



Dieter Stotz

Computergestützte Audio- und Videotechnik

Multimediatechnik in der Anwendung

3. Auflage



Springer Vieweg

Computergestützte Audio- und Videotechnik

Dieter Stotz

Computergestützte Audio- und Videotechnik

Multimediatechnik in der Anwendung

3. Auflage



Springer Vieweg

Dieter Stotz
Babenhausen, Deutschland

ISBN 978-3-662-58872-7 ISBN 978-3-662-58873-4 (eBook)
<https://doi.org/10.1007/978-3-662-58873-4>

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 GmbH Deutschland, ein Teil von Springer Nature 1995, 2011, 2019

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 allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag, noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Springer Vieweg ist ein Imprint der eingetragenen Gesellschaft Springer-Verlag GmbH, DE und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Heidelberger Platz 3, 14197 Berlin, Germany

Meiner Frau Karin Smiley gewidmet

Vorwort

Vorwort zur ersten Auflage

In der unglaublich schnell voranschreitenden Computertechnologie sind Neulinge und Profis fast gleichermaßen gefordert, sich nicht nur über den Stand der Technik, sondern auch über deren Zusammenhänge zu informieren. Da einschlägige Zeitschriftenartikel zumeist unzusammenhängend, unvollständig, oberflächlich und manchmal sogar sachlich falsch gehalten sind, liegt es nahe, ein Buch über die Thematik zu verfassen.

Die Handbücher von Hard- und Software-Herstellern für die bild- und tonverarbeitende Computertechnik sind oftmals so knapp gehalten, dass sie nur spezifische Handgriffe erläutern. Der Anwender hingegen ist sehr stark angewiesen auf die Beleuchtung der Hintergründe und das technische Zusammenspiel der Komponenten.

Zur Verwirklichung künstlerischer Werke sind mit den heutigen Voraussetzungen immer weniger technischer Arbeitsschritte zwischen künstlerischem Schaffen und endgültigem Werk notwendig; deshalb werden auch Künstler immer mehr direkt mit der Computertechnik konfrontiert – ein Umstand, der ein gewisses Maß an technischem Know-how erfordert.

Ziel des Buches ist daher, einer möglichst breiten Schar von Anwendern der computergestützten Bild- und Tonverarbeitung eine Hilfe zu sein, die sich auf der Vermittlung von Grundlagen stützt und die sich am praxisnahen Geschehen orientiert. Dilettantisches Vorgehen bei der Qualitätsverbesserung des Endprodukts bzw. bei der Vermeidung von zeitraubenden Fehlern soll ersetzt werden durch Sachverstand und Kompetenz, der sich bei der kreativen und produktiven Tätigkeit in Zeitsparnis und Qualitätssteigerung äußert.

Das Buch besteht aus zwei Teilen, dem Audio- und dem Videoteil. Diese wiederum gliedern sich in mehrere Kapitel, die sich den wichtigsten und brennendsten Fragen widmen. Jedem Teil ist ein Kapitel vorangestellt, das zunächst allgemeine Grundlagen zum Verständnis der Audio- bzw. Videotechnik wiedergibt. Dies ist für technisch Interessierte ohne Vorkenntnisse sehr hilfreich. Die einzelnen Kapitel beschreiben einerseits Verfahrensweisen und andererseits grundsätzliche technische Vorgänge bzw. Festlegungen. Der umfangreiche Anhang dient als Nachschlageteil für oft benötigte technische Informationen, die im Hauptteil keinen Platz fanden oder schwerer zu finden sind.

Es wurde weitgehend darauf verzichtet, spezifische Hardware oder gar Software einzubeziehen, da diese laufend erneuert bzw. ersetzt wird. Stattdessen erfolgt eine Konzentration auf bewährte Verfahren bzw. Technologien, welche in absehbarer Zukunft beibehalten werden – wenn auch unter wechselnder Hardware bzw. Software.

Bei der Darlegung von komplexer und zum Teil etwas trockener Themen findet oft eine Auflockerung durch Praxisbezüge statt, die in eingerückter Sonderschrift erscheinen. Zum einen soll damit die Thematik etwas transparenter werden – außerdem spornt der beabsichtigte Aha-Effekt zum Weiterlesen an.

Das Literaturverzeichnis am Ende des Buches korrespondiert mit Stellen im Text, die am Ende von Absätzen durch [nn] bezeichnet sind. Dabei handelt es sich teils um Basisliteratur und teils um weiterführende Werke (da ja manche Details hier nur ein wenig tangiert werden).

Das Buch entstand unter Aufbringung großer Sorgfalt. Trotzdem sind natürlich Fehler oder Unklarheiten nicht ganz auszuschließen. Für die Aufdeckung von Sach- und Rechtschreibfehlern bzw. für die Darlegung von Verbesserungen sind Verlag und Autor dankbar.

Bei allen Firmen, die bei der Erstellung des Manuskriptes hilfreich zur Seite standen, möchte ich mich ganz herzlich bedanken. Besonders erwähnenswert sind dies (in alphabatischer Reihenfolge):

Microsoft GmbH, Matrox GmbH, Q-Team Dr. Knabe Korschenbroich, Roland Elektronische Musikinstrumente Handelsgesellschaft mbH, Turtle Beach Systems, Yamaha Europa GmbH.

Ferner gilt auch besonderer Dank dem Springer-Verlag, der hiermit wieder einmal dem Anspruch gerecht wird, professionell Bücher für Wissenschaft und Technik herzustellen.

München, Herbst 1995

Dieter Stotz

Vorwort zur zweiten, verbesserten und überarbeiteten Auflage

Nach nunmehr bereits fünfzehn Jahren sind mit der neuen Auflage nicht nur Korrekturen der (und die Anpassung an die neue) Rechtschreibung fällig, sondern die technischen Neuerungen sind so drastisch, dass viele Ergänzungen notwendig wurden. So sind zum Beispiel Themen wie Raumklang (Surround), Kompression (MP3) und DSP-Echtzeitbearbeitung im Teil Audio hinzu gekommen. Der modernen Audiomesstechnik kommt ein komplettes Kapitel zu, während auf das Kapitel CD-ROM-Laufwerke verzichtet worden ist – dieses Thema hat heute keine spezifische Assoziation Computer/Audio mehr.

In der Videotechnik sind obsolete Aufzeichnungsverfahren dem Rotstift zum Opfer gefallen. Dagegen sind aktuellen Themen wie hochauflösendes Video und 3D-Technik besonderes Augenmerk verliehen worden.

Die Aufteilung in zwei Teile Audio- und Videotechnik erschien mir nach wie vor sinnvoll, obwohl natürlich manche Bereiche durchaus beides berühren.

Der technische Aspekt steht im Vordergrund, und viele Erläuterungen sollen helfen, die technischen Bedingungen zu optimieren und Fehler am Endprodukt zu vermeiden. Deshalb ist das Buch kein bloßes Grundlagenwerk, sondern es ist durchaus an der Praxis orientiert. Im Anschluss an viele Kapitel folgen daher Beispiele aus der praktischen Arbeit.

Vieles bei der Digitaltechnik wird anhand (durchaus gründlicher) anderer Literatur über Signaltheorie beschrieben. Meist fehlt dem jedoch ein intuitiver Impuls – diese Lücke zu schließen ist ebenfalls ein Anliegen dieses Buches.

Obwohl der Fokus auf professionelle Technik ausgerichtet ist, bereichert der Buchinhalt durchaus Semiprofessionelle oder auch Leute, die hobbymäßig Audio und Video verarbeiten wollen – oder ebenso solche, die einfach ihr technisches Verständnis erweitern wollen. Das Angebot an Hardware und Software ist glücklicherweise so reichhaltig, dass auch im Heimbereich hochwertige Arbeiten mit kleinem Geld entstehen können. Hobbymusiker konnten sich beispielsweise vor dreißig Jahren kaum ein Mehrspur-Aufnahmegerät leisten, heute ist eine einfache Software zur Digitalisierung für viele Spuren sogar kostenlos erhältlich.

Besonders erwähnenswert ist die Tatsache, dass allem Verständnis der Computer-Audioteknik und -Videoteknik die Kenntnis der Grundlagen voransteht. Deshalb sind in diesem Buch auch sehr viele – teilweise möglicherweise als weitschweifend anmutende – Grundlagen zu beiden Bereichen erörtert, die nicht immer modernste Technologien widerspiegeln.

Mein Dank gilt nachstehenden Personen und Firmen (in alphabetischer Reihenfolge), die mir mit ihren Bereitstellungen große Dienste geleistet haben. Natürlich habe ich auch nicht zuletzt dem Springer-Verlag zu danken, denn ohne ihn wäre es kaum möglich, solch qualitative Werke der Wissenschaft und Technik herzustellen.

- Guidorzi, Paolo – Software *Sample Champion PRO*
- Heidenreich, Florian – Software *Mp3tag*
- Hepp, Dr. Ulrich – Software *Magix Video deluxe*
- Knoll, Hans M. – Bildquelle Klirrfaktormessbrücke
- Mazzoni, Dominic, Andrews, Gale – Software *Audacity®*
- Montgomery, Chris, Xiph.Org Foundation – Vorbis-Format und -Verfahren
- Müller, Ulrich – Software *audioTester 3.0*

Babenhausen (Schwaben), Mai 2011

Dieter Stotz

Vorwort zur dritten, verbesserten und erweiterten Auflage

Nach weiteren Jahren – nun schon im 24. Jahr nach Entstehen der Erstauflage – ist die multimediale Entwicklung immer noch nicht zum Stillstand gekommen. Besonders erwähnenswert sind Neuerungen in der hochauflösenden Wiedergabe, der virtuellen

Realität und der 3D-Darstellung von Bildern und Videos. Hier ist noch längst kein Ende abzusehen.

Das Kapitel *Abtastung und Digitalisierung* wurde einer gründlichen Neustrukturierung unterzogen, da beide Vorgänge zum besseren Verständnis deutlicher voneinander abzugrenzen sind. Obwohl hinter dem Ganzen eine Menge Signaltheorie steckt, war beabsichtigt, neben technischer Exaktheit zugunsten einem intuitiven Verständnis auf die oftmals etwas komplex erscheinende Mathematik zu verzichten oder sie wenigstens auf das nötige Maß zu beschränken.

An dieser Stelle ist zu erwähnen, dass alte und teilweise auch ausgediente Technik in dieser Buchfassung nicht herausgefallen ist. Der Grund liegt darin, dass nicht nur aktuelle Technologie-Stände und die Grundlagen dazu von Interesse sind, sondern auch die Entwicklung dorthin. Um dem Buch einen wissenschaftlichen Wert zu verleihen, ist es nach Auffassung des Autors unerlässlich, evolutionäre Technologie darzustellen und das Prinzipverständnis zu vermitteln.

Dem Springer-Verlag habe ich zu danken, namentlich Frau Bromby, für die Lenkung des Organisatorischen und die hochqualitative Aufbereitung des Buches.

Babenhausen (Schwaben), Dezember 2018

Dieter Stotz

Inhaltsverzeichnis

Teil I Audiotechnik

1	Grundlagen zur Audiotechnik	3
1.1	Pegelmaße in der Audiotechnik	3
1.2	Lineare Verzerrungen und Frequenzgang	4
1.3	Nichtlineare Verzerrungen	6
1.3.1	Klirrfaktor	7
1.3.2	Intermodulation	8
1.4	Störabstand	11
1.4.1	Geräuschspannungsabstand	11
1.4.2	Fremdspannungsabstand	12
1.5	Dynamikbereich	12
1.6	Übersprechdämpfung	13
1.7	Phasenverzerrungen	14
1.8	Raumakustik	16
1.8.1	Schallreflexion	17
1.8.2	Diffraction	18
1.8.3	Refraktion (Brechung)	19
1.8.4	Doppler-Effekt	20
1.9	Räumliches Hören	22
1.9.1	Intensitäts-Stereofonie	22
1.9.2	Stereofonie durch Laufzeit	24
1.9.3	Kunstkopf-Stereofonie	24
1.9.4	Quadrofonie	25
1.9.5	Dolby Surround-Sound 5.1 (Dolby Digital)	25
1.9.6	Dolby Pro Logic	27
1.9.7	Dolby Pro Logic II	28
1.9.8	DTS	30
1.9.9	Wellenfeld-Synthese	30
1.9.10	Weitere Raumklang-Systeme – Raumklang mit nur zwei Lautsprechern	31

1.10	Speicherung von Audiosignalen	32
1.10.1	Analoge Magnetbandaufzeichnung	33
1.10.2	Digitale Magnetbandaufzeichnung	34
1.10.3	Digitale Aufzeichnung auf MiniDisc (MD)	35
1.10.4	Hard-Disk-Recording	36
1.10.5	Compact Disc (CD)	36
1.10.6	Unbewegte Speichermedien	40
1.10.7	Archivierung und Datenbanken	42
	Literatur	42
2	Abtastung und Digitalisierung	43
2.1	Einleitung	43
2.2	Abtastung	44
2.2.1	Notwendigkeit und technische Ausführung	44
2.2.2	Abtastspektren	45
2.2.3	Aliasing	48
2.2.4	Überabtastung	51
2.3	Quantisierung – Wandlung in Digitalwerte	54
2.3.1	Digitalformate nach der Wandlung	55
2.3.2	Methoden der A/D-Wandlung im Überblick	58
2.3.3	Quantisierungsrauschen und Linearitätsfehler	62
2.4	Verfahren zur Datenreduktion	66
2.4.1	Datenreduzierung durch Parameteränderung	67
2.4.2	Datenreduzierung durch Kompression	67
2.4.3	MP3-Verfahren	70
2.4.4	Sonstige Verfahren	75
2.5	Rückwandlung und Wiedergabe	78
2.5.1	D/A-Wandlungsarten	78
2.5.2	D/A-Quantisierungsrauschen	80
2.5.3	Überabtastung auf der D/A-Seite	80
2.6	Signalaufbereitung auf der Analogseite	80
2.6.1	Aussteuerung	80
2.6.2	Mikrofonwahl und -aufstellung	81
2.6.3	Echtzeitbeeinflussung des Signals	88
	Literatur	97
3	Bearbeitung von Sampling-Dateien	99
3.1	Schneidevorgänge	99
3.1.1	Bearbeitungsschritte	99
3.1.2	Nahtstellen nach dem Schneidevorgang	102

