produziert werden. Up-konvertiertes SD-Material darf bis zu maximal 25% der Programmdauer enthalten sein. Als Up-Konverter sollten hochwertige Geräte zum Einsatz kommen die bewegungskompensiert arbeiten. Die Art der Einpassung des 4:3-SD-Bildes in das 16:9-HD-Format wird den Redaktionen überlassen. Es wird darauf hingewiesen, dass für die SD-HD-Konversion die Veränderung des Farbraums beachtet werden muss.

Für die Verteilung unkomprimierter Signale sind die Schnittstellen HD-SDI, Dual link und 3G-Schnittstellen erlaubt. Bezüglich der Pegel ist diesbezüglich festgelegt, dass der dort definierte Headroom nicht zur Ausweitung der Signaldynamik genutzt werden darf, wie es zum Teil bei der Postproduktion von digitalen Filmdaten gemacht wird, sondern dass dieser Bereich nur für kurzfristige Pegelüberschreitungen vorgesehen ist, wie sie z. B. aus dem Einschwingverhalten von Filtern oder plötzlichen Kameraschwenks resultieren. Es wird angegeben, dass die Kontrolle des RGB-Signals mit einem Waveformmonitor die Gültigkeit des Signals ausreichend sicherstellen kann.

Die Produktionsvorgaben TPRF erlauben auch den Einsatz von Video-Datenreduktion. Zur Differenzierung der diesbezüglichen und weiterer Anforderungen werden zunächst zwei Produktionsplattformen definiert: die Mainstream-Produktion und die Premium-Produktion. Mainstream-Produktionen betreffen das Fernsehtagesgeschäft, bei dem neben der Qualität wirtschaftliche Aspekte eine größere Rolle spielen als bei Premium-Produktionen. Der Begriff Premium-Produktion bezeichnet Beiträge, für die ein größerer Aufwand betrieben wird, da sie im Programmablauf sogenannte Highlights darstellen. Dazu gehören szenische Produktionen wie der ARD-Tatort oder große Unterhaltungsshows und internationale Sportveranstaltungen (Tabelle 10.8).

Für Mainstream-Produktionen werden zwei Datenreduktionsverfahren festgelegt: der Codec AVC-I 100, der mit 4:2.2 Abtastung im I-Frame-only-Modus bei 112 Mbit/s arbeitet, sowie das Long-GOP-Format XDCAM HD422, das mit 8

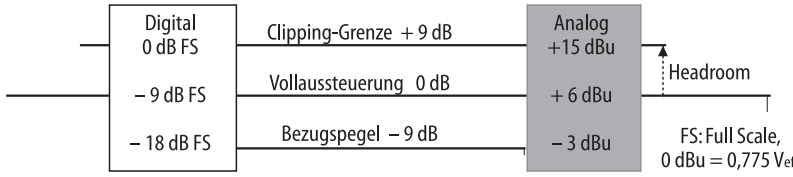


Abb. 10.57. Aussteuerung von digitalen Audiosignalen

Bit/Sample und 50 Mbit/s auskommt. Kameras für diese Produktionsplattform müssen nach TPRF mit Bildwandlern von mindestens 2/3"-Größe ausgestattet sein, die nativ mindestens 1920 x 1080 Pixel aufweisen. In Frage kommen z. B. die typischen Systemkameras aus dem Studiobereich mit drei 2/3"-Sensoren, es sind aber auch Single-Sensor-Kameras erlaubt.

Premium-Produktionen erfordern in der Regel 10 Bit pro Abtastwert und eine geringe oder keine Datenreduktion. Als Aufzeichnungsformat ist HDCAM SR spezifiziert, das eine Datenrate von 440 Mbit/s bietet. Wenn auf dieser Basis gearbeitet wird, ist die parallele Bereitstellung der Programme in einem der Mainstreamformate gefordert. Für den zukünftig zu definierenden File-Austausch bei Premium-Produktionen sind AVC-Intra und -Ultra sowie die Intermediate-Codecs DNxHD und Apple ProRes in der Diskussion.

Auch die Kameras sollten den höheren Ansprüchen an die Bildgestaltung genügen. Hier kommen vor allem Kameras, wie z. B. die Arri Alexa, in Frage, die mittels sehr guter Objektive eine herausragende Bildqualität erzielen und mit dem großen Bildsensor, der dem 35 mm-Filmformat entspricht, eine hohe selektive Schärfe bzw. eine viel geringere Tiefenschärfe erzeugen als Kameras mit 2/3"-Sensoren. Meist wird in p/25 oder bei internationalen Kooperation in p/24 gedreht. Als Abtastmuster ist 4:4:4 oder wenigstens 4:2:2 vorgesehen.

10.4.2.2 Richtlinien für den Audiobereich

Zunächst wird festgelegt, dass an die Audioproduktion ebenso hohe Ansprüche gestellt werden wie an die Produktion des Videosignals. Sie müssen in allen Qualitätsaspekten dem Stand der professionellen Studiotechnik genügen. Ton- und Bildgestaltung sollen einander entsprechen. Die verschiedenen Audioteile wie Sprache, Atmosphäre, Musik sollen ein ausgeglichenes Mischungsverhältnis aufweisen, wobei der Sprachverständlichkeit Priorität eingeräumt wird. Pegel- und Lautheitssprünge sind auch beim Wechsel zwischen verschiedenen Programmteilen zu vermeiden.

Abbildung 10.57 zeigt, wie im Digitalbereich, in dem die Verzerrungen sehr deutlich wahrnehmbar sind, Übersteuerungen vermieden werden, indem ein Headroom gegenüber analogen Signalen definiert wird und welchen Betrag dieser nach den TPRF haben soll. Im Zusammenhang mit der Signalaussteuerung wird zudem insbesondere auf die Richtlinie EBU R128 verwiesen in der festgelegt ist, wie die Pegelung neuerdings orientiert an der Lautheit, anstatt wie vorher, gemäß dem Spitzenpegel, stattfinden soll [112].

Bei der Audioproduktion soll ohne Datenreduktion gearbeitet werden. Das Standard-Audioformat ist das 2-kanalige stereofone Signal. Es ist jedoch, gera-

Tabelle 10.9. Tonspurbelegung für Programmaustausch und Anlieferung von Sendebeiträgen

Variante	1	2	3	4	5	6
Bezeich-nung	4-Kanal, IT	4-Kanal, 2. Sprachfass.	4-Kanal, Dolby E	8-Kanal, Standard	8-Kanal, 5.1 Diskret	8-Kanal, 2. Sprachfass.
Audio 1	---------- Stereo Sendeton Links (L) für alle sechs Varianten ----------					
Audio 2	---------- Stereo Sendeton Rechts (R) für alle sechs Varianten ----------					
Audio 3	IT-L	2. Sprach-L	Dolby E	IT-L	Mehrkan-L	Dolby E
Audio 4	IT-R	2. Sprach-R	Dolby E	IT-R	Mehrkan-R	Dolby E
Audio 5				2. Sprach-L	Mehrkan-C	2. Sprach-L
Audio 6				2. Sprach-R	Mehrkan-LFE	2. Sprach-R
Audio 7				Dolby E	Mehrkan-LS	2. Sprach-Dolby
Audio 8				Dolby E	Mehrkan-RS	2. Sprach-Dolby

de für hochwertige HD-Bilder erwünscht, einen 5.1-Mehrkanalton (s. Kap. 5), vorzugsweise parallel zum 2.0-Stereosignal anzubieten. Bei diskreter Übertragung erfordert das 5.1-Surroundsignal jedoch sechs Tonkanäle. Da sehr viele Übertragungs- und Programmaustauschsysteme nur vier Spuren bereitstellen, wurde entschieden, dass für diesen Bereich das Dolby E-Verfahren angewandt werden darf, das es, unter Verwendung einer schwachen Datenreduktion, erlaubt, den Bedarf von sechs auf zwei Audiokanäle zu senken. D. h. der Mehrkanalton kann mittels Dolby E kanalsparend über nur eine AES/EBU-Schnittstelle übertragen, bzw. in ein Stereopaar abgebildet werden.

Für den Programmaustausch ist natürlich von großer Bedeutung welcher Inhalt den verschiedenen Spuren zugeordnet wird. Für die Tonspurbelegung wurden sechs Varianten festgelegt, jeweils 3 für eine 4- und eine 8-kanalige Übertragung. Tabelle 10.9 zeigt die Belegung für alle sechs Fälle. Darin steht IT als Abkürzung für internationalen Ton, d. h. eine Tonmischung, in der kein Sprachkommentar enthalten ist. So kann z. B. bei einem Fußballspiel nur die akustische Atmosphäre des Stadions übertragen werden, und die Kommentare können separat in verschiedenen Sprachen überlagert werden. Die 2. Sprachfassung hat einen ähnlichen Zweck. Sie ermöglicht die Verteilung zweier Beiträge, die einschließlich Sprache fertig abgemischt sind, damit z. B. bei einem fremdsprachigen Film wahlweise der Originalton oder die synchronisierte Fassung gehört werden können. Der Begriff Mehrkanal steht für die diskrete Übertragung der sechs Audiokanäle einer 5.1-Sourroundtonproduktion.

Die TPRF legen fest, dass die Audiodaten mit den Videodaten synchronisiert sein müssen und weisen dabei auf die Empfehlung EBU R83 hin. Das beinhaltet, dass die Audiosignale phasenstarr mit der Taktfrequenz des Bildsignals verkoppelt sein müssen. Bei nicht komprimierten Tonsignalen geschieht dies über den Timecode, bei anderen Audioformaten werden Zeitmarken verwendet.

Bezüglich eines möglichen Bild/Tonversatzes, der besonders störend bei mangelnder Lippensynchronität erlebt wird, gilt, dass ein dem Bild voreilender Ton störender wahrgenommen wird als ein nacheilender. Die Toleranzgrenzen sind so festgelegt, dass, über die gesamte Produktionskette betrachtet, der Ton dem Bild nicht mehr als um 40 ms vor- und nicht mehr als 60 ms nacheilen darf. In Einzelstufen sollten für Ton, der dem Bild vor- bzw. nacheilt die Werte 5 ms bzw. 15 ms eingehalten werden. Dabei sollten signifikante Abweichungen

direkt in jeder Stufe ausgeglichen werden. Kodierte Signale, wie Dolby E, dürfen gegenüber den diskreten Tonkanälen keinen Versatz aufweisen. Untertitel, die ggf. den Ton ergänzen, müssen 2-streifig, d. h. separat vom Bildsignal filebasiert angeliefert werden.

10.4.2.3 Programmaustausch

Der Austausch der Programme kann klassisch erfolgen, indem wie früher Bänder oder bandlose Datenträger benutzt und transportiert werden. Daneben gibt es jedoch heute die Möglichkeit, die Programme als Files über Netzwerke oder per Satellit auszutauschen, wobei bezüglich der Echtzeitanforderungen der File Transfer vom Streaming unterschieden wird.

Für den Austausch mittels Datenträgern legen die TPRF fest, dass HD-Mainstream-Produktionen auf Professional Discs (XDCAM) angeliefert werden können. Hier ist als Codec XDCAM HD422 mit 50 Mbit/s implementiert. Als Alternative können P2-Speicherkarten mit dem AVC-I-100-Codec verwendet werden. Für Premium-Produktionen ist gemäß der TPRF von 2011 nur das HDCAM-SR-Band erlaubt. Für mobile Datenträger ist generell festgelegt, dass im Regelfall das HD-Abtastformat 1080i/25 verwendet werden soll. Die Datenträger sind nur über kurze Zeiträume zu verwenden und stellen kein Archivmedium dar.

Der Datenträgeraustausch soll jedoch die Ausnahme sein, der filebasierte Austausch dagegen der Regelfall. Dabei ist als Fileformat MXF vorgesehen (s. Abschn. 10.1.3). Die Files sollen hier nur das eigentliche Programm enthalten. Ein Vor- und Nachlauf oder ein technischer Vorspann dürfen nicht integriert sein, sie sind nur noch bei bandgestützten Speichersystemen erlaubt. Das File beginnt mit dem ersten Bild und endet mit dem letzten. Die in Tabelle 10.9 dargestellte Tonspurzuordnung muss auch innerhalb der MXF-Files gelten.

Die Implementierung von MXF muss nach dem einfachsten Operational Pattern OP 1a erfolgen. Für die Einbettung der Video- und Audiodaten sind dabei die Vorschriften für das Mapping von XDCAM HD422- und AVC-I-Daten zu beachten. Dabei sollen die Audio- und Videodaten in einem zeitlichen Multiplex so angeordnet sein, dass der MXF-File-Austausch per streaming erleichtert wird [112].

Metadaten, insbesondere Timecode, müssen in das MXF-File eingebettet sein. Die Verfügbarkeit des Timecodes zu Beginn Übertragung wird dadurch erzielt, dass die Header-Partition im Status „Abgeschlossen, closed“ verwendet wird.

10.4.2.4 Live-Zuspielung

Bei Live-Zuspielungen ist der zentrale Problembereich die Latenz, die so gering wie möglich gehalten werden soll. Sie hängt von der Laufzeit des Signals selbst ab und ergibt sich hier aus der Lichtgeschwindigkeit im Medium. Hinzu kommen die Verarbeitungszeiten beim Durchlaufen der Signalverarbeitungs- und Codierungsstufen die insbesondere bei der Verwendung von Long-GOP-Codecs zu erheblichen Werten führen können. Um der Latenzproblematik ganz aus dem Wege zu gehen, wird empfohlen, unkomprimiert zu arbeiten, die Übertragung also in transparenter Form über die dafür entwickelten HD-SDI-Schnittstellen abzuwickeln. Daneben gibt es aber auch die Bestrebung, den Filetransfer als generelle Austauschform auszuweiten. Die Betrachtungen

Tabelle 10.10. Anwendungsfälle für die SNG-Übertragung und empfohlene SNG-HD-Profile

Name	Anwendung	Videoqualität	Latenz	Audio-spuren	SNG-Profil MPEG-2	SNG-Profil H.264
SNG1	Live-Sendeleitung	sehr gut 422	< 400 ms	8	60HD	32HD
SNG2	hochw. Überspielung	sehr gut 422	< 800 ms	8	42HD	32HD
SNG3	hochw. Schaltgespräch	gut 422	< 400 ms	4	32HD	22HD
SNG4	Überspielung	gut 422	< 800 ms	4	32HD	22HD
SNG5	gew. Schaltgespräch	gut 422/420	< 400 ms	4	32HD	16HD
SNG6	Rückton und -bild	mittel 422/420	< 400 ms	4	32HD	11HD

des ARD-Hybnet im vorigen Abschnitt zeigen dass auch dieser Bereich bezüglich der Latenz relativ unkritisch ist und im Wesentlichen nur von den Codec-Laufzeiten bestimmt wird. In Tabelle 10.6 wird deutlich, dass in Netzwerken Laufzeiten erreicht werden, die unter der Dauer eines Halbbildes liegen.

Insgesamt ist also die leitungsgebundene Live-Übertragung als unkritisch anzusehen. Kritisch ist dagegen vor allem die Live-Zuspielung via Satellit in Form des Satellite News Gathering. Hier sind aus Effizienzgründen fast ausschließlich Long-GOP-Codecs im Einsatz. Hinzu kommt die Signallaufzeit zum Satelliten und zurück, die bei ca. 37000 km Abstand zwischen Satellit und Erdoberfläche allein schon mehr als 240 ms beträgt.

Die TPRF befinden für den SNG-Bereich eine Gesamtlaufzeit von weniger als 400 ms als wünschenswert. Für die Übertragung von HD-Programmen sind bisher vor allem MPEG-2-codierte Signale verwendet worden. Für diese und auch für H.264-codierte Signale wurden SNG-HD-Profile definiert, die in der vorangestellten Zahl die ungefähr erforderliche Datenrate für ein Programm in Mbit/s angeben. Für MPEG-2 wurden oft die Profile 60 HD, 42 HD und 32 HD genutzt, um sehr gute bis mittlere Bildqualitäten zu realisieren.

Mit H.264/AVC steht seit einiger Zeit ein etwa doppelt so effizienter Codec zur Verfügung, der jetzt und in weiterer Zukunft im SNG-Bereich immer mehr Einzug halten wird. Im Mittel wird für die SNG-Zuspielung im Produktionsbereich im Vergleich zur DVB-S-Ausstrahlung zum Zuschauer die dreifache Datenrate aufgewendet.

Für die Implementierung der H.264-basierten SNG-Profile sind in den TPRF der öffentlich rechtlichen Rundfunkanstalten in Deutschland sechs exemplarische Anwendungsfälle definiert, die in Tabelle 10.10 dargestellt sind und in denen die Anforderungen an die Laufzeit eine große Rolle spielen. Bei den Profilen SNG 1, 3 und SNG 5, 6 ist festgelegt, dass die Laufzeit < 400 ms betragen sollte, da sie dezidiert für die Live-Übertragung konzipiert sind. SNG1 ist für die Live-Übertragung mit höchster Bildqualität vorgesehen, wie sie z. B. bei großen internationalen Sportveranstaltungen gefordert wird. SNG3 ist auf die Live-Zuspielung mit guter Qualität, z. B. für Magazinsendungen ausgerichtet. Bei den Profilen 5 und 6 darf die Bildqualität noch geringer sein, da sie für sogenannte Schaltgespräche bestimmt sind, wie sie typischerweise in Nachrichtensendungen vorkommen, in denen ein Moderator mit einem Korrespondenten in einem anderen Land spricht. Bei SNG 6 wird nur noch eine Bild- und Tonqualität gefordert, die für Kontrollzwecke ausreicht.

DVB-S2 MPEG4 H.264		single encoding			re-encoding	
Festlegungen von ARD, ZDF, ORF und SRG in Anlehnung an die EBU Eurovision H.264-Profile		HDTV Quality Very Low 11 HD	HDTV Quality Low 16 HD	HDTV Quality Medium-1 22 HD	HDTV Quality Medium-2 32 HD	HDTV Quality HBR 42 HD
Color Sampling		4:2:0	4:2:0	4:2:2	4:2:2	4:2:2
Video	Mbits/s	9,556	14,014	19,022	27,772	38,224
Audio 1	Mbits/s	0,192	0,384	0,384	0,384	0,384
Audio 2 (or Dolby E)	Mbits/s	0,192	0,384	0,384	2,304	2,304
Audio 3	Mbits/s	0,192	0,384	0,384	0,384	0,384
Audio 4	Mbits/s	0,192	0,384	0,384	0,384	0,384
SID-Data	Mbits/s	0,128	0,128	0,128	0,128	0,128
Delay Mode - GOP 24/0 BF	Low Delay	0	0	0	-	-
Delay Mode - GOP 24/2 BF	Standard	2	2	2	2	2
Entropy Coding		CABAC	CABAC	CABAC	CABAC	CABAC
Info Bit Rate (188)	Mbit/s	**10,452**	**15,678**	**20,686**	**31,356**	**41,808**
BCH (factor)		1,03328588	1,03328588	1,03328588	1,03328588	1,03328588
Overall Bit Rate	Mbits/s	10,800	16,200	21,375	32,400	43,200
FEC		3/4	3/4	3/4	3/4	3/4
Pilot		On	On	On	On	On
Frame (64800 Bits)		normal	normal	normal	normal	normal
Modulation 8PSK	3 Bit/Symb.	3	3	3	3	3
Channel Rate (Symbolrate)	Msymb/s	4,80	7,20	9,50	14,40	19,20
Roll-off factor		25%	25%	25%	25%	25%
Occupied bandwith	MHz	6,00	9,00	11,88	18,00	24,00
Allocated bandwith	MHz	6	9	12	18	24

Abb. 10.58. ARD/ZDF-Profile für die H.264-basierte HD-SNG-Übertragung [112]

Die Profile SNG 2 und SNG 4 sind für Überspielungen mit sehr guter bzw. guter Bild- und Tonqualität vorgesehen, hier darf die Laufzeit zwischen 400 ms und 800 ms liegen. Abbildung 10.58 zeigt detailliert SNG-Parameter, die von ARD/ZDF für die H.264-basierte Übertragung empfohlen sind.

10.4.2.5 Technische Abnahme Bild und Ton

Die Begriffe Bild- und Tonabnahme bezeichnen eine systematische Überprüfung der Qualität von Programmbeiträgen. Sie erfolgt durch geschultes Personal; eine automatische Auswertung ist nur hinsichtlich einiger Parameter aber nicht bezüglich der Bild- und Tonqualität möglich.

Die Abnahme erfordert nach den TPRF bestimmte Darstellungswerkzeuge und definierte räumliche Bedingungen, die in der DIN 15996 beschrieben sind. Für die Bildabnahme ist festgelegt, dass der Betrachtungsabstand HD-gemäß a = 3H, also das Dreifache der Bildhöhe betragen muss. Das verwendete Display soll die Qualitätsstufe Grade 1 haben. Solange derartige Geräte nur bis zu einer Bilddiagonalen bis ca. 60 cm verfügbar sind, wird empfohlen, parallel ein bekanntes, gutes Display mit der doppelten Bilddiagonalen zu verwenden. Da moderne Flachdisplays intern progressiv arbeiten, sollte ein besonderes Augenmerk auf Interlaced-Artefakte gelegt werden, die insbesondere bei Lauftiteln etc. deutlich sichtbar werden.

Für die Abnahme von Tonprogrammmaterial ist der oben angesprochene adäquate Abhörraum von besonderer Bedeutung. Darin sollten hochwertige Lautsprecher mit linearem Frequenzgang betrieben werden, Fernsehlautsprecher oder Lautsprecher aus dem Heimbereich werden als ungeeignet eingestuft. Für weitere Vorgaben verweisen die TPRF auf die Festlegungen in EBU Tech Doc 3276. Das Ergebnis der Abnahme muss für Bild und Ton in einem Abnahmeprotokoll (MAZ-Karte) festgehalten werden [112].

10.5 Produktionseinheiten

Die Anforderungen an die technische Ausstattung von Videoproduktionsstätten erschließen sich leichter, wenn die Arbeitsabläufe bekannt sind. Zur Orientierung seien daher zunächst kurz die Produktionsbereiche dargestellt.

10.5.1 Produktionsbereiche

Die technisch einfachste Form ist die elektronische Berichterstattung (EB), die Grundform der Beistragserstellung für den Nachrichtenbereich (News). Ein EB-Team besteht gewöhnlich nur aus zwei bis drei Personen: Eine Redakteurin oder ein Reporter begleitet jeweils eine Person für die Kamera und den Ton. Bild- und Tonsignale werden auf eine transportable Kamera mit integriertem Recorder (z. B. Formate XDCAM oder P2-Karte) aufgezeichnet. Die Aufzeichnungseinheit wird auf einem leichten Stativ befestigt oder kann auf der Schulter getragen werden (Abb. 10.59). Das Mikrofon befindet sich oft an einer Angel, damit es möglichst nahe an Interview-Partner herangeführt werden kann, alternativ wird das Mikrofon vom Reporter gehalten oder es werden drahtlose Ansteckmikrofone verwendet. Der Signalpegel wird dauernd über einen separaten Audiomischer kontrolliert. Bei 4-spuriger Aufnahme werden die Tonkanäle 1 und 2 mit dem Signal der Angel und 3 und 4 mit dem Kameramikrofon belegt. Zur Videosignalkontrolle stehen beim EB-Einsatz außer der Zebra-Funktion keine Hilfsmittel zur Verfügung. Oft müssen Automatikfunktionen genutzt werden, die Bildqualität kann daher etwas eingeschränkt sein.

Die aufwendigere Produktionsweise mit nur einer Kamera ist die Electronic Field Production (EFP). Oft stehen Kamera und Recorder separat zur Verfügung, in jedem Fall aber Video- und Audiosignalkontrollgeräte, um eine möglichst hohe Qualität gewährleisten zu können. Entsprechender Aufwand wird auch bei der Lichtgestaltung getrieben. EFP entspricht der Arbeitsweise bei der Spielfilmherstellung (s. nächste Abschnitt) und ist entsprechend zeitaufwendig und teuer (Abb. 10.60). Jede Szene wird Einstellung für Einstellung eingerichtet und gedreht. Wenn z. B. eine Zwei-Personen-Szene mit einer Kamera in zwei Close Ups, zwei Over Shoulder Shots und eine Haupteinstellung aufgelöst werden soll, muss die Szene fünfmal aufgenommen werden.

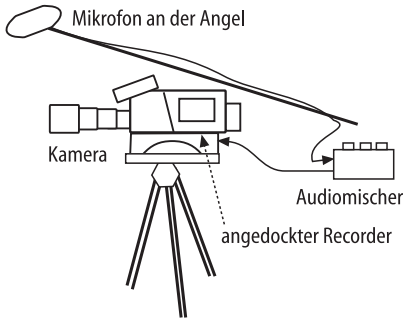


Abb. 10.59. Ausrüstung zur elektronischen Berichterstattung (EB)

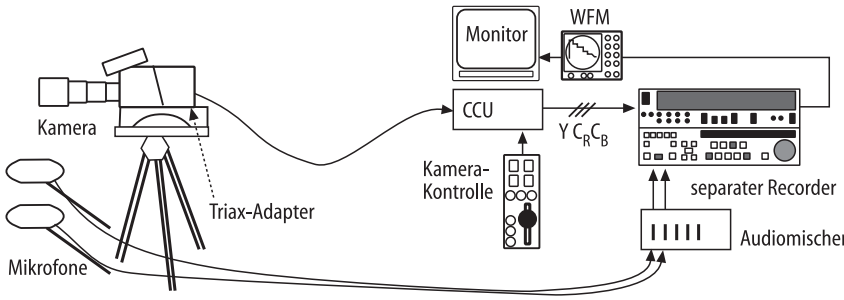


Abb. 10.60. Ausrüstung zur Electronic Field Production (EFP)

10.5.1.1 Fernsehfilm- und Filmproduktion

Diese Form zeichnet sich gegenüber der EB-Produktion ggf. durch den Einsatz von 2k-, 4k- oder Film-Equipment, vor allem aber durch den hohen Produktionsaufwand aus, der hinsichtlich Kamerabewegung, Lichtsetzung und Ausstattung der Szene getrieben wird. Besonders der personelle Aufwand ist beträchtlich. Die gesamte Produktion wird vom Produktionsteam verantwortet. Der Produzent selbst ist dabei vor allem der Geldgeber und Träger des wirtschaftlichen Risikos. Um die effiziente Verwendung der Mittel kümmert sich die Produktionsleitung, die u. a. die Drehpläne erstellt. Für den reibungslosen Ablauf der Produktion (Drehgenehmigung, Absperrung, Transport) sorgt die Aufnahmeleitung. Die inhaltliche Seite der Produktion wird vom Regieteam bestimmt, dessen zentrale Person der Regisseur bzw. die Regisseurin ist. Die Regieassistenz dient der Umsetzung der kreativen Idee und arbeitet eng mit der Produktionsleitung zusammen.

Die technische Seite der Bildaufnahme bestimmt das Kamerateam, das i. d. R. aus dem Kameramann und dem Kameraassistenten besteht. Der Kameramann oder die Kamerafrau, im englischen Sprachraum als Director of Photography (DoP) bezeichnet, ist für die Bildgestaltung, d. h. die Kadrierung, die Bildausschnittsfestlegung und das Setzen von Licht und Schatten verantwortlich. Die Kameraführung übernimmt der DoP selbst oder eine separate Person, der sogenannte Schwenker. Der Kameraassistent ist für die technische Prüfung und Säuberung der Kamera zuständig, um z. B. zu vermeiden, dass Staub und Fussel in das Bildfenster geraten. Während der Dreharbeiten hilft er vor allem bei der Einstellung der Schärfe. Ein Materialassistent ist für den sorgfältigen Umgang mit dem Filmmaterial einschließlich Ein- und Auslegen in die Kassette verantwortlich.

Die bisher dominanten 16- oder 35-mm-Kameras werden fast vollständig von den Digital-Cinematography-Kameras der Firmen, Arri, Red, Sony etc. abgelöst. Wenn sie zum Einsatz kommen übernimmt die Kameraassistenz vor allem Aufgaben als DIT (Digital Image Technician) und Data Wrangler und ist somit für die Einstellung der Kameraparameter und die Datensicherung zuständig.

Kameras aus diesem Sektor werden in der EBU-Empfehlung R118 in den obersten Rang eingestuft der dort als Tier LS bezeichnet wird. LS bedeutet large sensor. Weitere Ränge werden als Tier SP (Special Effects) und Tier 1

Tabelle 10.11. Klassifizierung für Kameras im Produktionsbereich nach EBU R118

Rang	min. Sensorgröße	min. Bildpunktzahl	Wechselobjektiv	Abtastformat	min. Bitrate onboard	sonstiges
Tier LS	1“	2880 x 1620	gefordert	4:4:4/4:2:2	100 Mbit/s	TC
Tier SP	3 x 1/4“	1280 x 720	-	4:2:0	35 Mbit/s	speziell
Tier 1	3 x 1/2“	1920 x 1080	gefordert	4:2:2/4:4:4	50 Mbit/s	genlock
Tier 2L	3 x 1/2“	1920 x 1080	-	4:2:2	50 Mbit/s	Audio in
Tier 2J	3 x 1/3“	1440 x 1080	-	4:2:0	35 Mbit/s	Audio In
Tier 3	3 x 1/4“	1280 x 720	-	4:2:0	24 Mbit/s	AVCHD
Tier 4	16:9-Format	kann von der Rundfunksanstalt spezifiziert werden				

–4 bezeichnet. Die Klassifizierung ist in Tabelle 10.11 dargestellt, die Rangfolge ergibt sich aus den Qualitätsparametern Bildauflösung, Sensorempfindlichkeit, Rauschen, Aliasing, Belichtungsumfang und Aufzeichnungscodec. Tier LS beinhaltet vor allem die Digital Cinematography-Kameras wie die Arri Alexa oder Red Epic. Tier SP ist eine Klasse für Kameras, die in einzelnen Parametern besonders herausragende Aspekte, wie Kompaktheit, sehr hohe Bildraten oder Markroobjektive aufweisen. Hier werden die Ansprüche an das Bildsignale ggf. reduziert und Auflösungen von 720p oder Codecs mit 35 Mbit/s erlaubt.

Tier 1 ist die Klasse für hochwertige HD-Studiokameras wie die Sony HDC 2500 oder GV LDK 8000, die meist mit 2/3“-Sensoren bestückt sind. Die Signale sollte hier mit 10-Bit-Auflösung vorliegen, es werden wechselbare Objektive und Genlock-Funktionen gefordert.

Rang 2 teilt sich in Tier 2L und Tier 2J. Diese Klasse ist für den EB-Bereich vorgesehen. Das L steht für Langformate mit höheren Ansprüchen an die Kamera. Hier sollten 3-Chip-Kameras mindestens mit 1/2“-Sensoren zum Einsatz kommen die jeweils 1920 x 1080 aktive Bildpunkte aufweisen. Integrierte Aufzeichnungssysteme sollten bei Long-GOP-Codecs mindestens mit 50 Mbit/s und bei I-frame-only-Codecs mit mindestens 100 Mbit/s arbeiten. Das J steht bei Tier 2 für den journalistischen Bereich. Hier dürfen die Ansprüche gesenkt werden. Die Bildsensoren dürfen ggf. auch nur 1440 x 1080 Pixels aufweisen und bei den Long-GOP-Codecs werden auch 35 Mbit/s akzeptiert. Für den Tonbereich ist in Tier 2 die Verfügbarkeit von mindestens zwei symmetrischen XLR-Audioeingängen gefordert.

Tier 3 ist eine Klasse für semiprofessionelle Kameras, die mit Sensoren bis hinab zu 1/4“ und 24 Mbit/s-Codecs arbeiten dürfen. Material von Kameras dieser Qualität sollten nur zu maximal 33% der Beitragslänge in professionellen Produktionen enthalten sein. Tier 4 ist nicht genauer spezifiziert. Die Festlegungen sind den Rundfunkanstalten überlassen.

Für die Kamerabewegung stehen Kräne und Kamerawagen (Dolly) zur Verfügung (s. Kap. 6), die der ruhigeren Fahrt wegen oft auf Schienen laufen und die von Assistenten (Crane Operator und Dolly Grip) in enger Zusammenarbeit mit dem Kameramann bedient werden. Wenn lange und komplizierte Bewegungen erforderlich sind, kann auch die SteadyCam, ein Kamerastabilisierungssystem verwendet werden, das am Körper getragen wird. Die Bildkontrolle wird hier mit Hilfe eines mobilen Monitors vorgenommen (s. Kap. 6).

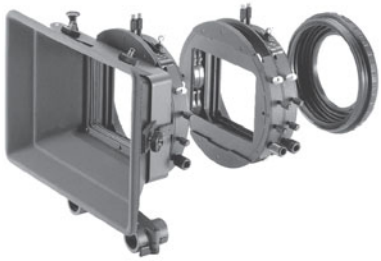


Abb. 10.61. Mattebox und Filmklappe

Wenn keine besonderen Effekte erzielt werden sollen, wie z. B. Zeitlupeneffekte, die eine Veränderung der Filmgeschwindigkeit erfordern, sind die wesentlichen technischen Parameter bei der Filmaufnahme am Objektiv einzustellen. Hier kann zunächst die Brennweite gewählt werden, die durch die Einstellung der Szene bestimmt ist. Zweitens muss auf das wesentliche Objekt der Szene fokussiert werden. Dabei verlässt man sich nicht auf die visuelle Beurteilung der Schärfe im Sucher, sondern es wird der Fokus auf die Entfernungsmarkierung eingestellt, die der mit dem Maßband ausgemessenen Entfernung zwischen Kamera und Objekt entspricht. Kamerafahrten erfordern dabei Schärfeverlagerungen, die mit Hilfe des Assistenten durchgeführt werden. In schwierigen Situationen können in diesem Zusammenhang elektronische Lens-Control-Systeme hilfreich sein, die es erlauben, die genannten Parameter fernzusteuern, was auch drahtlos geschehen kann (Funkschärfe).