3.2	Filter	103
3.2.1	FIR-Filter	104
3.2.2	IIR-Filter	107
3.3	Effekte	109
3.3.1	Echo	109
3.3.2	Hall	112
3.3.3	Phasing und Flanging	114
3.3.4	Verzerrung	115
3.3.5	Veränderung der Abspielgeschwindigkeit	118
3.3.6	Veränderung der Tonlage (Pitch)	120
3.3.7	Zeitdehnung	122
3.4	Pegelbearbeitung	122
3.4.1	Einstellung von Einzelbereichen	122
3.4.2	Hüllkurvenkontrolle	123
3.4.3	Dynamikkompression	123
3.5	Unterdrückung von Knackgeräuschen	125
3.5.1	Slewrate-Filterung	125
3.5.2	Einzelwert-Bearbeitung	126
3.6	Echtzeitbearbeitung per DSP	127
3.6.1	Grundlagen zum DSP	127
3.6.2	Bearbeitungs-Mechanismen	128
3.7	Praxisbeispiel einer Audio-Bearbeitung	129
3.7.1	Inversecho	130
3.7.2	Erzeugung von Dolby Pro Logic	131
3.7.3	Hüllkurven-Bearbeitung	131
	Literatur	132
4	MIDI	133
4.1	MIDI-Hardware	133
4.1.1	Anschlüsse bei MIDI-Instrumenten	133
4.1.2	Verschaltung von MIDI-Instrumenten und -Geräten	134
4.1.3	Ausführung der Schnittstelle	134
4.2	MIDI-Software-Standard	136
4.2.1	Begriffserläuterungen	136
4.2.2	Überblick zu den MIDI-Messages	139
4.2.3	Channel Message	139
4.2.4	System Exclusive	147
4.2.5	System Common	149
4.2.6	System Realtime	151
4.3	General MIDI	151
4.3.1	Festgelegte Programm-Nummern	152
4.3.2	Festlegungen für Schlaginstrumente	152

4.3.3	Weitere Festlegungen	154
4.3.4	Richtlinien für General MIDI Level 2	155
4.3.5	Das GS-Format	155
4.4	MIDI-Mapping	157
4.4.1	MIDI-Setup	157
4.4.2	MIDI-Patchmaps	158
4.4.3	MIDI-Keymaps	159
4.5	Verschiedene Prinzipien der Klangerzeugung	159
4.5.1	Analoge Klangsynthese	159
4.5.2	FM-Klangsynthese	162
4.5.3	Waveshaping	164
4.5.4	Fourier-Synthese	164
4.5.5	Lineare Arithmetik	165
4.5.6	Sampling-Synthese	165
4.6	Synchronisation von MIDI-Geräten	168
4.6.1	Zweck einer Synchronisation	168
4.6.2	Realisierung über Timecode	169
4.6.3	MIDI-Spezifikation für den Timecode	170
4.7	MIDI-Geräte, Audiokarten und Zubehör	172
4.7.1	Ansteuerung externer MIDI-Geräte durch den Computer	172
4.7.2	Audiokarten und Betrieb unter MIDI	173
4.7.3	MIDI-Steuergeräte	174
4.7.4	MIDI-Software	175
4.7.5	MIDI-Softwareinterface	177
4.8	MIDI-Applikationen und Problemlösungen	177
4.8.1	Nachvertonung von Videosequenzen	177
4.8.2	Synchronisation mehrerer MIDI-Applikationen	179
4.8.3	MIDI als Playback-Quelle für Musiker	180
4.8.4	MIDI-Arrangement und Sampling-Dateien	183
4.9	MIDI-Praxisbeispiel	185
	Literatur	187
5	Audio-Messtechnik	189
5.1	Herkömmliche analoge Messtechnik	189
5.1.1	Frequenzgang	189
5.1.2	Verzerrungen	189
5.1.3	Pegelvergleichsmessungen	191
5.1.4	Gleichlauf	192
5.1.5	Impuls-Messungen	193
5.1.6	Korrelationsgrad	193
5.2	Digitale Messtechnik und PC-Systeme	194
5.2.1	Voraussetzungen für das Gelingen der Messungen	195
5.2.2	Frequenzgang nach digitalem Verfahren	195

5.2.3	Verzerrungen nach digitalem Verfahren	195
5.2.4	Pegelvergleichsmessungen nach digitalem Verfahren	197
5.2.5	Korrelations-Messungen nach digitalem Verfahren	198
5.2.6	Gleichlauf nach digitalem Verfahren	199
5.2.7	Impuls-Messungen nach digitalem Verfahren	199
5.2.8	Raumakustik-Messungen nach digitalem Verfahren	200
5.2.9	Kompaktes Messlabor Red Pitaya	201
	Literatur	204
6	Vergleich mit herkömmlicher Audiotechnik	205
6.1	Schneidetechniken	205
6.1.1	Aufzeichnungen auf Analogband	205
6.1.2	Aufzeichnungen auf Computer-Festplatte	206
6.1.3	Digitale Bandaufzeichnungen	206
6.2	Filtertechniken	207
6.2.1	Filtern mit Spezialgeräten	207
6.2.2	Software-Filterung	208
6.2.3	Digitale Hardware-Filterung	208
6.3	Einsatz von Effekten	209
6.4	Sonstige Signalmanipulationen	209
6.4.1	Rechenintensive Veränderungen	209
6.4.2	Diskrete Signalverformung	210
6.4.3	Zeitinversion	210
6.5	Wartung der Systeme	211
6.5.1	Wartung bei Analogsystemen	211
6.5.2	Wartung bei Digitalsystemen	212
6.6	Latenz-Zeit	212
6.6.1	Monitoring-Latency-Zeit	212
6.6.2	MIDI-Latency	213
6.6.3	Latency-Jitter	213
6.6.4	Durchlauf-Latency	214
6.7	Gegenüberstellung Analogtechnik/Digitaltechnik	215
	Literatur	216

Teil II Videotechnik

7	Grundlagen und Verfahren der Videotechnik	219
7.1	Bildabtastung und -wiedergabe	219
7.1.1	Zeilensprungverfahren (Interlaced)	220
7.1.2	Non-Interlaced-Verfahren	222
7.1.3	Synchronisation	223
7.1.4	BAS-Signal	224
7.1.5	Probleme und Fehlersymptome	226

7.2	Farbübertragung	227
7.2.1	Physikalische und physiologische Aspekte	228
7.2.2	Prinzip der Farbbildaufnahme	231
7.2.3	Prinzip der Farbbildwiedergabe	234
7.2.4	RGB-Signal	240
7.2.5	YUV-Signal	242
7.2.6	Y/C-Signal	244
7.2.7	FBAS-Signal	249
7.2.8	Viertelzeilenoffset-Verfahren	251
7.2.9	Digitales Signal	252
7.2.10	Probleme bei der Farbübertragung	253
7.3	Aufzeichnungsverfahren	254
7.3.1	VHS-Verfahren	254
7.3.2	S-VHS-Verfahren	256
7.3.3	U-matic	257
7.3.4	Betacam	258
7.3.5	Betamax	259
7.3.6	Video 2000	259
7.3.7	Video Hi-8	259
7.3.8	1-Zoll-MAZ	260
7.3.9	Laserdisc	260
7.3.10	Digitale Aufzeichnungsverfahren	261
7.4	Weitere Farbübertragungsverfahren	262
7.4.1	NTSC	262
7.4.2	SECAM	264
7.5	Hochauflösende Standards	264
7.5.1	PALplus und dessen Verbesserungen	264
7.5.2	HDTV-Historie und nach EUREKA 95	266
7.5.3	Moderne HDTV-Technologie	266
7.5.4	UHDTV und 4K-Standard	267
7.6	3D-Verfahren und Standards	268
7.6.1	Stereoskopie	268
7.6.2	3D-Technik durch Monochromfilter-Brille	269
7.6.3	3D-Technik mit Shutter-Brille	271
7.6.4	3D-Technik mit Polarisationsbrille	271
7.6.5	Weitere Verfahren für 3D-Wiedergabe	274
7.6.6	Reales 3D-Sehen – Stand der Technik	275
7.6.7	Virtual Reality – virtuelle Realität	277
7.7	Speichermedien	280
7.7.1	DVD – Digital Versatile Disc	280
7.7.2	Blu-ray	281
	Literatur	281

8	Parameter beim digitalen Videobild	283
8.1	Farbcodierung	283
8.1.1	Farbtiefe	283
8.1.2	Verschlüsselung der Farbdaten	284
8.1.3	Datenanordnung der Farbinformation	285
8.2	Bildformat und Auflösung	287
8.2.1	Bildseitenformat	287
8.2.2	Auflösung	287
8.2.3	Overscan	288
8.2.4	Underscan	289
8.3	Aspect Ratio	290
8.4	Bildabtastung und Pixelraten	292
	Literatur	293
9	Videoausgabe einer Computergrafik	295
9.1	Anlegen von neuen Grafiken	295
9.1.1	Auflösung	295
9.1.2	Farbtiefe	297
9.1.3	Farbpegel	298
9.1.4	Grafik-Datenformat	299
9.2	Veränderungen an bestehenden Grafiken	300
9.2.1	Konvertierung der Auflösung	300
9.2.2	Änderung der Farbtiefe	303
9.2.3	Farbpegelanpassung	304
9.2.4	Export in andere Grafik-Datenformate	306
9.3	Wiedergabeeigenschaften von Grafikkarten	307
9.3.1	Speicherorganisation	307
9.3.2	Signalarten der Ausgänge	308
9.3.3	LUT-Manipulation	310
9.3.4	Durchschleifen von Signalen anderer Grafikkarten	311
9.3.5	Wiedergabe der Benutzeroberfläche	312
9.4	Genlock-Wiedergabebetrieb	313
9.4.1	Synchronisation des Pixeltaktes	313
9.4.2	H-Phasen-Einstellung	314
9.4.3	SC-Phasen-Einstellung	315
9.5	Sichtbare Störungen	317
9.5.1	Systembedingte Störungseffekte	317
9.5.2	Bildlagefehler	319
9.5.3	Synchronisationsfehler	320
9.6	Grafiken als Testbilder	321
9.6.1	Prüfung der Bildgeometrie und Aspect Ratio	321
9.6.2	Prüfung der Farbwiedergabe	322

9.6.3	Prüfung der Helligkeitsauflösung	323
Literatur		324
10	Einfrieren von Videobildern	325
10.1	Besonderheiten der Signalaufbereitung beim Live-Modus	325
10.1.1	Timebase-Corrector	325
10.1.2	Decodierung	327
10.2	Anzeige des Live-Bildes	328
10.2.1	Vollbildanzeige	328
10.2.2	Window-Anzeige	329
10.2.3	Durchschleifen des Live-Bildes	331
10.3	Vorgänge beim Einfrieren	331
10.4	Signalfluss	331
10.4.1	Einfrierzeitpunkt	333
10.4.2	Unterschiede zwischen Live- und Capture-Bild	333
Literatur		335
11	Spezialeffekte	337
11.1	Overlay-Technik	337
11.1.1	Overlay bei einer Farbtiefe von 16 Bit	337
11.1.2	Overlay bei einer Farbtiefe von 32 Bit	339
11.1.3	Kriterien für das Keying	340
11.2	Chromakeying	341
11.3	LUT-Effekte	342
11.4	Mischen und Überblenden	343
11.4.1	Homogene Blenden	344
11.4.2	Strukturblenden	344
11.5	Filtertechniken	345
11.5.1	Punktoperationen	345
11.5.2	Ortsbereichsoperationen	346
11.6	Hardware-Zoom und -Panning	347
11.7	Hardware-Flip	348
11.8	Bild-im-Bild beim Fernsehen	348
11.9	Videotext	349
Literatur		350
12	Animationen	351
12.1	Bewegliche 2D-Modelle	351
12.2	Animation über Drahtgittermodelle	352
12.2.1	Konstruktion eines Drahtgittermodells	352
12.2.2	Anfangs- und Endpunkt der Bewegung	353
12.2.3	Rendering	354

12.2.4 Raytracing	355
12.3 Morphing	355
12.4 Aufzeichnungstechniken	356
12.4.1 Echtzeit-Animation	356
12.4.2 Einzelbild-Aufzeichnungen	357
12.5 Rotoscoping	361
12.5.1 Effektbeschreibung	361
12.5.2 Arbeitsweise	361
12.6 Schrift-Animation	363
12.7 Kritische Parameter bei Echtzeit-Animationen	364
12.7.1 Anforderungen an das System	364
12.7.2 Auflösung	365
12.7.3 Farbtiefe	365
12.7.4 Bildinhaltsveränderungen	366
12.8 Problemlösungen zur Datenreduktion	366
12.8.1 JPEG-Verfahren	366
12.8.2 MPEG-Verfahren	370
12.8.3 QPEG-Verfahren	371
12.9 Neue Animationstechniken	372
12.9.1 Motion Capture	372
12.9.2 Performance Capture	372
Literatur	372
13 Schnittsysteme	373
13.1 Online-Schnittsysteme	373
13.1.1 Arbeitsweise	373
13.1.2 Endgültiger Schnitt	374
13.1.3 Nachteile des Verfahrens	375
13.1.4 Besondere Schnittverfahren	375
13.2 Offline-Schnittsysteme	376
13.2.1 Digitalisieren des Videomaterials	377
13.2.2 Durchführung des Offline-Schnitts	378
13.2.3 Rohschnittfassung	378
13.2.4 Schnittliste und endgültiger Schnitt	379
13.3 Audionachbearbeitung	379
13.3.1 Asynchrone Mischung	379
13.3.2 Synchrone Mischung	380
13.4 Praxisbeispiel mit einfachem Schnittsystem	381
14 Spezielle Zusatzgeräte	385
14.1 Signalarten-Wandler	385
14.1.1 Wandlung von RGB nach Y/C bzw. FBAS	385

14.1.2 Wandlung von FBAS bzw. Y/C nach RGB	387
14.1.3 Komponenten-Transcoder	388
14.1.4 Spezielle Sync-Einkopplungen	389
14.2 Normenwandler	390
14.3 Timebase-Corrector	392
14.4 Verzögerungsglieder	393
14.5 Synchrongenerator	395
Literatur	395
Anhang A: Audio-Datenformate	397
Anhang B: Tabellen und Diagramme	407
Anhang C: Beschaltungen	413
Anhang D: Grafikformate	419
Glossar	433
Stichwortverzeichnis	445

Teil I

Audioteknik

Als in den 70er-Jahren die digitale Aufzeichnung eines Tonsignals ermöglicht wurde, dachte noch niemand an die Computer-Anbindung. Selbst noch in den Kinderschuhen, beschleunigte die Computertechnik jedoch schnell ihre Entwicklung, sodass die digitale Weiterverarbeitung am PC ein erstrebenswertes Ziel wurde.

Der Vorteil einer solchen Medienverarbeitung lag und liegt auf der Hand: Editierfunktionen sind schnell und reproduzierbar zu erledigen, auch Rückwärtsschritte (Undo-Funktion) sind fast unbegrenzt möglich, der visuelle Schnitt ist gegenüber dem realen Bandschnitt wesentlich genauer, sicherer und schneller zu bewältigen. Die Werkzeuge wie Filter, Effekte und Geräuschgeneratoren sind komplett auf Software-Ebene realisierbar – ein Gerätepark ist somit entbehrlich.

Während die Bearbeitung anfangs praktisch nur offline möglich war, sind moderne Systeme in der Lage, auch ohne DSP (Digital Signal Processor) in Echtzeit Effekte und Geräusche zu generieren.

Neben Grundlagen der Audioteknik kommen auch Verfahren zur Sprache, beispielsweise *Datenkompression*. Es interessieren dabei nicht nur prinzipielle Arbeitsweisen, sondern auch die Faktoren der Parametereinstellung. Gerade bei dem sehr populär gewordenen MP3-Format ist es relevant, welche Einstellungen zu welchem Datenaufkommen führen und welche Qualität bieten. Ebenfalls ist *Raumklang* ein vielschichtiges Thema, und die Kenntnis über die Verfahren und Richtlinien werden hilfreich dabei sein, ein optimales Resultat zu erhalten.

Ein großes Kapitel beschäftigt sich mit MIDI, einem Datenstandard, der neben Kommunikation elektronischer Instrumente auch Steuerungsaufgaben von Audiogeräten ermöglicht. Der Aufbau des Protokolls kommt ebenso zur Sprache wie Hardware-Voraussetzungen. Auch Störeffekte wie Latenz und deren Vermeidung ist Inhalt dieses Abschnitts.

Da allgemeine Messaufgaben schon lange Einzug gehalten haben beim PC, ist es nur konsequent, auch für spezielle Audiomessungen Möglichkeiten zu schaffen. Es geht darum, die Qualität eines Systems zu bewerten, die Raumakustik zu analysieren,

Lautsprecher und Mikrofone auszumessen oder in kriminaltechnischen Untersuchungen Stimmen und Geräusche zu identifizieren. Die praktische Durchführung von Messungen wird anhand einfacher Beispiele beschrieben.

Abschließend gibt das Buch eine übersichtliche Gegenüberstellung zu herkömmlicher Audiotechnik wieder. Dieses kurze Kapitel dient lediglich als Quintessenz.



Grundlagen zur Audiotechnik

1

1.1 Pegelmaße in der Audiotechnik

Es gibt in der Audiotechnik eine wichtige Festlegung, die alle Pegel-, Verstärkungs- und *Dämpfungsmaße* betrifft. Hiernach werden meist keine Signalspannungswerte bzw. Dämpfungs- oder Verstärkungsfaktoren angegeben, sondern ein logarithmisches Maß hiervon. Der Grund liegt hauptsächlich in der Physiologie des menschlichen Gehörs. S. 42 [2,4].

Die Angabe eines *Spannungspegels* bezieht sich auf diejenige Effektivspannung U_0 , welche an einem Widerstand R von 600Ω die Leistung $P = 1 \text{ mW}$ umsetzt. Die Spannung ergibt sich dann zu:

$$U_0 = \sqrt{P \cdot R} \approx 0,775 \text{ V} \text{ (Effektive Bezugsspannung)} \quad (1.1)$$

Alle anzugebenden Spannungen bezieht man auf diesen Wert und erhält durch Logarithmierung den Spannungspegel:

$$a_u = 20 \cdot \log \frac{U}{U_0} \text{ [dBu]} \quad (1.2)$$

Ist also $U = U_0$, so handelt es sich um den *Bezugspegel* 0 dBu. Das u steht für *unloaded*, also unbelastete Situation. Das heißt, die Messpunkte sind keineswegs mit einem Widerstand von 600Ω abgeschlossen. Auf diese Weise wird die Angabe eines absoluten Spannungspegels möglich. Die Rückrechnung von Pegel auf Spannung erfolgt nach *De-logarithmierung* und Umstellung obiger Gleichung:

$$U = U_0 \cdot 10^{a_u/20} \text{ [V]} \quad (1.3)$$

Ebenfalls sind Pegelmaße in dBV anzutreffen. Bei diesem Maß wurde als Bezugsspannung 1 V verwendet.

Was in der Audiotechnik jedoch wesentlich häufiger vorkommt als ein Absolutpegel, sind relative Pegelmaße, die Abweichungen, Verstärkungen und Dämpfungen ausdrücken. Dabei kommt kein fester Bezugswert zur Anwendung, sondern man erhält situationsbedingt das Verhältnis zweier Spannungen. Soll beispielsweise das Verstärkungsmaß a_v eines Systems angegeben werden, so genügt die Kenntnis der Eingangsspannung U_e und die dazugehörende Ausgangsspannung U_a :

$$a_v = 20 \cdot \log \frac{U_a}{U_e} \text{ [dB]} \quad (1.4)$$

Bei negativem Vorzeichen des ermittelten Wertes handelt es sich eigentlich um eine Dämpfung – wird sie als solche jedoch angegeben, muss das Vorzeichen in Gl. 1.4 gewechselt werden, oder vom Bruch ist der Kehrwert zu nehmen:

$$a_d = -20 \cdot \log \frac{U_a}{U_e} = 20 \cdot \log \frac{U_e}{U_a} \text{ [dB]} \quad (1.5)$$

Pegelabweichungen bzw. -toleranzen können ebenfalls auf diese Art definiert werden; meist findet hier ein Bezug zu einem willkürlichen Wert statt, z.B. dem Pegel bei 1000 Hz. Den Pegel am Bezugspunkt setzt man mit 0 dB fest, unabhängig von der tatsächlichen Spannung. Abweichungen zu diesem Pegel erhalten dann einfach die nach Gl. 1.2 ermittelten Werte, wobei hier die Spannung U_0 wie gesagt von 0,775 V abweichen kann.

Wie zu sehen ist, erfolgt die Berechnung der logarithmischen Größen immer nach demselben Schema. Korrekte Angaben erfordern jedoch die Einhaltung der Definitionen für Verstärkung und Dämpfung (ein negatives Dämpfungsmaß entspricht einem positiven Verstärkungsmaß und umgekehrt) und somit die richtige Wahl von Zähler und Nenner in den oben genannten Gleichungen.

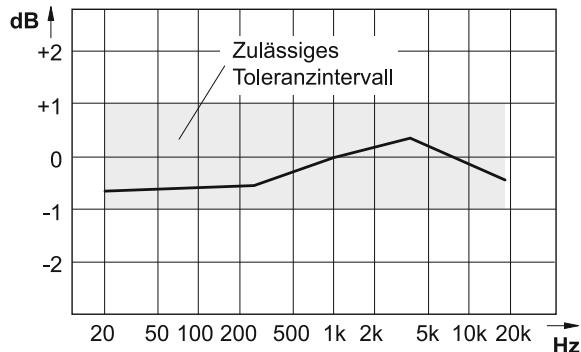
Die Maßeinheit *Dezibel* ist ein Zehntel der Einheit *Bel*, welche dem amerikanischen Erfinder A. Bell zu Ehren geschaffen wurde.

Aufgrund der einfachen Logarithmierung eines Verhältnisses entsteht eigentlich eine Pseudoeinheit, also keine echte Maßeinheit. Trotzdem ist es natürlich sinnvoll und notwendig, dB bzw. dBu anzuhängen, weil sonst keinerlei Rückschlüsse über das Zustandekommen der Angabe möglich sind.

1.2 Lineare Verzerrungen und Frequenzgang

Im Datenblatt von Audiogeräten ist der *Frequenzgang* eine häufig zu findende technische Spezifikation. Leider erfolgt nicht immer eine korrekte Angabe, und oft wird anstatt des Frequenzgangs eine Angabe über den Frequenzbereich gemacht, der allerdings so gut wie

Abb. 1.1 Frequenzgang mit erlaubtem Toleranzintervall



nichts ausdrückt. Zur Qualifikation der spektralen *Übertragungsqualität* sind neben den Eckfrequenzen auch die Toleranzen innerhalb des Intervalls anzugeben. Beispielsweise bezeichnet die Aussage

Frequenzgang 20 ... 18,000 Hz \pm 1 dB

einen Pegelverlauf zwischen den Frequenzgrenzen 20 und 18,000 Hz, der sich innerhalb eines Toleranzbereiches von 2 dB aufhält. Abb. 1.1 zeigt diese Situation.

Für manche Übertragungssysteme sind jedoch mehrere Frequenzbereiche spezifiziert, also z. B.

40 ... 1250 Hz \pm 3 dB

Beide Felder überlappen sich teilweise; beide Bedingungen müssen jedoch getrennt erfüllt sein. Abweichungen vom idealen, horizontal verlaufenden Frequenzgang röhren von sog. linearen Verzerrungen bzw. werden von diesen definiert. Der Begriff der linearen Verzerrungen tritt jedoch normalerweise nicht in einem Datenblatt auf – vielmehr handelt es sich um eine spezielle Kategorie von Verzerrungen. Schicken wir durch ein Übertragungssystem ein Sinussignal, also ein Signal mit einer diskreten Frequenz, so wird am Ausgang ebenfalls ein Sinussignal anstehen – mit den durch den Frequenzgang Amplitudenabweichungen. Die Beibehaltung der Signalform kommt von einer ungekrümmten, linearen statischen Kennlinie des Übertragungssystems, die bei nur *linearen Verzerrungen* im Idealfall vorliegt. Beschickt man das System jedoch mit einem Signal, das Obertöne (Harmonische) enthält, also z. B. mit einem Sägezahn- oder Rechtecksignal, so steht am Ausgang trotz linearer statischer Kennlinie eine veränderte Signalform an. Durch Abweichungen vom linearen, flachen Frequenzgang werden natürlich die Frequenzkomponenten (Grundton und Obertöne) nicht mit einheitlichem Verstärkungsfaktor behandelt.

HiFi-Verstärker der Spitzenklasse weisen oft extrem kleine Klirrfaktoren aus; dennoch erfährt ein Sägezahnsignal eine starke Signalverlaufs-Änderung, wenn die Klangregelung außer Mitte gebracht wird. Dagegen verändern sich Sinussignale dadurch nicht in ihrer Form. Lediglich die Amplitude wird beeinflusst durch die aktuelle Einstellung des Frequenzgangs.

Der Frequenzgang eines Übertragungssystems kann über die Messung von Ausgangsspannungen bei diskreten Frequenzen ermittelt werden; die Eingangsspannung ist hierbei bekannt und wird normalerweise konstant gehalten. Außerdem ist letztere deutlich unterhalb einer Übersteuerungsgrenze einzustellen, da sonst Messfehler entstehen können. Elegantere Methoden zur Messung des Frequenzgangs setzen *Weißes Rauschen* als Messsignal ein, bei dem ja bekanntlich alle Frequenzen (im hörbaren Spektrum) mit gleicher Amplitude vertreten sind (gleichmäßige Spannungsverteilung). Bei der Auswertung kommt ein *Spektrumanalysator* zu Einsatz, der die Spannungen einzelner Frequenzbänder (meist *Terzbänder*) gleich optisch zur Anzeige bringt. Hochmoderne Methoden bedienen sich eines kurzen *Rauschimpulses*, dessen Ausgangsantwort mit dem Originalimpuls über ein Computerprogramm korreliert wird; das Ergebnis bringt unmittelbar (nach kurzer Rechenzeit) den Frequenzgang zutage.

1.3 Nichtlineare Verzerrungen

Bei den *nichtlinearen Verzerrungen* hängt die Verformung des Signals nicht von der gewählten Frequenz ab; selbst ein Sinussignal erfährt eine Verformung. Dies geschieht, wenn das Übertragungssystem (Verstärker, Recorder, Sound-Karte usw.) eine statische Kennlinie aufweist, die nicht durch eine Gerade repräsentiert wird und die folglich Unstetigkeiten oder keine konstante Steigung hat. Abb. 1.2 zeigt ein Beispiel einer verzerrenden Kennlinie und die daraus folgende Signalform. Es handelt sich um den sehr häufig vorkommenden Fall der *Begrenzung (Clipping)*.

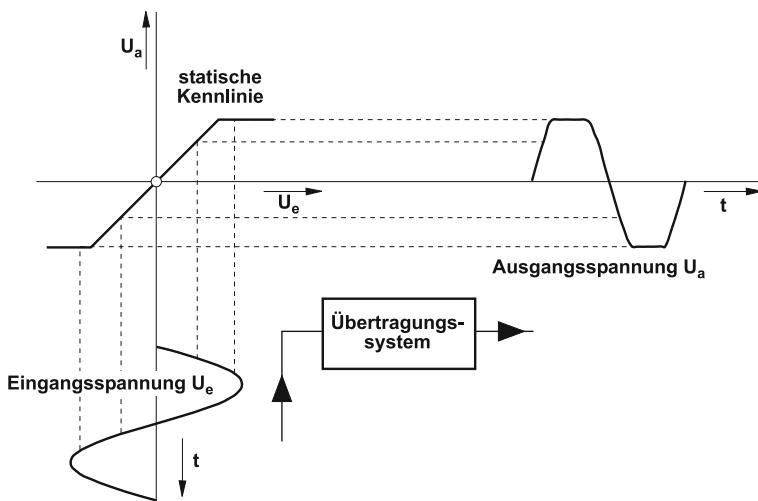


Abb. 1.2 Verformung einer Sinusspannung aufgrund einer gekrümmten bzw. geknickten Kennlinie

1.3.1 Klirrfaktor

Die deformierte Sinuslinie, die das Übertragungssystem verlässt, besteht nicht mehr nur aus einer diskreten Frequenz (wie das bei einem reinen Sinussignal der Fall wäre), sondern zur Grundschwingung addieren sich *Obertöne (Harmonische)* mit ganzzahligen Verhältnissen zur Grundfrequenz.

Bleiben wir bei der oben erwähnten (symmetrischen) Begrenzung, so ergibt sich eine Frequenzanalyse nach Abb. 1.3. Der Klirrfaktor ist ein Maß für die Intensität der Harmonischen und ergibt sich aus:

$$k = \sqrt{\frac{U_2^2 + U_3^2}{U_1^2 + U_2^2 + U_3^2 \dots}} \quad (1.6)$$

Die Indizes legen die Ordnung der Harmonischen fest. Somit entspricht obige Gleichung dem Verhältnis aus dem Effektivwert aller Obertöne (*Klirrkomponente*) zum Effektivwert der Gesamtspannung.