Ein relativ frei wählbarer Parameter ist die Blende. Die Bezugsblende wird für den mittleren Grauwert der Szene eingestellt, der mit 18% Remission angenommen wird. Die Blende hat neben der verwendeten Brennweite erheblichen Einfluss auf die erreichte Schärfentiefe. Da die Brennweite über die Bildgestaltung oft weitgehend festgelegt ist, kann die Schärfentiefe am ehesten über die Blende beeinflusst werden, soweit es die Lichtintensität zulässt.

Neben der mittleren Blende interessieren auch die gemessenen Blendenwerte, die an den hellsten bzw. dunkelsten Stellen im Bild auftauchen. Falls die Differenz 6 Blendenstufen überschreitet, ist evtl. der Kontrast zu vermindern, indem mit Scheinwerfern Schatten aufgehellt werden. Auch natürliches Licht wird oft durch Scheinwerferlicht ergänzt, z. B. um die Szenen mit Schattenwürfen interessanter zu gestalten. Für hartes Licht werden Stufenlinsenscheinwerfer, Verfolgerspots und PAR verwendet. Für weiches Licht Flächenleuchten (Fluter), Weichstrahler und Horizontfluter. Um die Kamera abschatten zu können, wird auf dem Objektiv meist ein Kompendium befestigt, das mit Flügeltoren ausgestattet ist, die die Entstehung von Lichtreflexen vermeiden helfen (Abb. 10.61). Das Kompendium dient auch als Halter von Vorsatzfiltern und wird dann Mattebox genannt.

Am Produktionsort wird neben dem Bild auch der Originalton aufgenommen. Die Aufzeichnung von Bild und Ton am Set erfolgt getrennt. Daher muss während der Dreharbeiten natürlich auf die Synchronisation von Bild und Ton geachtet werden [167]. Im Schneideraum wird das Bild- und Tonmaterial synchron zusammengebracht, man spricht vom Anlegen der Bänder (s. Abschn. 9.4). Als wesentliches Hilfsmittel dazu dient immer die Filmklappe, die zum

Symbol für den ganzen Filmbereich geworden ist (Abb. 10.61). Auf die Klappe (engl.: Slate) wird nicht verzichtet, auch wenn der Timecode als weiteres Hilfsmittel zur Verfügung stehen sollte, der in identischer Form auf Bild- und Tonband aufgezeichnet wird.

Die Klappe hat den Vorteil, dass bei richtiger Handhabung das Zusammentreffen der Balken sowohl im Bild sichtbar als auch in der Tonaufzeichnung deutlich hörbar ist. Sie trägt darüber hinaus Zusatzinformationen, die im Timecode meist nicht vorhanden sind. Das sind im Besonderen der Name der Produktion und der Produktionsfirma, das Datum, die Rollennummer, die Nummer der Szene und des Takes und ggf. die Namen von Regisseur und Kameramann. Darüber hinaus wird vermerkt, ob eine Tag- oder Nachtaufnahme stattfindet und ob ggf. ohne Ton gedreht wird (Kennzeichnung durch das Kürzel st für stumm). Im letzten Falle sollte die Klappe geschlossen bleiben, damit beim Schnitt nicht angenommen wird, dass vergessen wurde, die Kennzeichnung wegzuwischen [191]. Die Beschriftung muss deutlich lesbar sein. Beim Schlagen der Klappe ist darauf zu achten, dass sie groß genug und auch im Schärfebereich der Kamera erscheint, bei einer Nahaufnahme kann sie z. B. vorsichtig in der Nähe des Gesichts des Schauspielers geschlagen werden.

In manchen Fällen ist es nicht möglich, die Klappe zu Beginn der Aufnahme zu schlagen, z. B. weil es keinen Ort für sie gibt oder weil Mensch oder Tier nicht erschreckt werden soll. In diesem Fall wird zu Beginn nur eine Schlussklappe angesagt, die dann auch erst am Schluss geschlagen wird, wobei sie auf den Kopf gehalten wird.

10.5.1.2 Studioproduktion

Als Beispiel für eine Studioproduktion kann eine Talkshow dienen. Die Produktion geschieht mit mehreren Kameras »Live on Tape« oder »Live on Air« vor Gästen im Atelier. Im Atelier befinden sich die Dekoration, das Beleuchtungssystem, das Tonaufnahmesystem und eine Beschallungsanlage (Abb. 10.62 oben). Zentrale videotechnische Einrichtung sind die Kameras auf Fahrstativen, die von Kameraleuten bedient werden, die per Intercom von der Regie ihre Anweisung bekommen. Die Kameraleute haben die Aufgabe, den Bildausschnitt nach Regievorgabe oder selbstständig zu wählen. Sie bewegen die Kamera und bedienen Zoom- und Fokussteuerungen, die an den Stativgriffen angebracht sind (s. Kap. 6). Die Kameras sollten möglichst mit großen Suchern ausgestattet sein. Bei Nachrichtensendungen müssen die Kameras meist nicht schnell und flexibel bewegt werden, daher können sie auf Roboterstativen montiert und fernsteuerbar sein.

Zur Verbindung von Kamerakopf und zugehöriger Elektronik wird meist die Triax-Technik (s. Abschn. 6.2) eingesetzt, die die Verwendung relativ dünner Kabel erlaubt. Über das Kamerakabel und die Kontrolleinheit muss eine Vielzahl von Signalen, u. a. das Intercomsignal zwischen Regie und Kamerapersonal, übertragbar sein sowie ein am Videomischer generiertes Rotlichtsignal (Tally), damit signalisiert werden kann, welche Kamera gerade vom Bildmischer angewählt ist. Das Rotlicht erscheint auf der gewählten Kamera und im zugehörigen Sucher. So wird im Studio angezeigt, in welche Kamera ein Moderator schauen muss, während die Kameraleute wissen, dass sie den Bild-

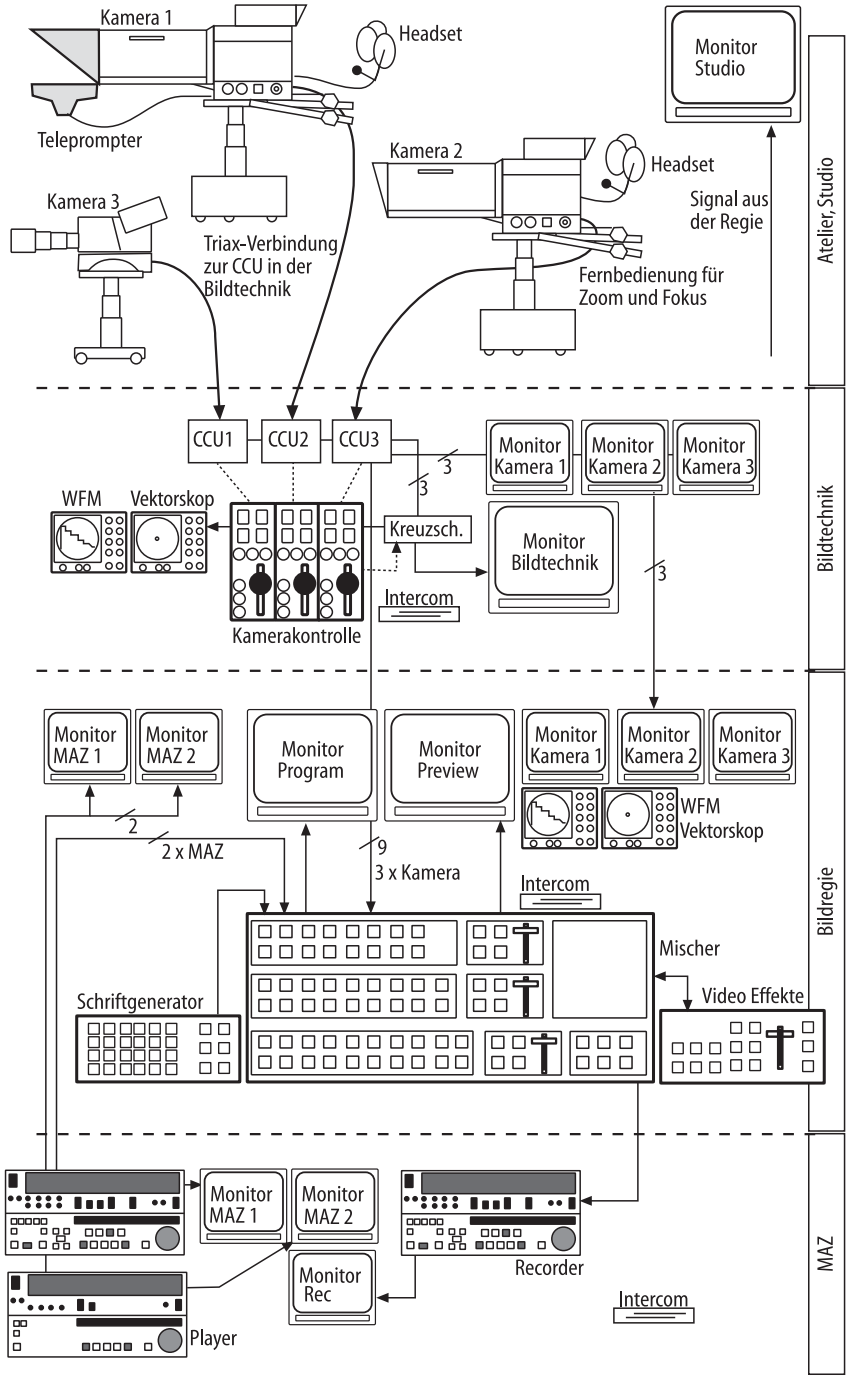


Abb. 10.62. Ausstattung zur Studioproduktion

ausschnitt höchstens auf ausdrückliche Anweisung verändern dürfen. Weiterhin werden Signale von der Regie zur Kamera zurück übertragen, damit sich die Kameraleute das PGM-Signal auf den Suchermonitor schalten können und damit das Teleprompterssignal die Kamera erreicht.

Teleprompter werden oft bei Nachrichtenproduktionen eingesetzt. Ein Teleprompter ist ein Kameravorsatzgerät mit einem im rechten Winkel zur Kameraachse montierten Monitor und einem Spiegel, der auf der Diagonalen dazwischen angebracht ist (Abb. 10.63). Der Spiegel ist so konstruiert, dass ein Moderator einen auf dem Monitor dargestellten Text lesen kann, während der Spiegel von der Kamera aus völlig transparent erscheint, sodass hier nur der Moderator zu sehen ist. Zur Textzuführung gibt es die herkömmliche und einfachste Form mit einer Papierrolle, die in der gewünschten Lesegeschwindigkeit unter einer Kamera abgerollt wird, deren Signal auf den Prompter-Monitor gegeben wird. Der Vorteil ist hier, dass sehr einfach individuelle Markierungen vorgenommen werden können. Aufgrund der Texterstellung im Computer wird aber meist direkt von PC aus eingespielt. Hierzu gibt es spezielle Teleprompter-Software, die eine Geschwindigkeitsregelung des Lauftextes per separater Handregelung erlaubt. Bekannte Prompter-System sind Autocue und Autoscript.

Neben dem Aufnahmestudio gibt es weitere Arbeitsbereiche für die Studioproduktion (Abb. 10.62). Die bildtechnischen Parameter der Kameras, vor allem die Blende und der Schwarzwert des Signals werden von der Bildtechnik aus gesteuert. Vor der Produktion findet hier das sogenannte Kameramatching statt, bei dem u. a. der Weißabgleich der Einzelkameras und die farbliche Angleichung der Kameras untereinander vorgenommen wird. Während der Produktion werden vor allem die Blende und der Schwarzwert ständig so eingestellt, dass damit die Maximal- und Minimalwerte des Videosignals an den Grenzen bleiben und diese nur wenig über- bzw. unterschreiten.

Der Arbeitsplatz für die Bildtechnik ist in der Bildregie oder im separaten Kamerakontrollraum untergebracht. Die Signalaussteuerung erfolgt über die hier installierten Remote-Control-Einheiten anhand eines Monitors sehr guter Qualität und mit Hilfe von Waveformmonitor und Vektorskop in der Art, wie

Abb. 10.63. Teleprompter von der Seite und von vorn

Abb. 10.64. Bildtechnik [222]

es in Abschnitt. 6.2 beschrieben wurde (Abb. 10.64). Als zweiter Arbeitsplatz kann in diesem Raum der Produktionsingenieursplatz untergebracht sein. Neben einer übergeordneten Kamerakontrolleinheit (Master Control Panel, MCP) stehen hier weitere Geräte zur Verfügung, mit denen ein Bildsignal beeinflusst werden kann, wie z. B. ein Farbkorrekturgerät und Signalwandler.

Die in der Bildtechnik optimierten Signale laufen zusammen mit den Grafik-, MAZ- und DVE-Zuspielungen in der Bildregie zusammen (Abb. 10.65). Hier werden die Szenenabfolgen mit Hilfe großer Produktionsmischer durch Umschnitt, Über- und Einblendungen nach Vorgaben der Regie von einer Bildmischerin erstellt. Die Arbeitsplätze befinden sich in einem eigenen Bildregieraum, in zentraler Position mit guter Sicht auf alle Monitore. Außer dem Bedienfeld des Videomischers ist hier auch eine Havariekreuzschiene installiert, sodass bei Ausfall von Signalen direkt reagiert werden kann. Daneben befinden sich die Arbeitsplätze am Schriftgenerator und am Effektgerät.

Die Regie braucht eine Vielzahl von Monitoren für möglichst alle Signale. Sie werden in einem Monitorgestell untergebracht (Abb. 10.65). Damit die An-

Abb. 10.65. Bildregie

zahl nicht zu groß wird, kann statt einer festen Zuordnung die Belegung der Monitore frei gestaltet werden. Dies geschieht mittels einer eigenen Monitorkreuzschiene, die auch ein Schriftdisplay unter jedem Monitor steuert, welches die jeweilige Belegung anzeigt (Under Monitor Display, UMD, Abb. 10.64).

Eine Monitorausstattung könnte z. B wie folgt aussehen: In der Bildregie zwei hochwertige, Grade 1-Monitore (20“) für die Mischerausgänge, ein 14“-Monitor für das Signal, das zur Kontrolle über den Modulationsweg (HF) zurückkommt, so wie elf kleine Vorschaumonitore für 4 Kameras, 4 MAZ-Geräte, Mischerausgänge und Grafik. Die Bildtechnik benötigt zwei 14“-Grade 1-Monitore zur visuellen Signalkontrolle, davon einer frei schaltbar. Der MAZ-Raum enthält z. B. einen 14“-Monitor (frei zuzuordnen) und diverse Vorschaumonitore, ebenso wie die Lichtregie und die Tonregie.

Für die HD-Produktion sollten die Monitore ein 16:9-Bildseitenverhältnis haben und aus einem Abstand betrachtet werden können, der den HD-Vorgaben entspricht. Dabei sind ggf. größere Studioräume erforderlich [211]. Vermehrt werden auch Multiviewer-Monitore eingesetzt, die die vielen kleinen Vorschaubilder als Teile auf einer großen Gesamtfläche darstellen (Abb. 10.66) [184].

Die Anordnung der Arbeitsplätze variiert [215]. In kleineren Studios können sich die Bildtechnik sowie die Bild- und Tonregie in einem Raum befinden. Die Bedieneinheiten sind im Pult angeordnet, die zugehörige Elektronik möglichst in einem separaten Maschinenraum. Hier werden neben Aufzeichnungsgeräten die CCU der Kameras ebenso untergebracht wie das Mainframe des Bildmischers. So können die Geräte mit kurzen Leitungen verbunden werden. Im Maschinenraum befinden sich auch die Steckfelder für die Audio- und Videosignale und Kontrollmonitore für die MAZ-Maschinen, mit deren Hilfe der Beginn von Zuspielsequenzen bestimmt wird (Abb. 10.67).

Eine sehr wichtige Einheit im Studio ist die Intercom-Einrichtung [129]. Weiterhin befinden sich im Produktionsbereich oft Framestore-Synchronizer zur Einbindung externer Signale, die nicht an den Studiotakt gekoppelt sind, und Einrichtungen zur Fernsteuerung von Kameras.

Abb. 10.66. Multiview-Monitoring

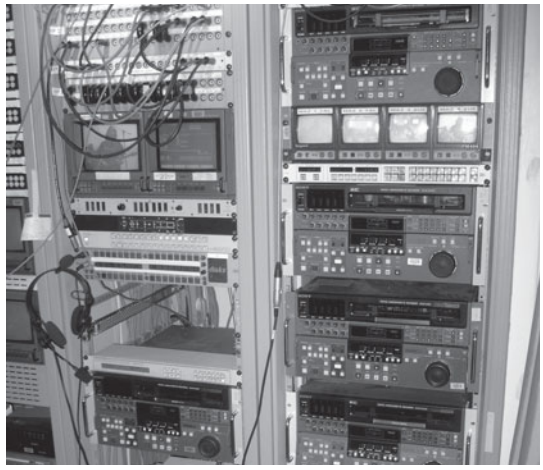


Abb. 10.67. Maschinenraum

Alle Produktionsstudios werden heute in Digitaltechnik aufgebaut. Dies geschieht aus Gründen der Signalqualität, aber auch aufgrund des verringerten Aufwands für die Verkabelung und den Service. Ein Digitalstudio basiert meist auf der Verwendung des digital-seriellen Komponenten-Signals (DSK), mit den zugehörigen Schnittstellen SDI bzw. HD-SDI. Dabei sind die üblichen BNC-Armaturen ausreichend, obwohl die Verwendung besonders hochwertiger Ausführungen empfohlen wird.

Ein Problem, das in digitalen Studios stärker auftritt als in analogen, ist die Laufzeitproblematik. Die digitalen Geräte brauchen oft erheblich mehr Zeit zur Signalverarbeitung als analoge, sodass durch Einsatz von Verzögerungsgeräten verschiedene Signale sorgfältig aneinander angepasst werden müssen.

Ein Beispiel für eine Studioausstattung mit digitalen Geräten könnte so aussehen: Vier Kamerazüge, deren Kameraparameter über RCP/OCP direkt gesteuert werden oder über ein Steuer-Netzwerk, in das ein Master Control Panel (MCP) integriert ist, an dem ein übergeordneter Bildingenieur sich Priorität über die Operational Control (OCP)-Steuerung verschaffen und umfassend in die Parameter eingreifen kann (Abb. 10.68). Das MCP erlaubt die Kontrolle

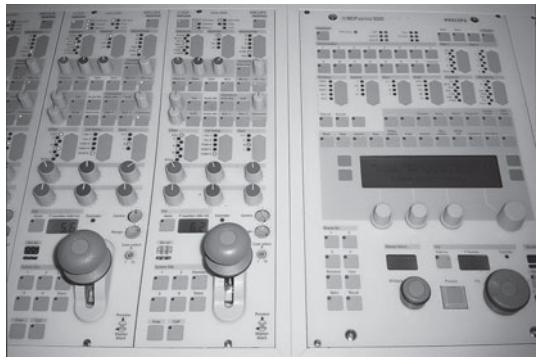


Abb. 10.68. Operational und Master Control Panel

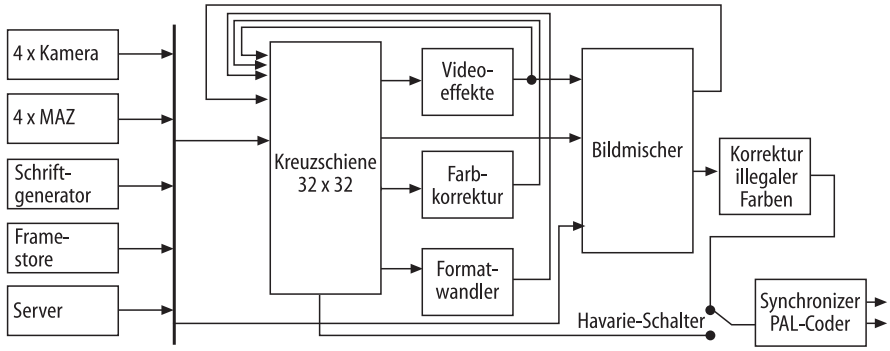


Abb. 10.69. Blockschaltbild des digitalen Produktionsstudios

mehrerer Kameras von einem Platz aus, ohne dass jeweils das zugehörige OCP bedient werden muss.

Die DSK-Signale werden über eine Kreuzschiene geführt, die es ermöglicht, softwaregesteuert z. B. 32 Eingänge flexibel mit 32 Ausgängen zu verbinden (Abb. 10.69). Die Steuerung geschieht gewöhnlich über serielle Schnittstellen nach RS 232 mit separaten Geräten oder per PC. Um Erweiterungen und unvorhergesehenes Routing zu ermöglichen, werden Steckfelder verwendet, weiterhin kommen viele Videowandler zum Einsatz, die eine Umsetzung verschiedener analoger Signalformate nach SDI und umgekehrt bewirken.

Neben der Kreuzschiene ist der Videomischer Kernbestandteil des Studios, hier sind Geräte der Firmen Sony und Thomson/GrassValley verbreitet. Diese Geräte können in der Regel skaliert werden, sodass z. B. die Anzahl der benötigten Keystufen vom Anwender bestimmt werden kann. Außerdem können meist mehrere Bedienpanel an die Haupt-Processing-Einheit angeschlossen werden. Auf diese Weise ist die Produktion verschiedener Programmvarianten möglich.

Die Produktion mit hochaufgelösten HD-Signalen geschieht, ähnlich wie mit SD-Signalen, auf Basis des HD-Komponentensignals. Wesentliche Voraussetzung ist die Verfügbarkeit von Produktionsgeräten, die möglichst als Multi-Standardgeräte Bitraten mit 1,5 Gbit/s, möglichst aber mit 3 Gbit/s verarbeiten. HD-Kameras haben meist mit 2/3"-Wandlern und 14-Bit-Processing, sie unterscheiden sich äußerlich nicht von Standardkameras. Da auch HD-Filmabtaster existieren, steht somit an der Signalquelle die höchste Auflösung zur Verfügung.

Zum Übergang auf andere HD- oder SD-Formate werden Downconverter eingesetzt die mit aufwendigen Filtern ausgestattet sind, sodass ein von HDTV nach SDTV konvertiertes Signal subjektiv besser beurteilt wird als ein direkt gewonnenes Standardsignal. Zur Signalverteilung und -bearbeitung stehen Kreuzschienen mit Bandbreiten von 1,5 oder 3 Gbit/s ebenso zur Verfügung wie Bildmischer und Effektgeräte. Alle Multiformatgeräte müssen neben der Skalierung der Ortsauflösung auch eine Umrechnung der Farbräume vornehmen, die in den ITU-Empfehlungen BT 601 und BT 709 beschrieben sind.

10.5.2 Postproduktion

Die Grundbegriffe und -funktionen elektronischer Schnittsysteme sind in Abschnitt 9.3 erklärt. An dieser Stelle geht es um die Zusammenführung zu Postproduktionseinheiten. Postproduktions-Studios lassen sich hinsichtlich der Qualität und bezüglich des Kundenkreises (Industriefilm, Broadcast, Werbung) unterscheiden [213]. Ein zweites Merkmal ist die Arbeitsweise, entweder linear mit MAZ-Geräten oder nichtlinear mit Computersystemen. Die klassischen Postproduction-Studios arbeiteten mit MAZ-Geräten auf Basis des Linear Editing, heute sind sie aber gegenüber den nichtlinearen Editingsystemen kaum noch verbreitet, da die nichtlinearen Systeme flexibler sind und keine teuren MAZ-Maschinen erfordern.

10.5.2.1 Linear Postproduction

In Abbildung 10.70 ist der Grundaufbau der linearen Postproduktionseinheit dargestellt. Hier steht das Editor-Gerät im Zentrum von dem aus die Zuspieler, die Mischer und der Recorder gesteuert werden. Der Einsatz des Linear Editing ist eigentlich nur noch bei kurzen Beiträgen von Vorteil, wenn bandbasiert gearbeitet wird und die Originalbänder direkt, z. B. aus dem EB-Recorder, im Zuspielgerät verwendet werden können, sodass das Material nicht erst in das Editingsystem eingespielt werden muss. Von Nachteil sind vor allem die hohen Kosten hochwertiger Bandmaschinen, die auch hohe Abnutzungskosten mit sich bringen, da ca. alle 2000 Betriebsstunden ein teurer Kopfwechsel notwendig wird. Entscheidend ist auch die Arbeitsweise. Bei aufwendigen Postproduktionen werden die Freiheiten der nichtlinearen Arbeitsweise bevorzugt.

Im Postproduktionsbereich gibt es Studios unterschiedlicher Qualitätsklassen. Die Qualität, mit der das Videosignal bearbeitet werden kann, hängt entscheidend von der Qualität der Aufzeichnungsmaschinen ab. So ist der Einsatz analoger Geräte wie Betacam SP und MII ebenso möglich wie der digitaler

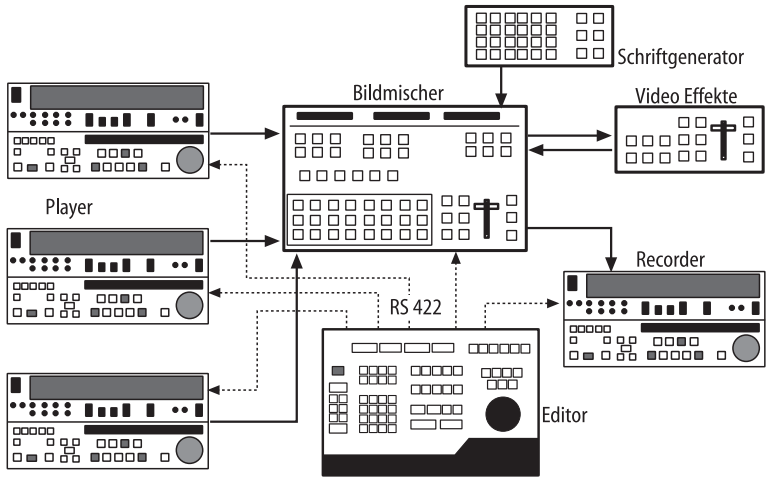


Abb. 10.70. Lineare Postproduktionseinheit

SD- oder auch HD-MAZ-Maschinen, wie Digital Betacam oder HDCAM. Ungeachtet dieser Einteilung können auf allen Ebenen einfache und komplexe Studios aufgebaut werden.

Für einfache Aufgaben gab es den sogenannten Laptop-Editor. Er enthält einen kompletten Zwei-Maschinen-Schnittplatz mit zwei Aufzeichnungsgeräten, zwei LCD- und zwei Audiomonitoren und einem Editcontroller. Er hat aber heute, angesichts von Schnittmöglichkeiten im Laptop-PC, keine Bedeutung mehr. Während der Laptop-Editor alle Funktionen in einem Gehäuse integriert, besteht der klassische Schnittplatz aus Einzelgeräten. Das funktionelle Zentrum ist dabei das Schnittsteuersystem. Wesentliches Merkmal der MAZ-Maschinen ist hier die Auslegung als Schnittrecorder mit Timecode-Unterstützung und Fernsteuerbarkeit. Die einfachste Ausführung des linearen Schnittplatzes ist ein Cut-only-Platz mit zwei Maschinen (z. B. DigiBeta-Recordern), wobei der Zuspieler vom Recorder aus gesteuert wird. Die gehobene Ausstattung bezieht ein Schnittsteuersystem mit ein (Abb. 10.71). Die Maschinen werden von dem separaten Schnittgerät über eine RS-422-Schnittstelle ferngesteuert, die auch zum Austausch der Timecode-Daten dient. Weiterhin steht ein Audiomischer zur Verfügung, an dessen Ausgang der Recorder-Audioeingang angeschlossen ist und dessen Eingänge mit dem Playersignal und weiteren externen Audiosignalen versorgt werden. Schließlich kommen noch ein Audio-Pegelmessgerät und ein Monitor für den Zuspieler zum Einsatz.

Eine erhebliche Erweiterung der Funktionalität ist die Möglichkeit, nicht nur hart schneiden, sondern auch mit Blenden und Effekten arbeiten zu können. Dieses gelingt mit dem AB-Roll-Schnittplatz mit drei Maschinen, zwei Playern und einem Recorder (Abb. 10.70). Die wesentliche videotechnische Erweiterung ist die Einbeziehung eines Bildmischers, mit dem die Signale der Zuspieler geschaltet und gemischt werden. Eine einfache Möglichkeit zum Aufbau eines AB-Roll-Schnittplatzes ist die Integration des Bildmischers in das Schnittsteuergerät. Auch bereits auf diesem Komplexitätsniveau sollten die Geräte in einem separaten Schrank oder Maschinenraum untergebracht werden, damit die Editoren nicht von dem erheblichen Lärmpegel und der Abwärme der Maschinen gestört werden.

Abb. 10.71. Drei-Maschinen-Schnittplatz

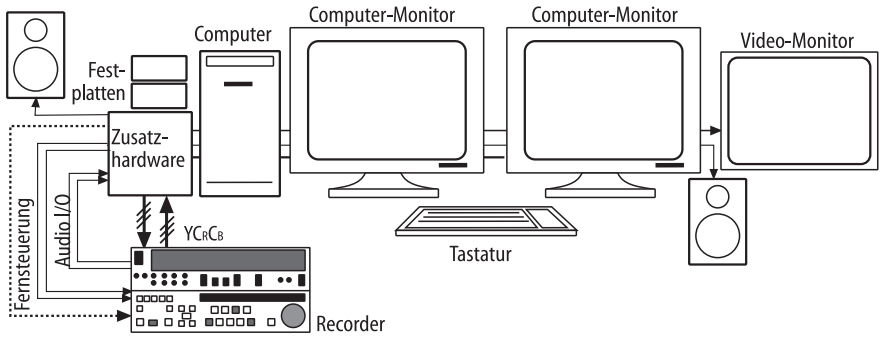


Abb. 10.72. Blockschaltbild des nonlinearen Schnittplatzes

10.5.2.2 Nichtlineare Schnittplätze

Nichtlineare Editingsysteme auf der Basis von Standard-PC sind in Kapitel 9.4 dargestellt. Die Schnittplätze erfordern einen vergleichsweise geringen Ausstattungsaufwand mit videospezifischen Geräten, trotzdem können, abhängig vom Funktionsumfang des Kernsystems, komplexe Bearbeitungen durchgeführt werden, da viele Funktionen der Einzelgeräte konventioneller Schnittplätze im PC-System integriert sind. Als videotechnische Peripheriegeräte sind im Wesentlichen nur ein Speichersysteme oder eine MAZ-Maschine für Zuspielungen sowie Kontrollmonitore erforderlich. Abbildung 10.72 zeigt einen sogenannten Desktop-Videoplatz. Auf dem Arbeitstisch der Cutter (Abb. 10.73) befinden sich nur zwei Computermonitore und ein Videomonitor, zwei Abhörlautsprecher, die Computertastatur und ggf. Audio- und Videomessgeräte

Große Postproduktionshäuser und Funkhäuser haben einen erheblichen Gewinn durch die Vernetzung der Arbeitsplätze. Wenn dafür Storage Area Networks (SAN) wie Avid Unity oder Editshare genutzt werden, kann das Material einmal zentral auf der zugehörigen großen Speichereinheit abgelegt werden und dann ist es möglich, von allen Workstations darauf zuzugreifen. Auf diese Weise kann am gleichen Material vernetzt von mehreren Arbeitsplätzen aus gearbeitet werden. Die Videodaten müssen nicht mehrfach auf die lokalen Speicher der einzelnen Workstations gespielt werden.

Abb. 10.73. Aufbau des nonlinearen Schnittplatzes

10.5.3 Sendekomplexe und -abwicklung

In Sendeanstalten ist das Produktionsstudio an den Sendekomplex und den Hauptschaltraum angebunden. Im Sendekomplex wird die Sendeabwicklung integriert, die die Aufgabe hat, durch eine dem Programmplan entsprechende exakte zeitliche Steuerung die Beiträge aus den verschiedenen Quellen, live oder vorproduziertes Material, zu einem durchgängigen Programm zu verbinden. Anschließend muss sie das Signal dem Netzanbieter in der gewünschten, und meist vom Zuspielweg bestimmten Form zur Verfügung stellen. Die Zusammenstellung der Beiträge geschieht heute über Playlists, die Server steuern, auf denen die Beiträge abgelegt sind.

Bei automatischem Betrieb ist die Hauptaufgabe der Sendeabwicklung die Signalkontrolle, die Abwicklung von Live-Beiträgen mit Hilfe des Sendeablaufmischers (Abb. 10.74) und ggf. die Einleitung von Havariemaßnahmen. Bei Live-Beiträgen muss die Sendeabwicklung mit der Sendeleitung kommunizieren, damit ein reibungsloser Übergang vom Live-Beitrag zum Hauptprogramm gewährleistet ist. Hier werden von der Abwicklung Cleanfeed-Mitschnitte, also die Aufzeichnung ohne Logos, Bauchbinden etc. verlangt, damit der Beitrag einer weiteren Verwertung zugeführt werden kann.

Abbildung 10.75 zeigt die Grundstruktur einer einfachen Sendeabwicklung, die natürlich auch das Audiosignal verarbeitet. Ein besonderes Charakteristikum für den Betrieb ist die herausragende Bedeutung der Betriebssicherheit [7]. Sendeabwicklungssysteme müssen mit aufwendigen Havariekonzepten, möglichst mit parallel laufenden Einspiel-Speichern, doppelten Signalwegen und unterbrechungsfreier Stromversorgung versehen sein.