Bei der konventionellen Bestimmung des Klirrfaktors wird auch in entsprechender Weise vorgegangen: zunächst erfolgt die Messung des Effektivwerts des Gesamtsignals,

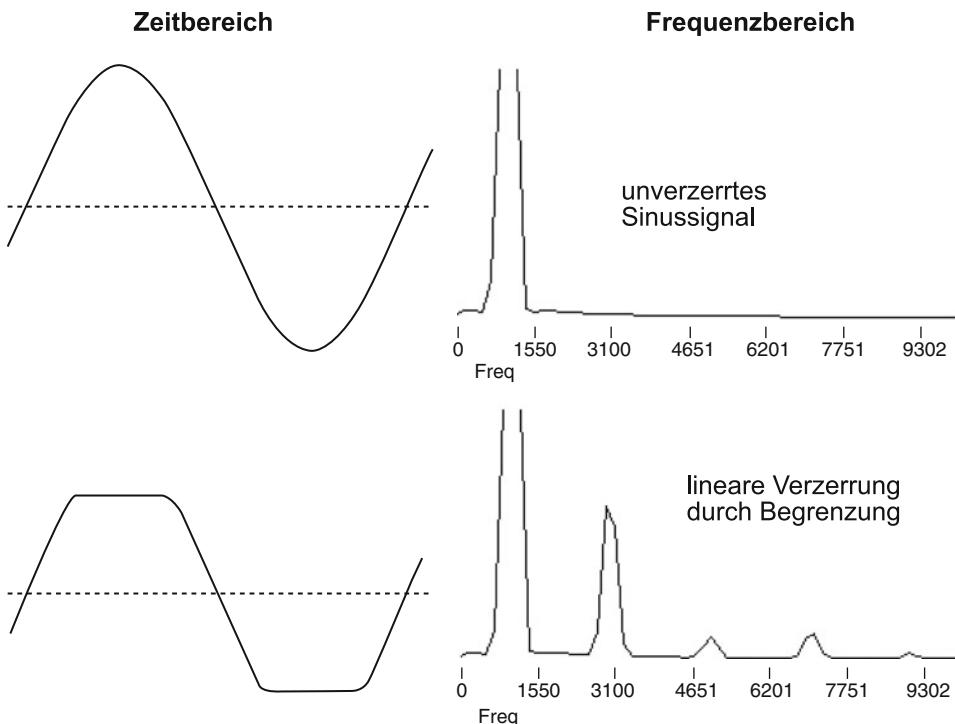


Abb. 1.3 Entstehung von Obertönen durch Verzerrung (Messtonfrequenz 1 kHz)

wobei das Ergebnis mit 100 % deklariert wird. Nun unterdrückt man über ein Filter die Grundfrequenz, so dass nur noch der Effektivwert der Obertöne zur Anzeige kommt. Durch die vorherige Festlegung kann sofort der Klirrfaktor in % abgelesen werden. Die Spezifikation der Messfrequenz ist für den Klirrfaktor, da in den meisten Fällen k eine Funktion von f ist (meist steigt k mit der Frequenz).

Für diese Methode existieren natürlich mehr oder weniger automatisierte Realisierungen – das Grundprinzip ist jedoch dasselbe. Daneben gibt es noch Verfahren, die wie bei der Frequenzgangmessung einen Rauschimpuls einsetzen, der mittels Algorithmus einer Korrelation eine Auswertung erfährt, welche am Ende auch zum Klirrfaktor führt.

Teilweise ist auch die Angabe einer *Klirrdämpfung* üblich. Sie ergibt sich aus:

$$a_k = 20 \cdot \log \frac{1}{k} = -20 \cdot \log k \text{ [dB]} \quad (1.7)$$

Der oben genannte Fall einer Begrenzung ist natürlich nur ein Beispiel für eine Verzerrung. Die Obertöne nach Abb. 1.3 zeigen hier nur ungeradzahlige Vielfache der Grundfrequenz. Es gibt jedoch sehr wohl Verzerrungsarten, bei denen beliebige Obertonkonstellationen entstehen (also auch Harmonische gerader Ordnung). Die Bezeichnung *THD* (Total Harmonic Distortion) berücksichtigt den Klirrfaktor nach Gl. 1.6, während teilweise auch nur bestimmte Harmonische zur Bestimmung herangezogen werden. So sind quadratische durch k_2 und kubische Verzerrungen durch k_3 gekennzeichnet – entsprechend findet auch nur die 2. bzw. 3. Harmonische Berücksichtigung bei der Bestimmung. Welcher der beiden Werte sinnvoller ist, hängt von der Art der Verzerrung und somit vom Übertragungssystem ab (für symmetrische Begrenzungen ist eine Angabe von k_2 nicht sinnvoll, da die 2. Harmonische nicht vorkommt).

Der Klang eines Sinussignals wird normalerweise als dumpf empfunden, da ja keinerlei Obertöne enthalten sind. Die harmonische Verzerrung ruft jedoch die Existenz solcher Obertöne hervor, so dass das verzerrte Signal recht schrill klingen kann. Der Unterschied zwischen unverzerrtem und verzerrtem Signal kann jedoch physiologisch manchmal nur mit Mühe registriert werden, wenn das Signal nur aus einem Ton besteht, weil erstens Harmonische gegenüber einem Grundton nicht besonders auffallen und zweitens höherfrequente Obertöne durch das Gehör mitunter schlechter oder überhaupt nicht mehr wahrgenommen werden können.

1.3.2 Intermodulation

Die *Intermodulation* bestimmt nicht die durch Verzerrung entstehenden Obertöne, sondern die Störsignale, die sich bei der *Mischung* ergeben. (Grundsätzlich hat beides, das Klirren und die Intermodulation, dieselbe Ursache, nämlich eine nichtlineare statische Kennlinie beim Übertragungssystem. Beides tritt daher nur gemeinsam oder überhaupt nicht auf; die Unterschiede liegen deshalb nur in der Messmethode und natürlich in der Aussagekraft).

Weil im höheren Frequenzbereich das Entstehen von Obertönen infolge von Verzerrungen zumindest teilweise durch die *Bandbegrenzung* wieder rückgängig gemacht wird, eignen sich solche Messmethoden besser, die zwei Prüftöne einsetzen – sie heißen deshalb auch *Zweitonverfahren*.

Soll der Klirrfaktor bei einer Frequenz von 10 kHz bestimmt werden, so kann dies bereits Probleme bereiten, da die zweite Harmonische eine Frequenz von 20 kHz und die dritte gar 30 kHz aufweist. Treten bevorzugt kubische Verzerrungen auf, so liegen diese tatsächlich bei 30 kHz. Wenn der Proband jedoch am Ausgang ein Tiefpassfilter einsetzt, das Frequenzen ab 20 kHz unterdrückt, so sind die zur Messung des Klirrfaktors relevanten Obertöne überhaupt nicht mehr vertreten, obwohl irgendeine Einheit des Gerätes dennoch möglicherweise Verzerrungen verursacht. Aber eine Messung des Klirrfaktors bringt die Verzerrungseigenschaften nicht zutage.

Es existieren verschiedene solcher Zweiton-Messsignale, die nach Frequenz und Amplitudenverhältnis von einigen Gremien festgelegt wurden. Die Festlegung ist sinnvoll – ohne sie wären Messungen nicht reproduzierbar und somit auch ohne Aussagekraft. So schreibt z. B. die DIN 45500 (*HiFi-Norm*) die Frequenzen 250 und 8000 Hz und ein Amplitudenverhältnis von 4 : 1 vor, während CCIF die Frequenzen 13 und 14 kHz bei gleichen Amplituden einsetzt. Abb. 1.4 zeigt die Auswirkungen bei einem verzerrenden System, wobei wie in Abb. 1.3 eine Darstellung des Zeit- und Frequenzbereiches erfolgt.

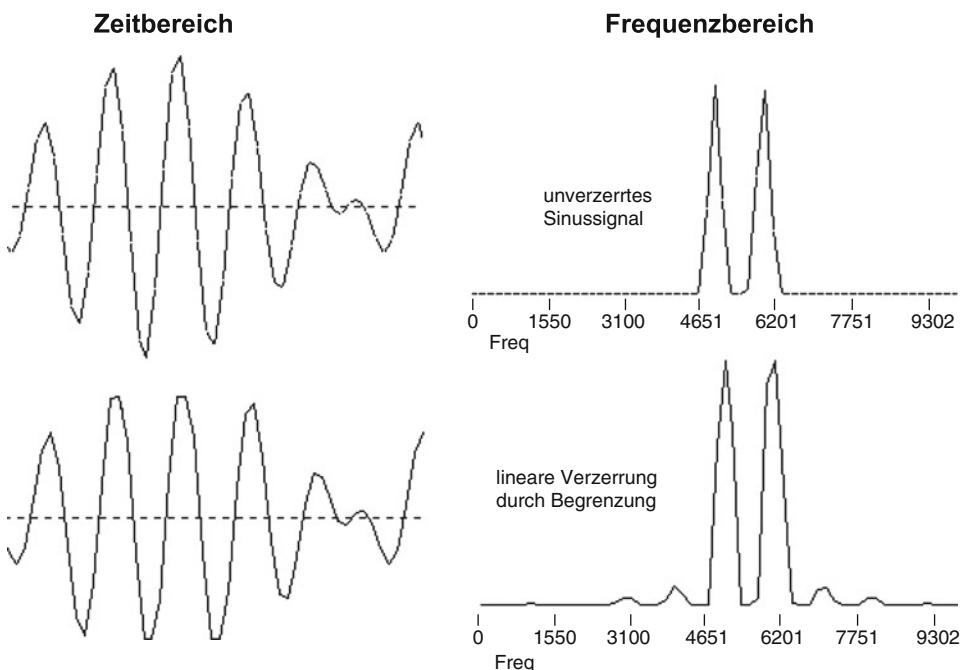


Abb. 1.4 Entstehung von Differenztönen durch Verzerrung (Messtonfrequenzen 5 und 6 kHz, Amplitudenverhältnis 1:1)

Es entstehen als *Mischprodukte* Signale mit den Frequenzen

$$\Delta f_{i2} = \Delta f = f_2 - f_1 \text{ (Komponente aus quadrat. Verzerrung)} \quad (1.8)$$

$$\Delta f_{i3} = f_2 + \Delta f = 2f_2 - f_1 \text{ (Komponente aus kub. Verzerrung)} \quad (1.9)$$

Die kubische Verzerrung, wie sie bei unserem Beispiel hauptsächlich besteht, verursacht also neben den Peaks der Hauptfrequenzen immer im Abstand Δf weitere Komponenten, die natürlich auch als *Spiegelfrequenzen* auftreten. Gl. 1.9 müsste also erweitert werden um die Komponenten $3f_2 - 2f_1$, $4f_2 - 3f_1$ usw. sowie $3f_1 - 2f_2$, $4f_1 - 3f_2$ usw. (Spiegelfrequenzen). Auch Abb. 1.4 zeigt deutlich diese Störanteile.

Für die Intermodulation ergibt sich folgender Zusammenhang:

$$D = \sqrt{\frac{U_{\Delta 1}^2 + U_{\Delta 2}^2 + U_{\Delta 3}^2 \dots}{U_1^2 + U_2^2}} \quad (1.10)$$

Im Zähler dieser Gleichung steht die Summe aller Spannungsquadrate der Differenztöne, während der Nenner nur die Quadrate des Messtonpaars berücksichtigt.

Oftmals werden jedoch nur die Komponenten nach den Gl. 1.8 und 1.9 bestimmt und deren Amplitudenverhältnis zum ursprünglichen Zweitonsignal gebildet. Daraus ergeben sich die *Differenztonfaktoren* D_2 und D_3 . Die geometrische Addition beider Werte ergibt näherungsweise ebenfalls den Intermodulationsfaktor, wenn die Glieder höherer Ordnung vernachlässigt werden sollen (was in den meisten Fällen geschehen kann).

Es leuchtet ein, dass vor allem diese Mischprodukte dem menschlichen Ohr auffallen, da beim Musikhören meist ein Frequenzgemisch vorkommt, welches dann bei einem verzerrenden Übertragungssystem zusätzlich Mischprodukte entstehen lässt, die teilweise wesentlich tieffrequenter und somit auffallender als die Originaltöne bzw. deren Obertöne sind. In bestimmten Situationen (oberer Frequenzbereich, Bandbegrenzung) ist also die Angabe von Intermodulationswerten günstiger und vor allem informativer als der Klirrfaktor. Wenn die Art der Intermodulations-Messung nicht eindeutig hervorgeht, muss sie angegeben werden.

Die Wahl des Frequenzpaars ist – falls kein festgelegtes Verfahren zur Anwendung kommt – so vorzunehmen, dass sich die oberen Mischprodukte aus kubischer Verzerrung und die unteren aus quadratischer Verzerrung klar voneinander unterscheidbar sind, also einen gewissen Mindestabstand haben.

Aus dem oben genannten geht bereits hervor, wie das Prinzip der Intermodulationsmessung aussieht. Entscheidend ist beim Messverfahren vor allem die Selektion der Intermodulationskomponenten mittels Filter, die einerseits die Amplitude der zu bestimmenden Anteile möglichst wenig beeinflusst und andererseits für die Originalsignale sperrend sind. S. 42 [2].

Umrechenbarkeit Klirrfaktor- und Intermodulationwert lassen sich ineinander umrechnen, vorausgesetzt, das Übertragungssystem ändert seine Kennlinie nicht mit der Frequenz. S. 42 [2].

1.4 Störabstand

Störsignale, die nicht vom Nutzsignal abhängig sind und infolgedessen auch nicht von diesem verursacht werden, tragen ebenfalls zur Verschlechterung der Übertragungsqualität bei. Das Störsignal kann unterschiedliche Bestandteile aufweisen und deswegen natürlich auch physiologisch spezifiziert und identifiziert werden.

Die Störgröße steht in einem Bezug zu einem definierten Nutzsignalpegel, welcher immer angegeben sein muss – ein Störabstand ohne die Angabe des Gesamtsignalpegels (der i. Allg. näherungsweise dem Nutzsignalpegel entspricht) ist wenig sinnvoll, da hier ja ein direkter Zusammenhang besteht, wie wir weiter unten sehen werden.

1.4.1 Geräuschspannungsabstand

Der *Geräuschspannungsabstand* berücksichtigt Störungen, die normalerweise nur durch Rauschen verursacht werden. Der willkürlich festgelegte Gesamtsignalpegel (Nutzsignal plus Störsignal) sei 0 dB. Fehlt dieser gänzlich, kann die Messung des Störgeräusches erfolgen. In beiden Fällen ist natürlich eine Effektivwertmessung erforderlich. Der Pegel des Störgeräusches liegt mit dem Maß des Geräuschspannungsabstands unterhalb vom Gesamtsignalpegel. Bei der Messung der Gesamtspannung U_{ges} und der *Störrauschspannung* U_r ergibt sich also mathematisch für den Geräuschspannungsabstand (vgl. Abschn. 1.1):

$$a_g = 20 \cdot \log \frac{U_{\text{ges}}}{U_r} \text{ [dB]} \quad (1.11)$$

Bewerteter Geräuschspannungsabstand Die Empfindlichkeit des menschlichen Ohres ist frequenzabhängig; somit benötigt man unterschiedliche Schalldrücke bzw. Schallleistungen, um bei verschiedenen Frequenzen das gleiche Lautstärkeempfinden zu erreichen. Nach umfangreichen Untersuchungen hat man demnach die *Kurven gleicher Lautstärke* ermittelt (siehe Anhang B.1). Dabei ergibt sich jedoch ein anderer Verlauf für z. B. 30 Phon als für 60 Phon. Eine Spiegelung an der Waagerechten ergibt die entsprechenden Empfindlichkeitsverläufe des Ohres. Geringe Lautstärken (kleiner als 30 Phon) – und darunter fallen normalerweise auch die Störgeräusche – werden vom Ohr mit abgeschwächten Tiefen und Höhen vernommen. Es ist also sinnvoll, das Störsignal über ein Filter zu schicken, das sozusagen denselben Frequenzgang hat wie das Ohr bei Lautstärken bis 30 Phon. Von dieser Möglichkeit macht man Gebrauch und setzt ein sog. *30-Phon-Ohrkurvenfilter* (auch dB(A)-Filter genannt) ein, welches die oben genannten Eigenschaften besitzt. Wichtig ist,

dass bei der Messung des bewerteten Geräuschspannungsabstandes das Filter lediglich zur Bestimmung des Störpegels zwischengeschaltet werden darf. Außerdem ist bei einer *Bewertung* die Art bzw. das Filter anzugeben, was meist vom Typ dB(A) oder einfach Typ A ist.