Ein gesamter Sendekomplex umfasst neben dem Sendestudio die Senderegie und die Maschinen- und Sendezuspielräume. Die Sendestudios sind relativ klein und oft mit Roboterstativen für die Kameras ausgestattet. Die Senderegie enthält mehrere Ausgänge, die über die Sendeablaufsteuerung bedient werden [76]. Neben dem Sendeablaufmischer stehen manchmal kleinere Produktionsmischer zur Verfügung. Hinzu kommen DVE, Standbildspeicher und Schriftgenerator sowie Audiomischer und Audiozuspielgeräte. Der Sendeablaufmischer ist technisch ein gewöhnlicher Digitalmischer, der der Sendeablaufsteuerung den Zugriff auf die notwendigen Funktionen erlaubt und den Mischerstatus

Abb. 10.74. Sendeablaufmischer [208]

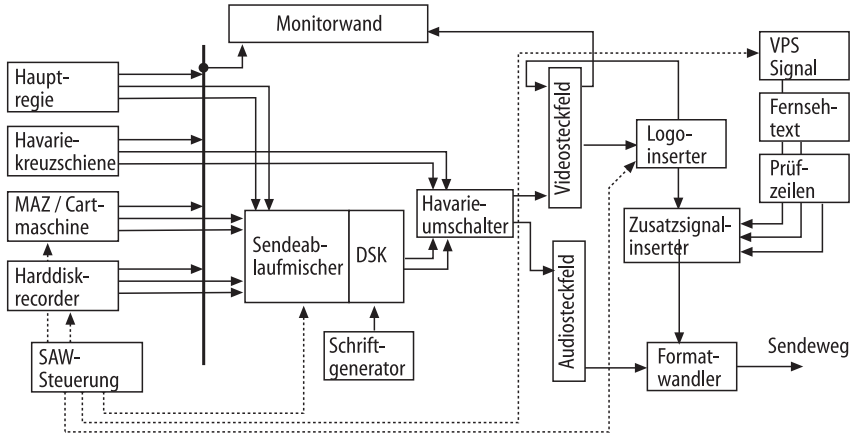


Abb. 10.75. Blockschaltbild der Sendeabwicklung

zurückmeldet. Somit kann der Mischer automatisch oder manuell gesteuert werden. Auch die Zuspielung der Beiträge geschieht meist automatisch, in der Regel vom Server aus. Früher stand dafür ein Cassettenautomat (Cartmaschine) zur Verfügung, der die Cassetten mit den Beiträgen mit einem Roboterarm einem von mehreren Playern zuführte. Die Cassetten mussten dafür an einem eigenen Arbeitsplatz vorbereitet und mit Labels versehen werden, die auf den Cassettenrücken geklebt wurden (Abb. 10.76).

Für die gesamte Steuerung existiert ein Sendeabwicklungssystem. Hierin werden die in der Sendeleitung erstellten Sendepläne mittels der Sendeablaufbearbeitung in eine Folge von Datensätzen umgesetzt und mit Kennungen, Trailern, Hinweisen etc. verbunden. Die Datensätze werden dann der eigentlichen Sendeablaufsteuerung (SAST) zugeführt. Diese steuert im Wesentlichen den Server und den Sendeablaufmischer, dabei kann der Ablauf bis kurz vor der Sendung verändert werden.

Zunehmend werden diese Speicher-, Steuer- und Misch-Funktionen heute in Computersystemen integriert. Das System verwaltet dann nur noch digitale Daten auf Servern, steuert Zuspielgeräte und bietet die Steuerung per PC, wie z. B. ColumBus Station Automation der Firma OmniBus und Avid AirPlay. Damit entfällt der Sendeablaufmischer als eigenständiges Gerät, außerdem viele Kopiervorgänge und MAZ-Kontrollen. Als Neuerung kommt für die Sendeabwicklung hinzu, dass bei digitaler Ausstrahlung über die verwendeten Datenraten eines Einzelprogramms, unter Berücksichtigung der anderen Programme im Multiplex, entschieden werden muss und neben dem Bild- auch das Datenstrom-Monitoring erforderlich ist.

Abb. 10.76. Label für Cassettenautomaten

10.5.3.1 Hauptschaltraum

Zentraler Knotenpunkt für die Verbindung der Produktionskomplexe, der Sendekomplexe und der Außenwelt ist der Hauptschaltraum (HSR). Hier laufen alle Signale zusammen, einschließlich Audio- und Kommandoleitungen. Die Hauptaufgaben der Mitarbeiter im HSR sind, die schaltungstechnischen Verbindungen herzustellen, die Signale zu überwachen und ggf. zu korrigieren, sowie bei Bedarf Signalwandlungen durchzuführen.

Kern des HSR ist die Zentralmatrix-Kreuzschiene für Audio und Videosignale, über die die Sendeabwicklung, Sende-MAZ und Senderegie untereinander und mit der Außenwelt verbunden werden. Mit dem Einsatz digitaler Signale erlauben neuere Konzepte die Verwendung dezentraler Kreuzschienen, die den einzelnen Produktionseinheiten fest zugeordnet sind, sodass jede Einheit die von ihr benötigten Signale selbst schalten kann. Dieses Konzept erhöht die Flexibilität und die Sicherheit des Sendebetriebs. Die Abteilungen können die Signale über bidirektionale Verbindungen zwischen den einzelnen Kreuzschienen oder mittels Netzwerken untereinander austauschen.

Ein großer HSR ist auch für das Schalten von Kommando- und Kommunikationsverbindungen zuständig und muss Konferenzschaltungen abwickeln können. Wenn die Beiträge über Satelliten empfangen und Sendungen zum Großteil über Satelliten abgestrahlt werden ist es sinnvoll, auch die Satelliten-Steuerung mit in den HSR zu integrieren [209].

10.5.4 Vernetzte Produktion

Im Kapitel 10.3 wurde bereits deutlich, dass Videoproduktionen durch den Einsatz von Netzwerken erheblich schneller und kostengünstiger werden können. Die Entwicklung filebasierter Produktion mittels Netzwerken begann mit den Newsroom-Systemen. Nachrichtensendungen stellen im Vergleich zu anderen Produktionen besondere Anforderungen. Oft ist die Aktualität wichtiger als die höchste Qualität. In diesem Bereich war es daher bereits früh möglich, mit stark datenreduzierten Signalen zu arbeiten, die die damals weniger leistungsfähigen Netze nicht über die Maßen beanspruchten.

Mit der Einführung von Newsroom-Systemen fand ein erheblicher Wandel der Produktion im Nachrichtenbereich statt, der weit über die Veränderungen der Arbeitsabläufe durch Digitalisierung bei gewöhnlicher Produktion von Magazinen, Talkshows, Musiksendungen etc. hinausgeht. Die Veränderungen im News-Bereich sind in besonderem Maß von dem Wunsch nach Steigerung der Geschwindigkeit der Nachrichtenerstellung getrieben [15].

Um die neuen Dimensionen zu erfassen, ist ein kurzer Ausblick auf die bisher üblichen Arbeitsabläufe, z. B. bei der Nachrichtenproduktion, erforderlich: Nach Recherchen, Materialsammlung und Vorarbeiten auf journalistischer Seite wird gewöhnlich ein EB-Drehteam (Elektronische Berichterstattung) aus 2 bis 3 Personen zusammengestellt, das Bild- und Tonmaterial für einen aktuellen Beitrag sammelt. In der Redaktion wird das Material evtl. mit Hilfe von Arbeitskopien, gesichtet und es wird ein Kommentar oder ein Text für den Moderator im Studio erstellt.

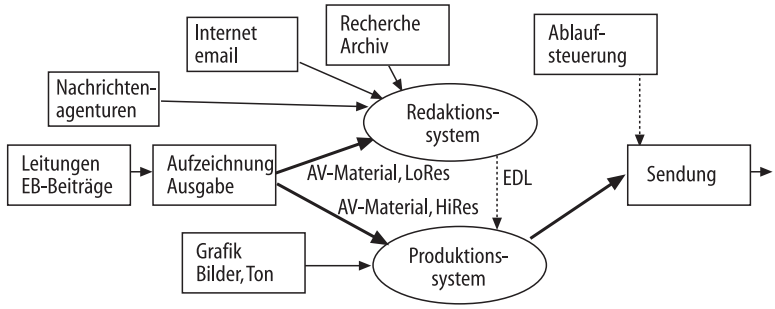


Abb. 10.77. Newsroom-System

Anschließend wird das Material mit Hilfe einer Cutterin nach den redaktionellen Vorgaben editiert Der fertige Beitrag wird wiederum gespeichert und in der Grafikabteilung weiter bearbeitet. Die Anmoderation des Beitrags erfordert zusätzlich einen eigenen Text, der dem Moderator meist per Teleprompter präsentiert wird. Darüber hinaus müssen Info-Texte eingegeben und genau so wie Standbilder abgerufen werden. Schließlich gelangt der Beitrag in das Archiv. Die Gesamtarbeit ist von viel Textarbeit geprägt.

Erste Änderungen in diesem Bereich entstanden durch Textverarbeitung in PC-Systemen, hinzu kamen moderne, digitale Kommunikationsmedien, die die Recherche vereinfachten und beschleunigten. Mit den Newsroom-Systemen (Abb. 10.77) konnte dann eine völlig neue Qualität durch effizientere und schnellere Arbeitsabläufe mit erweiterten Recherchemöglichkeiten zu günstigeren Preisen erreicht werden. Bekannte Anbieter für Newsroom-Systeme sind Avid und Sony. Die dominante Innovation ist die Vernetzung aller Abteilungen und Komponenten. Über das Netz können sowohl die Videodaten als auch Informationen zu und über die Daten verteilt werden. Dazu gehört auch die Verbindung der Fernsehproduktion mit den Hörfunk- und Internetangeboten der Sender, die mit dem Stichwort trimediales Arbeiten gekennzeichnet ist [36].

Die Server können verschieden, z. B. als sogenannte Daily Server, On Air Buffer und Archiv-Server, aufgeteilt sein. Die Einspielung der Beiträge (Ingest) geschieht über SDI, SDTI oder Videonetzwerke z. T. mit mehr als Echtzeitgeschwindigkeit oder auch über Satellitenzuspielung.

10.5.4.1 Browsing und Editing

Ein großer Vorteil der Digitaltechnik ist in diesem Zusammenhang die Möglichkeit der Bereitstellung verschiedener Quality of Service (QoS) und damit verschiedener Datenraten. Hier ist das z. B. flexible MPEG-2- oder auch das H.264-Signal sehr geeignet. Der Einsatz variabler QoS ist eine einfache Möglichkeit zum Umgang mit begrenzter Netz-Bandbreite. Die Videodaten befinden sich in hochqualitativer Form auf einem Media- oder Sendeserver und können von dort ausgespielt werden. Von diesen Daten können automatisch Kopien mit minderer Qualität und erheblich verringerter Datenrate erzeugt werden, die dann auf einem separaten, sogenannten Browse-Server abgelegt werden und über das Netz zu den einzelnen Redakteursarbeitsplätzen gelangen

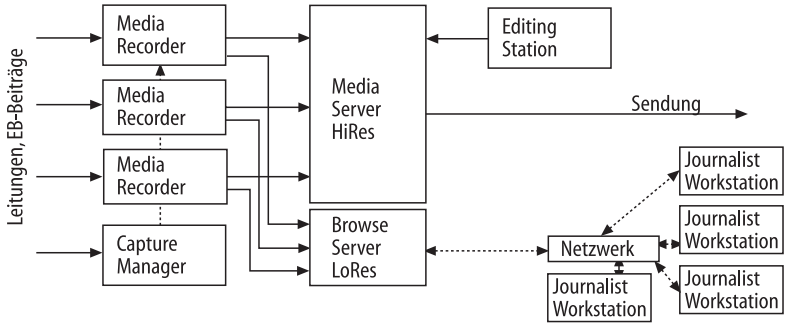


Abb. 10.78. Datenverteilung auf Browse- und Editing-Server

(Abb. 10.78 und 10.80). Die Redakteure können alles Material direkt sichten, auswählen und kommentieren [240]. Die Auswahl kann in Form einer Grob-Schnittliste vom Editor abgerufen werden, der dann anhand des hochqualitativen Signals den Feinschnitt durchführt, wobei dieser Prozess wiederum über das Netz von den Redaktionen beurteilt und beeinflusst werden kann (Abb. 10.79). Das Editing geschieht mit Hilfe von Schnittlisten, die auf das Online-Material verweisen.

10.5.4.2 News Room Automation

Durch Verweise auf die fertigen Bestandteile einer Sendung kann dann auch der Sendeablauf erstellt werden. Dabei ist es bis zur letzten Minute vor oder gar während der Sendung möglich, in den Ablauf einzugreifen. Die Redaktionen erstellen dazu Run-Down-Listen für die Auswahl der Beiträge auf dem Server. Mit Hilfe dieser Listen lassen sich auch Peripheriegeräte, wie Schriftgeneratoren, Grafiksysteme, Teleprompter-Einspielungen bis hin zu Roboterstativen für die Kameras steuern.

10.5.4.3 Nachrichtenverteilung

Während die Server und das Netzwerk die technischen Kernpunkte des Systems sind, ist der redaktionelle Kontenpunkt die Journalist Workstation und das Nachrichtenverteilsystem, das vor allem dazu dient, Agenturmeldungen von DPA, Reuters, AP etc. empfangen zu können [232]. Für die Recherche und Erstellung des jeweils aktuellen Berichts hat die Journalist Workstation auch Zugriff auf viele Kommunikationssysteme wie Internet, Fax und Archive. Wichtig ist dabei die Qualität der Suchfunktionen und die Darstellung der Ergebnisse.

Je nach Zugriffsberechtigung kann von der Workstation aus auch auf die Systemkomponenten und die Server zugegriffen und es können einfache Audio- und Videobearbeitungen durchgeführt werden. Ein echtes Newsroom-System zeichnet sich dadurch aus, dass vor allem eine umfassende Vernetzung der Textredaktion mit den Browsing/Editing-Systemen und den Maschinen-Automationssystemen gegeben ist. Die Journalist Workstation enthält natürlich einen Texteditor und evtl. die Möglichkeit, Beiträge direkt zu vertonen (Voice Over).

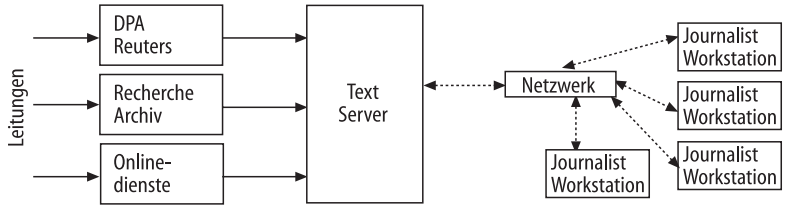


Abb. 10.79. Zugriffsmöglichkeiten für die Journalisten

10.5.4.4 Archivierung

Zum Schluss gelangen die Beiträge in das Archiv das auch als Quelle für die Erstellung neuer Beiträge sehr wichtig ist. Das Archiv hat auch unabhängig von Newsroom-Systemen große Bedeutung, denn hochwertig produzierte Beiträge werden zunehmend zu Kapital für die Produzenten. Der Grad der Wiederverwertung von Beiträgen steigt, denn die Erstausstrahlung deckt nur ca. 60% der Produktionskosten [160].

Traditionelle Archive enthalten ggf. auch sehr viele alte Filme und Videobänder, die bei Anforderung manuell gefunden und entnommen werden müssen. Dabei sind für alle Formate Abspielgeräte vorzuhalten. Aufgrund von Verschleiß von Bandmaterial und Maschinen müssen relativ häufig Umkopierungen vorgenommen werden.

Zum Erhalt der Informationen ist es wichtig, alle Beiträge digital zu speichern, um weitere Qualitätsverluste zu minimieren. Zusätzliche wichtige Punkte sind die Verbesserung der Suchmöglichkeiten und die Steigerung der Zugriffsgeschwindigkeit auf die Beiträge. Der Benutzerführung im Archiv kommt dabei sehr große Bedeutung zu. Auch die Wahl der Kompressionsart ist wichtig und unterliegt besonderen Bedingungen – z. B. wenig Artefakte, dafür kein Echtzeiterfordernis. Die neuen Funktionen erfordern ein vernetztes Archiv bei dem die Informationen nicht mehr physikalisch auf Datenträgern ausgegeben werden.

Abb. 10.80. Media-Server und Netzwerkverwaltung [240]

Auch das Archiv sollte zum Teil als Server ausgelegt und im Netzverbund integriert sein, sodass die Journalisten wiederum auf einfache Weise auf diese Beiträge zugreifen können. Um sich schnell einen Überblick über das Material verschaffen zu können, ist es günstig, wenn die Beiträge auch hier auf Browse-Servern gespiegelt werden. Die Langzeitarchivierung eines großen Bestandes überfordert ggf. die Kapazität von Servern, sodass für diesen Bereich auch noch digitale Bandformate zum Einsatz kommen, die mit Hilfe von Robotersystemen die gewünschten Beiträge automatisiert auf den Server übertragen. In diesem Zusammenhang zeichnet sich als neues Berufsfeld das des sogenannten Media-Managers ab, der das Einspielen, Löschen und Archivieren steuert und überwacht.

10.5.4.5 Content Management Systeme

Dieser Begriff bezeichnet ein umfassendes System zur Verwaltung der produzierten und archivierten Inhalte ohne besondere Rücksicht auf Signale und Datenformate. Zum Content gehören die eigentlichen Audio- und Videodaten sowie die Metadaten, also Beschreibungstexte, Beispielbilder, Timecode etc. Das System erleichtert die Recherche und die Wiederverwertung der Produktionen. Dabei geht es auch um den Verkauf von Programmen. In diesem Zusammenhang ist die Frage der Verwertungsrechte bedeutsam. Content und Verwertungsrechte werden zusammen als Asset bezeichnet.

10.5.5 SNG-Fahrzeuge

Die Satellitenübertragung ist in Kapitel 4 dargestellt. Sie wird als ein wichtiger Fernsehverteilweg für die Distribution an die Zuschauer eingesetzt, aber, im Rahmen des Satellit News Gathering (SNG), auch in vielfältiger Weise für die Contribution, also die Zuspielung zu den Funkhäusern.

SNG-Fahrzeuge kommen den immer weiter steigenden Ansprüchen an Aktualität bei der Nachrichtenproduktion entgegen. Die Sendeanstalten möchten den Zuschauern ein Gefühl der unmittelbaren Nähe zum Geschehen vermitteln und schicken möglichst schnell Reporter zum Schauplatz der Ereignisse, deren Berichte dann live in Nachrichtensendungen eingespielt werden. Die dazu eingesetzten SNG-Fahrzeuge sind spezielle Ü-Wagen (s. u.), die für die besonders zeitsparende Übertragung ausgelegt sind. Dazu muss das Fahrzeug schnell genug sein, d. h. möglichst eine PKW- statt einer LKW-Zulassung haben und das technische System muss ohne lange Rüstzeiten in Betrieb zu nehmen, sowie autark sein, d. h. unabhängig von technischer Infrastruktur und Energieversorgung vor Ort.

Aufgrund des Einsatzes kleiner, leichter und wendiger Fahrzeuge unterliegt die technische Ausstattung vielen Beschränkungen. Unverzichtbar sind aber ein Generator zur Stromversorgung, eine Klimaanlage und ein Satellitenübertragungssystem für Sendung und Empfang. Als A/V-Ausstattung müssen wenigstens eine Kamera, ein Speichersystem und ein drahtloses Mikrofon sowie ein Kommandosystem zur Verfügung stehen. Die größte Flexibilität bezüglich der Signalübertragung zum Funkhaus ergibt sich wenn eine Zuspielung über

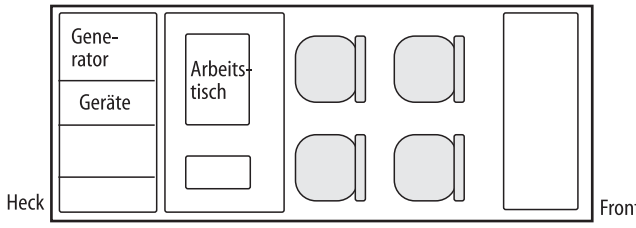


Abb. 10.81. Raumaufteilung im SNG-Fahrzeug

eine Satellitenstrecke gewählt wird. Die Richtfunkübertragung (RiFu) ist zwar auch möglich und teilweise kostengünstiger, doch ist sie nur auf kurzen Strecken mit entsprechender Infrastruktur, z. B. in Städten mit ansässigen Funkhäusern einsetzbar [133].

10.5.5.1 Fahrzeugtechnik

Eine häufige Forderung ist die Beschränkung des zulässigen Gesamtgewichts, damit das Fahrzeug als PKW zugelassen werden kann. Die größeren Typen dieser Kategorie haben ca. 4 m Länge. Die Satellitenantenne ist auf dem Dach angebracht, dort nimmt sie kaum Platz in Anspruch und braucht für Sendezwecke nur ausgefahren zu werden. Der Generator und die A/V-Anschlussfelder werden im hinteren Wagenteil installiert. Die Geräte innerhalb des Fahrzeugs werden oft in quer stehende Gestelle eingebaut, ähnlich wie es bei den kleinen Ü-Wagen geschieht, damit sind die Geräterückseiten von der Hecktür aus erreichbar. Vor der Gestellfront befinden sich zwei Sitzplätze. Zwei weitere können dahinter verfügbar sein, wenn die Fahrer- und Beifahrersitze drehbar eingebaut werden (Abb. 10.81).

10.5.5.2 SNG-Video- und Audiotechnik

Für die Bildaufnahme sollten nach Möglichkeit zwei oder mehrere Kameras verfügbar sein, sodass die Zuschauer z. B. neben dem Bild vom Moderator mittels einer Totaleinstellung der zweiten Kamera einen Gesamteindruck vom Ort des Geschehens bekommen können. Die Kameras sollten den Studiobetrieb ermöglichen. Dazu werden sie mit einem Multiplexsystem ausgestattet, das in der Lage ist, wie im Studio Audio- und Videosignale bidirektional über nur eine Leitung zwischen der Kamera und der im Wagen befindlichen CCU zu übertragen. Hierüber kann dann auch das Rückbild für den Reporter geleitet werden, sodass der zugehörige Rückbildmonitor am Kameraadapter angeschlossen wird. Eine Kamera sollte als Camcorder ausgelegt sein.

Weitere videotechnische Einrichtungen sind Steckfeld, Taktgenerator, Verteilverstärker und ein Framestore-Synchronizer zur Anpassung des Signals vom Satellitenempfänger an den Zentraltakt des Wagens. Zur Aufzeichnung sollte ein Recorder oder Server vorhanden sein, als Format wird z. B. XDCAM HD422 gewählt. Als Messsysteme stehen wie üblich Waveformmonitor und Vektorskop zur Verfügung. Das Monitoring umfasst z. B. einen hochwertigen 14“- und mehrere kleine 6“-Monitore.

Die gesamte Signalführung sollte so gestaltet sein, dass die Audio- und Videogeräte auch weitgehend umgangen werden können, damit das SNG-Fahr-

zeug auch als reine Uplink-Station verwendet und beispielsweise einem Ü-Wagen beigestellt werden kann (Abb. 10.82).

Auch das Vorhandensein eines Nachbearbeitungssystems ist wünschenswert. Dann kann der Beitrag vor Ort auch vorproduziert werden oder der Reporter kann live zum aufgezeichneten Material sprechen (News in Film, NIF). Hier bieten sich sehr leichte Laptop-Editoren an. Abbildung 10.83 zeigt ein Blockschaltbild der videotechnischen Ausstattung.

Die Audioausstattung umfasst mehrere, teilweise drahtlose Mikrofone, einen Audiomischer, einen Kompressor/Limiter, Audiopegelmesser und eine Abhöranlage. Die Audiosignalführung sollte zweikanalig sein, damit der Beitrag sowohl mit als auch ohne Kommentar, also als sogenannter IT-Ton (international) übertragen werden und so das Bildmaterial für andere Beiträge oder andere Anstalten verwendet werden kann. In das Audiosystem wird auch die Kommunikations- und Kommandoeinrichtung (Intercom) eingebunden. Besonders wichtig ist hier die Kommunikation zwischen Senderegie und dem Reporter vor Ort. Diese kann über die Satellitenstrecke oder über gewöhnliche Telekom-Einrichtungen laufen, sodass auch Festnetz- und drahtlose Telefone eingebunden werden müssen. Der Reporter wird über das Kamera-Anschlusskabel mit einem sogenannten (n-1)-Signal versorgt, sodass er die Summe aller notwendigen Intercom- und Audiosignale hört, nur nicht sich selbst.

10.5.5.3 Satellitenübertragungstechnik

Die Grundlagen der Satellitenübertragungstechnik sind in Abschnitt 4.3 beschrieben. Das SNG-Fahrzeug ist sowohl Sender als auch Empfänger des Satellitensignals. Für Up- und Downlink werden verschiedene Frequenzbe-

Abb. 10.82. Arbeitsraum im SNG-Fahrzeug [133]

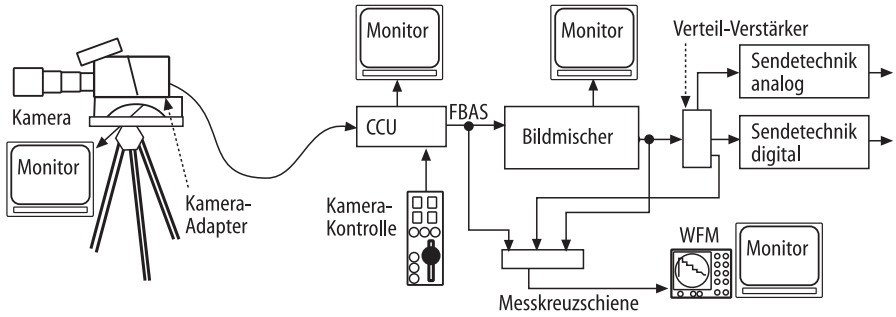


Abb. 10.83. Videotechnische Ausstattung im SNG-Fahrzeug

reiche verwendet, z. B. 14...14,5 GHz für den Uplink und 11,5...11,7 GHz sowie 12,5...12,7 GHz für den Downlink (Kopernikus Sat-System).

Standardmäßig werden ein Bild- und zwei Tonsignale übertragen. Bei Analogübertragung ist das Bildsignal PAL-codiert, als Modulationsart wird aus Störabstandsgründen FM verwendet. Das Tonsignal wird auf einen Tonunterträger moduliert, der einen Abstand von 6,6...7,2 MHz zum Bildträger hat. Die Analogübertragung erfordert eine Bandbreite von 27 MHz.

Bei der Digitalübertragung werden Audio- und Videosignale datenreduziert, als MPEG-2- oder H.264-Signale codiert und mit einem Fehlerschutz versehen. Als Modulationsart wird QPSK verwendet (s. Abschn. 4.5) und als Datenrate für ein SD-Signal z. B. 8,44 Mbit/s gewählt [133]. So belegt der Programmbeitrag nur 8 MHz und damit erheblich weniger Bandbreite als bei Analogübertragung. Für die HD-Übertragung kommt auch noch oft die MPEG-2-Codierung zum Einsatz, hier vor allem die drei, durch die EBU festgelegten MPEG-2-HD-Profile, 60HD, 42HD und 32HD mit Datenraten von jeweils 60, 42 oder 32 Mbit/s pro Programm. In Zukunft wird die Übertragung vermehrt auf den datenratensparenden H.264-Codes der Klassen 11HD, 16HD, 22HD und 32HD beruhen, die in Abbildung 10.58 genauer dargestellt sind.

Der Modulator liefert ein 70 MHz-Signal, welches im ersten Upkonverter in das L-Band um 1 GHz heraufgesetzt wird. Auf dieser Ebene findet die Frequenzkontrolle und Leistungsregelung statt. Ein zweiter Upkonverter setzt dann die Frequenz in das Ku-Band um 14 GHz herauf und dieses Signal wird dann über den Leistungsverstärker der Antenne zugeführt (Abb. 10.84). Die HF-Leistungsverstärker für den Up-Link liefern oft mehr als 100 W. Es werden zwei Verstärker parallel über einen Phasenanpasser (Phase Combiner) betrieben, um bei einem Ausfall den Betrieb ungestört fortführen zu können.

Für Sendung und Empfang kann dieselbe Antenne verwendet werden. Aufgrund des besseren Antennengewinns werden relativ große Reflektoren von z. B. 1,5 m Durchmesser verwendet. Große Spiegel müssen exakter ausgerichtet werden als kleine, im SNG-Fahrzeug stehen dafür motorische Steuerungen zur Verfügung. Die exakte Ausrichtung erfolgt mit Hilfe von Spektrum-Analysatoren, dabei wird die Antenne geschwenkt und geneigt bis das Bakensignal des Satelliten maximale Amplitude aufweist. Damit die Antenne die Ausrichtung beibehält, kann das gesamte Fahrzeug durch Stützen stabilisiert werden.

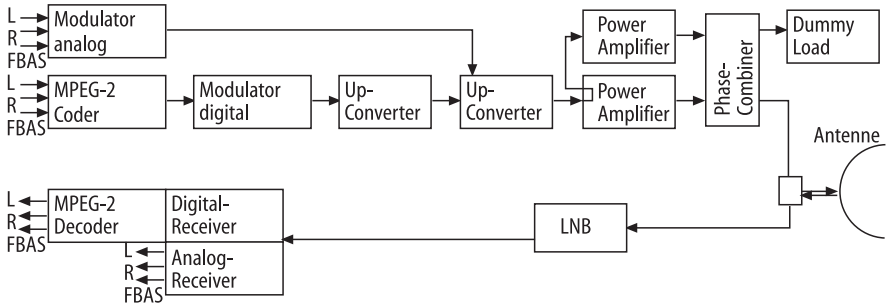


Abb. 10.84. Blockschaltbild der Sende- und Empfangstechnik im SNG-Fahrzeug

Für den Empfang des Down-Link-Signals stehen analoge und digitale Satellitenreceiver bereit (Abb. 10.84). Es ist günstig, wenn zur Kontrolle des über das Funkhaus abgestrahlten Beitrags das Signal über eine separate Flachantenne und einen zweiten Sat-Receiver nochmals vom Broadcast-Satelliten (z. B. Astra) empfangen werden kann.

Für die Überspielung des Nachrichtenbeitrags wird die Satellitenstrecke für eine bestimmte Zeit angemietet. Die Mietkosten sind bei digitaler Übertragung erheblich günstiger als bei analoger. Die Sendeanlage braucht eine Zulassung für die jeweiligen Satelliten (Earth Station Code). Die Verwaltung der deutschen Satellitensysteme geschieht durch die deutsche Telekom, zum Aufbau der Verbindung muss der SNG-Operator mit der Sendetechnik der Erdefunkstelle, z. B. der Telekom in Usingen kommunizieren. Dabei werden als Testsignale ein 100/75 Farbbalken mit Kennung der SNG-Einheit und ein 1 kHz-Pegelton mit –3 dBu, also 9 dB unter Vollaussteuerung ausgestrahlt. Die Sendeleistung wird zunächst ohne Modulation langsam erhöht, bis in Usingen eine optimale Aussteuerung festgestellt wird, anschließend wird die Modulation zugeschaltet. Das Downlink-Signal wird vom SNG-System empfangen und kontrolliert, das gleiche geschieht im Funkhaus, wo das Signal im Schaltraum überprüft wird. Auch dorthin muss vom SNG-Fahrzeug aus eine Kommunikationsverbindung bestehen. Diese wird meist über das drahtlose Telefonnetz realisiert, dessen Signal in die Intercom eingebunden wird.

Die Bereitstellung von Übertragungswegen wird in Funkhäusern oft durch sogenannte Leitungsbüros unterstützt. Für eine Transponderbuchung sind folgende Informationen erforderlich: Datum und Zeit, Kennung (Earth Station Code) und Telefonnummer am Uplink, Standort der Uplink-Anlage und ggf. Wunschsatellit -slot und -datenrate.

Nachdem die Verbindung steht muss bei Beginn der Sendung die Laufzeitdifferenz zwischen den Funkhaus-Regiekommandos und dem Rückbild auf dem Monitor des Moderators beachtet werden. Das Rückbild kommt oft vom Broadcastsatelliten und durchläuft damit zwei Satellitenstrecken womit erhebliche Verzögerungen gegenüber der Telefonübertragung auftreten können, insbesondere wenn digital gesendet wird. Ein Moderator darf sich also nicht auf das Rückbild verlassen, sondern muss bei Beitragsbeginn auf das Funkhaus-Kommando hören.

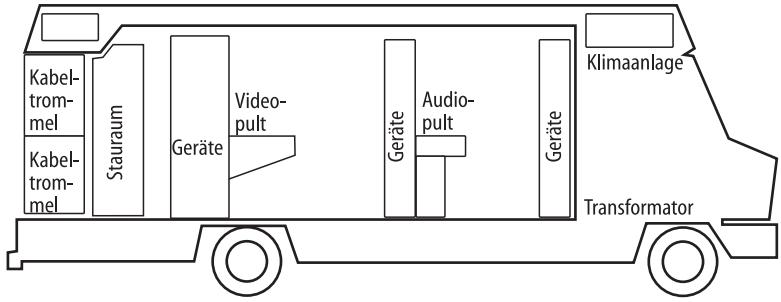


Abb. 10.85. Raumaufteilung in kleinem Ü-Wagen

10.5.6 Ü-Wagen

Übertragungswagen dienen der Abwicklung von Außenübertragungen (AÜ) (engl.: Outside Broadcasting, OB), wie z. B. bei Sportberichterstattungen. Die Übertragungswagen (Ü-Wagen, engl.: OB-Van) können in den meisten Fällen nicht wie Festregien oder SNG-Fahrzeuge für spezielle Aufgaben ausgelegt werden, sondern müssen universell für eine Vielzahl verschiedener Produktionen einsetzbar sein. Daher ist eine umfassende Ausstattung der Wagen erforderlich, die der in einer Festregie entspricht und oft noch darüber hinausgeht.