1.4.2 Fremdspannungsabstand

Der *Fremdspannungsabstand* ist der Störabstand zu allen Störsignalen; insbesondere kommen hier auch Brummen, *Rumpeln* usw. zum Tragen. Ein Bemessungsfilter wie beim Geräuschspannungsabstand ist hier i. Allg. nicht vorgesehen. Damit ähnelt die mathematische Bestimmung der Gl. 1.11:

$$a_{fr} = 20 \cdot \log \frac{U_{ges}}{U_{st}} \text{ [dB]} \quad (1.12)$$

Aufgrund der Tatsache, dass hier alle Störungen gemessen und nicht bewertet werden, fällt der Fremdspannungsabstand zahlenmäßig grundsätzlich ungünstiger (kleiner) aus als der Geräuschspannungsabstand. Letzterer hat jedoch mehr mit der physiologisch Wirksamkeit zu tun und ist deshalb der etwas häufiger anzutreffende Wert bei den technischen Spezifikationen eines Gerätes. Dennoch findet man den Fremdspannungsabstand vor allem bei Geräten, bei denen Rumpel- oder Brummgeräusche zu erwarten sind, was z. B. bei Plattenspielern oder Magnettongeräten der Fall ist.

1.5 Dynamikbereich

Der Pegelabstand zwischen leisensten und lautesten Signalen, die vom System gerade noch verarbeitet werden können, gibt den nutzbaren Dynamikbereich an. Die obere Pegelgrenze wird hierbei von der Übersteuerungsgrenze bestimmt, während die untere Pegelgrenze vom Störsignal und bei digitalen Systemen auch von der Quantisierungsauflösung abhängt. Zum Störsignal muss ein gewisser Sicherheitsabstand vorhanden sein (z. B. 6 dB), damit sich das Nutzsignal noch deutlich abhebt und vom Störpegel nicht überdeckt wird. Nach dieser Aussage muss die *Dynamik* immer kleiner sein als der Fremdspannungsabstand. Abb. 1.5 zeigt die qualitativen Verhältnisse einiger Qualitätskriterien.

Bei digitalen Systemen lässt sich der theoretische Dynamikbereich veranschlagen durch:

$$a_d \approx 6 \cdot N + C_f \quad (1.13)$$

Hierin sind N die Auflösung in Bit und C_f eine Signalformkonstante, die bei Sinusignalen etwa 1,7 beträgt. Hierdurch ist jedoch die Obergrenze für die Dynamik

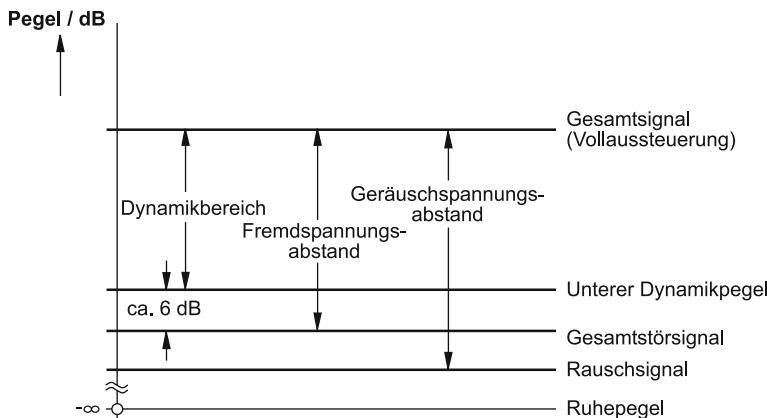


Abb. 1.5 Qualitative Darstellung der Störabstände von Übertragungssystemen

gegeben – die nutzbare Dynamik wird dann wie bereits erwähnt durch den Fremdspannungsabstand beschränkt, der von Eigenschaften der Analogseite abhängt.

1.6 Übersprechdämpfung

Bei mehrkanaligen Übertragungssystemen spielt auch die *Übersprechdämpfung* (teilweise auch *Kanaltrennung* genannt) eine gewisse Rolle. Man versteht darunter die Fähigkeit, Signale auf nur einem Kanal zu übertragen, während auf dem anderen Kanal ein möglichst kleiner Pegel dieses Testsignals erscheint.

Da es sich um ein Dämpfungsmaß handelt, erhalten wir wieder folgenden Zusammenhang mit den gemessenen Spannungen:

$$a_{\ddot{u}} = 20 \cdot \log \frac{U_0}{U_{\ddot{u}}} \text{ [dB]} \quad (1.14)$$

U_0 ist hier wieder eine Bezugsspannung, die angegeben werden sollte, da nicht sichergestellt ist, ob zwischen U_0 und der Übersprechspannung $U_{\ddot{u}}$ ein proportionaler Zusammenhang besteht. Das System gibt auf einem Kanal einen Testton mit Bezugspegel ab. Der andere Kanal soll dagegen im Idealfall stumm sein; das dennoch vorhandene Übersprechsignal der Spannung $U_{\ddot{u}}$ zeugt von Übersprecheffekten.

Erfahrungsgemäß reduziert sich die Übersprechdämpfung mit der Frequenz, d. h. der Störeffekt des Übersprechens wird bei höheren Frequenzen begünstigt. Aus diesem Grund ist die Angabe einer Frequenz, für die die Kanaltrennung gilt, erforderlich.

Eine Kanaltrennung mit Werten über 40 dB ist normalerweise nicht mehr auffällig, wenn auf beiden Kanälen ein zusammengehörendes Musiksignal zu übertragen ist. Hier bewirkt eine schlechte Kanaltrennung lediglich, dass die auf die Kanäle platzierten Stimmen bzw.

Instrumente etwas näher „zusammenrücken“. Dies fällt i. Allg. kaum auf, weil die Instrumente sowieso meist nicht ganz auf einer Seite liegen. Kritisch kann die Kanaltrennung werden, wenn beide Kanäle nicht direkt zusammengehörende Signale übertragen sollen. Ein Beispiel ist eine Mehrsprachensendung, bei der auf einem Kanal der Ton in Landessprache und auf dem anderen Kanal der Originalton liegt. In dieser Situation wäre ein „Durchklingen“ des unerwünschten Kanals sehr störend, und deshalb ist hier die Forderung einer hohen Kanaltrennung durchaus angebracht und wünschenswert.

1.7 Phasenverzerrungen

Ein selteneres Phänomen der Beeinträchtigung der Signalqualität ist das der Phasenverzerrungen. In diesem Zusammenhang fällt sehr häufig auch der Begriff der Gruppenlaufzeit, welche mathematisch die Ableitung des *Phasenfrequenzgangs* $\varphi(\omega)$ nach der Kreisfrequenz ω entspricht:

$$T_g(\omega) = -\frac{d\varphi(\omega)}{d\omega} \quad (1.15)$$

Ein ideales Übertragungssystem weist einen proportionalen (also auch linearen) Zusammenhang zwischen Phasengang und Frequenz auf. In diesem Falle hätten wir sozusagen ein Verzögerungsglied mit konstanter Zeitverzögerung. Die Gruppenlaufzeit kommt in diesem Spezialfall der Zeitverzögerung gleich. In realen Systemen ist die Gruppenlaufzeit wie in obiger Gleichung von der Frequenz abhängig, und die Art und der Umfang dieser Abhängigkeit bestimmt die Phasenverzerrung. In Abb. 1.6 ist das ideale und ein reales Übertragungssystem gegenübergestellt.

Wir stellen uns ein ideales Übertragungssystem vor. Die akustische Wiedergabe über Lautsprecher beinhaltet keinerlei Verzerrungen. Verändern wir unsere Position, so variieren wir damit auch die Laufzeit, die der Schall bis zu unseren Ohren benötigt. Hohe und tiefe Frequenzen unterliegen jedoch gleichen Laufzeiten, so dass der Phasengang linear und die Gruppenlaufzeit konstant ist. Durch den Abstand Ohr/Lautsprecher entstehen also keine Phasenverzerrungen. Anders sieht es aus, wenn wir getrennte Lautsprechersysteme für verschiedene Frequenzbereiche hätten und diese auch räumlich voneinander entfernt aufstellen. Beim Anklingen eines Instruments kommen die Phasenverzerrungen besonders zum Tragen, wenn die hohen Töne zu einem anderen Zeitpunkt als die tiefen beim Ohr ankommen.

Obwohl es sich nicht um nichtlineare Verzerrungen handelt, kann bei extremen Phasenverzerrungen ein Sprachsignal bis zur Unkenntlichkeit verändert werden.

Insbesondere Filtereinheiten tragen zu Phasenverzerrungen bei. Besonders günstig verhalten sich hierbei sog. digitale FIR-Filter, weil sie eine frequenzunabhängige (also konstante) Gruppenlaufzeit aufweisen. Das besondere Merkmal solcher Filter zeigt sich u. a. auch in der *symmetrischen Sprungantwort*. Abb. 1.7 stellt ein solches Signal einem

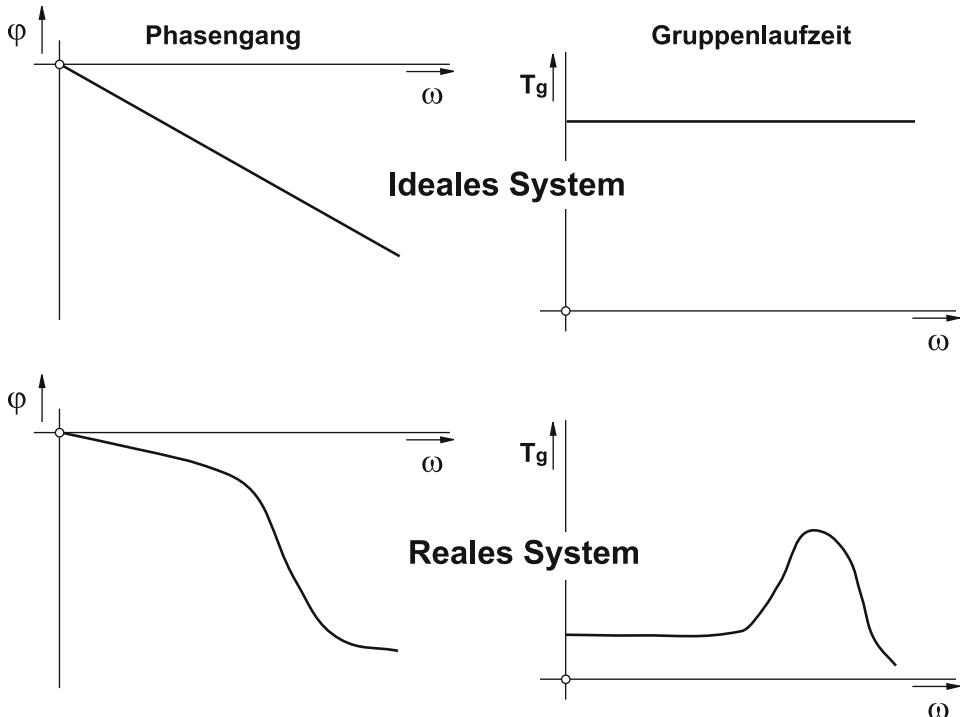


Abb. 1.6 Phasen- und Gruppenlaufzeitgänge eines idealen und eines realen Übertragungssystems

Rechtecksignal gegenüber, welches mit einem phasenverzerrenden Tiefpass behandelt wurde und somit eine asymmetrische Sprungantwort zeigt.

Phasenverzerrungsfreie FIR-Filter kommen bei digitalen Geräten, also auch bei Audiokarten, bevorzugt zum Einsatz. Sie dienen dort der Interpolation beim Oversampling (siehe nächstes Kap.). Auch in manchen DSPs (Digital Signal Processor) ist ein FIR-Filter bereits integriert – zumindest ist er für die Operationen eines FIR-Filters ausgelegt.

Bei modernen Geräten sind die störungsbedingten Phasenverzerrungen meist so gering, dass sie auch von empfindlichen Ohren nicht wahrgenommen werden können. Trotzdem sind teilweise Spezifikationen in technischen Datenblättern zu finden. Es handelt sich jedoch i. Allg. um Angaben über einen Phasenfehler, der in Altgrad angegeben ist und der als Abweichung gegenüber tiefen Frequenzen zu verstehen ist.

Bei hörbaren Phasenverzerrungen – was wie gesagt selten vorkommt – ist große Erfahrung notwendig, um sie als solche zu identifizieren. Es ist zwar zu hören, dass etwas nicht in Ordnung ist, jedoch ist meist eine andere Einordnung als „schlechte Aufzeichnung“ zunächst nicht möglich.

Zu erwähnen wäre noch ein (frequenzunabhängiger) Phasenfehler eines Kanals bezüglich des anderen, was oft durch Verpolung eines Lautsprechersystems verursacht wird.