Die Ausstattung wird sowohl in den kleinen als auch in den größten Wagen von Platzproblemen dominiert, mit denen von den verschiedenen Herstellern sehr spezifisch umgegangen wird. Grundsätzlich werden nach Möglichkeit separate Abteilungen für die Bildregie, die Tonregie, für Bildtechnik und Aufzeichnung gebildet. Dabei wird angestrebt, dass zwischen Bild- und Tonregie eine Sichtverbindung bei gleichzeitig guter akustischer Abschirmung besteht.

Ein kleinerer Ü-Wagen hat z. B. ein Gesamtgewicht von 10 t und eine Länge von 8 m. Abbildung 10.85 zeigt eine mögliche Anordnung der Arbeitsbereiche. Ein derartiger Wagen kann mit bis zu vier Kameras ausgestattet werden. Es steht ein eigener Taktgeber zur Verfügung, der extern synchronisierbar ist. In einem solchen Wagen können auch Nachbearbeitungen durchgeführt werden. Das Führerhaus wird oft als Kommentatorplatz benutzt. Weiterhin stehen auch Intercom-Einrichtungen, Steckfelder, Framestoresynchronizer und digitale Videoeffektgeräte etc. zur Verfügung.

Große Ü-Wagen werden bei Großveranstaltungen und häufig für die immer aufwendigere Sportberichterstattung eingesetzt. Sie sind oft mit mehr als zwanzig Kameras, entsprechend großen Bildmischern und einer großen tontechnischen Abteilung ausgestattet. Die größten Ü-Wagen basieren auf Sattelschleppern. Sie werden so konstruiert, dass die Grenzen, die von der deutschen Straßenverkehrsordnung für Fahrzeuge gesetzt sind, gerade erreicht werden, d. h. der gesamte Sattelzug darf höchstens 16,5 m lang, 4 m hoch und 2,55 m breit sein und ein Gesamtgewicht von 40t aufweisen.

Die gerätetechnische Ausstattung ist so umfangreich, dass die Sattelzüge meist von Rüstwagen der gleichen Größe begleitet werden (Abb. 10.87). Die gestiegenen Ansprüche an die im Wagen vorhandenen Arbeitsplätze, wie z. B.

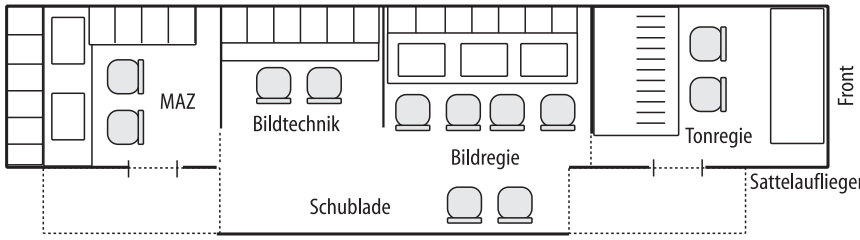


Abb. 10.86. Beispiel für eine Raumaufteilung in einem großen Ü-Wagen

die Einrichtung einer zweiten Regie für die internationale Berichterstattung [101], führte im Laufe der Zeit dazu, dass die Fahrzeuge nach dem Konzept der Schublade verbreitert werden können, wenn sie ihren Zielort erreicht haben (Abb. 10.86). Die größten Fahrzeuge sind beidseitig ausfahrbar. Ein solches Fahrzeug bietet komfortable Produktionsbedingungen, bringt aber manchmal auch Probleme hinsichtlich der verfügbaren Stellplatzflächen an Fußballstadien etc. mit sich. Daher wird heute versucht, Ü-Wagen mit möglichst flexibler Größe zu konstruieren. Ein Beispiel ist der Ü3 der Firma TVN, der hier exemplarisch näher betrachtet werden soll. Der Sattelauflieger ist nach dem genannten Schubladenkonzept konzipiert und zwar so, dass der Wagen sowohl mit einals auch mit zweiseitigem Auszug nutzbar ist. Abbildung 10.88 zeigt die dabei mögliche Vergrößerung der zur Verfügung stehenden Arbeitsbereiche [222].

Der Auszug der in den Schubladen befindlichen Geräte ist nicht unproblematisch, denn er erfordert eine bewegliche Führung der Leitungen über Kabelketten (Abb. 10.87 rechts). Zudem haben die Ketten nur eine begrenzte Tragfähigkeit [98]. Der Auszug ist daher meist nicht vollständig vorgesehen, die Orte der Geräte, die über sehr viele Leitungen angeschlossen werden, wie die externe Signalzuführung am XKT (externe Kabeltafel) und der zentrale Gerä-

Abb. 10.87. Beidseitig ausfahrbarer Ü-Wagen mit Rüstwagen, Kabelführung unter dem Auszug

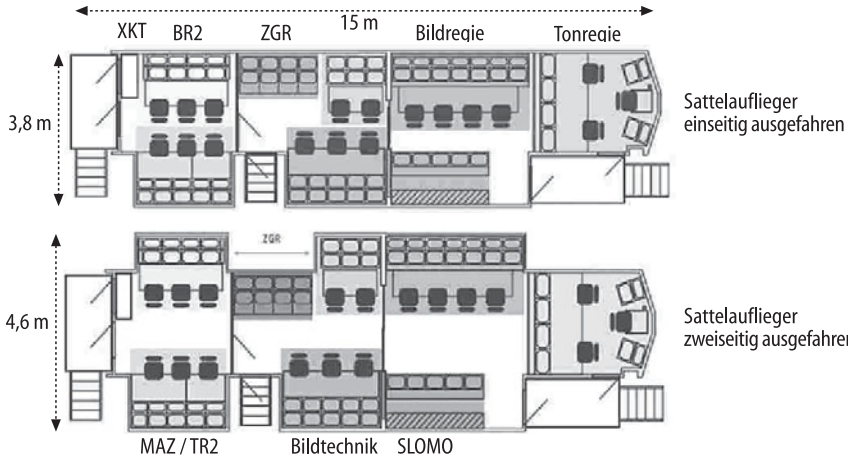


Abb. 10.88. Raumaufteilung in flexiblem Ü-Wagen [222]

teraum (ZGR) behalten eine feste Position, ebenso wie die Tonabteilung (Abb. 10.88). Die Tonregie (Abb. 10.89), die für Surroundproduktionen ausgelegt ist, ist schallisoliert im vorderen Teil des Sattelaufliegers untergebracht. Die Tonproduktion erfolgt mit dem Kreuzschienen- und Mischsystem der Firma Lawo. Die Tonabteilung ist auch für die Konfiguration der Kommandowege zuständig. Hinter der Tonregie befinden sich die Bildregie und die Bildtechnik sowie die Slomo-Plätze, an denen die Steuereinheiten für die Super-Slow-Motion-Systeme untergebracht sind. Mit diesen werden die im Sportbereich erforderlichen sehr weichen Zeitlupeneffekte erreicht. Am Heck des Wagens ist eine zweite Bild- und Tonregie und der MAZ-Bereich eingerichtet. Die XKT ist am Heck von außen zugänglich. Die folgenden Abbildungen zeigen die Räume.

Abb. 10.89. Tonregie im Ü-Wagen [222]

Abb. 10.90. Bildregie im Ü-Wagen [222]

Abbildung 10.90 zeigt die Bildregie. Hier kommt der Bildmischer DVS 8000 der Firma Sony zum Einsatz. Das Monitoring ist flexibel angelegt. Es geschieht mit drei Grade 1-Röhrenmonitoren. Weitere Signale werden auf Teiflächen von großen LC-Flachdisplays dargestellt, die in Abbildung 10.90 dunkel unter den Röhrenmonitoren sichtbar sind. Auf einem der LCD können z. B. neun Vorschausignale dargestellt und frei skaliert werden, u. a. auch bezüglich des Bildseitenverhältnisses. Das gleiche gilt für die Darstellung der Vorschausignale in der Bildtechnik, die zusammen mit den RCP für die Kamerakontrolle in Abbildung 10.91 dargestellt ist. Videotechnisch ist der Ü-Wagen für Multinormproduktionen ausgelegt, er kann im SD- und HD-Format und bei letzteren

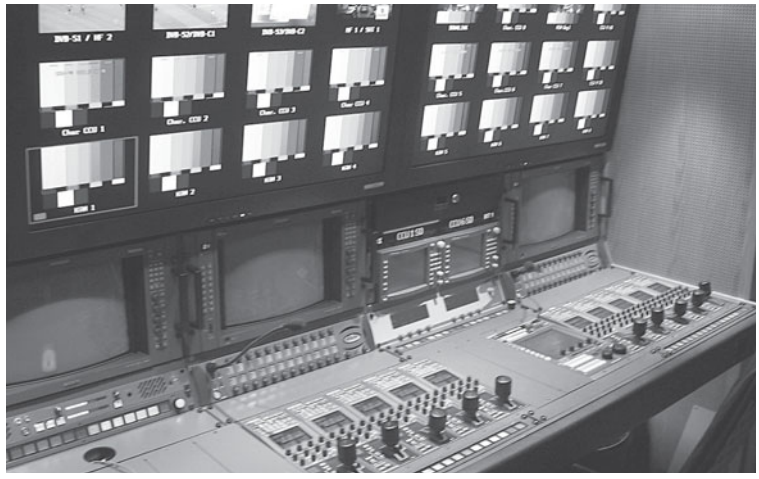


Abb. 10.91. Bildtechnik-Bereich im Ü-Wagen [222]

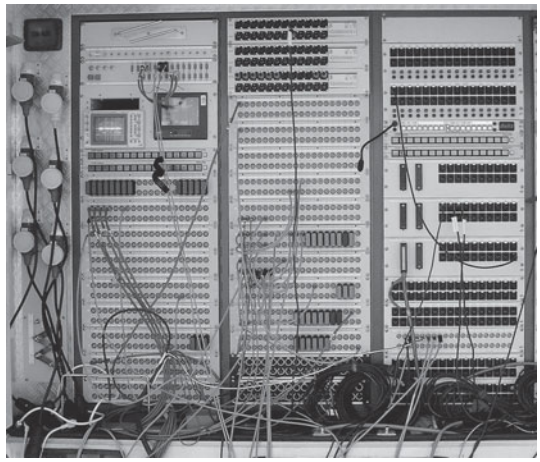


Abb. 10.92. Beispiel für ein Anschlussfeld im Heck des Ü-Wagens und Hybridstecker

wahlweise mit 720p oder 1080i produzieren. Die Kamerazüge basieren auf der HD-Systemkamerareihe HDC-xxx von Sony, die direkt an der Kamera auch 1080p-Signale bereitstellt. Es kommen Studioobjektive von Canon mit 25-, 86- und 100-fach Zoom zum Einsatz.

Aufgrund der langen Kabelwege, die insbesondere bei der Sportberichterstattung auftreten können, ist eine Signalübertragung über Hybridkabel gemäß SMPTE 311 vorgesehen. Diese enthalten eine Glasfaser und mindestens zwei Kupferleitungen. Die Steckverbindung ist in Abbildung 10.92 neben dem Anschlussfeld (XKT) zu sehen.

Die Aufzeichnung erfolgt in den Formten HDCAM und HDCAM SR und XDCAM HD422. Abbildung 10.93 zeigt den MAZ-Bereich im Ü-Wagen, in

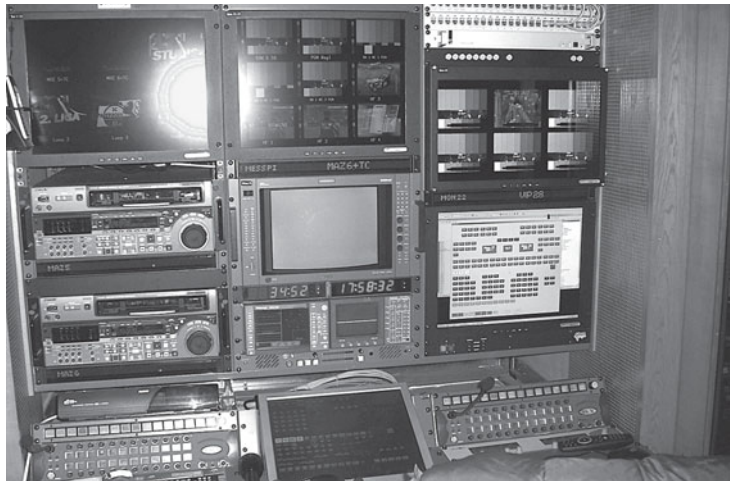


Abb. 10.93. MAZ-Bereich im Ü-Wagen [222]

Abbildung 10.94 ist der Raum mit den Slomo-Steuereinheiten zu sehen. Die Signalverteilung erfolgt mittels Videosteckfeldern (VSF) und einer digitalen, multiformatfähigen digitalen Kreuzschiene (DKS), die auch zur Zuordnung analoger Signale verwendet wird die dafür A/D-Umsetzer durchlaufen. Die Kreuzschiene kann modular auf bis zu 512 x 512 Punkte erweitert werden und ist in der Lage, auch mit Glasfasermodulen zu arbeiten.

Die meisten der genannten Geräte können sehr umfangreich konfiguriert werden. Damit dies zentral geschehen kann, werden sie von einem System namens Virtual Studio Manager (VSM) der Firma Broadcast Solutions verwaltet. Das gilt insbesondere für die Bild- und Tonmischer, die Absetzkreuzschiene, Monitor- und Under-Monitor-Displyaverwaltung etc. Aus Gründen der Havariesicherheit wird das System mit zwei parallel laufenden Rechnern realisiert.

Der Wagen enthält Synchronizer und hochstabile Taktgeneratoren. Aufgrund der Multiformatfähigkeit müssen sowohl HD-Geräte mit Tri Level Sync (TLS) als auch SD-Geräte mit Black Burst-Sync (BB) versorgt werden (s. Abschn. 10.2.1). Die parallele Verwendung führt dabei manchmal zu Timing-Problemen, die geschickt ausgeglichen werden müssen [98].

Die Energieversorgung eines großen Ü-Wagens erfordert einen Dreiphasen-Wechselstromanschluss mit 3 x 400 V/125 A, weil nicht nur die A/V-Geräte, sondern vor allem auch die Klimaanlage mit Energie versorgt werden muss.

10.5.6.1 Super Slow Motion

Für den häufig auftretenden Einsatz bei Sportveranstaltungen ist es wichtig, dass im Ü-Wagen Slow-Motion-Controller für Zeitlupeneinspielungen verfügbar sind, sowie entsprechende Kameras (Supermotion) und Aufzeichnungsgeräte, die mit dreifacher Bildwechselfrequenz arbeiten (z. B. die Kameras Grass-Valley LDK 23 oder LDK 8300 für HD oder von Sony die HD-Systemkamera HDC-3300). Damit wird eine echte Zeitlupenwiedergabe ermöglicht, bei der nicht einfach Bilder wiederholt werden müssen. Auch bei dreifach verlangsamter Bewegung stehen immer noch 25 verschiedene Vollbilder pro Sekunde zur Verfügung sodass sehr weiche Zeitlupenwiedergaben möglich werden.

Für diesen speziellen Bereich sind als Aufzeichnungssysteme Diskrecorder verfügbar, die mehrkanalig mit Random Access arbeiten, sodass einerseits

Abb. 10.94. Slomo-Bereich im Ü-Wagen und EVS-Controller [222]

das Auffinden und Ausspielen markierter Slow-Motion-Abschnitte erheblich schneller geht und andererseits nur ein Recorder Aufnahme und Wiedergabe gleichzeitig realisieren kann.

Derartige Geräte werden von der Firma EVS als Super Live Slow Motion (SLSM) angeboten (Abb. 10.94 und 10.95). Sie sind insbesondere gefragt, wenn bereits sehr kurz nach aktuellen Ereignissen, wie z. B. bei Fußballspielen, eine Kurzfassung der Abläufe, der sogenannte Highlight-Schnitt erfolgen soll. Die Geräte sind so weit verbreitet, dass der Name EVS zum Synonym für Live Slow Motion und Servereinspielung geworden ist. Geräte wie die EVS XT2 haben sechs HD-SDI-Kanäle, die frei belegbar sind. Der Diskrecorder arbeitet intern mit Datenreduktion. Beim Starten des Systems kann die Konfiguration bestimmt werden, sodass z. B. vier Ein- und zwei Ausgänge zur Verfügung stehen. Über eine Kreuzschiene kann die Kanalbelegung beliebig geändert werden. Die Maschine zeichnet die Signale nach dem Start durchgängig auf und überschreibt die am Anfang aufgezeichneten Sequenzen, wenn sie nicht als Clips markiert und abgelegt werden. Ein Operator kann sehr schnell interessante Szenen aufsuchen und sie mit Geschwindigkeit zwischen + 400% und – 400% mit über den T-Bar gesteuerter, variabler Geschwindigkeit direkt abspielen, wobei auch Material verwendet werden kann, das nur 3 Bilder Abstand vom Livesignal hat. Die Clips können auch zu einem Gesamtsignal verbunden werden, somit steht eine einfache Schnittfunktion zur Verfügung, die mit voreingestellten Blenden von z. B. einer halben Sekunde verbunden sein kann.

Im Super-Slow-Motion-Betrieb werden drei Eingänge der EVS belegt, da das Videosignal von der Kamera mit der dreifachen Bildfrequenz als drei parallele HD-SDI-Signale ausgegeben wird, die auf die entsprechenden Kanäle des Diskrecorders aufgezeichnet werden. Bei der Wiedergabe mit 33% der Originalgeschwindigkeit werden die verschiedenen Bilder der Kanäle nacheinander ausgelesen und es entsteht ein gewöhnliches 50/60Hz-Halbbildsignal.

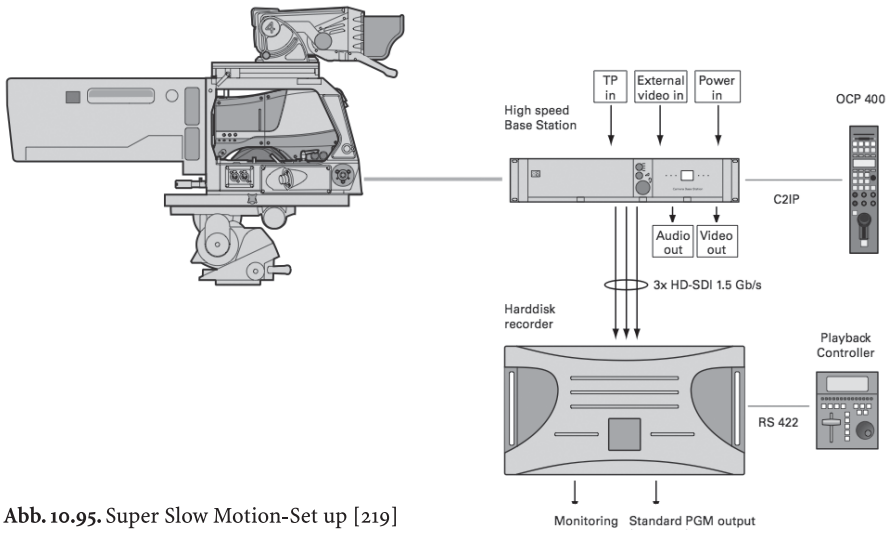


Abb. 10.95. Super Slow Motion-Set up [219]

10.6 Das virtuelle Studio

Der Begriff virtuelles Studio (VS) im weiteren Sinne steht für die Erzeugung künstlicher Szenen als Kombination realer Vordergrundobjekte mit nur scheinbar echten Hintergrundbildern, in der Art, dass die Kombination echt wirkt und die Täuschung nicht auffällt. Die künstliche Szene wird interessant, wenn sich die Kameraperspektive ändern kann oder ein Akteur im Vordergrund auf Veränderungen im Hintergrundbild reagiert – z. B. wenn ein Auto in der Hintergrundszene auf den Akteur zu fährt und er ausweicht. Letzteres erfordert wenig technischen Aufwand, ist aber aufwendig einzustudieren und in den meisten Fällen wird sich die gewünschte Illusion nicht einstellen (s. alte amerikanische Spielfilme).

Beim virtuellen Studio im engeren Sinne werden belebtere Bilder durch die Änderung der Kameraperspektive erzeugt. Dabei muss bei jeder Positionsveränderung der Kamera die Perspektive des Hintergrundbildes angepasst werden, damit die Illusion eines zusammengehörigen Bildes nicht zerstört wird.

Das virtuelle Studio beruht auf üblichen videotechnischen Geräten. Diese werden schon seit längerem nicht nur zur Abbildung der realen Umwelt eingesetzt sondern auch zur Erzeugung bestimmter Atmosphären und künstlicher Umgebungen. Bei vielen Videoproduktionen wird zudem mit Elementen wie Studiodekorationen, Rückprojektionen, Bildeinblendungen etc. gearbeitet, um das Bild interessanter, ansprechender oder informativer zu gestalten. Typische Beispiele sind Nachrichtensendungen und Wahl- und Sportberichterstattungen bei denen Fakten und Zahlen ansprechend präsentiert werden sollen. In einer Show bewegen sich Akteure in einer Studiodekoration, bei Nachrichtensendungen befindet sich ein Moderator oder die Nachrichtensprecherin vor einem festen Hintergrund, in dem Informationen eingeblendet werden. Das Hintergrundbild zeigt hier nicht die Dekoration, sondern kommt von einer fremden Videoquelle. Die Kombination von Vorder- und Hintergrundbild geschieht mit Chroma-Key-Technik. Die Person im Vordergrund befindet sich vor einer gleichmäßig ausgeleuchteten Wand (Abb. 10.97), deren Farbe im Vordergrund nicht vorkommt (oft Blau oder Grün). Eine Key-Einheit ersetzt Flächen von diese Farbe durch die externe Videoquelle (s. Abschn. 9.7).

Der Nachteil der klassischen Chroma-Key-Technik ist, dass es keine Verbindung zwischen dem von der Studiokamera aufgenommenen Vordergrundsignal

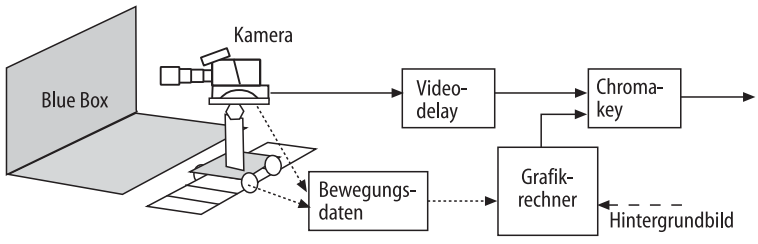


Abb. 10.96. Funktionsprinzip des virtuellen Studios

und dem eingeblendeten Hintergrundbild gibt, sodass der Chroma-Key-Trick sehr künstlich und störend erschiene bzw. die Illusion eines zusammenhängenden Bilder zerstört wäre, wenn die Kamera bei unverändertem Bildhintergrund bewegt werden würde.

Virtuelle Studios verändern die beschriebenen Vorgehensweisen. Sie ermöglichen vor allem die Änderungen der Kameraperspektive, machen Studiodekorationen überflüssig und führen zu völlig neuartigen Präsentationsformen. Künstliche Hintergrund-Szenen werden komplett in einem Grafikrechner erzeugt, wobei die Besonderheit beim virtuellen Studio im engeren Sinne darin besteht, dass die Bilder in Abhängigkeit von der Kameraposition in Echtzeit neu berechnet werden (Abb. 10.96). Die Vorteile des virtuellen Studios sind einerseits praktischer Art: Aufwendige Kulissen müssen nicht mehr real gebaut werden und es reicht ein blau ausgekleidetes Studio aus, um mehrere verschiedene Produktionen durchzuführen, da innerhalb kurzer Zeit eine Umschaltung zwischen den zugehörigen virtuellen Räumen vorgenommen werden kann. Andererseits gibt es Veränderungen inhaltlicher Art: Die Sehgewohnheiten verändern sich, die Präsentationen können lockerer und dynamischer wirken, was z. B. der Darstellung trockener Daten (Sportergebnisse) zugute kommen kann.

Die Echtzeitberechnung der Hintergründe ist streng genommen nur im sogenannten Online-Betrieb erforderlich, der praktisch nur realisierbar ist, wenn hochleistungsfähige Grafikrechner zur Verfügung stehen. Trotz des damit verbundenen Aufwands wird in den meisten Fällen der On- dem Offline-Betrieb aus Gründen der Flexibilität vorgezogen, da die richtige Perspektive jeder Zeit zur Verfügung steht und somit auch Live-Sendungen realisierbar sind, die offline nicht durchgeführt werden können.

Der Offline-Betrieb bietet dagegen den Vorteil, dass die erforderlichen Perspektiven in beliebig langer Zeit auch mit einfacheren Computern berechnet

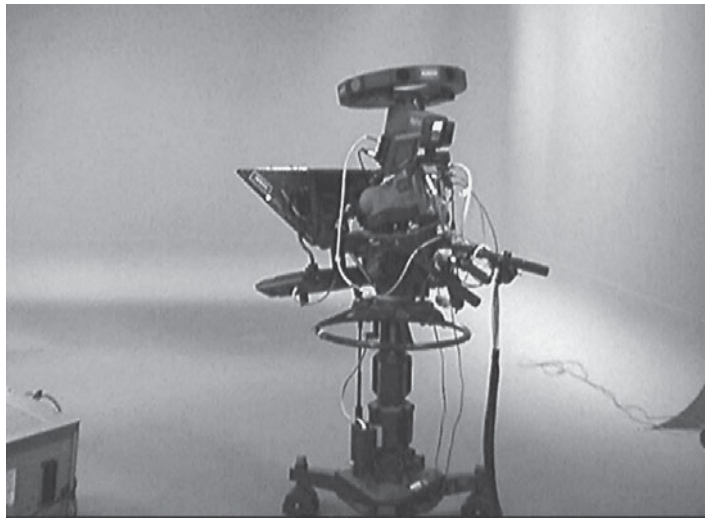


Abb. 10.97. Blauraum (Blue Box) mit Hohlkehlen als Basis des virtuellen Studios

werden können, sodass auf die Komplexität der Bilder wenig Rücksicht genommen zu werden braucht und auch Schärfentiefen gut simulierbar sind. Die Einsatzmöglichkeiten sind dadurch eingeschränkt, dass bei Kamerafahrten die Bewegungsdaten aufgezeichnet werden müssen. Anhand dieser Daten wird anschließend die Perspektive für das computergenerierte Bild (virtuelle Kamera) so berechnet, dass sie mit der Bewegung der realen Kamera übereinstimmt.

Aus ähnlichen Gründen wie bei der Entscheidung zwischen On- oder Offline-Betrieb, wird auch bei der Entscheidung zwischen dem Einsatz von 2- oder 3-dimensionalen Hintergrundbildern gewöhnlich die rechenintensive 3D-Variante gewählt. Der 2D-Betrieb ist erheblich preiswerter, die Funktionalität ist aber stark eingeschränkt. Es sind nur Zoom und Schwenkbewegungen der Kamera erlaubt, jedoch mit dem Vorteil, dass das Hintergrundbild einfach per DVE oder rechnergestützt in Abhängigkeit von diesen Parametern nachgeführt werden kann. Dabei wird beim Schwenk z. B. ein 1920 x 1080 Pixel großer Ausschnitt aus einem hochaufgelösten Bild mit z. B. 4096 x 3072 Pixel gebildet. Der 2D-Betrieb stellt quasi ein Mittelding zwischen der klassischen, statischen Chroma Key-Produktion und einem echten virtuellen Studio dar.

10.6.1 Technische Ausstattung des virtuellen Studios

Der Kern des virtuellen Studios ist der Hochleistungsrechner mit einer speziellen Software zur Echtzeitberechnung der Hintergründe, die Besonderheiten der weiteren Ausstattung beziehen sich vor allem auf die Verfahren zur Bestimmung der Kameraposition, die als »Tracking« bezeichnet werden. Die Studiokamera ist meist wie üblich bedienbar, sie ist aber mit speziellen Sensoren ausgestattet, die die Positionsdaten für die Schwenk- und Neigungswinkel, Zoom- und Fokuseinstellung sowie die Kameraposition im Raum in Echtzeit erfassen und an den Grafikrechner weitergeben. Videotechnische Grundlage ist das Chroma-Key-Stanzverfahren. Die Moderatoren können sich relativ frei in einem großen Raum bewegen, der einfarbig ausgekleidet und gleichmäßig ausgeleuchtet ist.

10.6.1.1 Das Studio

Das typische virtuelle Studio ist ein Raum der als Blue- oder Green-Box großflächig in Blau oder Grün ausgekleidet ist und bei dem Ecken und Kanten durch Hohlkehlen möglichst unsichtbar gemacht werden (Abb. 10.97). Wenn diese Voraussetzung erfüllt ist, ergibt sich die Größenwirkung des virtuellen Raums nur aus dem Set-Design im Rechner. Die absolute Größe des Studios ist dann allein in Relation zur Blue- oder Greenscreen-Fläche, zu den Beleuchtungseinrichtungen und der Größe der Vordergrundobjekte relevant.

Eine besondere, aber bisher kaum eingesetzte Blue-Box-Variante verwendet hochreflektierende Materialien anstelle des Blue Screen. Das blaue Licht des Hintergrunds kommt von einem Kranz blauer Leuchtdioden, die um das Kameraobjektiv angebracht sind und das an dem Hintergrundmaterial reflektiert wird. Das Verfahren erlaubt eine flexiblere Ausleuchtung des Studios und ein einfacheres Keying in schwierigen Situationen. Das Problem der starken Win-

kelabhängigkeit derartiger Reflexionen wurde gelöst: die BBC ließ ein Material entwickeln, das auch bei einem Winkel von 60° zur Vertikalen mehr als 60% des Lichts reflektiert.

10.6.1.2 Chroma Key

Der videotechnische Trick des virtuellen Studios ist das Stanzverfahren. Um hohen Ansprüchen gerecht werden zu können, kommen oft Chromakeyer der Firma Ultimatte zum Einsatz. Beim Ultimatte-Verfahren werden Masken für den Vorder- und Hintergrund gebildet, die separat bearbeitet werden können und am Schluss kombiniert werden. Damit können Schatten und auch teiltransparente Objekte wie Glas in die Szene einbezogen werden.

Das Verfahren sei am Beispiel einer Person und eines Glases auf einem Tisch beschrieben (Abb. 10.98). Der Vordergrund wird so maskiert, dass die Person und das Glas statt vor der Hintergrundfarbe vor Schwarz erscheinen. Je nach Transmissionsgrad des Glases wird der Hintergrund an der Stelle des

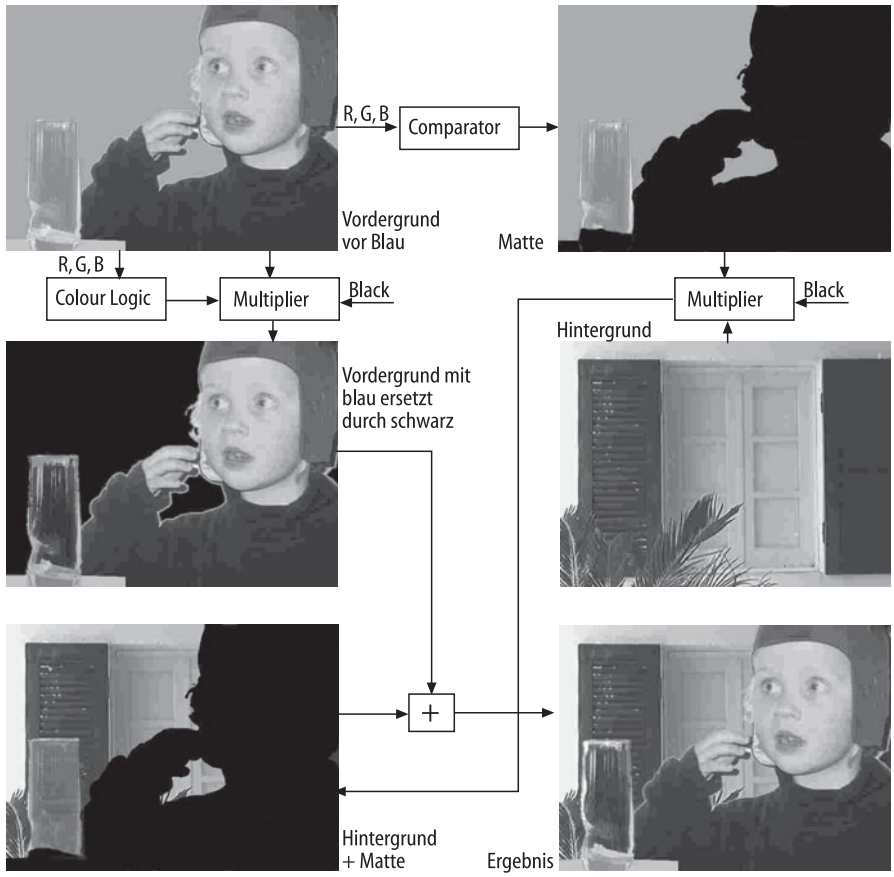


Abb. 10.98. Funktionsprinzip des Ultimatte Chromakey-Verfahrens

Glases mehr oder weniger grau aussehen. In einem zweiten Kanal ist dieses Bild quasi invertiert, wobei Elemente, die die Stanzfarbe nicht enthalten – wie die Person im Vordergrund – vollständig schwarz erscheinen. Je nach Transmissionsgrad des Glases wird der Hintergrund an der Stelle des Glases wieder mehr oder weniger grau erscheinen. In dieses Bild wird nun anhand der Helligkeit (linear Key, s. Kap. 9) mehr oder weniger intensiv das gewünschte Hintergrundbild des Endergebnisses eingesetzt, sodass der Hintergrund durch das Glas hindurch schimmert, die Person aber weiterhin als schwarze Silhouette erscheint. Das Vordergrundbild im zweiten Kanal enthält die Reflexionen. Nachdem schließlich das Hintergrundbild des Kanals 1 mit dem Vordergrundbild aus Kanal 2 verbunden wurde, erscheint die Person an Stelle der Silhouette und das Glas erscheint nicht mehr nur transparent, sondern zeigt auch die Lichtreflexe des Originals [154]. Mit diesem Verfahren können neben Transparenzen und Reflexionen viele weitere kritische Bildvorlagen verarbeitet werden, wie z. B. Rauch, Wasser, feine Haare etc.