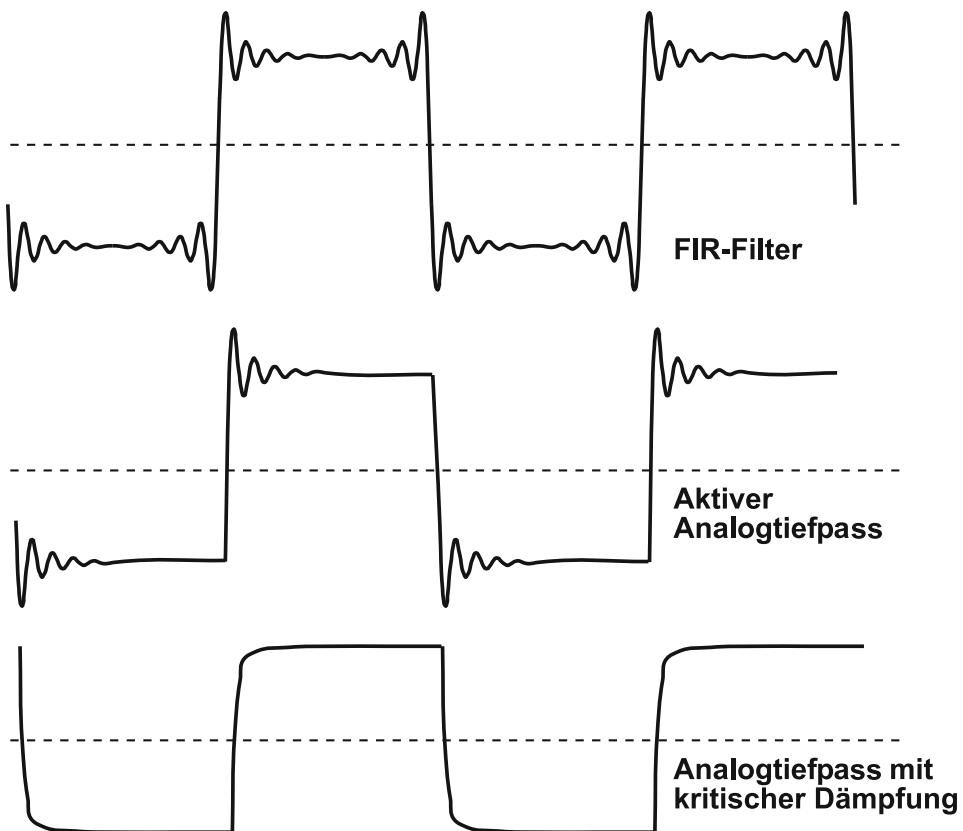


Abb. 1.7 Sprungantworten verschiedener Tiefpässe

Stereoübertragungen klingen dadurch relativ „flach“, und der Klang ändert sich stark bei Variation des Abstands zwischen Ohr und Lautsprecher, was sich durch extreme Interferenzen erklären lässt.

1.8 Raumakustik

Zur Erzielung optimaler Beschallungsergebnisse sind Kenntnisse über Raumakustik und natürlich über den Raum selbst nötig. Grundsätzlich gibt es den idealen Raum überhaupt nicht. Um die von Lautsprechern wiedergegebene Übertragung jedoch möglichst wenig zu verfälschen, sind einige Dinge zu berücksichtigen, die nicht nur vom Frequenzgang der Lautsprecher abhängen. S. 42 [2,5].

Voraussetzung, jedoch nicht hinreichend ist also der Frequenzgang der Schallwandler bzw. Lautsprecher. Ausführungen für gehobene Ansprüche liefern meist ein Datenblatt mit, aus dem zumindest ein Toleranzfrequenzgang hervorgeht. Die Nennleistung (Sinusdauerleistung)

ist maßgebend für die Belastbarkeit; sie muss deutlich über der maximal genutzten Leistung liegen. Für mittlere Räume kann man dabei von einem Richtwert von ca. 20 W ausgehen, wenn es sich um Monitorboxen im Studio handelt. Eine Diskotheken-Beschallung erfordert allerdings größere Leistungsstufen.

Die Zimmerwände beeinflussen den Schall nicht unerheblich. Die Ursachen hierfür sind Reflexion, Absorption und teilweise auch Resonanz. Keiner dieser drei Effekte, die physikalisch bedingt sind, lässt sich gänzlich ausschalten; es sind lediglich Maßnahmen zur Minimierung möglich.

1.8.1 Schallreflexion

Die Reflexion von Schallwellen an einer Wand führt zu weiteren Folgeeffekten, die dann endgültig den Klangeindruck verändern.

Interferenz Dadurch, dass der Schall unterschiedliche Wege bis zum Ohr zurücklegen kann, nämlich den direkten und den Reflexionsweg, entstehen auch Laufzeitunterschiede, was wiederum zu Interferenzen und somit Auslöschungen bei diskreten Frequenzen führt. Abb. 1.8 verdeutlicht das Entstehen solcher Laufzeitunterschiede. Ist x_0 der direkte und x_r der Reflexionsweg sowie c_s die Schallgeschwindigkeit, so ergeben sich Auslöschungen für folgende Frequenzen:

$$f_{\text{aus}} = (n + 0,5) \cdot \frac{c_s}{x_r - x_0} \quad \text{für } n = 1, 2, 3, \dots \quad (1.16)$$

Daraus geht hervor, dass in die Wegdifferenz immer eine ungerade Anzahl von Halbwellenzügen passen muss, damit Auslöschung vorliegt.

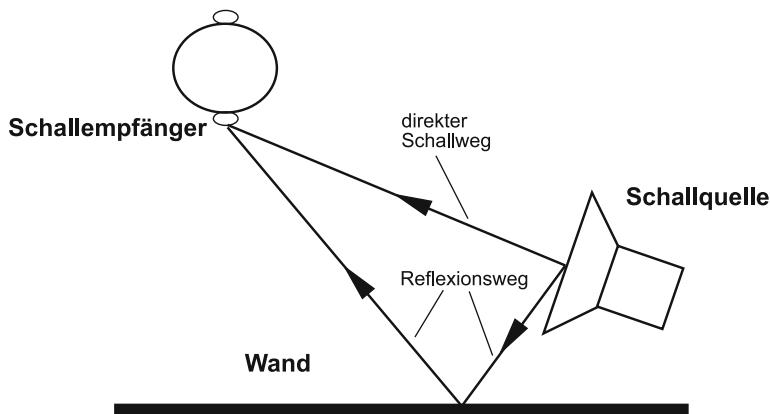


Abb. 1.8 Überlagerung von Direktschall und Reflexionsschall

Zur Vermeidung von Reflexionen sind spezielle *Dämmstoffe* notwendig, die sowohl in stofflicher als auch in struktureller Beschaffenheit geeignet sein müssen, den an kommenden Schall zu absorbieren, was in diesem Fall heißt, ihn in Wärme umzuwandeln. Es handelt sich um weiches und faseriges Material, welches jedoch auch eine grobe Profilstruktur (Berge und Täler) aufweist, damit auch noch möglichst tiefe Frequenzen wirkungsvoll absorbiert werden können. Die Interferenzen spielen sich nunmehr nicht mehr inmitten des Schallraumes ab, sondern beschränken sich auf die nahe Oberfläche des Dämmmaterials.

Hall und Echo Die Reflexion führt jedoch zu einem uns vielleicht bekannteren Effekt, nämlich dem Hall bzw. Echo. Während unter letzterem das diskrete Wiederholen eines Schallereignisses nach einer gewissen Zeit zu verstehen ist, entsteht der Hall durch wesentlich diffusere Reflexionsvorgänge. Er lässt sozusagen unendlich viele Echos erklingen, die als Summe wie das kontinuierliche Abklingen des zuletzt aktuellen Schalles zu hören sind. Die Eigenschaften eines Raumes mit glatten Wänden bewirken jedoch eher den Effekt des Halls. Echos entstehen meist im Freien, wenn eine reflektierende Wand den Schall zu spiegeln vermag.

Es steht außer Frage, dass schalldämmende Maßnahmen auch dem Entstehen von Hall- und Echoeffekten entgegenwirken können. Jedoch sind in manchen Fällen gewisse Raumeigenschaften sogar sinnvoll, wenn es darum geht, dem Klang eines Instruments oder einer Stimme eine klangliche Atmosphäre zu verleihen. Im Falle einer bloßen Wiedergabe mittels Lautsprecher wäre ein möglichst schalltoter Raum erstrebenswert.

Wichtige praktische Anwendungen von akustischen Reflexionen sind Hallräume in Studios, die eine kontrollierte Zumischung von Raumeffekten ermöglichen. Ferner arbeiten besonders empfindliche Richtmikrofone mit akustischen Parabolspiegeln, welche eine Konzentration von Schallwellen bewerkstelligen (Bündelung der Schalleistung einer großen auf eine kleine Fläche).

1.8.2 Diffraktion

Was man beim Licht beobachten kann, wenn es sich nicht ungehindert ausbreiten kann, sondern durch Öffnungen (Spalt, Blende oder Gitter) gelangt, kann in ähnlicher Form auch beim Schall auftreten. Ist der Öffnungs durchmesser von der Größenordnung der Wellenlänge des Schalls, so interferieren die einzelnen Kugelwellen, die von dieser Öffnung bei Beschallung von der einen Seite ausgehen, besonders wirkungsvoll. Auf der anderen Seite der Beugungsöffnung sind dann Gebiete mit Verstärkungen und Auslöschungen festzustellen, wie dies aus Abb. 1.9 hervorgeht.

Die Begrenzungen der Zonen der Verstärkung und der Auslöschung sind natürlich nicht so scharf zu verstehen, wie sie dargestellt wurden. Es ist vielmehr ein kontinuierlicher Übergang der Zustände zwischen den Zonenachsen vorhanden. In Hauptrichtung (also

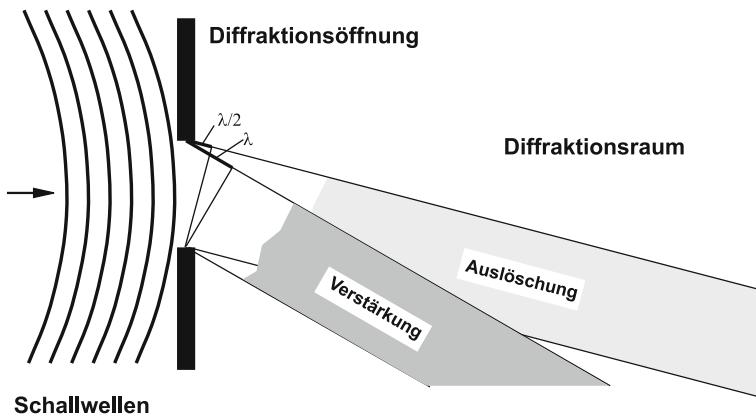


Abb. 1.9 Beugung einer Schallwelle an einer Öffnung

senkrecht zur Öffnung) bildet sich eine neutrale Zone, in der theoretische keine Veränderung der einfallenden Wellen stattfindet.

In der Lautsprecherboxen-Technik werden teilweise akustische Linsen eingesetzt, die teilweise als Lamellengitter und teilweise als Trichter ausgeführt sind. In beiden Fällen handelt es sich um die Nutzung von Wegdifferenzen – man kann den Effekt deshalb zur akustischen Beugung rechnen. Hauptnutzen sind dabei bessere *Streuung* (Erhöhung der Divergenz) bei Hochtönen, die sonst die Tendenz zur *Schallbündelung* zeigen, und die Vermeidung von *akustischen Kurzschlüssen* bei Tief tönen. S. 42 [2,5]

1.8.3 Refraktion (Brechung)

Unter Refraktion sind diejenigen Vorgänge zu verstehen, die sich an der Trennzone zwischen unterschiedlichen Medien für Wellen ergeben. Treffen diese Wellen nämlich nicht senkrecht auf die Trennebene der (durchlässigen) Medien, so ändern sie ihre weitere Ausbreitungsrichtung. Die Ursache hierfür ist in den unterschiedlichen Ausbreitungsgeschwindigkeiten der Wellen in den Medien zu suchen. Abb. 1.10 soll dies etwas veranschaulichen.

Laut Abb. 1.10 gilt dabei für die Winkel (mit c als den Ausbreitungsgeschwindigkeiten und n den *Brechungsindizes*):

$$\frac{\sin \alpha}{\sin \beta} = \frac{a}{b} = \frac{c_1}{c_2} = \frac{n_2}{n_1} \quad (1.17)$$

Diese in der Optik bekannten Beziehungen gelten mit gewissen Einschränkungen auch für das Verhalten des Schalls. Eine praktische Nutzung solcher Gegebenheiten fand bisher jedoch kaum statt. Die Effekte stellen sich lediglich als Störung bei der Schallwiedergabe ein.

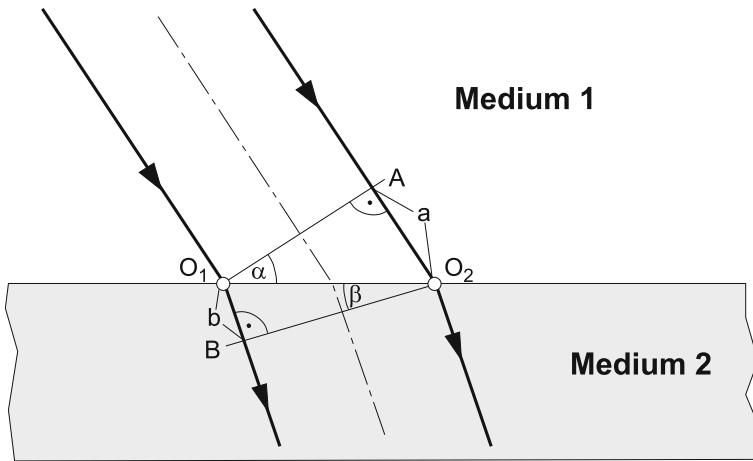


Abb. 1.10 Richtungsänderung einer Welle durch Verlangsamung der Ausbreitungsgeschwindigkeit. Eine Wellenfront A gelangt zunächst auf den Punkt O_1 der Mediengrenzfläche. Zu diesem Zeitpunkt befindet sich die andere Begrenzung des Strahls noch um a entfernt von Punkt O_2 . Während diese Entfernung durchschritten wird, pflanzen sich von Punkt O_1 Wellen mit der Distanz b fort – aufgrund der kleineren Ausbreitungsgeschwindigkeit eine kleinere Strecke als a . Die Strahlrichtung wird somit zum Einfallslot hin gebrochen

1.8.4 Doppler-Effekt

Der nach dem österreichischen Physiker benannte Effekt tritt immer dann in Erscheinung, wenn eine relative Bewegung zwischen Schallsender und -empfänger besteht. Da der Schall ein Medium zur Ausbreitung benötigt (im Gegensatz zu Licht), unterscheidet man zwei Fälle von Bewegungen:

Sender bewegt sich bezüglich Empfänger (+/-: Wegbewegung/Hinbewegung):

$$\lambda = \frac{c_s \pm v_t}{f_0} \Rightarrow f_e = \frac{c_s}{\lambda} = f_0 \cdot \frac{1}{1 \pm v_t/c_s} \quad (1.18)$$

Empfänger bewegt sich bezüglich Sender (-/+: Wegbewegung/Hinbewegung):

$$c_{\text{eff}} = c_s \pm v_t \Rightarrow f_e = \frac{c_{\text{eff}}}{\lambda} = f_0 \cdot (1 \pm v_t/c_s) \quad (1.19)$$

c_s Schallgeschwindigkeit, v_t Geschwindigkeit des Senders, f_0 Signalfrequenz des Senders, f_e Signalfrequenz am Empfänger, λ Wellenlänge des Schalls und c_{eff} effektive Schallgeschwindigkeit (vom Empfänger aus betrachtet).

Beide Arten unterscheiden sich also in Bezug auf Quantität der Frequenzveränderung. Eine Bewegung des Senders kann interpretiert werden als Wellenlängenveränderung,

während die Empfängerbewegung auf eine Änderung der Schallgeschwindigkeit (bezüglich des Empfängers) hinausläuft.

Qualitativ erhöht sich die empfangene Frequenz immer bei einer Hinbewegung; eine Wegbewegung senkt die empfangene Frequenz dagegen ab. Nach dieser Aussage tritt also sofort eine Veränderung des Klangeindrucks ein, sobald sich der Zuhörer im Akustikraum bewegt. Der Effekt hält sich bei langsamen Bewegungen einigermaßen in Grenzen, so dass nur kleine Beeinträchtigungen bzw. Verfälschungen zu erwarten sind.

Zur Erzeugung eines besonderen Klanges ist in der Vergangenheit in der Studio- und Bühnentechnik der Doppler-Effekt praktisch genutzt worden. Dabei kommen sowohl sich bewegende Lautsprecher als auch Mikrofone zur Anwendung. Der zustande kommende Klangeffekt ist auch bekannt unter dem Begriff *Phasing*, was heute fast ausschließlich elektronisch erzeugt wird. Man gewinnt vom Originalsignal mehrere Nebensignale, die zeitlich gegeneinander verschoben sind. Dies allein kann einen Raumeffekt bewirken; unterliegen die Zeitverschiebungen einer ständigen Variation, so entsteht der Phasing-Effekt.

Ein unerwünschter Doppler-Effekt wird systembedingt von jedem Lautsprechersystem hervorgerufen, wenn es mehr als eine diskrete Frequenz wiedergeben soll. Wir nehmen an, dass ein Lautsprecher zwei Signale wiedergeben soll, die sich stark in ihrer Frequenz voneinander unterscheiden. Der tiefe Ton wird dabei die Lautsprechermembran zwar stark auslenken, jedoch geschieht dies mit der niedrigen Frequenz, so dass die Geschwindigkeit der Membran beim idealen Lautsprecher nicht von der Frequenz abhängt (bei gleichbleibender Ansteuerung). Gleichzeitig soll das System einen hohen Ton der Frequenz wiedergeben. Jedes mal, wenn der tiefe Ton die Membran um die Ruhelage bewegt, erreicht sie die höchste Geschwindigkeit. In dieser Situation ist sie als sich bewegender Schallsender zu betrachten, der den hohen Ton aussendet. Die wahrgenommene hohe Frequenz ergibt sich dann aus der abgewandelten Gl. 1.18:

$$f_h^* = f_h \cdot \frac{c_s}{c_s \pm 2\pi f_t A_t} \quad (1.20)$$

c_s Schallgeschwindigkeit, f_h hohe Signalfrequenz, f_t niedrige Signalfrequenz, A_t Membranamplitude bei Ansteuerung durch niedrige Signalfrequenz.

Die vorliegende *Frequenzmodulation* kann zu extrem verwaschenem Klang führen. Ausweg aus dieser Situation ist nur die Verwendung von *Mehrwegsystemen*, bei denen das Signal in verschiedene Frequenzbereiche aufgeteilt wird und daher jeder Lautsprecher nur einen Teilbereich des Spektrums verarbeiten muss. Dies bringt neben einer Verminderung des Doppler-Effekts auch eine Verbesserung der Eigenschaften der einzelnen Lautsprecher, die ein auf ihren Frequenzbereich zugeschnittenes Design erhalten können.

1.9 Räumliches Hören

Ein ebenfalls wichtiger Punkt bei der Vermittlung des Originalhöreindrucks ist die Reproduktion der Aufnahmearkustik mit all ihren räumlichen Differenzierungen. Hierzu gehört natürlich auch das *Richtungshören*, das uns die *Ortung einer Schallquelle* ermöglicht. Um dieses Vermögen auch bei einem übertragenen Signal (also künstlich wiedergegebenen) Signal zu erzielen, sind verschiedene Verfahren der *Stereofonie* entwickelt worden.

Die Philosophie all dieser Verfahren beruht darauf, der Originalakustik möglichst nahezukommen. Um dies zu erreichen, ist vor allem bei Mikrofonaufnahmen ein großes Maß an Kenntnissen der Akustiklehre und auch Erfahrung vorauszusetzen. Ferner ist bei künstlich erzeugten (oder veränderten) Audiosignalen, wie das ja mit dem Computer vielfach praktiziert wird, ebenfalls einiges Grundwissen erforderlich, um die gewünschte Klangatmosphäre herzustellen.

1.9.1 Intensitäts-Stereofonie

MS-Verfahren Ein sehr gebräuchliches Verfahren bildet eine Art der Intensitäts-Stereofonie, welches zwei Mikrofone mit unterschiedlichen Richtcharakteristika einsetzt, die möglichst nahe beieinander liegen müssen (*Koinzidenz-Mikrofone*). Das erste Mikrofon besitzt *Nierencharakteristik*, während das zweite *Acht-Charakteristik* zeigt und gegenüber dem ersten um 90° um die senkrechte Achse verdreht ist. Abb. 1.11 zeigt die Anordnung beider Richtcharakteristika.

Die MS-Technik setzt – daher der Name – durch die Mikrofonsignale ein Mitten- und ein Seitensignal zusammen. Dadurch ist auch eine Kompatibilität zu Einkanalsystemen leicht zu realisieren, indem man hier nämlich nur das M-Signal auswertet, welches vom ersten Mikrofon kommt. Das Seitensignal S des zweiten Mikrofons wird zur Erzeugung des L-Signals einerseits zum Mittensignal M addiert und zur Erzeugung des

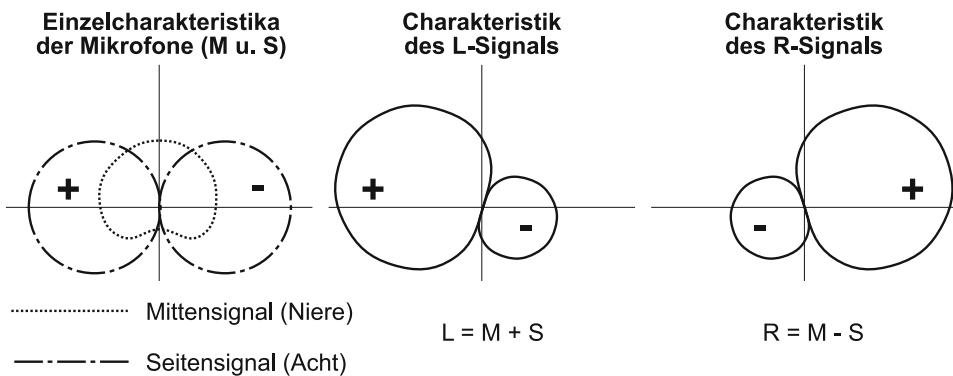


Abb. 1.11 System einer Mikrofananordnung bei der MS-Intensitäts-Stereofonie

R-Signals von ihm subtrahiert (soll die oben erwähnte Kompatibilität gewahrt bleiben, darf eine derartige Zusammensetzung natürlich erst am Ende der Übertragungstrecke bzw. beim Empfänger stattfinden). Daneben ist natürlich beachtenswert, dass das Mikrofon mit Acht-Charakteristik beide Richtungskomponenten in unterschiedlichen Phasenlagen wiedergibt, was durch die Vorzeichen angedeutet ist (diese sind nicht mit den Vorzeichen in den Synthesegleichungen zu verwechseln). Man kann sich das so vorstellen, dass eine von beiden Seiten zugängliche Membran eine dem Druckunterschied (Druckgradient) zwischen Vorder- und Hinterseite entsprechende Schwingung durchführt und mittels Schwingsspule eine Spannung induziert wird. Gleichartiger Schall – abwechselnd auf beide Hauptrichtungen gerichtet – bewegt die Membran jedoch gegenphasig und lässt somit auch gegenphasige Spannungen entstehen.

Aus der Abbildung ist ebenfalls zu entnehmen, welche Richtungscharakteristika die Synthesesignale L und R aufweisen. Sie sind offensichtlich vollkommen symmetrisch; außerdem zeigen sie durch die Anordnung der Mikrofone zueinander keinerlei Laufzeitunterschied. Die Signale L und R unterscheiden sich somit nur in ihrer Intensität, wenn eine Schallquelle von einer Seite her einwirkt.

XY-Verfahren Der Nachteil des MS-Verfahrens besteht darin, dass unterschiedliche Mikrofontypen verwendet werden müssen; außerdem ist die Auswahl der Mikrofone mit Acht-Charakteristik nicht sehr groß. Eine sehr einfache Methode, die lediglich zwei Kapseln mit Nierencharakteristik benötigt, ist die XY-Stereofonie. Beide Kapseln zeigen dabei in diametrale Richtungen. Abb. 1.12 zeigt eine solche Anordnung.

Die Einzelsignale beider Mikrofone werden direkt als Kanalsignale herangezogen. Auch hier ergeben sich keinerlei Laufzeitunterschiede, da die Mikrofone sich annähernd an der gleichen Stelle befinden. Selbstverständlich gibt es auch komplett Stereomikrofone, die bereits beide Kapseln in richtiger Anordnung beinhalten.

Sind ebenfalls Kompatibilitätsrichtlinien zu erfüllen, so lassen sich auch bei diesem Verfahren das Mittensignal durch $M = X + Y$ und das Seitensignal durch $S = X - Y$ generieren.

Optimale Voraussetzungen ergeben sich bei Lautsprecherwiedergabe, wenn die Lautsprecher und der Zuhörer etwa ein gleichseitiges Dreieck bilden.

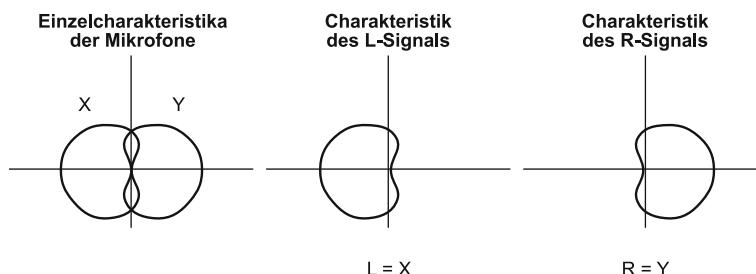


Abb. 1.12 XY-Anordnung und Verlauf der Richtungscharakteristika der Signale

1.9.2 Stereofonie durch Laufzeit

Zur Erzeugung ausgeprägter Raumakustiken und eindrucksvollerer Klangbildern wird oft eine Stereofonie mittels *Laufzeitunterschiede* der Mikrofonsignale gewählt. Dieser Laufzeitunterschied kann im einfachsten Fall dadurch realisiert werden, dass die Mikrofone einen Abstand von über 0,5 m zueinander aufweisen. Der Schall einer Quelle benötigt somit unterschiedliche Zeiten zu beiden Mikrofonen, was bei der Wiedergabe ebenfalls als Raumeindruck bzw. Richtungshören empfunden wird.

Der Aufbau entspricht dem in Abb. 1.13, jedoch werden eben beide Mikrofone räumlich voneinander getrennt. Dadurch entstehen bei einem einseitig einfallenden Schallsignal nicht nur Pegel-, sondern auch Laufzeitunterschiede.

Der Laufzeitunterschied kann ohne weiteres auch auf elektronischem Wege realisiert werden (analoge oder digitale Verzögerungsstrecken). Vorteilig ist dann eine schnelle und einfache Änderung der Laufzeit selbst, ohne an der Position der Mikrofone etwas verändern zu müssen.

Bei der Wiedergabe unterscheidet sich das Laufzeitverfahren hauptsächlich dann von der Intensitäts-Stereofonie, wenn keine Lautsprecher-, sondern Kopfhörerwiedergabe gewählt wird.

1.9.3 Kunstkopf-Stereofonie

Das neueste Verfahren der räumlichen Aufnahme und Wiedergabe stellt die Kunstkopf-Stereofonie dar. Die Grundidee hierbei ist die Verwendung eines Mikrofons, das den Eigenschaften des menschlichen Kopfes und seines Gehörorganes möglichst genau entspricht. Ein Kopf aus Kunststoff dient deshalb der geometrischen und klanglichen Nachbildung des menschlichen Kopfes, wobei ebenfalls (vereinfachte) Gehörgänge

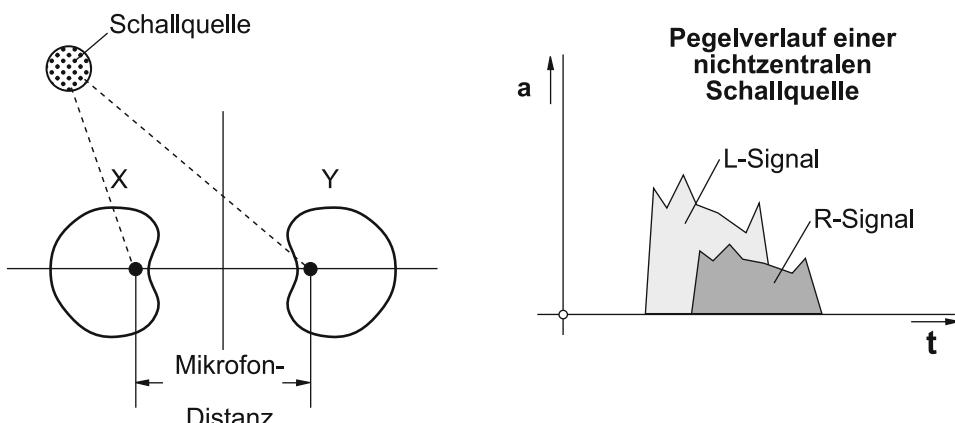


Abb. 1.13 Aufbau und Signale bei der Laufzeit-Stereofonie

ausgebildet sind, an deren Enden jeweils eine Mikrofonkapsel (Druckempfänger) angeordnet ist.

Eine auf diese Weise aufgezeichnete Übertragung erfordert natürlich die Wiedergabe über einen Kopfhörer, damit das Originalsignal nicht durch Reflexionen, Laufzeiten usw. verfälscht wird. Der Höreindruck ist systembedingt demjenigen sehr ähnlich, den auch ein Zuhörer in der Nähe des Kunstkopfes hatte. Das Richtungshören ist ausgezeichnet – so können z. B. Schallquellen auch oberhalb und hinter dem Zuhörer geortet werden. Lediglich in der Meridianebene gibt es teilweise Ortungsprobleme; ebenfalls treten gelegentlich sog. *Im-Kopf-Lokalisationen* auf. Der Grund dafür könnte sein, dass das ursprüngliche Hörerlebnis mit einem ständigen Kopfbewegen verknüpft ist – der Klangindruck hängt dann nicht nur mit der physikalischen Aufnahme des Schalls zusammen, sondern auch mit der Fähigkeit des Gehirns, aus der Empfindung des Organs zusammen mit den *Kopfbewegungen* ein Klangbild formen zu können (ähnliches geschieht ja auch bei visuellen Eindrücken). Ein interaktives Hören, bei dem sich sozusagen im Nachhinein bei Kopfbewegungen des Hörers auch der Kunstkopf bewegt oder sich sein Signal entsprechend ändert, ist derzeit noch nicht möglich.