Moderne Ultimatte-Systeme arbeiten vollständig digital mit RGBA-Signalen und Abtastratenverhältnissen von 4:4:4:4, um maximale Farbauflösung zu erhalten. Das A steht in diesem Zusammenhang für den Alphakanal, in dem der Grad der Transparenz durchscheinender Objekte codiert wird. Dabei kann das System mit dem Automatic Background Defocusing auch dazu eingesetzt werden, realistischere Bilder durch Unschärfen zu erzeugen, wie sie bei geringer Schärfentiefe z. B. im Hintergrund auftreten. So muss der Grafikrechner mit dieser Funktion nicht belastet werden.

Hochwertige Chromakeyer sind einerseits erforderlich, um die Illusion eines Gesamtbildes möglichst perfekt erscheinen zu lassen, andererseits auch, um Schattenwürfe auf virtuelle Objekte zu realisieren. Schattenwürfe geben dem Bild nicht nur Tiefe, sondern steigern die Echtheit der Illusion und unterstreichen das Gefühl für den Zuschauer, dass es einen festen Bezug zwischen virtuellen Raum und einer agierenden Person gibt und sie nicht im Raum »schwebt«. Dabei muss darauf geachtet werden, dass der Schattenwurf im Vordergrundbild nicht mit dem im virtuellen Set im Widerspruch steht. Hohe Ansprüche stellt auch der Umstand, dass sich die Personen und die Kamera im virtuellen Studio bewegen dürfen und die Relation der Ausleuchtung von Vorder- und Hintergrund nicht so optimal sein kann wie bei der klassischen statischen Nachrichtenproduktion.

10.6.1.3 Extern Key

Gute Chromakeyer wie Ultimatte erlauben die Verarbeitung eines Extern Key-Signals, mit dem sich die Funktionalität des Virtuellen Studios erheblich erweitern lässt und der virtuelle Raum durch Tiefe besonders echt erscheint. Mit dem 3D-Programm des Grafikrechners werden dazu separate Objekte im Raum erzeugt für die ein Extern Key Signal ausgegeben wird. Der Grafikrechner generiert dann nicht nur ein Backgroundsignal (Fill), sondern daneben ein sogenanntes Matte Out-Signal welches dem Matte-In oder Extern Key-Eingang des Chromakeyers zugeführt wird.

Bezogen auf o. g. Beispiel wäre es möglich, dass neben der Person mit dem Glas auf dem Tisch ein virtuelles Spielzeug am Ort des Tisches erzeugt wird.

Abb. 10.99. Anordnung virtueller Objekte mit Extern Key

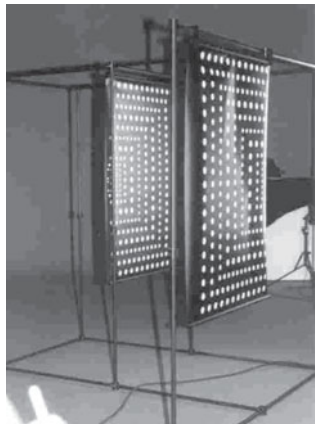
Mittels Ein- und Ausschalten des Extern Keys könnte das Spielzeug vor oder hinter dem Glas erscheinen (Abb. 10.99). Wenn in der Bildregie der Extern Key zum richtigen Zeitpunkt umgeschaltet wird, kann die Szene so gestaltet werden, dass das Spielzeug zunächst vor dem Glas sichtbar wird und nachdem die Person das Glas erhoben und wieder abgesetzt hat, dahinter. Auf diese Weise gelingt es auch, Akteure auf der Bühne scheinbar um virtuelle Objekte herumgehen zu lassen.

10.6.1.4 Ausleuchtung des virtuellen Studios

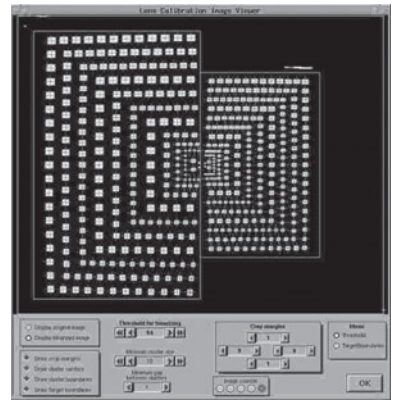
Beim Einsatz des Chroma Key-Verfahrens ist darauf zu achten dass der Hintergrund möglichst gleichmäßig farbig erscheint, da so der Stanzvorgang erleichtert wird. Die Lichtgestaltung für die Ausleuchtung von Personen und Objekten muss dagegen zurücktreten, d. h. es muss ein Kompromiss zwischen technisch notwendigem und künstlerisch anspruchsvollem Licht gefunden werden. Für den Hintergrund werden oft Flächenleuchten und Diffusoren verwendet, die Personen und Objekte werden mit üblichen Stufenlinsen-Scheinwerfern ausgeleuchtet. Um interessante Bilder zu erzielen, sollten sie eine Tiefe durch Schatten enthalten, jedoch sind Schattenwürfe auch bei hochqualitativen Keyern nicht unproblematisch.

10.6.1.5 Kameras und Objektive

Im virtuellen Studio werden die üblichen Studiokameras eingesetzt (s. Kap. 6). Sie können aus konventionellen Studios übernommen werden. Üblicherweise sind sie mit Zoomobjektiven ausgestattet und Zoomfahrten werden auch bei der Produktion eingesetzt. Diesbezüglich ergibt sich das Problem, dass kein linearer Zusammenhang zwischen der mechanischen Einstellung am Objektiv und dem resultierenden Öffnungswinkel besteht. Das Objektiv muss daher einmalig exakt vermessen und der nichtlineare Zusammenhang muss bei der Berechnung des Bildes der virtuellen Kamera berücksichtigt werden [234]. Dazu gibt es besondere Kalibrierungssysteme für das optische System, welche meist auf einer Bilderkennungs-Software beruhen (Abb. 10.100). Dabei werden bekannte Muster in allen Zoom- und Fokusstellungen analysiert und die Kalibrierungsparameter in einer Datenbank abgelegt, wobei auch Objektiv-Verzeichnungen korrigiert werden können.



Musterwände im Studio



Muster in der Kalibrierungssoftware

Abb. 10.100. Einrichtungen zur Kalibrierung der Optik [234]

Ein zweites Problem ist der Knotenpunktversatz, der Nodal Point Shift. Gerade bei Zoomobjektiven kann nicht davon ausgegangen werden, dass der Knotenpunkt für die Dreh- und Neige-Bewegung (Pan und Tilt) mit einem festen Punkt des Objektivs zusammenfällt. Um eine Veränderung der Perspektive bei Drehbewegungen zu vermeiden, muss aber der Knotenpunkt auf der objektseitigen Hauptebene des Objektivs liegen [207]. Die Lage der Hauptebene verändert sich jedoch bei der Zoomfahrt (s. Abschn. 6.2) und dem entsprechend muss der Versatz rechnerisch für die virtuelle Kamera berücksichtigt werden. Falls die Software nicht in der Lage ist, den Nodal Point Shift auszugleichen, kann er auch in die Kalibrierungsdaten eingerechnet werden.

10.6.2 Kameratracking

Der kritische Punkt im virtuellen Studio und damit der Ansatz für die meisten Weiterentwicklungen ist neben der Rechenleistung die Erfassung der veränderlichen Kameraperspektive. Diese Erfassung muss sehr genau sein, damit die in der Software etablierte virtuelle Kamera exakt der realen Kamera folgen kann und das virtuelle Set immer als Einheit erscheint. Andernfalls scheinen Objekte oder Personen bei einer Kamerabewegung im virtuellen Raum zu schweben oder befinden sich nach einer Kamerafahrt und zurück an einem anderen Ort als vor dem Start. Auch in zeitlicher Dimension unterscheiden sich reale und virtuelle Räume, denn es dauert i. d. R. einige Zeit, bis die Trackingdaten bestimmt sind und dann das entsprechende Hintergrundbild berechnet ist. Das Problem wird gelöst, indem durch eine Hardwareeinheit auch das Vordergrundsignal verzögert wird. Dieses Delay darf die Dauer einiger Frames (meist Werte zwischen 2 und 8 Frames) jedoch nicht überschreiten, da sonst die Produktion zu sehr erschwert wird.

Trackingsysteme sind in den meisten Fällen so konstruiert, dass die üblichen Studiokameras verwendet werden. Die Perspektive der realen Kamera

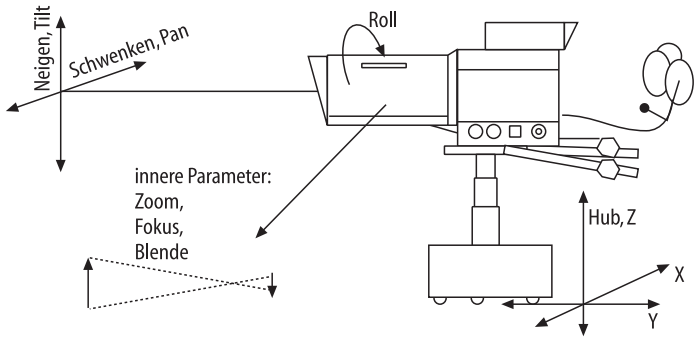


Abb. 10.101. Freiheitsgrade der Kamerabewegung

ist durch verschiedene Freiheitsgrade bestimmt. Sie kann in die inneren Parameter Zoom, Focus und Blende und die äußeren Parameter der ortsfesten Kamera Schwenk-, Neige- und Rollwinkel (Pan, Tilt und Roll) aufgeteilt werden. Hinzu kommen die drei Raumkoordinaten der beweglichen Kamera (Abb. 10.101). Nicht alle Freiheitsgrade werden von den verschiedenen Trackingsystemen erfasst.

Die Positionserkennung und Verfolgung der Kameraparameter ist ein Problem, das in ähnlicher Form in vielen technischen Bereichen auftritt und sich meist auf Winkel- und Längenmessungen zurückführen lässt. Kameraspezifisch sind lediglich die inneren Parameter wie Zoom und Fokus. Es gibt bereits viele Lösungsansätze zur Lage- und Orientierungsbestimmung, denn ähnliche Probleme gibt es auch im Bereich virtueller Realität, bei der Entwicklung mobiler Roboter oder in der Regelungstechnik und bei der Werkkontrolle. Allgemein lassen sich berührungslose Entfernungs- und Richtungsmessung mittels Ultraschall oder elektromagnetischer Wellen von Weg- und Winkelmessung durch mechanische Verbindung oder Trägheitssysteme unterscheiden. Als dritter Bereich kommen bildbasierte, videometrische Verfahren hinzu, die die Lage von Referenzpunkten an Objekten auswerten. Die Ansätze, die auf akustischen, magnetischen oder elektrischen Feldern beruhen, scheiden aus Gründen mangelnder Genauigkeit oder fehlender Praktikabilität aus. Für den Einsatz im virtuellen Studio haben sich sensorgestützte Systeme und optische Systeme etabliert sowie sogenannte Hybridsysteme, bei denen die Vorteile beider Verfahren kombiniert werden.

Die erforderliche Genauigkeit der Positionierung ergibt sich aus dem Bildeindruck. In Abhängigkeit vom Bildinhalt können Abweichungen zwischen realer und virtueller Perspektive mehr oder weniger auffallen. Als strenger Richtwert kann gelten, dass Verschiebungen von maximal 1 Pixel erlaubt sind. Die daraus resultierende Winkelgenauigkeit hängt vom Blickwinkel und damit von der Zoomstellung ab. Für eine Horizontalauflösung von 720 Pixeln heißt das bei einer Abbildung mit 40° Öffnungswinkel, dass eine Genauigkeit von 0,05° erforderlich ist. Wird die Forderung von 1/720 auch für die seitliche lineare Verschiebung erhoben, so ergibt sich bei einer Breite von 3 m eine maximale Abweichung von ca. 4 mm. Grob lässt sich sagen, dass für Standardpro-

duktionen mit Videoauflösung und Personen im Virtuellen Studio eine Ortsgenauigkeit besser als 0,5 mm und eine Winkelgenauigkeit von weniger als 1/10 Grad gefordert werden muss. Generell ist die Winkelgenauigkeit kritischer als die Translationsgenauigkeit. Neben der absoluten Abweichung kann auch das Rauschen zum Problem werden das den Trackingdaten überlagert sein kann. Bei nicht ausreichender Glättung kann dies im Extremfall zu einem instabilen, ständig schwankenden Bildeindruck führen.

10.6.2.1 Sensorgestützte Systeme

Als Sensoren zur Bewegungserfassung kommen verschiedene Varianten in Frage, sie können induktiv, kapazitiv, magnetisch (Hallsensor) oder optisch arbeiten. Die sensorgestützen Systeme können in aktive und passive Typen unterschieden werden. Bei den aktiven Systemen werden die beweglichen Elemente mit Motoren getrieben (Roboter) [61]. Der Antrieb muss sehr exakt sein, die Antriebsimpulse werden zum Motor und auch zum Rechner geschickt und liegen dort vor bevor die Bewegung ausgeführt wird. Damit kann die Berechnung des Hintergrundbildes sehr schnell beginnen, sodass eine Verzögerung des Vordergrundbildes, die bei anderen Verfahren unumgänglich ist, hier ggf. vermieden werden kann. Beispiele für aktive Systeme sind der Ultimatte Memory Head, der für die Reproduktion von Kamerafahrten zur Realisierung von Filmtricks konzipiert wurde und Systeme von Vinten/Radamec.

Bei den passiven mechanischen Sensorsystemen ermittelt ein mechanisches Zahn- oder Reibrad den Wegunterschied zwischen festen und beweglichen Supportelementen sowie die Drehung der Zoom- und Fokusringe am Objektiv. Dabei werden lineare Bewegungen in Drehbewegungen umgesetzt, sodass für alle Parameter Winkelcodierer zum Einsatz kommen können. Bei jeder Bewegung werden an den Drehachsen Impulse erzeugt, die Winkel ergeben sich aus der Impulszählung. Bei ortsfesten Kameras wird meist mit 4-Achsen-Systemen (Zoom, Focus, Pan, Tilt) gearbeitet. Bewegliche Kameras sind an eine Schiene gebunden, hier werden zusätzlich Stativhub und Schienenstrecke erfasst.

Häufig werden optische Impulse gezählt. Optische Winkelcodierer bestehen aus einer lichtdurchlässigen Scheibe mit einem Muster, welches kontinuierlich einfallendes Licht rhythmisch unterbricht. Die addierten oder subtrahierten

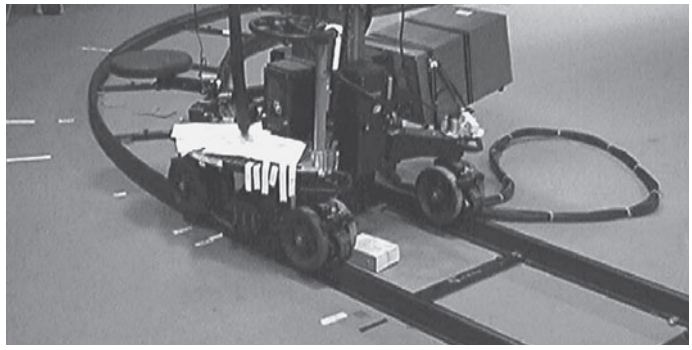


Abb. 10.102. Sensoren am schienengestützten System

Impulse werden einem separaten Tracking-PC zugeführt, wo sie in serielle Datenblöcke umgerechnet und dann über serielle Schnittstellen dem Grafikrechner übergeben werden. Inkrementalgeber liefern dabei relative Werte, was eine aufwendige Kalibrierung des Systems erfordert, bei dem ein frei definierter Bezugspunkt (Home-Point) im realen Raum mit dem Bezugspunkt im virtuellen Raum in festen Bezug gebracht wird. Die Kalibrierung ist entscheidend für die Übereinstimmung von Vorder- und Hintergrund und muss vor und teilweise auch während der Produktion erneut durchgeführt werden. Die ermittelten Verläufe können in lineare und nichtlineare unterschieden werden. Letztere müssen eigens vermessen werden.

Der Nachteil sensorbasierter Systeme ist die nicht frei bewegliche Kamera und die fehlende feste Verkopplung zwischen realem und virtuellem Bild, mit der Gefahr, dass diese nach längerer Betriebszeit auseinanderlaufen. Sensorbasierte Systeme sind alltagstauglich, aber aufgrund der Stativbindung sperrig und unflexibel (Abb. 10.102). Um diesen Nachteil zu vermeiden, wird mit bildbasierten Systemen versucht, die Positionsdaten aus der Bildveränderung aufgrund der Änderung der Perspektive zu ermitteln.

10.6.2.2 Bildbasierte Trackingsysteme

Die Änderung der Kameraperspektive ist mit einer Änderung des Bildinhalts verbunden. So rücken z. B. Linien im Bild näher zusammen, wenn die Kamera zurückfährt, oder sie laufen bei seitlicher Betrachtung auseinander (Abb. 10.103). Die Grundidee bildbasierter Systeme ist, dass bei bekanntem Bildinhalt auf die veränderte Perspektive bzw. die Trackingparameter zurück geschlossen werden kann. Dazu muss ein bekanntes, nicht zu einfach strukturiertes Muster im Bild vorhanden sein. Bei konsequenter Anwendung dieser Idee wird das Muster durch das Studioobjektiv betrachtet, sodass dieses keinen verfälschenden Einfluss mehr haben kann. Allerdings darf das Muster im Vordergrundbild nicht störend in Erscheinung treten.

Als rein bildbasierte Verfahren ist das ORAD Cyberset bekannt geworden. Bei ORAD wird ein großflächiges Streifenmuster verwendet. Dieses ist auf der Blauwand des virtuellen Studios aufgebracht. Die Streifen müssen für das Tracking sichtbar sein und sind daher heller als die Blauwand, andererseits müssen sie aber auch blau sein, damit sie vom Chroma Key-System unterdrückt werden können (Abb. 10.104). Beim Orad-System wird die Positions-

Abb. 10.103. Musterveränderung durch Änderung der Kameraperspektive

erkennung von separaten Digital Video Prozessoren unabhängig vom eigentlichen Grafikrechner durchgeführt, wobei mit Verzögerungszeiten von der Dauer einiger Frames gerechnet werden muss. Ähnlich arbeitet das System RT-Set, bei dem das Muster durch einen Projektor von schräg oben auf die Blauwand projiziert wird. Diese bildbasierenden Trackingsysteme bieten theoretisch den Vorteil einer frei beweglichen Kamera, die Produktionsbedingungen unterliegen jedoch einigen Einschränkungen. Das größte Problem ist, dass immer ein großer Teil des Musters im Bild sichtbar sein muss. Das heißt, dass sich ein Moderator nicht in einem Raumbereich befinden darf, in dem kein Muster ist. Ein Moderator darf auch das Muster nicht verdecken, was allerdings bei Teleeinstellungen des Objektivs leicht geschehen kann. Dieser Umstand ist für die Betriebssicherheit ein großes Problem, denn der Verlust des Musters führt zu Orientierungslosigkeit.

Ein generelles Problem bildbasierter Verfahren ist auch, dass die Positionsdaten aus verrauschten Halbbildern der Kamera gewonnen werden. Die Positionsbestimmung kann zwangsläufig nur so gut sein wie es die Lokalisation des Referenzmusters oder -punktes erlaubt, sodass aufwendige und den Bewegungen angepasste Glättungsalgorithmen zur Stabilisierung des resultierenden Jitter verwendet werden müssen. Weiterhin gilt als Einschränkung, dass das Muster nicht verzerrt werden darf, d. h. es kann nicht mit Hohlkehlen gearbeitet werden. Das Muster muss auch ausreichend scharf abgebildet sein, in der Praxis wird deshalb in einem kleinen Blauraum gearbeitet und die Abbildung auf hohe Schärfentiefe eingestellt. Weiterhin gilt es, Bewegungsunschärfen durch zu schnelle Kameraschwenks zu vermeiden. Für Orad kommt hinzu, dass die Einschränkungen für die Ausleuchtung stärker sind als in gewöhnlichen Blauräumen, da der Chromakeyer so eng eingestellt werden muss, dass er beide Blautöne unterdrückt. Häufig muss dann ein Moderator so hell ausgeleuchtet werden, dass die Ablesung eines Teleprompters erschwert wird.

Schließlich ist noch das Problem zu nennen, dass in vielen Fällen die Musterveränderung aufgrund einer Zoomfahrt nicht von der Veränderung aufgrund einer Kamerafahrt unterschieden werden kann. Daher wird das Orad-System oft so modifiziert, dass LED an der Hauptkamera durch externe Kame-

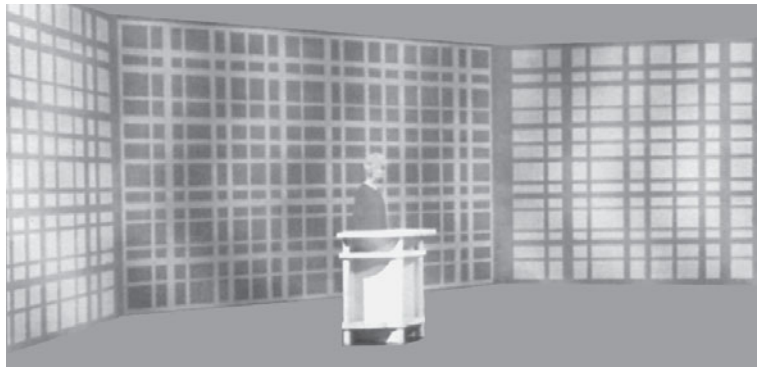


Abb. 10.104. Streifenmuster im Blauraum für das Orad-System

ras (in der Blauwand) registriert werden, sodass eine Zusatzinformation über die Kameraposition zur Verfügung steht. Die erforderliche Modifikation des Orad-Systems zeigt, dass Systeme auf Basis einfacher Bilderkennung sehr viele praktische Probleme mit sich bringen. Aus diesem Grund wurden im Laufe der Zeit vermehrt Hybridsysteme entwickelt.

10.6.2.3 Hybridsysteme

Die meisten Systeme dieser Kategorie nutzen Sensoren für die inneren Parameter (Zoom, Focus) und bildbasierte Verfahren für die Raumorientierung. Diesem Ansatz kommt die Entwicklung von digital gesteuerten Objektiven entgegen, die nicht mehr mechanisch betrieben werden. Bei diesen Typen können die Steuerdaten für Zoom, Focus und Blende direkt für das Tracking-System übernommen werden.

Die Raumorientierung basiert wegen der oben genannten Probleme nicht mehr auf einem Muster, das über die Studiokamera aufgenommen wird. Es liegt außerhalb des Blickfeldes der Hauptkamera und wird durch externe Hilfs-Kameras aufgenommen, die an eigene Bildverarbeitungssysteme angeschlossen werden. Die Gestaltung des Musters unterliegt damit weniger Beschränkungen und kann nicht vom Moderator verdeckt werden.

Bei dem von der BBC entwickelten System Free-d wird zu diesem Zweck an der Studiokamera eine zweite Kamera befestigt, die zur Studiodecke gerichtet ist, wo eine Vielzahl von Markern von ca. 20 cm Durchmesser in bekannter Position in verschiedenen Höhen angebracht wurden. Die Marker verfügen über konzentrische schwarze und weiße Ringe, die eine Art Barcode ergeben und damit die Identifikation jedes einzelnen Markers gestatten. Die Marker sind stark reflektierend und werden durch hell strahlende LEDs an der Hilfskamera beleuchtet, damit sie in der Studiobeleuchtung deutlich identifiziert werden können. Trotzdem bleibt die Studiobeleuchtung ein Problem, denn Scheinwerfer verdecken viele Marker, sodass insgesamt sehr viele zum Einsatz kommen müssen, damit immer mindestens vier, besser jedoch 10-20 Marker im Bild sind. Bei anderen Systemen wird das Verfahren umgekehrt und ein Marker-

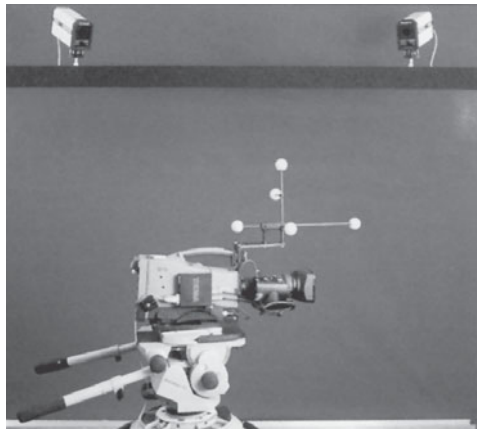


Abb. 10.105. Hybridsystem

system aus Reflektoren bzw. Infrarotsendern an der Studiokamera angebracht, was von externen Kameras an der Studiodecke beobachtet wird (Abb. 10.105). Das Problem der gegenseitigen Verdeckung der Marker wird durch den Einsatz mehrerer Beobachtungskameras gelöst. Durch die Vielzahl der Beobachtungsrichtungen kann das System sehr stabil und betriebssicher werden. Die Bildverarbeitung und Positionsbestimmung wird in separaten Rechnern parallel zur Berechnung des Hintergrundbildes ausgeführt.

Die Kalibrierung der Hybridsysteme erfordert meist erheblich weniger Aufwand als bei sensorgestützten Verfahren. Der Vorgang teilt sich in die studio-orientierte Kalibrierung und die Vermessung des Kameratargets. Bei der Studio-Kalibrierung werden mit Hilfe von Markern, die auf dem Studioboden angebracht sind, die Positionen der Beobachtungskameras und die Studio-Koordinaten festgelegt. Diese Kalibrierung kann software-unterstützt durchgeführt werden, wobei ein IR-Lichtpunkt frei durch den Raum geführt wird, in der Art, dass er mindestens in jeder Überwachungskamera einmal erscheint.

10.6.3 Produktion im virtuellen Studio

Die Produktionsmöglichkeiten sind von der Art des verwendeten Systems bestimmt. Bei der Auswahl sollten folgende Punkte beachtet werden:

- Betriebssicherheit, Redundanz der Trackingdaten,
- Genauigkeit, und Rauschfreiheit der Trackingdaten,
- Erfassungsdauer der Position, d. h. Größe des Delays für das Vordergrundsignal,
- Unabhängigkeit vom Kamerasupport, Verwendbarkeit bestehender Supportsysteme, Einsetzbarkeit als Schulterkamera,
- Unabhängigkeit von spezieller VS-Software,
- Aufwand bei der Kalibrierung,
- Beleuchtungseinschränkungen,
- Erfassung der Objektivparameter, Einschränkung des Schärfebereichs,
- Variationsmöglichkeit für die Schärfentiefe,
- Verwendbarkeit für Außenproduktionen,
- Verwendbarkeit verschiedener Objektive.

Die Produktion im virtuellen Studio beginnt mit dem Entwurf von Szenenbildern anhand eines Storyboards. Die meist zweidimensionalen Szenenbilder werden anschließend dreidimensional umgesetzt, was entweder in separaten CAD-Programmen, meist aber direkt mit den Tools geschieht, die die VS-Software zur Verfügung stellt. Mit den Sets können nun Stellproben und die Ausleuchtung im Studio durchgeführt werden, anschließend folgen die feinere Ausgestaltung und die genaue Kalibrierung.

Virtuelle Sets basieren i. d. R. auf einfachen geometrischen Objekten, den Wireframes. Diese werden skaliert, rotiert und mit Farben, Texturen und Beleuchtungsattributen oder Videosignalen versehen. Die Set-Designer können jedoch nicht nur nach ästhetischen Gesichtspunkten arbeiten, sondern müssen die beschränkte Rechenleistung berücksichtigen. Die Echtzeitanforderung erlaubt eine maximale Rechenzeit von der Dauer eines Halbbildes. Bei der

Abb. 10.106. Moderator vor Grün und mit verschiedenen virtuellen Hintergründen [246]

Optimierung des Sets bezüglich der Rechenzeit kommt es einerseits auf die Komplexität des Sets, vor allem die Anzahl der zu berechnenden Polygone an, andererseits ist aber auch die Art der Umsetzung sehr entscheidend, d. h. es müssen die speziellen Gegebenheiten von Hard- und Software geschickt berücksichtigt werden.

Wichtige Aspekte für die Rechenzeit sind neben der Anzahl der Polygone der Speicherbedarf der Textur, die mehrfache Definition von Texturkoordinaten, die Verwendung gestaffelter Transparenzen und die Wahl des Beleuchtungsmodells. Allein der Einsatz von Live-Videoquellen erfordert drei der 20 ms, die zur Berechnung eines Bildes zur Verfügung stehen [44]. Eine verbreitete Lösung für das virtuelle Studio ist die Software VizRT, die auch bei der Produktion der Heute-Nachrichten des ZDF und des Wetterberichts der ARD verwendet wird (Abb. 10.106). Als Trackingsystem kommt hier, wie bei vielen aktuellen Installationen, ein aktives Robotersystem zu Einsatz [61].

Abb. 10.107. Virtuelles Set mit Schattenwurf

Ein besonders realitätsähnliches Set-Design wird durch Einsatz von Schatten und Reflexionen erreicht. Schattenwürfe können mit guten Chromakeyern, wie Ultimatte, verarbeitet werden, es muss aber darauf geachtet werden, dass sich die Richtungen der Schatten in realen und virtuellen Räumen gleichen (Abb. 10.107). Soll sich ein Schattenwurf über ein virtuelles Objekt bewegen, so muss im Studio eine blaue Stellfläche an dem Ort aufgestellt werden, dort wo sich das beschattete Objekt im Gesamtbild befindet. Die Form der Stellfläche muss dabei der Form des virtuellen Objekts angepasst sein. Falls Reflexionen, z. B. am Boden des virtuellen Sets auch im realen Bild auftauchen sollen, kann dies durch Verwendung von Plexiglasscheiben auf dem Studioboden erreicht werden. Eventuelle Spiegelungen der Objekte im Studio müssen dann im virtuellen Raum eingerechnet werden.

Die Verwendung mehrerer Kameras im virtuellen Studio stellt einen erheblichen Aufwand dar, denn für jede Kamera müssen synchronisierte virtuelle Sets aus einer anderen Perspektive in Echtzeit verändert werden können, was im Prinzip für jede Kamera einen teuren Grafikrechner erfordert. Der Aufwand kann aber mit Hilfe des Virtual Camera Switch VCS verringert werden (Abb. 10.108). Hier wird die Tatsache genutzt, dass sich i. d. R. nur eine Kamera tatsächlich im Program-Mode befindet während die andere ein Preview-Bild liefert. Dieses Bild kann aber mit verminderter Qualität und verringerter zeitlicher Auflösung errechnet werden, so weit, dass die Orientierung des Kamerapersonals nicht eingeschränkt wird. Auf diese Weise entsteht der Vorteil, dass für das Preview-Bild Low Cost-Computer zum Einsatz kommen können. Ein VCS hat nun die Aufgabe, die Signalumschaltung zwischen Program und Preview für die Vorder- und Hintergrundsignale zu steuern und die richtigen Computer anzusprechen. Dabei werden direkt die Steuersignale der im Studio befindlichen Videomischer oder Kreuzschienen manipuliert, indem das Gerät zwischen Hauptelektronik und Bedienpanel des Videomischers geschaltet wird, sodass die gewöhnliche Bedienung des Bildmischers beibehalten werden kann [44]. Das VCS-System erlaubt die Verwaltung von bis zu 16 Kameras bei gleichzeitiger Darstellung je eines Program- und Preview-Bildes.

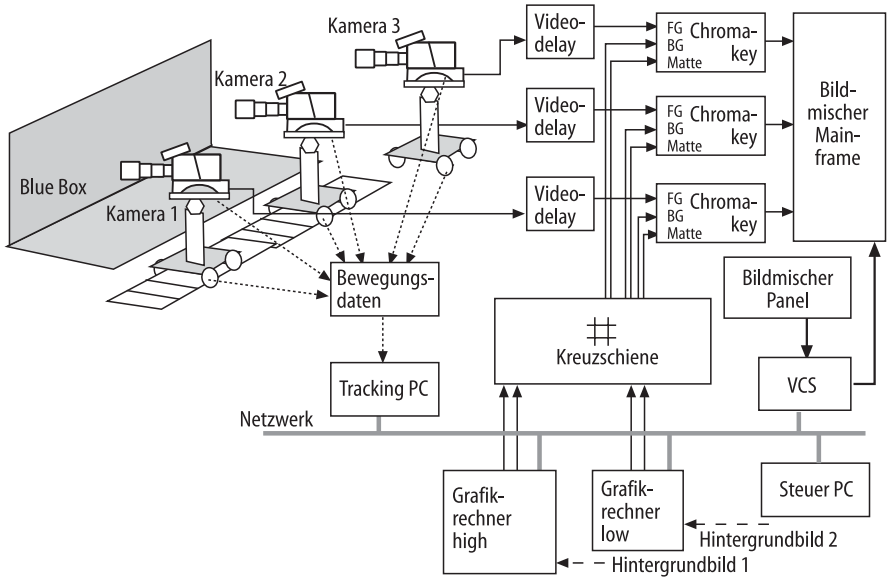


Abb. 10.108. Produktion mit mehreren Kameras im virtuellen Studio

Ein großes Problem für die realitätsnahe Wirkung des virtuellen Sets ist die Schärfentiefe, d. h. dass Objekte, die sich weit entfernt von der gewünschten Schärfeebene befinden, wie bei einer echten Kamera unscharf erscheinen. Die naheliegende Lösung der Verwendung mehrerer Bilder in verschiedener Auflösung ist zu aufwendig und übliche Softwareprogramme unterstützen noch keine Defokussierung. Aber es gibt ein Rechensystem das als Input das gewöhnliche, scharfe Bild nimmt und in Abhängigkeit von den Objektiv-Parametern und einer Information über die Tiefenstaffelung jedes Bildpunktes in Relation zur gewünschten Schärfeebene ein Bild mit unschärferem Hintergrund simuliert. Auch die neueste Version des Ultimatte-Systems bietet Ansätze zur Simulation geringer Schärfentiefe.

Das virtuelle Studio bietet noch viele Ansätze zur Weiterentwicklung dies gilt hinsichtlich der Trackingsysteme, ein weiterer Aspekt ist aber auch die Erkennung der Moderatorposition, damit auch die Umschaltung des Extern-Key automatisch erfolgen kann, falls es so aussehen soll, als ginge eine Person im Studio um eine virtuelles Objekt herum (s. Extern Key). Auch die realitätsnahe Einbindung von Schatten muss noch optimiert werden.

An den in diesem letzten Kapitel dargestellten Beispielen aus verschiedenen Produktionsbereichen wird deutlich, dass der Fortschritt, vor allem durch den Einsatz digitaler Verfahren, nicht nur auf technischem Sektor von Bedeutung ist, sondern sich vermehrt auf Produktionsmethoden und Inhalte auswirkt. Zum Abschluss lässt sich feststellen dass der gesamte Produktions- und Postproduktionsbereich von einem derart schnellen Wandel betroffen ist, dass insbesondere die letzten Kapitel dieses Buches vermutlich schon bald umgeschrieben und erweitert werden müssen.

Literaturverzeichnis

[1] Abekas: Produktinformation
[2] Adams, G.: Erweiterte Werkzeuge und Möglichkeiten im DI-Prozess. Fernseh- und Kinotechnik 60, Nr. 10, 2006
[3] Ahrens, N.: Bewertung von Farben. Medien Bulletin 10/97
[4] Allendorf, M., Schäfer, A.: Der Einsatz des nonlinearen Editing-Systems Avid im Kine-Film und Videobereich. Diplomarbeit, FH Hamburg 1995
[5] Allenspach, R.: Magnetismus in ultradünnen Schichten: Auf dem Weg zum Datenspeicher der Zukunft. Physik in unserer Zeit 27, Nr. 3, 1996
[6] Ampex: Produktinformation
[7] Andresen, T.: Varianten einer Senderegie/Sendeabwicklung für den Bereich Ballungsraumfernsehen. Diplomarbeit, FH Hamburg 1998
[8] ARD: Chronik 50 Jahre ARD, website
[9] Arri: Produktinformation
[10] Avid: Produktinformation
[11] Autodesk: Produktinformation
[12] Barclay, S.: The Motion Picture Image. Focal Press, Boston 2000
[13] Below, M.: Innerstädtische zellulare Netzwerke für digitale Drathlos-Kamerasysteme. Fernseh- und Kinotechnik 60, Nr. 8–9, 2006
[14] Bentrup, M.: Messtechnische Untersuchung digitaler Projektoren in der Postproduktion im Bereich Digital Cinema. Bachelorarbeit, HAW Hamburg 2011
[15] Bernd, M., Fettig, U., Glänzel, P., Montagne, P., Tettenborn, N., Weingarten, J., Wieland, H.: Funkhausneubau SWR Stuttgart. Fernseh- und Kinotechnik 67, Nr. 3 und 4, 2013
[16] Bertsch, D.: Bilddatenreduktion mit JPEG2000. Diplomarbeit, HAW Hamburg 2007
[17] Biaesch-Wiebke, C.: Videosysteme. Vogel Verlag, Würzburg 1991
[18] Blu-ray Disc Association: Technical Whitepapers
[19] Bock, G.: Übergang vom analogen zum digitalen terrestrischen Fernsehen. Fernseh- und Kinotechnik 53, Nr. 8–9, 1999
[20] Bohlmann, J.: Nachrichtentechnik. Vorlesungsskripte, FH Hamburg 1996
[21] Bolewski, N.: BTS: Neue Kameras mit Dynamic Pixels Management. Fernseh- und Kinotechnik 48, Nr. 3, 1994
[22] Bolewski, N.: Die DVD im Broadcastbereich. Fernseh- und Kinotechnik 51, Nr. 11, 1997
[23] Bolewski, N.: Langzeitarchivierung über „“Bits on Film““ Fernseh- und Kinotechnik 62, Nr. 3, 2008
[24] Bolewski, N.: Neues Lichtventil-Projektionssystem. Fernseh- und Kinotechnik 47, Nr. 5, 1993
[25] Bolewski, N.: Schaulust: Technik- und Wahrnehmungsgeschichte von der Renaissance bis zur Neuzeit. Fernseh- und Kinotechnik 60, Nr. 4, 2006
[26] Brendel, H.: The ARRI DI Companion, Arri, Müchen 2007, website
[27] Briede, M.: Analyse aktueller stereoskopischer Bewegtbildsysteme und Entwicklung eines entsprechenden Experimentier und Postproduktionssystems. Diplomarbeit, HAW Hamburg 2006

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4, © Springer-Verlag Berlin Heidelberg 2013

[28] Bross, B. Marpe, T., Schwarz, H.: „High-Efficiency Video Coding (HEVC-)-Standard. Fernseh- und Kinotechnik, Nr. 1-2, 2013
[29] Brosthaus, M.: Einsatzmöglichkeiten drahtloser Videoübertragungssysteme unter besonderer Berücksichtigung zellularer Empfangsnetze. Diplomarbeit, HAW Hamburg 2007
[30] Brosthaus, M.: Zelluläre Empfangsnetze - Erweiterte Einsatzmöglichkeiten digitaler Drahtloskamerasysteme. Fernseh- und Kinotechnik 62, Nr. 4, 2008
[31] Brugger, R.: 3D-Computergrafik und -animation. Addison Wesley, 1993
[32] BTS: Produktinformation
[33] Burghardt, J.: Handbuch der professionellen Videorecorder. Edition Filmwerkstatt, Essen 1994
[34] Burghardt, J.: MPEG-2 für Fernseh-Anwendungen. Fernseh- und Kinot. 53, Nr. 3, 1999
[35] Burk, H., Heimann, R.. Janßen, S., Thomas, S.: ARD HYBNET-Neu. Fernseh- und Kinotechnik, Nr. 3 und 4, 2012
[36] Busjäger, A., Nickelsen, B.: Radio Bremen neu. Fernseh- und Kinotechnik Nr. 6, 2012
[37] Canon: Produktinformation
[38] Conrod, R; Briscoe, P.: Datenorganisation für eine 3-Gbit/s-Infrastruktur. Fernseh- und Kinotechnik 63, Nr. 1-2, 2009
[39] Costard, H.: Film: Der kleine Godard
[40] Dambacher, P.: Digitale Technik für Hörfunk und Fernsehen. R. v. Deckers`s Verlag, Heidelberg 1995
[41] Daniels, K.: Analyse von Digital-Cinematography-Produktionen mit HD-Kameras aus drei Preissegmenten. Diplomarbeit, HAW Hamburg 2009
[42] De Lameillieure, Schäfer, R.: MPEG-2-Bildcodierung für das digitale Fernsehen. Fernseh- und Kinotechnik 48, Nr. 3, 1994
[43] Deutsches Fernsehmuseum Wiesbaden, website
[44] Dierßen, H.: Über die Grundlagen des virtuellen Studios zur Entwicklung einer Multikameralösung. Diplomarbeit, FH Hamburg 1999
[45] Digital Cinema Initiatives: DCI-Spezifikationen 1.2
[46] Discreet, Produktinformation und Tutorials
[47] Dorn, E.: Fachsprache Regie, Kamera, Bildschnitt. NDR 1995
[48] Douven, P.: HD-Kameraqualität - Multinormkameras und ihre Auflösungseigenschaften. Fernseh- und Kinotechnik 62, Nr. 10, 2008
[49] DVB-Project: Fact Sheets, website
[50] Ebner, A. et al.: PALplus: Übertragung von 16:9-Bildern im terrestrischen PAL-Kanal. Fernseh- und Kinotechnik 46, Nr. 11, 1992
[51] Ebner, A.: PALplus: Ein neuer Sendestandard für das Breitbildformat. Fernseh- und Kinotechnik 49, Nr. 7-8, 1995, S. 401 ff.
[52] EBU: EBU-Empfehlungen
[53] Eich, H.-W.: Argumente pro Kupferkabel: 10GBase-TX und höher. Fernseh- und Kinotechnik Nr. 4, 2012
[54] Faulstich, W.: Grundwissen Medien. UTB, Stuttgart 2004
[55] Fell-Bosenbeck, F.: DV-Kompression: Grundlagen und Anwendungen. Fernseh- und Kinotechnik 53, Nr. 6, 1999
[56] Fiedler, M., Scheller, G., Wessely, U.: Fraktale Bild- und Videokompression. Fernseh- und Kinotechnik 52, Nr. 1–2, 1998
[57] Fischer, W.: Digitale Fernseh- und Rundfunktechnik. Springer, Berlin 2009
[58] Fößel, S.; Miller, Th.; Widera, A.: Diskbasierte Speichersysteme für HDTV und Digital-Cinema-Produktionen. Fernseh- und Kinotechnik 60, Nr. 10, 2006
[59] Freyer, U.: DVB – Digitales Fernsehen. Verlag Technik, Berlin 1997
[60] Fröhlich, J.; Bäuerle, M.: Digitales Kopierwerk - heute und morgen. Fernseh- und Kinotechnik 62, Nr. 5, 2008
[61] Frühling, S., Yuan, J.: Roboter-Technik in VR-Studios. Fernseh- und Kinot. Nr. 3, 2012

[62] Fuji Produktinformation
[63] Gebhard, C., Voigt-Müller, G.: Marktübersicht Nonlineare Systeme. Film- und TV-Kameramann 1998
[64] Geis, A.: Mobilempfang von DVB-T – aktuelle Analyse und Optimierungspotenziale. Diplomarbeit, HAW Hamburg 2008
[65] Gerhard-Multhaupt, R., Röder, H.: Lichtventil-Großbildprojektion: Eine Übersicht. Fernseh- und Kinotechnik 45, Nr. 9, 1991
[66] Gerhard-Multhaupt, R.: Amplitudengesteuerte Beugungsgitter für die Großbildprojektion. Physikalische Blätter 51, Nr. 2, 1995
[67] Gersdorf, C.: Diorama Illusion – Möglichkeiten und Anwendungen des Miniaturisierungseffekts im Film. Bachelorarbeit, HAW Hamburg 2012
[68] Gierlinger, F.: Display-Technik für HD. HDTV-Seminar, RBB November 2007
[69] Gierlinger, F.: Farbwiedergabe in HD. Fernseh- und Kinotechnik 60, Nr. 7, 2006
[70] Glabbart, H.: Digital Intermdiate – Technik und Workflow im digitalen Filmlabor. Diplomarbeit, HAW Hamburg 2006
[71] Görne, T.: Tontechnik. Hanser, München 2011
[72] Göttmann, K.: Fibre Channel, Technologie und Netzwerke. Fernseh- und Kinotechnik 52, Nr. 8–9, 1998
[73] Götz-Meyn, E., Neumann, W.: Grundlagen der Video- und Videoaufzeichnungstechnik. Hüthig Verlag, Heidelberg 1998
[74] Grambow. L.: Farbwiedergabeindex für Displays. Fernseh- und Kinotechnik 52, Nr. 3, 1998
[75] Gunkel, G.: DVB-T2 - Der neue Standard für den terrestrischen Rundfunk. Fernseh- und Kinotechnik 62, Nr. 8–9, 2008
[76] Haarstark, G.R. et al.: Das neue Produktions- und Sendezentrum des NDR. Fernseh- und Kinotechnik 49, Nr. 10, 1995
[77] Hahne, C.: Konfiguration einer digitalen Lichtfeldkamera. Bachelorarbeit, HAW Hamburg 2012
[78] Hartwig, S., Endemann, W.: Tutorial Digitale Bildcodierung (Folge 1–12). Fernseh- und Kinotechnik 46, Nr. 1, 1992 bis FKT 47, Nr. 1, 1993
[79] Hastedt, A.: Digitale Filmmuster – Eine Analyse von Konzepten daten- und bandbasierter Workflows. Diplomarbeit, HAW Hamburg 2008
[80] Heber, H., Teichner, D., Hedtke, R.: Special Tapeless Production & Broadcasting. Fernseh- und Kinotechnik 49, Nr. 6, 1995
[81] Heber, H.K.: `Random Access`-Speicher Eine Einführung. Fernseh- und Kinotechnik 49, Nr. 6, 1995
[82] Heber, H.K.: TFHS Beschreibung und Kommentierung des Final Reports Teil 1. Systembetrachtungen. Fernseh- und Kinotechnik 53, Nr. 4, 1999
[83] Hedtke, R., Schnöll, M.: Schnittbearbeitung von MPEG-2-codierten Videosequenzen. Fernseh- und Kinotechnik 50, Nr. 7, 1996
[84] Hedtke, R.: Netzwerktechnik im Videostudio. Fernseh- und Kinotechnik 51, Nr. 6, 1997
[85] Hedtke, R.: TFHS Beschreibung und Kommentierung des Final Reports Teil 2. Datenkompression. Fernseh- und Kinotechnik 53, Nr. 4, 1999
[86] Hedtke, R.: HDTV ... und was kommt danach? Fernseh- und Kinotechnik 67, Nr. 4, 2013
[87] Hedtke, R.: Zukünftige Anwendungen der MPEG-2-Codierung. Fernseh- und Kinotechnik 52, Nr. 10, 1998
[88] Hedtke, R.:Videocodierung für die HD-Produktion Fernseh- und Kinotechnik 62, Nr. 6–8, 2008
[89] Heitmann, J.: MPEG-2 im Fernsehstudio. Fernseh- und Kinotechnik 52, Nr. 11, 1998
[90] Henle, H.: Dolby-Mehrkanalton. Fernseh- und Kinotechnik 55, Nr. 1–2, 2001

[91] Henle, H.: Filmsound der nächsten Generation. Fernseh- und Kinotechnik Nr. 7, 2012
[92] Herberholz, G.: Optimierung der Erstellung eines DCI-konformen digitalen Films. Diplomarbeit, HAW Hamburg 2008
[93] Herbig, G. P.: Stereofotografie, website
[94] Herpel, C.: Der MPEG-2-Standard (Teil 3). Fernseh- und Kinotechnik 48, Nr. 6, 1994
[95] Herter, E., Lörcher. W.: Nachrichtentechnik. C. Hanser Verlag, München Wien 1992
[96] Heyna, A., Briede, M., Schmidt, U.: Datenformate im Medienbereich. Fachbuchverlag Leipzig 2003
[97] Heywang et al.: Physik für Fachhochschulen und technische Berufe. Verlag Handwerk und Technik, Hamburg
[98] Hinte, B.: Planungsschritte und technische Entwicklung eines Multiformat-Übertragungswagens für HD- und SD-Produktionen, Diplomarbeit, HAW Hamburg 2006
[99] Hochgürtel, G.: HDC 2500 – eine neue HD-Systemkameraserie. Fernseh- und Kinotechnik Nr. 8-9, 2012
[100] Hochmeister, G. v.: Handbuch für den Filmvorführer. Wirtschaftsverband der Filmtheater, München 1991
[101] Hoehler, D.: HDTV in Europa. HDTV-Seminar, RBB November 2007
[102] Hoffmann, H.: Der Weg zum SDTI – Serial Data Transport Interface. Fernseh- und Kinotechnik 53, Nr. 1–2, 1999
[103] Hoffmann, H.: TFHS Beschreibung und Kommentierung des Final Reports Teil 3. Netzwerke und Transferprotokolle. Fernseh- und Kinotechnik 53, Nr. 7, 1999
[104] Hofmann, H.: Neue Anforderungen an die Netztechnik im Umfeld Rundfunk. Fernseh- und Kinotechnik 51, Nr. 6, 1997
[105] Hofmann, S.: AVC-I - Konzept und Applikationen. Fernseh- und Kinotechnik 61, Nr. 8–9, 2007
[106] Hornbeck, L.J.: DLP (Digital Light Processing) für ein Display mit Mikrospiegel-Ablenkung. Fernseh- und Kinotechnik 50, Nr. 10, 1996
[107] Hubbes, H., Tauchnitz, H., Wagner, R.E.: Steuerung und Überwchung von Kontributions- und Verteilnetzen. Fernseh- und Kinotechnik, Nr. 8-9, 2012
[108] Hübscher, H. et al.: Elektrotechnik Fachstufe 2 Nachrichtentechnik. Westermann, Braunschweig 1986
[109] Hummel, R. C.: Vergleiche der Filmformate 1.85; anamorphotisch und Super 35. Fernseh- und Kinotechnik632, Nr. 1-2, 2009
[110] Huusmann, H.: Entwicklung eines stereoskopiswchen Videosystems für Unterwasserfahrzeuge. Dipomarbeit, HAW Hamburg 2008
[111] Institut für Microelectronics, IMS-Chips, Produktinformation, 2002
[112] Institut für Rundfunktechnik (Hrsg): Technische Richtlinien - HDTV zur Herstellung von Fernsehproduktionen für ARD, ZDF und ORF. IRT 2011
[113] Ishikawa, K. et al.: FIT-CCD-Kamera für HDTV mit 2 Millionen Pixeln. Fernseh- und Kinotechnik 46, Nr. 9, 1992
[114] ITU- Empfehlungen
[115] Jähne, B.: Objektive Charakterisierung von Bildsensoren und digitalen Kameras. Fernseh- und Kinotechnik, Nr. 5, 2012
[116] Jütte, F.: Umsetzung journalistischer Auslandsproduktionen am Beispiel des NDR. Diplomarbeit, HAW Hamburg 2011
[117] Jung, M., Rosendahl, U., Brormann, M.: Das neue Fernsehproduktionsstudio des SWR in Baden-Baden. Fernseh- und Kinotechnik 53, Nr. 3, 1999
[118] JVC: Produktinformation
[119] Kalb, H. W.: Erste digitale Produktionsinsel im SDR in Stuttgart. Fernseh- und Kinotechnik 48, Nr. 10, 1994
[120] Kalb, H.W.: Erfahrungen auf dem Weg von analogen zum digitalen Fernsehzentrum. Jahrestagung der FKTG, Tagungsband 1998

[121] Kammermayer, S.: OLED -Technologie und messtechnischer Vergleich von OLED-, LCD- und CRT- Monitoren. Bachelorarbeit, HAW Hamburg 2011
[122] Kauff, P., Müller, M., Schäfer, R., Tanger, R.: 3D-Konvertierung für autostereoskopische Displays, Fernseh- und Kinotechnik, Nr. 1-2, 2011
[123] Kaufmann, A.: Entwicklungstendenzen bei Farbfernsehkameras für Standard- und HDTV. Fernseh- und Kinotechnik 47, Nr. 3, 1993
[124] Kays, R.: Kanalcodierung und Modulation für die digitale Fernsehübertragung. Fernseh- und Kinotechnik 48, Nr. 3, 1994
[125] Kiening, H.: 2K oder 4K - Wie groß sind die Unterschiede. Fernseh- und Kinotechnik 61, Nr. 11, 2007
[126] Kinoton, Produktinformation, 2001
[127] Knör, R.: Kaskadierung und Formatwahl. HDTV-Seminar, RBB November 2007
[128] Kodak: Produktinformation und Kodak Motion Picture Film. Publ. H1, Rochester, 2000
[129] Koops, M.: Intercom-Migration beim NDR Fernseh- und Kinotechnik 62, Nr. 1-2, 2008
[130] Köpping, A.: Konzeption und Entwicklung einer SAN Management Software. Diplomarbeit, HAW Hamburg 2007
[131] Kraus, F.: Arri, Alexa und 4k. Fernseh- und Kinotechnik Nr. 4, 2013
[132] Kraus, H.: Plasma-Displays in Fernseh- und Multimedia-Endgeräten. Fernseh- und Kinotechnik 52, Nr. 11, 1998
[133] Kurz, A.: Konzeption und Einsatz eines Fahrzeugs für die satellitengestützte Nachrichtenübertragung. Diplomarbeit, FH Hamburg 1998
[134] Ladebusch, U.: Einführung in den DVB-Datenrundfunk (Teil 1). Fernseh- und Kinotechnik 52, Nr. 6, 1998
[135] Lanciloti, M.: Holographische Datenspeicher. Fernseh- und Kinotechnik 60, Nr. 8-9, 2006
[136] Lang, H.: Farbwiedergabe in der Medien: Fernsehen, Film, Druck. Muster-Schmidt Verlag, Göttingen Zürich 1995
[137] Larbier, P., Reuter, S.: AVC-I: Broadcast-Format für die Videoverarbeitung. Fernseh- und Kinotechnik, Nr. 4 und 5, 2012
[138] Lebowsky, F.; Vonneilich, K.: Multiprimary-Bildschirmarchitektur mit Präemphase. Fernseh- und Kinotechnik 62, Nr. 10, 2008
[139] Lehmann, H., Gierling, F.: „Audio/Video-Delay“ – ein unlösbares Problem? Fernseh- und Kinotechnik 60, Nr. 10, 2006
[140] Leske, S.: LCD-Mastermonitoren im professionellen Umfeld. Fernseh- und Kinotechnik 62, Nr. 11, 2008
[141] Leske, S.: XAVC - Advanced-Video- Coding. Fernseh- und Kinotechnik Nr. 3, 2013
[142] Mahler, G.: Die Grundlagen der Fernsehtechnik: Systemtheorie und Technik der Bildübertragung. Springer, Berlin 2006
[143] Marey, M.: Messen im digital-seriellen Studio. Fernseh- und Kinotechnik 51, Nr. 3, 1997
[144] Marey, M.: Vorteile und Nutzen einer digitalen Fernsehkamera mit Studioqualität. Fernseh- und Kinotechnik 51, Nr. 5, 1997
[145] Maßmann, V.: Der Weg zur digitalen Filmkopie am Beispiel des Spirit DataCine Filmabtasters. Fernseh- und Kinotechnik 52, Nr. 4, 1998
[146] MAT Produktinformation, website
[147] Mauch, R. H.: Stand der Displaytechnik. Fernseh- und Kinotechnik 49, Nr. 4, 1995
[148] Mäusl, R.: Digitale Modulationsverfahren. Hüthig Buch Verlag, Heidelberg 1991
[149] Mäusl, R.: Fernsehtechnik. Hüthig Verlag, Heidelberg 1990
[150] Mayr, S.: Aufbau eines modernen Kontributionsnetzes für Hörfunk und Fernsehen. Diplomarbeit, HAW Hamburg 2003

[151] Meier, W.: Film in der digitalen Ebene. Diplomarbeit an der FH Hamburg, 1998
[152] Messerschmid, U.: Abschied vom Zeilensprung. Fernseh- und Kinotechnik 60, Nr. 11, 2006
[153] Messerschmid, U.: Höhere Bildfrequenzen. Fernseh- und Kinotechnik, Nr. 5, 2012
[154] Meyer-Schwarzenberger, G.: Noch analog – schon digital. Fernseh- und Kinotechnische Gesellschaft und Schule für Rundfunktechnik (Hrsg), 1995
[155] Möllering, D., Slansky, P.: Handbuch der professionellen Videoaufnahme. Edition Filmwerkstatt, Essen 1993
[156] Morgenstern, B.: Farbfernsehtechnik. B. G. Teubner, Stuttgart 1989
[157] Mücher, M.: Broadcast Fachwörterbuch. BET, Hamburg 2008
[158] Müller, A.: Der elektronische Schnitt. HV&F-Verlag, Hamburg 1992
[159] Müller, G. O.: Flachdisplays für das hochauflösende Fernsehen. Fernseh- und Kinotechnik 45, Nr. 9, 1991
[160] Naurosi, P.: Analyse von Methoden zur Speicherplatzoptimierung bei der Archivierung von Sendemitschnitten hinsichtlich technischer Machbarkeit und wirtschaftlicher Effizienz. Diplomarbeit, HAW Hamburg 2007
[161] Nekes, W.: Media Magica, DVD-Sammlung 1995-1997
[162] Nowak, A.; Schmitt, A.; Fößel, S.: Digitale Kinoproduktion: Neue Arbeitsabläufe und Werkzeuge für den praktischen Einsatz. Fernseh- und Kinotechnik 62, Nr. 5, 2008
[163] P+S-Technik: Produktinformation
[164] Panasonic: Broadcast und Professional Media, Produktinformation
[165] Panasonic: Das Videokompressionsbuch 1999
[166] Panther: Produktinformation
[167] Petrasch, T., Zinke, J: Videofilm, Konzeption und Produktion. Fachbuchverlag Leipzig 2012
[168] Popp, H.: CMOS-Bildsensoren. Fernseh- und Kinotechnik 61, Nr.7, 2007
[169] Porchert, M.: Videokamera mit digitaler Signalverarbeitung. Fernseh- und Kinotechnik 45, Nr. 7, 1991
[170] Przybyla, H.: HDTV-Studiotechnik in den neuesten Generationen. Jahrestagung der FKTG, Tagungsband 1998
[171] Quantel: Produktinformation
[172] Quantel: An Introduction to Aces on Quantel, 2012
[173] Red Digital: Produktinformation
[174] Reimers, U.: DVB-Digitale Fernsehtechnik. Springer, Berlin 2007
[175] Reimers, U.: Systemkonzept für das Digitale Fernsehen in Europa. Fernseh- und Kinotechnik 47, Nr. 7-8, 1993
[176] Reimers, U.: Zugangsnetze zum Internet: ISDN, xDSL, Kabelmodems. Fernseh- und Kinotechnik 53, Nr. 6, 1999
[177] Reuber, C.: TV-Displays 100 Jahre nach Braun. Fernseh- und Kinotechnik 52, Nr. 3, 1998
[178] Richter, U.: Empfangsvoraussetzungen für digitales Fernsehen unter besonderer Berücksichtigung von Pay-TV. Diplomarbeit, FH Hamburg 1999
[179] Ricken, C.: Das 4:2:2 Profile von MPEG-2 – Ein ISO/IEC-Standard zur Bilddatenreduktion im Produktionsbereich. Fernseh- und Kinotechnik 50, Nr. 6, 1996
[180] Riesener, N.: Modulare Video- und Audiokreuzschienen. Fernseh- und Kinotechnik 52, Nr. 10, 1998
[181] Riesner, S.: Robotertechnik im Studio - Stand der Technik. Fernseh- und Kinotechnik 62, Nr. 7, 2008
[182] Rindtorff, H.: Bildstabilisation in Consumer-Camcordern. Fernseh- und Kinotechnik 49, Nr. 1-2, 1995
[183] Rohde & Schwarz: Videomesssystem VSA, technische Datenblätter
[184] Ruchti, S.: Monitorwände der Zukunft - was können Multi-Monitoring-Systeme leisten? Fernseh- und Kinotechnik 60, Nr. 10, 2006

[185] Ruelberg, K. D.: Adreßgenerierung in digitalen Video-Effektgeräten. Fernseh- und Kinotechnik 46, Nr. 7-8, 1992
[186] Ruhberg, F.: Möglichkeiten und Bedingungen zur Einführung hochauflösender Fernsehsysteme in Rundfunkanstalten. Diplomarbeit, HAW Hamburg 2005
[187] Sachtler, Produktinformationen
[188] Schäfer, J.: C-Reality – Ein neuer Filmabtaster von Cintel. Fernseh- und Kinotechnik 52, Nr. 4, 1998
[189] Schäfer, R.: HD-Ausstrahlungsformat 720p/50. Fernseh- und Kinotechnik 62, Nr. 10, 2008
[190] Schlienz, O.: H.265/HEVC – Eine messtechnische Untersuchung des Videokompressionsverfahrens im Verglech zu H.264/AVC. Bachelorarbeit, HAW Hamburg 2013
[191] Schmidt, U.: Digitale Film- und Videotechnik, Carl Hanser Verlag, München 2008
[192] Schmude, R.: Formatvergleich Film – HD – SD für Akquisition und Postproduktion. Diplomarbeit, HAW Hamburg 2005
[193] Schnepf, U.: IPTV: Fernsehen über das Internet-Protokoll. Fernseh- und Kinotechnik 61, Nr. 1-2, 2007
[194] Schönfelder, H.: Bildkommunikation. Springer Verlag, Berlin 1983
[195] Schönfelder, H.: Fernsehtechnik im Wandel. Springer-Verlag, Berlin Heidelberg 1996
[196] Schrey, O: Methoden zur Dynamikerweiterung in der zweidimensionalen CMOS-Bildsensorik. Dissertation, Gesamthochschule Duisburg
[197] Schröder, K., Gebhard, H.: Audio/Video-Streaming über IP (Internet Protokoll). Fernseh- und Kinotechnik 54, Nr. 1–2, 2000
[198] Schule für Rundfunktechnik (Hrsg.): Fernsehstudiotechnik. Nürnberg 1989
[199] Schwerdtner, A.: Holographische Displays. Fernseh- und Kinotechnik 62, Nr. 5, 2008
[200] Semar, W.: 4GOOD - Optischer Datenspeicher der 4. Generation nach CD; DVD und Blu-ray Disc. Fernseh- und Kinotechnik 61, Nr. 11, 2007
[201] SES-Astra Produktinformationen, website
[202] Siem, O..: Anregungen für die Auswahl passender HD-Objektive. Fernseh- und Kinotechnik, Nr. 5, 2011
[203] Sikora, Th.: Entwicklung eines MPEG-4 Video-Standards: Das MPEG-4 Video-Verifikationsmodell. Fernseh- und Kinotechnik 50, Nr. 8–9, 1996
[204] Slansky, P. C.: Die DCI-Specification und ihre Auswirkung auf die Filmproduktion. Fernseh- und Kinotechnik 60, Nr. 8-9, 2006
[205] Smith, D. A.: Video über IP. Fernseh- und Kinotechnik 62, Nr. 4, 2008
[206] SMPTE-Standards
[207] Sommerhäuser, F.: Das virtuelle Studio. Fernseh- und Kinotechnik 50, Nr. 1–2, 1996
[208] Sony Produktinformation
[209] Stanko, G., Correns, H.-J.: Neuer RTL-Schaltraum. Fernseh- und Kinotechnik 53, Nr. 3, 1999
[210] Steurer, J.: Digitale Filmbelichtung mit dem Arrilaser. Fernseh- und Kinotechnik 53, Nr. 5, 1999
[211] Stollenwerk, F.: Braucht man größere TV-Studios für HD? Fernseh- und Kinotechnik 62, Nr. 7, 2008
[212] Stolzmann, J.: DVD: Grundlagen, Technische Betrachtung. Fernseh- und Kinotechnik 51, Nr. 11, 1997
[213] Svatek, W., Lesowsky, S.: Professionelle Videobearbeitung. Franzis Verlag, München 1994
[214] Swinson, P.: Neue Entwicklungen bei Fimabtastern und Von der Quelle zur Bildwand. Fernseh- und Kinotechnik 44, Nr. 9, 1990 und FKT 55, Nr. 7, 2001
[215] Tacke, S.: Planung und Realisierung einer mobilen HDTV-Produktionseinheit. Diplomarbeit, HAW Hamburg 2005
[216] Tänzler, S.: Farbmetrische Untersuchung der Darstellung digitaler Bilder mittels Video und Film. Diplomarbeit, HAW Hamburg 2003

[217] Teichner, D.: Der MPEG-2-Standard (Teil 1 und 2). Fernseh- und Kinotechnik 48, Nr. 4 und 5, 1994
[218] Tektronix: Guide to Standard HD Digital Video Measurements, website
[219] Thomson Broadcast / GrassValley: Produktinformation
[220] Thorpe, L.: Whitepapers Canon C300, Canon 2012
[221] Tredup, L.: HD in 3D. Fernseh- und Kinotechnik 62, Nr. 1-2, 12008
[222] TVN-Television Programm und Nachrichten: Produktinformation Digitale Ü-Wagen. Hannover 2009
[223] Vogt, C.: Die DVB-Spezifikation für die Multimedia Home Plattform. Fernseh- und Kinotechnik 53, Nr. 1–2, 1999
[224] Waidson, M.: HD-Messtechnik. Fernseh- und Kinotechnik 61, Nr. 1-2, 2007
[225] Watkinson, J.: The Art of Digital Audio. Focal Press, Oxford 1994
[226] Watkinson, J.: The Art of Digital Video. Focal Press, Oxford 1994
[227] Weber, K.: CMOS-Sensoren bei Camcordern. Fernseh- und Kinotechnik 62, Nr. 6, 2008
[228] Webers, J.: Das Handbuch der Tonstudiotechnik. Franzis Verlag, München 1999
[229] Webers, J.: Handbuch der Film- und Videotechnik. Franzis Verlag, München 1991
[230] Weidner, S.: Avid Unity ISIS - ein Supercomputer für den Medienbereich. Fernseh- und Kinotechnik 60, Nr. 4, 12006
[231] Weith, M.: Vergleich und Einsatz komprimierend aufzeichnender Bandformate im gemischten Betrieb. Diplomarbeit, FH Hamburg 1998
[232] Wellerdick, M., Bancroft, D.: Vernetzte Redaktionsarbeitsplätze mit Zugriff auf Diskserver. Fernseh- und Kinotechnik 50, Nr. 11, 1996
[233] Wendland, B., Schröder, H.: Fernsehtechnik Bd. 1 und 2. Hüthig Verlag, Heidelberg 1991
[234] Wilke, H.: Virtuelle Effekte für Realtime- und Postproduktion. Diplomarbeit, FH Hamburg 1999
[235] Wilkinson, H.J., Przybyla, H.: Studio-Kompression für das neue Jahrtausend. Fernseh- und Kinotechnik 52, Nr. 5, 1998
[236] Windelband, G.: Beurteilung des filmtypischen Bewegungs- und Bildeindrucks einer Proscan Videokamera im Vergleich mit Film. Diplomarbeit, FH Hamburg 1999
[237] Witte, J.: Entwicklung und Realisation einer User Generated-Videoplattform für mobile Endgeräte. Diplomarbeit, HAW Hamburg 2007
[238] Wittenburg, J. P.: Flash-Memory-basierter mobiler Recorder für kinotaugliche digitale Kameras. Fernseh- und Kinotechnik 61, Nr. 1-2, 2007
[239] Wixforth, M.: Entwurf und Realisation einer elektronischen Schaltung zur Selektion der Luminanz der menschlichen Hautfarbe aus einem analogen Videosignal. Diplomarbeit, HAW Hamburg 2002
[240] Yousofy, S.N.: Der digitale Newsroom am Beispiel von AvidNews. Diplomarbeit, FH Hamburg 1999
[241] Zielinski, S.: Zur Geschichte des Videorecorders. Wissenschaftsverlag V. Spiess, Berlin
[242] Zieme, J.: Digitale Newssysteme – Heute und Morgen. Jahrestagung der FKTG, Tagungsband 1998
[243] Ziemer, A. (Hrsg): Digitales Fernsehen. R. v. Decker´s Verlag, Heidelberg 1994
[244] Zollner, M., Zwicker, H.: Elektroakustik. Springer-Verlag, Berlin Heidelberg New York 1993
[245] Zukowski, A.: Netzwerk-Speichersysteme in der Videobearbeitung am Beispiel von Avid Unity MediaNet. Diplomarbeit, HAW Hamburg 2000
[246] Zweites Deutsches Fernsehen, Website

Sachverzeichnis

1:1-Pixelmapping 485
1,485 Gbit/s 151
1“-B-Format 554
1“-C-Format 552
10 Bit logarithmische Daten 333
10 GbE 745
10-Bit-log 334
10-Bit-System 138
100 Hz-Technik 491
1000/1001 94
1080i-System 44
1080i/25 150
1080p/50 150
16 mm-Film 301
16-PSK 261
16/9-Vollformat 484
16/9-Vollformat Letterbox 484
2“-Quadruplex 550
20T-Impuls 108, 236
24 fps 683
270 Mbit/s 143
2D-Betrieb 796
2k 331
2T-Impuls 108, 235
2V-Sequenz 102
3 Gbit/s 151
3-D-Effekte 649
35 mm-Film 300
3D-Bildwiedergabe 492, 515
3D-Displays 492
3D-Editing 686
3D-Kino 516
3D-Liveübertragungen 459
3D-LUT 347
3D-Rig 457
3D-Sickness. 29
3G-SDI 155
3perf 304
4-PSK 260
4:1:1 140
4:2:0 140
4:2:2 139
4:2:2:4 139
4:3-Format 381
4:4:4 151
4/3-Letterbox 484
422P@ML 704
48 V-Phantomspeisung 427
4k 331
4k-UHD 153
4V-Sequenz 102
65 mm-Film 300
720p-HD-System 44
720p/50 150
74,25 MHz 149
8 mm-Film 301
8k-UHD 153
8V-Ident 571
8V-Sequenz 102
9 Pin-Schnittstelle 666

A

A-Format 552
AAC 68, 207
AAF 715
AAL 743
Aaton 430
Abberation 385
Abbildung 384
Abbildungsfehler 385
Abbildungsgleichung 385
ABC 415
Ablenkung 463
Ablenkwinkel 464
Abnahme 761
Abschattung 240
Abtastfrequenz 120
Abtastrate 150
ACE 198
ACES 451
ACK 558
ACM 277
Activity 741
Adaption 21
Adaptive coding and modulation 277
Adaptive Loop Filter 222
Additionssatz 57

U. Schmidt, *Professionelle Videotechnik,*
DOI 10.1007/978-3-642-38992-4, © Springer-Verlag Berlin Heidelberg 2013

Additive Farbmischung 70
ADSL 293
Advanced Audio Coding 207
Advanced Coding Efficiency 198
Advanced TV System Committee 266
Advanced Video Coding 216
Advanced-Authoring-Format 715, 718
AES/EBU-Format 146
AFM 547
After Effects 694
AGC 541
Agenturmeldungen 780
AIF 398
AIFF 671
Airbrush 651
AirPlay 777
Akkomodation 21, 29
Aktiv-Matrix-Display 474
Aktive Pixel Sensoren 369
Aktive Zeilendauer 44
Aktives Prisma 417
Alexa 444
ALF 222
Alias 33, 122
Alignment Layer 471
Alpha-Kanal 690
Aluminiumschicht 462
AM 228
AM negativ 240
AM-LCD 474
Amerikanische Einstellung 395
AMOLED 481
Amplitude Phase Shift Keying 262
Amplitude Shift Keying 260
Amplitudenmodulation 228
Amplitudenquantisierung 123
Amplitudensieb 469
Amplitudenspektrum 54
Anaglyphen-Verfahren 516
Analoge Modulationsverfahren 226
Analysator 470
Anamorphot 382
Anamorphotische Kompression 303
ANC 145
Ancillary Data 145
Angle 618
Animation 650
Anlegen 681
Anodenanschluß 462
ANSI-Lumen 510
Antennen 241
Anti-Scheimpflug 391
Anwendungsschicht 738
Apertur 33, 359
Aperturkorrektur 408
API 289
Application Layer 738
Application Programming Interface 289
APS 369
APSK 262
Arbeitskopie 682
Arbitrated Loop 749
Archiv 781
ARD 12
Arri 430
Arri-RAW 444
Arricode 305
Arriflex 432
Arrilaser 350
Arriscan 344
ASA-Wert 313
ASCII-Code 119
ASI 732
ASK 260
Assemble-Schnitt 663
Asset 782
Assisted Internal Focus 398
Astra-Satellitensystem 245
Asymmetrical Digital Subscriber Line 293
Asynchronous Serial Interface 732
Asynchronous Transfer Mode 742
ATA 606
ATAPI 606
ATF 536
ATM 742
ATM-Adaption Layer 743
ATSC 266
AÜ 787
Audio-Video-Interleave 723
Audiobearbeitung 678
Audiospezifikation 618
Aufführungslizenzen 513
Auflagemaß 400
Auflösung 39, 40, 48
Auflösungsvermögen 34, 315
Aufnahmeelektronik 541
Aufnahmeregeln 459
Aufnahmevorgang 527
Aufprojektion 494
Augendiagramm 130
Augmented Reality 698
Ausbreitungsgeschwindigkeit 64
Auslesemodus 366
Ausleuchtung 799

Außenübertragungen 787
äußerer Fehlerschutz 271
Aussteuerung 546
Austastlücken 43
Auszug 393
Auto Transition 641
Autocue 768
Autofocus 401
Automatic Black Control 415
Automatic Colour Kill 558
Automatic Gain Control 416, 541
Automatic Track Following 536, 567
Automatic White Control 414
Autostereoskopische Displays 492
Autosync 683
AUX Bus 640
Auxiliary Data 144
Avatare 195
AVC-I 221, 712
AVCHD 221
AVI 723
AVI-File-Format 723
Avid Media Composer 673
AWC 414
Azimut-Fehler 529
Azimut-Winkel 244

B

B – Y 83
B-Format 554
B-Frames 185
B/G-Standard 51
B/H 34
B4-Mount 400
Back Surround 328
Back Up 680
Backbone 294
Background 640
Backlight 476
Bakensignal 785
Band-Kopf-Abstand 528
Bandbreitenreduktion 86
Bandführung 534
Bandwellenlänge 526
BAS-Signal 41
Base plus fog 309
Basic UMID 715
Basisband 59, 127
BAT 269
Batch Digitize 674
Bauchbinde 633
Bayer-Pattern 380
Bayonet Neill-Concelman 62
BB 725
BD 620
Beleuchtungsstärke 20
Belichtung 307
Belichtungsfehler 310
Beobachtungsmonitore 486
BER 129
BES-OLED 480
Beschleunigungsspannung 462
Best Effort Transfer 737
Betacam 568
Betacam SP 571
Betacam SX 596
Betacart 572
Betamax 560
Betrachtungsabstand 33, 393
Betriebsmesstechnik 108
Betriebssysteme 289
Bewegungsadaptiv 258
Bewegungsauflösung 368
Bewegungskompensation 184
Bewegungsparallaxe 29
Bewegungsphasen 39
Bezugsblende 311
Bezugsebene 385
Bi-Phase-Mark-Code 132, 658
Bidirektionale Prädiktion 175
BIFS 195
Bild-Tonträgerabstand 233
Bild/Tonversatz 324, 758
Bildanmutung 368
Bildaufbau 30
Bildauflösung 331
Bildbasierte Trackingsysteme 803
Bildcodierung 159
Bildendstufe 469
Bildfeldgrößen 302
Bildfenster 338, 506
Bildgenerator 469
Bildhelligkeit 48
Bildhöhe 34
Bildmischer 624
Bildpunktabweichungen 707
Bildregie 769
Bildseitenverhältnis 34
Bildsignal 41
Bildstabilisierungssysteme 416
Bildstand 368
Bildtechnik 421, 768
Bildträger 233

Bildverkleinerung 649
Bildwähler 727
Bildwandlerabmessungen 361
Bildwandlerzeile 342
Bildweite 385
Bildwinkel 392
Bin 675
Binary Format for Scenes 195
Binary Shape 196
Binokular 27
Bi-Phase-Mark-Code 132
Bipolares Synchronsignal 45
Bit 119
Bit Error Rate 129
Bittaktfrequenz 128
BK-Netz 249
BK-Verstärkerstelle 251
BKGD 640
Black Burst-Signal 725
Black Matrix 465
Black-Burst 724
Blackgamma 409
Blackstretch 409
Blau 66
Bleichbad 299, 317
Blende 387
Blendeneinstellring 387
Blendenzahl 387
Blendfleisch 677
Blimp 432
Blinder Fleck 22
Block Coding 271
Block Typ 732
Blöcke 163
Blocking-Artefakte 168, 288
Blow up 317
Blu-ray Disc 620
Blue Box 633
Blue Screen 633
Blur 691
BNC-Stecker 62
Bobby 305
Bolex 430
Border 630
Bottom Emission Structure 480
Bouquet 290
Bouquet Association Table 269
Braunsche Röhre 31
Brechkraft 384
Brechung 383
Brechungsgesetz 18, 383
Brechungsindex 251
Breitbandkommunikationsnetz 249
Breitbild 303
Brennpunkt 384
Brennweite 384
Brennweitenverlängerungsfaktor 397
Brillen 517
Broadcast-Wave-Format 723
Browse-Server 779
Browser 293
Browsing 779
Bump Mapping 696
Burst 96
Burst-Blanking 101
Burstfehler 272
Bus 640
Buselektrode 478
Business-TV 291
Bust Shot 394
BWF 723
Byte 119

C

C 88
C-Format 552
C-MAC 256
C-Mount 401
C-Realtiy 340
CABAC 219
Camcorder 565
Camera Control Unit 421
Camera Network Unit 423
Camera obscura 3
Candela 19
Capstan 340, 530
Carrier Sense Multiple Access/Collision Detection 745
CAT 213
CAT 6 745
CAT5 745
Cathode Ray Tube 461
CAV 612
CB 84
CBR 162
CCD-Filmabtaster 342
CCD-Zeile 360
CCFL 476
CCIR B/G-Standard 52
CCIR-G 241
CCM 277
CD 612
CDi 612

CELP 208
Centerkanal 325
CF-Cards 623
Charge Coupled Device 359
Checksum 270
Chroma Key 633,797
Chroma Time Compressed Multiplex 570
Chromabandpass 105
Chromaspur 569
Chromatische Abberation 385
Chromatrack-M 568
Chromdioxyd 523
Chrominanzsignal 87, 88
Chrominanzvektor 88
Chunk 608
CIE 75
CIE-Diagramm 75
CIELAB 76
CIELUV 77
CIEXYZ 76
CIF 194
Cinemascope 303
Cineon 330
CIRC-Fehlerschutz 614
Clear Scan 376
Clip Level 631
Clipping 411
Close Up 394, 762
CLV 612
CMOS-Bildwandler 368, 437
CNU 423
Code Excited Linear Prediction 208
Coded OFDM 280
Coder-Kennimpuls 100
Coderate 271
COFDM 280
Color Analyser 320
Color Spill 691
Colour Frame Flag 656
Colour Framing 668
Colour Under-Aufzeichnung 556
Colour Under-Frequenz 557
COM-Verfahren 234
Combustion 694
COMMAG 322
Commission Internationale d'Eclairage 75
Common 2k/8K-Specification 281
Common Interchange Format 194
Common Interface 289
COMOPT 322
Compact Disc 612
Compact Disc interactive 612
Composite Sync 46
Composite-Signal 92
Compositingsysteme 687
Composition 716
Composition-Paket 716
Compressed Serial Data Interface 731
Compressed Time Division Multiplex 570
Computeranimation 695
Computermonitore 36
Conditional Access 289
Conditional Access Table 213
Constant Angular Velocity 612
Constant coding and modulation 277
Constant Linear Velocity 612
Content 736
Content Management System 782
Context Adaptive Binary Arithm. Coding 219
Continuous Contact Printer 318
Contouring 637
Contribution 242
Convolution Coding 271
Cooke 396
Cordband 322
Core-Profile 194
Countdown 680
CR 84
Crab and steer 455
Crash Record 662
CRC 659
Crispening 408
Cross Colour 93
Cross Luminanz 93
Crossconnect 743
CRT 461
CRT-Belichter 348
CSDI 731
CSMA/CD 745
CTCM 570
CTDM 570
CTL 655
CTL-Spur 535
CTL-Zeitcode 660
Cue 566
Curie-Temperatur 523
Cutter 670
Cyclic Redundancy Check 659

D

D-ILA 500
D-MAC 256
D-VHS 603

D-VITC 659
D/A-Wandler 126
D1-Format 583
D10 598
D2-Format 586
D2-MAC 256
D20 442
D3-Format 586
D5-Format 584
D6-Format 600
D65 69
D7-Format 591, 594
D8-Format 596
D9-Format 595
Da Vinci 346
DAB 285
Dachschräge 60, 108
Daguerreotypen 4
DAI 214
Darstellungsschicht 738
DAS 750
Data Broadcasting 284, 285
Data Carousel 286
Data Piping 286
Data Streaming 286
Data Wrangler 763
Dataplanes 741
Datei-Body 718
Dateisysteme 606
Datenformate 713
Datenpaket 736
Datentransferrate 606
DATV 258
dB 65
DBS 286
DC-Komponente 165
DC-PDP 478
DCDM 513
DCI 513
DCP 514
DCT 163
DCT-Basisfunktionen 166
DCT-Format 588
DCT-Koeffizienten 168
DDP 619
Deemphasis 540
Defokussierung 377
Degauss 465
Deinterlacing 40
Delay 800
Delivery Multimedia Integration Framework 193
Desktop Video 672
Detailing 408
Deutlicher Sehwinkel 392
deziBel 65
Dezimations-Tiefpassfilterung 254
Dezimeterbereich 241
DI 346
Dialnorm 209
Diamond-Display 117
Dichroitische Schichten 387
Dichteumfang 333
Dichtewert 308
Dictionaries 716
Differential Puls Code Modulation 172
DigiBeta 588
Digital 118
Digital 8 596
Digital Audio Broadcasting 285
Digital Betacam 588
Digital Cinema Distribution Master 513
Digital Cinema Initiatives 513
Digital Cinema Package 514
Digital Cinematographie 430
Digital Component Signal 135
Digital Component Technology 588
Digital Composite Signal 133
Digital Cut 680, 685
Digital Image Stabilizer 416
Digital Image Technican 763
Digital Light Processing-Einheit 503
Digital Micromirror Device 502
Digital Negative 451
Digital Rights Management 200
Digital S 595
Digital Source Master 513
Digital Theater System 329
Digital Versatile Disk 612
Digital Video Broadcasting 267
Digital Visual Interface 156
Digital-Intermediate-Prozess 346
Digitalaufzeichnung 576
Digitale Bildmischer 644
Digitale Effektgeräte 648
Digitale HD-Signale 148
Digitale Klemmung 637
Digitale Kreuzschiene 792
Digitale Modulationsverfahren 259
Digitale Fernsehempfänger 490
Digitales Negativ 335
Digitales Fernsehen 267
Digitales Positiv 335
Digitales Satellitenfernsehen 274

Digitales terrestrisches Fernsehen 279
Digitalisierung 120, 675
Digitalkamera 413
Digitally Assisted TV 258
DIN-Zahl 313
Dioptrie 384
Dioramaeffekt 391
Dirac-Impuls 56
Dirac-Pro 223
Direct Attached Storage 750
Direct Drive-ILA 500
Direct Read after Write 582
Director of Photography 763
DIS 416
Disc Description Protokoll 619
Discrete Multitone 294
DiSEqC 248
Diskrete Cosinustransformation 163
Displacement Mapping 696
Display-Port 735
Displayphase 478
Distribution 239
DIT 763
DivX 723
DKS 792
DLP 503
DMD 502
DMIF Application Interface 214
DMT 294
DNG 451
DNxHD 224
Dolby 546
Dolby A 326
Dolby AC3 618
Dolby B 326
Dolby Digital 328
Dolby E 209
Dolby Noise Reduction 326
Dolby SR 326
Dolby Stereo 326
Dolby Surround 327
Dolly 454
Donnereffekt 324
DoP 763
Doppeloptik 518
Doppelzackenschrift 323
Down Stream Key 641
Down-Link 245
Downstream 293
DPA 780
DPCM 172
DPM 382
DPX 721
DR 374
Dreidimensionale Wahrnehmung 26
Dreikopf-Recorder 562
DRM 200
Drop Frame 661
Drop Frame-Modus 656
Drop Out Compensation 544
DSK 143, 641
DSLR 357, 449
DSM 513
DT 537
DTS 213, 329
Dual Link 155
Dual-Layer 616
Dualzahlensystem 119
DUB 558
Dunkelfeldprojektion 501
Dunkelsack 306
Dunkelstromrauschen 373
Duplikat Films 307
DV-Algorithmus 178
DV-DIF 180
DV-Format 591
DV-Signale 703
DVB 267
DVB Endgeräte 286
DVB-C 277
DVB-C2 279
DVB-H 284
DVB-MC 279
DVB-MS 279
DVB-S 274
DVB-S2 276
DVB-SH 285
DVB-T 279
DVB-T2 284
DVCam 593
DVCAM 593
DVCPro 594
DVCPro50 595
DVCProHD 602
DVD 614
DVD-Authoring 618
DVD-R 619
DVD-RAM 619
DVD-RW 619
DVD-Video 617
DVD+RW 619
DVE 648
DVI 157
DVD-RW 619

DWT 170
DXB 285
Dynamic Pixel Management 382
Dynamic Range 374
Dynamic Tracking 537
Dynamikbereich 202

E

E-E-Betrieb 541
E'r 81
EAP-Kamera 425
Eastman/Kodak 296
EAV 140
EB 404
EB-Team 762
Ebene 627
EBU 78, 151
EBU R118 763
EBU R128 757
EBU R132 754
EBU-Norm N10 84
EC-Spur 685
Echosignal 264
Echtzeitübertragung 701
ECM 289
ECN-2 317
ECP-2B 317
Edge Code 305
EDH 147
Edit Decision List (EDL) 654, 715
Editing 652, 675
Editor 667
EDL 654, 669
EF-Mount 401
Effects Memory 643
Effects-Spur 676
Effekte 678
EFM 614
EFP 762
Eight to Fourteen Modulatio 614
Eimerkette 359
Einseitenband-Modulation 229
Einstrahlröhre 464
Eisenoxyd 523
EIT 269
EL 479
Elapsed Time 661
Electron Transport Layer 480
Electronic Cinema 512
Electronic Field Production 762
Electronic News Gathering 404
Electronic Program Guide 289
Electronics to Electronics 541
Elektromagnetische Strahlung 67
Elektronische Berichterstattung 404
Elevationswinkel 244
EM 229
Embedded Audio 146
Emitter Layer 479
EMM 289
Empfangsantennen 241
End of Active Video 140
End of Block 176
Energieversorgung 245
Energy Dispersal Scrambling 270
ENG 404
Entitlement Control Message 289
Entitlement Management Message 289
Entmagnetisierung 465
Entropiecodierung 161
Entspiegelung 386
Entwicklung 297
EOB 176
EOS C500 450
EOTF 485
EPG 289
Epic 447
Equalizer 679
Erdefunkstelle 245, 786
Error Detection and Handling 147
eSATA 607
ESB 229, 230
Essence 716, 741
Estar 296
Ethernet 744
ETL 480
Euro 1080 276
European Broadcasting Union 78, 151
Eutelsat 248
Event Information Table 269
EVS 793
Expansion 570
Exposure 307
Exposure Index, EI 313
ExpressCards 623
Extended UMID 714
Extern Key 635, 798
External Serial ATA 607
Externe Kabeltafel 788
Extrem Close Up 394

F

F-Drop 399
F-Stop 387
F/H-Phase 102
F23 446
F35 445
F55 448
F65 447
F900 438
Fabric 749
Facial Animation Object 197
Faderbar 625
Fahrstative 766
Faksimile 338
Faltungscode 271
Faltungssatz 56
Farb-, Bild-, Austast-, Synchronsignal 91
Farbabschalter 558
Farbart 70
Farbartsignal 88
Farbbalkensignal 82
Farbbildröhre 464
Farbdifferenzkomponenten 83
Farbe 66
Farbempfindung 69
Farbfernsehnormen 93
Farbfilm 299
Farbflächen 642
Farbgemisch 72
Farbhilfsträger 87, 88
Farbkorrektur 345
Farbmetrik 72
Farbmischkurven 80
Farbmischung 70, 299
Farbprüfsignal 82
Farbraum 73, 691
Farbreinheitsfehler 465
Farbreiz 66
Farbsättigung 70
Farbsignal 66
Farbsignalhierarchie 104
Farbstich 312
Farbstoffwolken 299, 314
Farbtemperatur 67, 414
Farbton 70
Farbtripel 464
Farbtüte 73
Farbvalenz 66
Farbwahrnehmung 69
Farbwerte 70
Farbwertsignale 66
FBAS-Direktaufzeichnung 550
FBAS-Signal 91
FCC 77
FCP 734
FDMA 274
FEC 270
FED 482
Feed Forward-Steuerung 179
Feedhorn 247
Fehlerschutz 270
Feldlinien 522
Fenster 252
Fernsehbegleitsignale 233
Fernsehbegleitton 233
Fernsehprogramm 226
Fernsehsignalübertragung 226
Fernsehtextsignal 237
Festbrennweite 393
Festplatten 605
Festwertspeicher 623
Field Emission Display 482
Field Jitter 107
Field Read Mode 366
Field Still 562
Field-Integration-Mode 366
Fields 37
File Transfer Protokoll 293
File-Paket (FP) 720
Filetransfer 736
Film Label 306
Film Rate 685
Film Settings 684
Film-Look 341
Filmabtaster 337
Filmantrieb 430
Filmbelichtung 348
Filmbildfrequenz 25
Filmeigenschaften 307
Filmformate 300
Filmkamera 368, 430
Filmkennlinie 308
Filmkorn 298, 314, 368
Filmlook 298, 368
Filmmusik 682
Filmoption 684
Filmproduktion 763
Filmprojektor 506
Filmschleife 324
Filmschnitt 681, 683
Filmtechnik 296
Filmtellereinrichtungen 507
Filmton 322

Filmtransport 430
Filter 402
Filterbank 203
Final Cut Pro 673
Finder 432
Firewire 733
First field 40
FIT 365
Fit to Fill 676
Fit-Funktion 668
Fixed Block Structure 732
Fixed Pattern Noise 343, 373
Fixed Satellite System, 244
Fixierbad 317
Fixiervorgang 297
Flächenleuchten 799
Flame 694
Flare 407
Flash 723
Flash-RAM 623
Flashing 318
Flaw-Korrektur 406
FlexMux 215
Flicker 434
Flimmererscheinung 22
Flimmerverschmelzungsgrenze 507
Flint 694
Flip-Flop-Betrieb 640
Floating Gate 623
Flügelblende 362
Fluorescent Multilayer Disk 622
Flüssigkristallanzeigen 470
Fluter 765
Flying-Spot 339
FM 231
FMD 622
Fokussierelektroden 462
Fokussiergruppe 397
Fokussierung 389
Fokussierungsfehler 614
Footer 718
Footprint 245
forcierte Entwicklung 318
Forward Error Correction 270
Fotodetektor 613
Fotodiode 340
Fotografie 296
fotometrische Größen 19
Fourierkoeffizienten 53
Fouriertransformation 53
Fovea centralis 21
FPN 343
Fps 337
Frame 37
Frame Read Mode 366
Frame Sequential 517
Frame Transfer 362
Frame-Interline-Transfer 365
Frame-Multiplexing 731
Frame-Reset-Mode 366
Frames per Second 337
Framestore 726
Framestore-Synchronizer 726
Fremdspannungsabstand 62
Frequency Division Multiple Access 274
Frequency Shift Keying 260
Frequenzhub 231, 233
Frequenzmodulation 231
Frequenzmultiplex 227
Frequenzspektrum 53, 102
Frontfokussierung 398
Frontlinse 398
FSK 260
FSS 244
FT 362
FT-CMOS 370
Ftp 293
Full apertur 331
Full Length Shot 395
Full screen contrast 485
Full-Frame-CCD 362
Full-well capacity 374
Füllsender 242
Functional Channel Protocol 734
Funkschärfe 765
Fury 349
Fusion 27
Fußnummer 305
Fußnummern 683

G

γ 310
γ-Vorentzerrung 42
Gain Up 407
Gain-Faktor 505
Gamma-Vorentzerrung 409
Gammawert 310, 409
Gamut 105
Ganged Packages 720
Gap 592
Gates 338
GbE 745
Gegenstandsweite 385

Geisterbilder 242
Gelbfilterschicht 299
Gen-Lock 405
General Purpose Interface 642, 665
General-eXchange-Format 720
Genesis 445
GEO 243
Geometrischen Optik 383
Geostationärer Orbit 243
Gerber-Norm 12
Gewichtete Quersumme 271
Glasfasern 250
Gleichabständige Farbwahrnehmung 76
Gleichwellenbetrieb 279
Gleichwellennetz 263
Global Shutter 370
GoCam 455
GOP 185
GoPro 450
GPI 642, 643, 665
Graceful Degradation 191
Gradation 310
Grade 1 485
Grade 2 486
Gradientenfaser 251
Grafiktablett 651
Grain 314
Granularität 314
Grauglas 462
Grautreppe 83, 236
Greiferwerk 430
Großaufnahme 394
Großflächenflimmern 25, 37
Group Management Protocol 295
Group of Pictures 185
Grün 66
Grundfarben 66
Grundgesetz der Nachrichtentechnik 227
Grundlagen der Optik 383
Grundnetzsender 242
GSE Rapid Timecode 660
Guard-Intervall 265
Gültiges Signal 105
GXF 720

H

H-Periodendauer 43
H-Phase 725
H-Phasenabgleich 65
H-Ruck 668
H-Synchronpuls 45
H-V Delay 486
H.264-Formate 712
H.264/AVC 216, 221
H.265/HEVC 221
H.26x 216
Halbbilddominanz 40
Halbbilder 27
Halbbildkennung 141
Halbzeilenoffset 93
Hard Disk 605
Harddisk-Recorder 609
Harmonische 54
Hauptebenen 385
Hauptschaltraum 778
Hausübergabepunkte 251
Havariefall 729
Havariekreuzschienen 728, 769
HDTV-Logo 288
HD-D5 602
HD-DVD 620
HD-Objektive 399
HD-Schnittstelle 154
HD-SDI 154
HD-Signale 148
HD1 602
HD2 602
HDCAM 601
HDCAM SR 599, 601
HDMI 157
HDTV 35, 253
HDTV analog 257
HDV 602
Head Mounted Display 516
Header 718
Header Partition 718
Headerdaten 736
Headroom 125, 549
Headset 419
Hellempfindlichkeitskurve 74
Helligkeitsempfindung 26
Hellsektor 431
Helpersignal 254
HF-Leistungsverstärker 785
HF-Vormagnetisierung 545
HFR 25, 519
Hi 8 567
HiFi-Ton 547
High 4:2:2 220
High Definition Multimedia Interface 157
High Definition Television 36
High Definition-MAC 258
High Efficiency Video Coding 221

High Frame Rate 25, 519
High Profile 191
High-Band 540
Hintergrundbeleuchtung 476
HIT 458
HMD 516
Hochpass 60
Hochspannungskaskade 469
Höhenparallaxe 457
Hohlkehlen 796
Hold Type Displays 475
Hollywood 7
Holographic Versatile Disc 622
Holographie 492
Home-Point 803
Hörfläche 201
Horizontal Image Translation 458
Horizontale Austastlücke 43
Horizontales Ausleseregister 362
Hörschwelle 201
Hosidenstecker 90
Hot Bird 248
Hot Swap 609
HP-Koeffizient 170
HSR 778
HTML 293
http 293
Hub 231, 749
Hubbereich 539
Hue 70
Huffman Code 161
Hüllkurvendemodulator 229
HVD 622
Hybnet 751
Hybridsysteme 805
Hyperstereoskopie 458
Hypertext Markup Language 293
Hypertext Transfer Protokoll 293
Hypostereoskopie 458
Hysteresekurve 524

I

I 94, 260
I und Q 94
I-Frames 185
i-Link 733
I-only 708
IA 457
IBK-Diagramm 75
ID3-Tag 206
IDCT 165
IDE 606
IDFT 265
IEEE 1394 733
IGMP 295
Ikonoskop 10, 356
ILA 499
Illegale Farben 105
Image Enhancement 407
Image Light Amplifier 499
IMAX 304
Impulsdauer 54
Impulskreuz 486
Impulszählung 802
IMX 598
In House Netze 741
In Plane Switching 473
Indoor Unit 247
Induktionsgesetz 528
Inferno 694
Infrarotimpulse 518
Inkohärente Demodulation 229
Innenfokussierung 398
innere Mischung 73
Innerer Fehlerschutz 271
Insert 676
Insert Track Information 592
Insertkamera 646
Instanz-Nummer 714
Integer-Transformation 219
Integrated Device Electronics 606
Integrated Services Digital Network 742
Intensitätsmodulator 350
Interaktion 291
Interaktive Regie 291
Interaktives Fernsehen 291
Interaxial 457
Intercarrier-Verfahren 490
Intercom 419, 766
Interlace-Faktor, 39
Interlaced 438
Interlaced-Mode 37
Interleaving 272, 581
Interline Transfer 363
Intermediate Codec 224
Intermediate-Material 307
Intermittierender Antrieb 324
Intermodulation 108, 237
Internationale Standards 51
Internationale Beleuchtungskommision 75
Internegative 307
Internet 292
Internet-Protokoll 292

Interpolation 254
Interpositiv 307
Intersymbol Interferenz 128
Intraframe-Prädiktion 218
Inverse diskrete Fouriertransformation 265
IP 292, 294
IPMP 215
IPS 473
IPTV 292
IPv4 294
IRE 49
Irrelevanzreduktion 158
ISDN 294, 742
ISI 128
ISO (International Organization for Standardisation 723
ISO/OSI-Schichtenmodell 737
IT 363
IT-Ton 784
ITI-Sektor 592
ITU-R BT. 601 135
ITU-R BT.709 148

J

JAVA 290
JBOD 608
JPEG2000 514
Jib-Arm 455
Jitter 119, 543, 728
Jog/Shuttle-Modus 668
Joint Photographics Expert Group 176
Joint Stereo 205
Journalist Workstation 780
JPEG 176
JPEG2000 177
Just a bunch of disks 608

K

K 68
k-Faktor 236
Kabelfernsehen 249
Kabelkanäle 249
Kabelkopfstation 251
Kabeltauglich 250
Kalibrationsprozess 335
Kalibrierung 113, 806
Kalibrierungssysteme 799
Kamera-Matching 108
Kameraassistenten 763
Kamerafahrt 393
Kamerakennlinie 408
Kamerakontrollraum 768
Kameramann 763
Kameramatching 768
Kameraperspektive 795, 800
Kamerasucher 418
Kamerasupport 453
Kameratracking 800
Kammfilter 106
Kanalaufteilung 241
Kaskadierung 627
Kassette 430
Kathode 461
Kathodenstrahlröhre 461
Kelvin 68
Kern 251
Key Bus 640
Key Fill 631, 635
Key Mask 636
Key Source 635
Key-Stoning 458
Keycode 305, 683
Keyframe 688
Keying 630, 690
Keystone 505
Kinematographie 6
Kinematoskop 6
Kinofilmproduktion 681
Kinora 5
Kippender Farbstich 312
Kissenverzerrung 464
Klappe 766
Kleinbildkamera 393
Klemm-Schaltung 61
Klirrfaktor 62,549
KLV (Key, Length, Value) 718
Knotenpunktversatz 800
Koaxial-Leitungen 724
Koerzitivfeldstärke 523
Kohärente Demodulation 231
Komparator 728
Kompendium 388, 765
Kompensator 397
Komplementärfarben 70, 299
Komponentenaufzeichnung 568
Komponentensignal 82
Komponentensignalübertragung 87
Kompression 158, 160
Kontaktkopie 318
Kontrast 388, 765
Kontrastempfindlichkeit 69
Kontraststeller 488

Kontrollmonitore 486
Konvergenz 29, 466
Konvergenzwinkel 27, 458
Konversionsfilter 402, 403
Kopernikus 785
Kopfpaare 569
Kopfspalt 525
Kopfspiegel 526
Kopftrommel 531, 533
Kopfüberstand 533
Kopfumschaltung 535
Kopfwechsel 773
Kopierwerk 317
Körperfarben 68
Korrelation 160
Korrespondierende Punkte 27
Kreuzschienen 727
Kristallgröße 298
Ku-Band 245
Kunstlicht 403
Küpfmüller-Nyquist-Bedingung 128

L

L, C, R 325
Labelseite 613
Laboratory Films 307
Lambertstrahler 19
Lampenhaus 508
Land 613
Landschaftsaufnahmen 458
Längsspur-Aufzeichnung 545
Laptop-Editor 774
Laser-Belichter 349
Laser-Lichttonkamera 324
Laserdioden 252
Laserprojektion 496
Last Mile 294
Latentes Bild 297
Laterale Hemmung 22
Laterna Magica 1
Lauftitel 647
Laufzeitdifferenzen 64
Laufzeiten 263
Laufzeitproblematik 771
Layer 186
Layer 2 203
Layer-3 204
LCD 470
LCD-Projektoren 498
LCEC 222
LCLV 498
LCOS 501
LDPC-Code 274
Lebensdauer 243
LED 479
Legal Range 352
Legales Signal 105
Leitungen 62
Leitungsabschluß 63
Leitungsentzerrer 62, 724
Leitungslänge 144
LEO 243
Letterbox 253, 304
Leuchtdichte 19
Leuchtdichtesignal 82
Leuchtstoffe 71, 77, 461
Level 190
Licht 16
Lichtbestimmer 320, 682
Lichtempfindliche Schicht 358
Lichtempfindlichkeit 313
Lichtenergie 17
Lichtfeldkamera 371
Lichtmodulator 496
Lichtreflexe 798
Lichtstärke 19, 387
Lichtstrom 19
Lichttonkamera 323
Lichttonverfahren 323
Lichtventile 320
Lichtwellenleiter 251
Ligthning-Darstellung 111
Line Jitter 107
Line Shuffling 258
Line-Multiplexing 731
Linear Editing 652
Linear Matrix 412
Lineare Polarisation 245
Lineare Verzerrungen 61
Linienpaare 315
Linienstrahler 67
link 749
Link A 155
Linsenvergütung 386
Liquid Crystal Display 470
Liquid Crystal Light Valve 498
Liquid Crystal on Silicon 501
Live on Air 766
Live-Zuspielung 759
LNB 247
LNC 247
Lochmaskenröhre 466
Logging 674

Lokalisation 325
Long GOP 709
Long-Play 562, 565
Longitudinal Timecode 657
Look up-Table 347
Löschkopf 529
Löschphase 478
Löschvorgang 529
Low Complexity Entropy Coding 222
Low Density Parity Check Code 274
Low Earth Orbit 243
Low Noise Converter 247
Lower field 40
LP 565
lp/mm 315
LTC 657
Lumen 19
Luminance 70
Luminanz-Key 631
Luminanzsignal 83
LUT 347
Lux 20

M

M-JPEG 177
M-Loading 532
MAC 256
MAC (Media Access Control) 714
MAC OS 673
Machsche Streifen 22
MACP 253
Magazin 238, 430
Magnetaufzeichnungsgerät 520
Magnetisch hart 524
Magnetisch weich 524
Magnetische Felder 522
Magnetisierung 522, 523
Magnetisierungskennlinie 523, 527, 545
Magnetkopf 525
Magnetowiderstandselement 605
Magnetpartikel 521
Magnetplatten 605
Magnetschicht 521
Magnetspuren 322
Main Profile 191
Mainstream-Produktion 756
Makroblöcke 179
Malterserkreuz 508
Mapping 731
Marker 805
Maske 465
Maskierung 687
Massenkopien 307
Master System Unit 423
Matched Filter 130
Material-eXchange-Format 718
Material-Paket (MP) 716, 719
Materialorganisation 675
Matrix 80, 412
Matrixcodierung 327
Matrizierung 83
Matte 636
Matte Fill 636
Matte-In 798
Mattebox 388, 765
Mattscheibe 432
Maximalpegel 115
Maya 697
MAZ 520
MAZ-Karte 761
McAdam-Ellipsen 76
MDCT 207
Media Block 515
Media-Manager 782
Mediaserver 610, 779
Mehrkanaltonverfahren 325
Mehrteilbildcodierung 258
Mehrteilnehmeranlagen 248
Mehrwegempfang 266
Mehrwertige Modulation 132, 259
Memory Head 802
MEO 243
Mergin Bits 614
Metadaten 718, 740, 741
Metall Oxid Semiconductor 359
Metamerie 71
Mezzanine 223
MFD 482
MFT-Sensor 357
MHP 290
Micro For Thirds 357
Micro Reciprocal Degrees 68
Microscanning 345
Microtip Fluorescent Display 482
Microwave Cable Based 279
Microwave Satellite Based 279
Mid Shot 395
MIDI 722
Midsize-Cassetten 594
Mini 35 441
Miniaturisierungseffekt 390
Minuspol 522
Mired-Verschiebungswert 403

Mirror-Rig 457
Mischatelier 682
Mithörschwelle 202
Mittenortung 325
Mix-Mode 237
Mix/Effektstufe 626
MMDS 279
Mobile and Calendar 187, 706
Modelling 695
Modulationsgrad 228, 315
Modulationsindex 231
Modulationstransferfunktion 23, 315
Modulationszustand 259
Monitorabgleich 488
Monitorkreuzschiene 770
Monokulare Tiefenhinweise 29
Monomodefasern 252
Montage 652
MOS-Kondensator 359
Mosaikfilter 380
Motion Adaptive Colour Plus 253
Motion Control-Systeme 456
Motion JPEG 177
Motion Picture Film 296
Motion Picture Matrix 327
Motorfeed 248
Mount 400
Moviecam 430
Moving Pictures Experts Group 182
MP@H14L 191
MPEG 182
MPEG-1 182
MPEG-2 182, 189
MPEG-2-Editing 709, 712
MPEG-21 200
MPEG-4 182, 192
MPEG-4-Systems 214
MPEG-7 199
MPEG-Audio 200
MPEG-Decoder 185
MPEG-H Part 2 222
MPEG-Systems 210
MPEG-Videocodierung 183
MPLS 747
MR-Code 305
MR-QAM 281
MSE 706
MSU 423
MTF 23, 315
Mulitimedia 698
Multi-Layering 721
Multiburst 236
Multiburst-Signal 58
Multicast-Technik 295
Multichannel/Multipoint Distrib. System 279
Multigenerationsverhalten 707
Multimedia 291
Multimedia Content Description Interf. 199
Multinormproduktionen 790
Multiple Subsampling Encoding 257
Multiplexed Analogue Components 256
Multiplizierer 638
Multiprotocol Encapsulation 286
Multiprotocol Label Switching 747
Multisymbolsignal 265
MUSE 257, 258
Muster 681
MVA 473
MXF 718

N

n-1-Signal 784
Nachbelichtung 311
Nachbildeffekt 9
Nachbildwirkung 2
Nachhallzeit 329
Nachrichtenquader 227
Nachrichtensendungen 778
Nachrichtenverteilung 780
Nachtaufnahme 766
Nadeltonverfahren 322
Nah 394
NAS 750
Nasskopiereinrichtung 319
National Television System Commitee 94
Navigation 290
ND-Filter 402
Nebendichten 312
Negative Parallaxe 28
Negativentwicklung 317
Negativschnitt 682
nematische Phase 471
Network Attached Storage 750
Network Information Table 213
Network Layer 738
Netzebenen 250
Netzhaut 21
Netzwerk-Speicher 750
Netzwerke 736, 737
Neutraldichtefilter 402
News 762
News in Film 784
News Room Automation 780

Newsroom Systeme 778
Next Effects 641
Nichtkopierern 681
Nichtlineare Schnittplätze 775
Nichtlineare Verzerrungen 61
NIF 784
Nipkowscheibe 10
NIT 213
NLE 670
Nodal Point Shift 800
Node 749
Non Drop Frame 661
Non Square Pixels 136
Nonlinear Editing 653
Nonlineare Editingsysteme 670
Nordpol 522
Normal-Bildwinkel 392
Normalabbildung 391
Normalbelichtungsumfang 310
Normalbeobachter 72
Normalbrennweite 392
Normalfilmformat 302
Normalized Primary Matrix 79
Normally Black Mode 471
Normfarbtafel 75
Normfarbwerte 74
Normlichtart 69, 75
Normspektralwertkurven 74
Normvalenzsystem 74
NP1-Akkus 426
NPM 79
NRZ 131
NRZ-Code 580
NRZI 131, 580
NRZI-Code 144
NTSC 94
NTSC-Coder 95
NTSC-Pulldown 337, 683
NTSC-Zeilen 98
Nullparallaxe 28
NXCAM 221
Nyquistflanke 230,490

O

OB 787
Oberes Seitenband 228
Oberschwingungen 53
Object Descriptors 195, 215
Object modelling technique 741
Objekt Carousel 286
Objektiv 396
Objektivbefestigungen 400
OCL 364
OCP 421
OETF 485
OFDM 263, 280
Offline-Betrieb 795
Offsetmontage 380
Okklusion 29
OLED 479
Omega-Gang 528
OMF 671, 715, 721
OMFI 721
On Air Buffer 779
On-Chip-Lens 364
Online-Betrieb 795
Online-Schnitt 671
Opazität 307
Open Media Framework 671, 721
Open Media Framework (OMF) 721
Open Range 352
OpenTV 289
Operational Control Panel, 421
Operationsverstärker 65
Optical Black 358
Optical Printer 321
Optische Abbildung 384
Optisches Tiefpaßfilter 377, 402
ORAD 803
Orbitalposition 275
Organisationscode 293
Orthogonal Frequency Division Multiplex 262, 263
Orthogonale Abtastsstruktur 135
Ortsabschnitt 251
Ortsauflösung 22
Ortsfrequenzen 32
OSB 228
Oszilloskop 109
OUT 668
Outside Broadcasting 787
Over Shoulder Shot 762
Overlay 111
Overscan 49
Overwrite 676

P

P-Frames 185
P-OLED 479
P2 623
Paarige Rasterstruktur 38
Packet Elementary Stream 210

Packet Identification 212
Page 293
Paintbox 650
PAL-4V-Sequenz 102
PAL-8V-Sequenz 102, 656
PAL-Burst 100
PAL-Frequenzspektrum 102
PAL-Kennimpuls 100
PAL-Kippstufe 468
PAL-Laufzeitdecoder 468
PAL-Plus-Kennung 239
PAL-Pulldown 683
PAL-Sequenzen 101, 102
PAL-Verfahren 87, 98
PAL-Zeilen 98
PALplus 253
Pan 453
Pan and Scan 253, 304
Panavision 430
Pandora 346
Parade-Darstellung 111
Parallaxenbarriere 492
Parallelausrichtung 458
Parallelschnittstelle 142
Parallelstrahl 384
Paralleltonverfahren 490
Partitionen 606
Passive Matrix 474
PAT 213
Pay per Channel 291
Pay per View, 291
Pay TV 291
Payload 732
PB 84
PCM-Signal 127
PCMCIA 623
PCR 212
PDH 742
PDP 477
Peak Program Meter 549
Peaking 418
Pedestal 410, 453
Pegel 373
Pegelreduktion 84
Pegelsenkung 679
Perfoband 322
Perforation 338
Perlschnurmuster 93
Permanent Virtual Circuit 742
Perpendicular Recordin 606
Perspektive 393
PES 210
PGM 639
Phantomschallquellen 325
Phantomspeisung 427
Phase Alternation Line 98
Phase Change-Technologie 619
Phase Combiner 785
Phase Locked Loop 232
Phase Shift Keying 260, 262
Phasenzustandsdiagrammen 261
Phenakistiskop 2
Phosphore 71, 461
Photoeffekt 339
Photopisches Sehen 19
Photoshop 651, 721
Photozelle 339
Physical Layer 737
Physical-Source-Paket 716
Physikalische Schicht 737
Pict 647, 671
Picture in Picture 491
Picture-Layer 187
PID 212
Piloten 282
Pilothilfsträger 234
Pin registered 344
Pit 613
Pitch 466
Pixel 32
Pixel Aspect Ratio 137
Pixelansteuerung 474
PL-Mount 400
Plasmadisplay 477, 479
Play 663, 669
PLED 479
Plenoptischer Bildwandler 371
Plesiochrone digitale Hierarchie 742
PLL-Schaltung 232
PLUGE-Signal 117
PMT 213
Pogle 346
Point to Point 736
Polarisation 18, 245
Polarisationsfilter 403
Polarisationsfolie 470
Polarisationsverfahren 493, 517, 518
Polarisator 470
Pole 522
Ports 733
Positioner 630
Postscript 647
PPM 549
PR 84

Prädiktion 173
Prädiktions-Unit 222
Präsenzfaktor 459
Pre Roll-Zeit 662
Pre- & Post-Equalizing Pulses 46
Precinct 177
Premiere 673
Premium-Produktion 756
Presentation Layer 738
Presentation Time Stamp 213
Preset 640, 661
Preset Pattern Key 636
Primärreizkurven 71
Primärvalenzen 70, 74
Primärvalenzfarbdreieck 73
Prime Lens 396, 440
Print 318
Pro Logic 327
Produktionsanforderungen 753
Produktionseinheiten 762
Professional Disc 621, 622
Professional Profile 704
Program 639
Program Association Table 213
Program Clock Reference 212
Program Map Table 213
Programm-Modus 346
Programm-Strom 211
Programmaustausch 759
Programmspezifische Informationen 213
Progressiv 37
Progressiver Bildaufbau 37
Projektion 494, 509, 512
Projektionsobjektiv 508
ProRes-Codec 224
Provider 293
Proxy AV-Daten 621
Prüfzeilen 235
PSD 721
PSI 213
PSK 260
PST 640
PTS 213
PU 222
Pull-Modell 513
Pulldown 337
Puls-Kreuz-Modus 46
Pumpen 453
Push-Betrieb 737
PVA 473
PVC 742

Q

Q 94, 260
QAM 88, 230
QoS 292, 737
QPSK 260
Quad-PSK 260
Quadraturamplitudenmodulation 88, 230, 261
Quadruplex 550
Quality of Service 292, 737
Quantisierung 166
Quantisierungsrauschen 124
Quantisierungsstufen 123
Quantisierungstabellen 167
QuarterCam 568
Quasi-Paralleltonverfahren 490
Quasi-Realzeit 661
Quelle 669
Quellencodierung 127
Querspuraufzeichnung 530
Quersumme 270
QuickTime 671, 721, 722
QuickTime VR 723

R

R – Y 83
Radio Manufacturers Association 50
Radiositity 697
RAID 608
Randlichtabfall 397
Random Access 670
Rasen 531
Rasterdeckung 379
Raumtiefenwahrnehmung 29
Raumwinkel 19
Rauschen 62, 374
Rauschunterdrückung 543
RAW-Codec 447
RAW-Modus 443
Raytracing 697
RC-Integrationsglied 46
RC-Kombinationen 60
RCP 421
RCTC 566, 660
Read after Write, 582
Read Modify Write 582
Real Time Control Protokoll 295
Real Time Protokoll 295
Real Time Streaming Protocol 295
RealD 517
Recherche 778

Rechteckimpuls 56
Rechts/links-zirkular 517
Reclocking 729
Record 663
Red One 446
RedCode 447
Redundant Array of Independent Disks 608
Redundanzbits 271
Redundanzreduktion 158
Reed-Solomon-Code 272
Reference External 111
Referenzmonitore 485
Reflektorantennen 247
Regelgröße 535
Regelkreise 534
Regenerate 661
Regions of Interest 171
Reineisen 523
Rekonstruktionsfilter 127
Relaisgruppe 398
Remanenz 523
Remission 18, 383
Remote Control Panel 421
Resource Interchange File Format 723
Restseitenband-AM 230, 241
Reuters 780
Revolverkopf 508
Rewritable Consumer Time Code 660
Rezeptive Felder 22
RGB-Matrix 467
RGB-Parade 117
RGB-Signal 81
RGBA 155
Richtantennen 241
Richtfunkanlagen 242
RIFF (Resource Interchange File Format) 723
RiFu 783
Rivalität 29
RLC 162
RM 230
RMA 50
RMS-Körnigkeit 314
Roboterstativ 766, 776
Rohschnitt 675, 682
ROI 171
Roll-off-Faktor 129
Rolling Shutter 370
Röntgenstrahlung 462
Rot 66
Rotlicht 418, 766
Router 749
RS 422-Schnittstelle 666
RSB-AM 230
RST 269
RTCP 295
RTP 295
RTSP 295
Rubberband 679
Rückbildmonitor 783
Rückkanal 291
Rückprojektion 494
Rücktransformation 55
Run Length Coding 162
Rundungsfehler 123
Running Status Table 269
Rushes 681
RZ-Code 579

S

S-Video 90
S/N 65
S/W-kompatibel 82
S1 151
S3D-Systeme 27
Safe Area 418, 646
Sägezahnsignal 38
Sammellinse 384
Sample Multiplexing 731
Samplingrate 121
SAN 672, 750
SAS 606
SAST 777
Sat-ZF 247
SATA 606, 607
Satellitentransponder 245
Satellitenübertragung 243
Sattelzug 787
Saturation 70
SAV 140
SBOF 669
SC 88
SC-H-Phase 102
SCART-Steckverbindung 62
Scattered Pilotes 282
Schaltgespräche 760
Schaltrolle 508
Schärfentiefe 388
Schärfeverlagerung 765
Schattenmaske 465
Schattenwürfe 798
Scheimpflugsche Regel 391
Scheinfenster 29
Schieberegister 272

Schleier 297
Schlitzmaskenröhre 466
Schlussklappe 766
Schmalband-FM 231
Schnitt 652
Schnittliste 654
Schnittmarkierung 667
Schnittseite 669
Schrägprojektion 505
Schrägspuraufzeichnung 530
Schreib/Leseköpfe 605
Schrittkopiermaschine 319
Schrotrauschen 373
Schublade 788
Schulter 309
Schutzintervall 264, 280
Schwarz/Weiß-Film 297
Schwarzabhebung 41
Schwarze Strahler 67
Schwarzschulter 45
Schwärzungskurve 308
Schwarzwert 49
Schwelle 309
Schwenker 763
Schwerpunktregel 73
Scrambled Synchro. Non Return to Zero 584
Scrambling 131, 580
Scrolling 647
Scrolling Color Mode 502
SCSI 606
SD-Card 623
SD-Systeme 36
SDDI 731
SDDS 329
SDHC 623
SDI 143, 730
SDT 269
SDTI 731
SDTI-CP 731
SDTI-Header 731
SDTV 35
SDXC 623
Seamless Splicing 710
SECAM 96
SED 482
Segmented Frames 438
Segmentierte Aufzeichnung 531
Segmentierung 581
Sehvermögen 21
Seitenbänder 228
Sektoren 605, 606
Selbstkonfiguration 733
Selektive Schärfe 389
Sendeablaufmischer 776
Sendeablaufsteuerung 777
Sendeabwicklung 776
Sendekomplex 776
Senkel 681
Sensorempfindlichkeit 372
Sensorgestützte Systeme 802
SEPMAG 322
Séquentiel couleur à mémoire 97
Serial Data Transport Interface 731
Serial Digital Data Interface 731
Serial Digital Interface 143
Serielle Schnittstellen 143, 665
Service Description Table 269
Session Layer 738
Set Top Box 286
SFN 263
Shading 407, 697
Shape-Adaptive DCT 198
Shift Keying 260
Shift/Tilt-Objektiv 391
Shockwave 723
Shuffling 581
Shutter 376
Shutterverfahren 493, 518
Shuttle 668
Shuttle Modus 668
si-Funktion 54
Siberbromid 297
Side by Side-Rig 457
Side Panel 253
Signal Frequency Network 263
Signal to Noise Ratio 314
Signal to Noise, S/N 65
Signalanalyse 53
Signalkontrolle 107
Signalreflexion 65
Signalverzerrungen 60
Silberleinwände 518
Silbersalze 297
Simple Mail Transfer Protokoll 293
Single Side Band 229
Single-Sensor-Kamera 380
SIS 234
SIS-Verfahren 234
SK 260
Skin-Detailing 408
Skineffekt 63
Slice 631
Slip Mode 676
Slomo 792

SLS 380
SLSM 793
Small Computers System Interface 606
Smartcard 289
SMOLED 480
SMPTE 151, 657, 714
SMPTE C 79
SMPTE VC-1 222
SMPTE VC-3 224
SNG-HD-Profile 760
SNG-Operator 786
SNHC 193
SNRZI 131
Society of Motion Picture and Television Engineers 151, 657
Soft Edge 629
Softimage 697
Soll-Größe 535
Sonderkanalbereich 249
Sonnenlicht 67
Sony 9 Pin-Schnittstelle 666
Sony Dynamic Digital Sound 329
Sound in Sync 234
Source Pack 714
Spaltbreite 527
Spalten 32
Spaltenpaare 49
Spaltlänge 526
Spartenprogramme 291
Spatial Offset 380
Speed 298, 313
Speicherröhrenabtaster 344
Speicherspalten 364
Spektraldichte 55
Spektrallinien 68
Spektralwerte 74
Spektralwertkurven 72
Sperrgreifer 431
SPG 726
Sphärische Abberation 385
SpiderCam 455
Spiegelbarren 504
Spiegeldurchmessern 247
Spiegelprojektion 502, 504, 505
Spielhandlung 681
Spirit Datacine 342, 343
Spitzlichter 309
Splice-Point 710
Splicing 710
Splitrechner 496
Spot-Fotometer 20
Spulentürmen 507
Spulenwindungszahl 522
Spurzwischenraum 531
SRW 1 446
SSB 229
SSNRZ 584
Stäbchen 21
STAN 459
Stand by off 669
Standard Definition TV 35
Standard M 52
Standardvideosignal 92
Standbildabtastung 537
Stanze 690
Stanzverfahren 630
Start of Active Video 140
StaticSlot 716
Stative 453
Status A 313
Status M 312
STB 286
STC 212
Steady Shot 417
Steadycam 454, 764
Steenbeck 682
Stereobasis 457
Stereographer 459
Stereoscopic Analyser 459
Stereosendungen 233
Stereoskop 516
Stereoskopie 27
Stereoskopische Aufnahmen 457
Stereowiedergabe 234
Sternpunkt Frankfurt 751
Stoppbad 317
Storage Area Network 750
Strahlrücksprung 43
Strahlströme 462
Strahlteiler 614
Streaming 295, 736
Streifen 254
Streifenmuster 803
Stripe-Pattern 445
Striping 608
Stroboskopeffekt 2
Studiodekorationen 794
Studiokamera 419, 420
Studiopegel 201
Studioproduktion 766
Studiotakt 624, 724
Stufenfaser 251
Stumm 766
Stummfilm 7

Sub Master 675
Subcarrier 88
Subclips 675
Subcode-Bereich 592
Subjektive Beurteilung 706
Subpixel 382
Subträger 265
Subtraktive Farbmischung 71
Sucher 432
Sucherlupe 432
Super 16 304
Super 35 302
Super 8 304
Super Live Slow Motion 793
Super Slow Motion 424
Super Vertical Mode 366
Super-Slow-Motion 554
Super-VHS 563
Superblock 179
Superhet 490
Superscope 303
Supportsystem 457
Surface Cond. Electron Emitter Display 482
Surround 206, 327
SVC 742
Sweet Spot 493
Switchable Barrier 493
Switched Virtual Circuit 742
Switcher 624
Switches 745
SXRD-Projektor 511
SXS 623
Symbol 132, 259
Symboldauer 264, 280
Symmetrische Signalführung 65
Sync-Bytes 578
Sync-Köpfe 552
Synchronboden 48, 539
Synchrondemodulator 468
Synchronisation 724
Synchronizer 726
Synchronous Shutter 370
Synchronsignal 46, 140
Synthetic Natural Hybrid Codierung 193
System Time Clock 212
Systemdynamik 546
Systemkamera 420
Systemmodell 741

T

T-Stop 387
Tageslicht 68
Tageslichtspulen 306
Taktgeber 724
Tally 418, 642, 766
Tapetime 655
Target 356
Task Force for Harmonized Standards 740
TBC 543
TCP 294
TDMA 274
TDT 269
Technicolor-Verfahren 8
Technirama 303
Technische Richtlinien 755
Technischer Vorspann 114, 667
Techniscope 303
Telcom 546
Teleobjektiv 392
Telepräsenz 36
Teleprompter 768
Teletext 237
Televisionssysteme 34
Temperaturstrahler 67
Terrestrische Ausstrahlung 240
TES-OLED 480
Testbesen 50
Textur 695
TFHS 740
TFT-Display 474
THD 549
Theatersystem 515
Thight Shot 395
Thin Film Transistor 474
Thunderbolt 735
THX 329
Tiefenbudget 458
Tiefenscript 459
Tiefenwahrnehmung 27
Tiefpass 60
Tiefpassfilterung 121
Tier 763
Tiff 647
Tilt 453
Time and Data Table 269
Time Division Multiple Access 274
Time to Live 294
Timecode 305, 655
Timeline 676
TimelineSlot 716

Timing Reference Signal, 153
TLS 45, 725, 792
TN 473
Toleranzfelder 113
Tonfilm 7
Tonsignalverarbeitung 490
Tonspurbelegung 758
Tonträger 233
Top Emission Structure 480
Total Harmonic Distortion 549
Totale 395
Totalreflexion 18, 251, 384
TPRF 755
Trabanten 46
Tracking 796
Trägersignal 228
Trägerunterdrückung 229
Transfomation Unit 222
Transformation 687
Transformationscodierung 163
Transition Type 641
Transmission Control Protokol 294
TransMux 214
Transparency Mapping 696
Transparenz 307, 798
Transpondersharing 274
Transpondertechnik 242
Transport-Strom 211
Transportschicht 738
Transversalschrift 323
Trapezkorrektur 505
Trellis-Diagramm 273
Tri Level Sync 45, 725, 792
Triax-Kabel 422
Trickblende 629
Trickkopiermaschine 321
Trickmischer 625
Tricktisch 321
Triggerstufe 109
Trim-Mode 677
Trimediales Arbeiten 779
Tripod 453
Trocknung 317
TRS 140
TTL 294
TU 222
Tuner 490
Tungsten 313
Tunneleffekt 482
TV-Auswertung 304
Twin-LNB 248
Twisted Pairs 745

U

U 85
U und V 85
U-Core 306
U-Loading 532
U-Matic 559
Ü-Wagen 787
Überblendungen 321
Übermodulation 91
Übersteuerungskontrolle 411
Übertragungsprotokoll 295
Übertragungswagen 787
UDF 617
UDP 295
UGC 292
UHDTV 153
UHF-Bereich 240
Ultimatte 635, 797
Ultra High Frequency 240
Ultraschall-Laufzeitglied 468
Ultraschall-Verzögerungsleitung 96
Umfeldleuchtdichte 488
Umicast-Technik 295
UMID 714
Umkehrverfahren 297
Umlaufblende 431
Umschlingungsgrad 532
Umtastung 260
Unbunt-Abgleich 70
Underscan 49
Uniform Ressource Locator 293
Unique Material Identifier 713, 715
Universal Serial Bus 735
Unschärfeblenden 321
Unterabtastung 159
Unterbelichtung 318
Unteres Seitenband 228
Untertitel 237
Up-Konverter 756
Up-Link 245, 274
Upper field 40
Upstream 293
URL 293
USB 228, 735
User Bits 656
User Datagram Protocol 295
User generated Content 292
Usingen 786

V

V 85
V-Synchronpuls 46
Variable Bitrate 162
Variable Blocklänge 732
Variable Längencodierung 161
Variator 397
Varicam 440
Varioobjektive 399
VBR 162
VC-1 222
VC-2 223
VC-3 224
VC-I 712
VCI 743
VCM 279
VCO 232, 541
VCR 524, 567
VCS 808
VDSL 294
Vektorskop 99, 112
Verbreitungsgebiete 103
Verdeckung 29
Verdeckungseffekt 202
Verlaufsfilter 402
Vermittlungsschicht 738
Vernetzte Produktion 778
Verrundung 108
Verschlüsselung 289
Verschlüsselungsdaten 513
Verstärker 406
Verstärkerabstände 250
Vertauschungssatz 56
Verteilwege 239
Vertical Interval Subcarrier 571
Vertical Interval Timecode 659, 660
Vertikalaustastlücke 43, 46, 235
Vertikalsynchronisation 45
Verwackelung 416
Very High Frequency 240
Verzerrungen 60
Vestigal Side Band 230
VGA 156
VHF 240
VHS 561
Video 2000 567
Video 8 565
Video for Windows 671
Video near Demand 291
Video Object Plane 196
Video Objects 195
Video on Demand 291
Video Program System Information 238
Video Rate 684
Video_TS 617
Video-Cassettenrecorder 524
Video-Streaming 292
Video-Wand-Systeme 496
Videoausspiegelung 433
Videodienste 291
Videografik 646
Videogrenzfrequenz 49
Videokamera 404
Videokopf 525
Videomesssystem 107
Videoserver 610
Videosignal 92
Videosignalbearbeitung 624
Videosignalübertragung 59
Videosteckfelder 792
Videostudio 699
Videotext 237
Videowandler 772
Vidikon 356
Vier-Röhren-Kameras 379
Vierkopf-Recorder 562
Viertelzeilen-Offset 100
Vierzeilen-Kammfilter 106
Vignettierung 397
Virtual Camera Switch 808
Virtual Channel Identifier 743
Virtual Path Identifier 743
Virtual Reality Modeling Language 195
Virtual Studio Manager 792
Virtuelle Bilder 385
Virtuelle Kamera 796
virtuelle Primärvalenzen 74
Virtuelles Studio 794
VISC 571
Vistavision 303
Visuelle Kontinuität 393
Visueller Transparenz 706
VITC 659
VLC 161
VO 195
Voice Over 679, 780
Vollbildmodus 178
Vollduplex 734
Vollformat-Senso 357
Voltage Controlled Oscillator 232, 541
Volume Unit 549
Volumendisplay 492
Voodoo 600

VOP 196
Vor- und Nachtrabanten 46
Vorcodiert 664
Vorfilterung 32
Vormagnetisierung 545
Vorverarbeitung 159
VPI 743
VPN 778
VPS 238
VRML 195
VSB 230
VSF 792
VSM 792
VTR 520
VU-Meter 549

W

W-VHS 564
Wässerung 297
WAV 671
WAVE 723
Waveformmonitor 109
Wavelet-Transformation 170
Webbrowser 293
Website 293
Wehneltzylinder 461
Weichzeichner 403
Weißabgleich 414
Weißsche Bezirke 522
Weißwert 48
Weit 395
Weitwinkelobjektiv 392
Wellenwiderstand 63
Wide-Screen-Signalling-Signal 484
Wiedergabeelektronik 542
Wiedergabevorgang 528
Windows Media 9 723
Windows NT 673
Winkelcodierer 802
Winkelgenauigkeit 802
WM9 223
WMA 723
WMV 723
Wobbled Groove 619
World Wide Web 293
WRB-Kamera 379
WWW 293

X

XAVC 221, 713
XDCAM 621
XDCAM EX 623
XDCAM HD 621
XDCAM HD422 622
xDSL 294
XKT 788
XLR-Stecker 549
XLR-Steckverbindungen 65
Xpand 519
XY-Betrieb 109
xyY 76

Y

Y 83
Y/C-Signal 90
Y/C-Trennung 106
Yagi 241
YC 443 558
YC 623 558
YRGB 379

Z

Z-Screen 517
Zackenschrift 323
Zähldiskriminator 232, 542
Zapfen 21
ZDF 13, 242
Zebra 411
Zeichnung 115, 411
Zeile 32
Zeilenendstufe 469
Zeilenoszillator 469
Zeilensprungartefakte 39
Zeilensprungverfahren 37
Zeilentransformator 469
Zeilenzahl 33
Zeilenzählung 47
Zeiss 396
Zeitauflösung 25
Zeitdiskretisierung 120
Zeitfehler 543
Zeitkonstante 60
Zeitlupe 433, 537

Zeitlupenwiedergabe 792
Zeitmultiplexverfahren 227
Zeitraffer 433, 537
Zeitreferenz 140
Zentraler Geräteraum 788
Zentralmatrix 778
Zentralstrahl 384
Zero Parallax Setting 28
ZF-Durchlaßkurve 490
ZGR 789
Zieladresse 294
Zirkulare Polarisation 245
Zoetrop 3
Zoned CLV 619
Zoomfahrt 394
Zoomobjektiv 397
Zoopraxiskop 5
ZPS 28
Zugriffszeit 606
Zündspannung 478
Zuspielung 242
Zwei-Maschinen-Schnittplatz 774
Zweikanalbetrieb 233
Zweiten Deutschen Fernsehen 241
Zwischenfrequenz 490
Zwischenkopien 307
Zylinder 605