1.9.4 Quadrofonie

Eine besondere Form des räumlichen Hörens bildet die Quadrofonie. Sie bezieht insbesondere den *Reflexionsschall* bei der Aufnahme mit in das zu speichernde Signal ein. Somit ergeben sich vier zu übertragende Kanäle, die auch über vier Lautsprechersysteme wiedergegeben werden (*Matrixverfahren* können jedoch auch eine quadrofone Übertragung via zwei Kanälen ermöglichen). Der Hörer befindet sich dabei im Zentrum.

Die Quadrofonie hat sich aus verschiedenen Gründen nicht besonders durchsetzen können. Zum einen liegt dies vielleicht an der recht kritischen Position des Zuhörers, bei dem sich der Klangindruck wesentlich verändert, sobald er sich nur wenig vom optimalen Punkt entfernt. Dann hängt es ferner damit zusammen, dass das Gros der Konsumenten den größeren Aufwand der Lautsprecheranordnung nicht betreiben will. Das größte Handikap jedoch besteht wohl in der nur schwer erfüllbaren Kompatibilitätsforderung mit den herkömmlichen Stereooverfahren.

1.9.5 Dolby Surround-Sound 5.1 (Dolby Digital)

Bei dem von Dolby entwickelten 6-kanaligen Raumklangsystem sind laut Abb. 1.14 bei der Wiedergabe zwei Lautsprecher für die vorderen beiden Seiten (L und R), einer für das Mittensignal (Center), zwei für die rückwärtigen Seiten (Rear L und R) sowie ein Subwoofer für Effekte im Niedrigfrequenzbereich nötig. Letzterer wird häufig zur Vereinfachung weggelassen – falls die beiden Seitenlautsprecher genügend Tiefenreserve

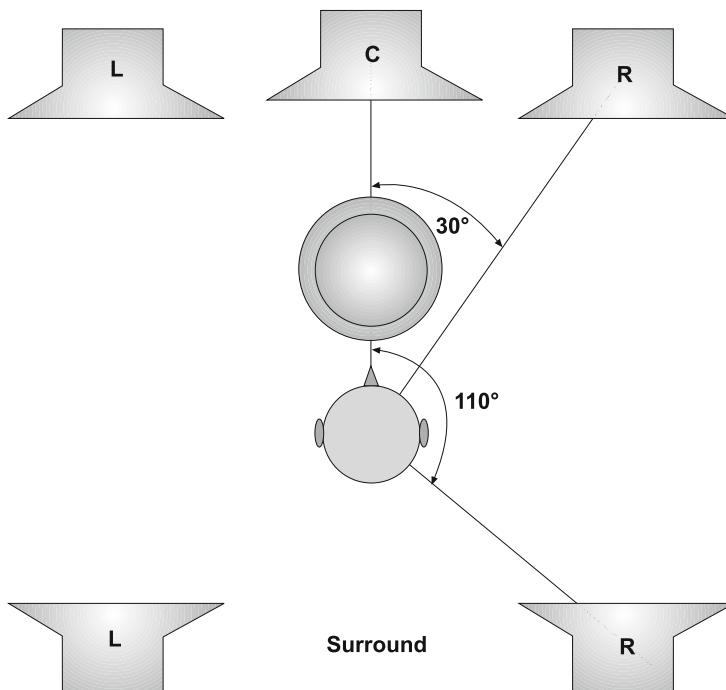


Abb. 1.14 Lautsprecher-Anordnung bei Dolby Surround

besitzen, können diese bei geeignetem Verstärker/Decoder die Aufgabe des Subwoofers übernehmen.

In kommerziellen Kino-Systemen sind üblicherweise zur Verteilung des Klangs nicht nur einzelne Lautsprecher aufgestellt, sondern ganze Bänke.

Der Begriff *Dolby Digital*, der häufig auch für diesen Sachverhalt fällt, bedeutet nicht notwendigerweise eine solche Auflösung in der genannten Weise, sondern steht nur als Oberbegriff, der eine echte 5.1-Wiedergabe erst möglich macht (bei Dolby-Systemen).

In der Begriffsvielfalt begegnet man in diesem Zusammenhang auch häufig dem Ausdruck *AC3*, welcher jedoch ein Audiocodec ist für verlustbehaftete Kompression, die für die Abspeicherung der Daten zum Einsatz kommt.

Viele Geräte im Heimaudio- bzw. Heimvideo-Bereich können heute mehrkanalige Audioformate decodieren. Auch und vor allem im PC-Bereich sind Mehrkanal-Soundsysteme schon fast zum Standard geworden. Hierzu gehört natürlich auch die Unterstützung auf Software-Seite durch Programme wie Cubase oder ProTools. Doch auch einfache Freeware-Programme wie Audacity sind in der Lage, über einen eigenen Codec mehrkanaliges Audio zu unterstützen. Siehe hierzu auch Abschn. 3.7.

1.9.6 Dolby Pro Logic

Während das oben beschriebene Dolby Surround die sechs Audiokanäle diskret oder innerhalb eines Datenformats gepackt vorliegen hat, arbeitet das einfache Dolby Pro Logic aus dem Jahre 1987 anders. Es existieren hier nämlich lediglich zwei Audiosignale wie bei einem herkömmlichen Stereosignal. Um es an dieser Stelle gleich vorwegzunehmen: Pro Logic macht keinen Surround-Sound aus einem herkömmlichen Stereosignal. Das Verfahren ist lediglich in der Lage, zusätzliche Surround-Information in ein normales Stereosignal hineinzupacken. Dies geschieht durch eine spezielle Matrix-Technik, die aus Abb. 1.15 hervorgeht.

Bei dieser Anordnung ist eine kompatible 2-Kanal-Stereo-Wiedergabe einfach über die beiden Signale **Lt** und **Rt** möglich, weil durch die 180° -Phasenlage der partiellen Surround-Signale eine Auslöschung im Raumzentrum stattfindet. Soll hingegen tatsächlich eine Surround-Wiedergabe stattfinden, so gelangen die Signale **Lt** und **Rt** auf eine Subtrahierstufe, und der Phasenwinkel von 180° zwischen den Signalen bewirkt eine Extrahierung des Surround-Anteils und eine komplette Auslöschung des normalen Stereo-Mittensignals (Basisbreitensignal). Die unabhängigen Anteile im Stereoraum werden allerdings nicht unterdrückt.

Formal betrachtet arbeitet das Verfahren recht einfach. Man darf dabei aber nicht übersehen, dass die Generierung der 90° verschobenen Signale keinesfalls trivial ist. Schließlich muss die Phasenverschiebung unabhängig von der Frequenz sein. Die Implementierung erfordert mehr Aufwand, denn eine einfache Phasenschieberkette würde eine starke Amplitudenabhängigkeit von der Frequenz aufweisen. Die Lösung wäre eine Aufteilung in viele Frequenzbereiche mit jeweils einem Phasenschieber oder eine Anordnung, die *Hilbert-Transformator* heißt. Letzteres lässt sich per Software an einem

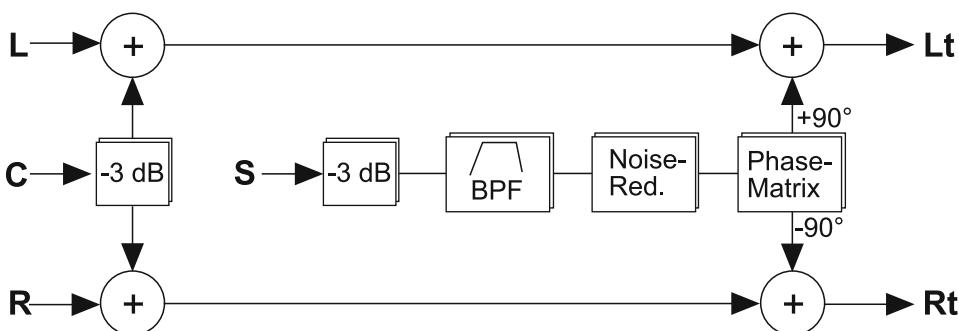


Abb. 1.15 Encoder für Dolby Pro Logic. Das Center-Signal wird anteilsrichtig zu beiden Seitensignalen hinzu addiert, während das Surround-Signal nach der Abschwächung zunächst ein Bandpassfilter und dann ein dynamisches Rauschfilter passiert. Die Matrix splittet dann in zwei Signale, die zum vorherigen Signal um 90° bzw. -90° in der Phase verschoben sind. Somit sind beide die Matrix verlassenden Signale um 180° gegeneinander in der Phase verschoben. Das endgültige zweikanalige Signal ist dann **Lt** und **Rt**.

zeitdiskreten (also digitalen) Signal einfach anwenden. Diese aufwändige oder komplexe Schaltung ist allerdings nur auf Seite der Signalgenerierung, also dem Encoder nötig. Der Decoder kommt prinzipiell mit wenigen Bauteilen aus.

Es gibt neben „passiven“ Decodern gibt es auch solche, die dynamisch arbeiten. Man stelle sich vor, die Anteile des Center-Signals sind ungleich groß, so gibt es auch ein Differenzsignal, das von null verschieden ist. Und dies bedeutet konsequenterweise, dass diese Differenz als Übersprechen im Surround-Kanal erscheint. Um dem zu begegnen, hat man dynamisch arbeitende Decoder entwickelt, die gleichphasige Pegel in beiden Kanälen durch VCAs angeleichen. Das Übersprechen von etwas seitlich liegenden Center-Signalen auf den Surround-Kanal ist somit verhindert.

Es gibt bei Pro Logic nur ein monaurales Surround-Signal. Außerdem ist dieses – wegen des technischen Aufwands als Analogschaltung – bandbegrenzt bei 7 kHz.

1.9.7 Dolby Pro Logic II

Der oben genannten Bandbegrenzung unterliegt Pro Logic II nicht. Ein weiterer großer Fortschritt ist die Realisierung von zwei Surround-Kanälen, so dass die rückwärtige Klangraum ebenfalls eine Stereo-Aufteilung erfährt. Das Herzstück ist eine modifizierte Matrix, die in Form eines Zeigerdiagramms aus Abb. 1.16 hervorgeht.

Neben der Matrizierung der Einzelsignale befinden sich im Decoder auch noch Verzögerungsglieder für die Surround-Kanäle, so dass deren Signal stets um 10 ms gegenüber den Frontsignalen später kommen. Dies entspricht einem Abstand von ca. 3,3 m.

Weitere Verbesserungen gegenüber dem einfachen Pro Logic sind drei Modi, die verschiedene Einstellmöglichkeiten bieten:

- Dimension Control. Diese Möglichkeit erlaubt die Verschiebung des Klangfelds nach vorne oder hinten. Somit lassen sich auch Aufstellungsgeometrien der Lautsprecher in gewissen Grenzen ausgleichen.
- Center Width Control. Der Einstellvergang verlagert den tatsächlichen Anteil des Center-Signals zu den beiden Seitensignalen, so dass diese Lautsprecher eine virtuelle Mitte erzeugen.
- Panorama Mode. Durch diesen Einstellmodus kann der Surround-Effekt intensiviert werden.

Viele Geräte mit der Technologie von Pro Logic bieten ferner noch spezielle durch DSP gestützte Effekte wie z. B. Stadium, Church, Hall oder dergleichen. Diese Zusätze haben jedoch mit dem eigentlichen Dolby-Verfahren nichts zu tun.

Pro Logic IIx Dieses stark erweiterte Verfahren ermöglicht die Bereitstellung von zwei zusätzlichen Surround-Kanälen. Somit kann es Formate mit bis zu 7.1 Kanälen codieren

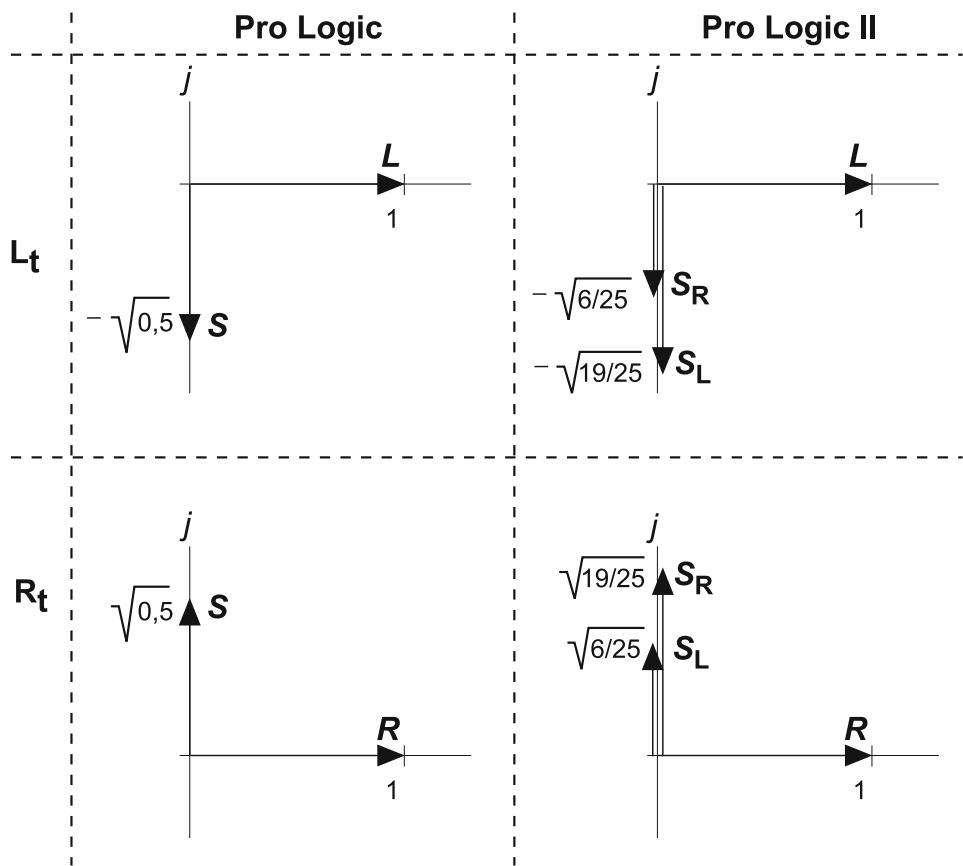


Abb. 1.16 Zeigerdiagramme als Vergleich zwischen Pro Logic und Pro Logic II. Bei Pro Logic gibt es nur einen Surround-Kanal. Seine Signalanteile gelangen einmal mit -90° und einmal mit $+90^\circ$ auf die L_t - (Left total) bzw. R_t -Zweige (Right total). Hingegen liefert Pro Logic II zwei Surround-Signale, welche mit unterschiedlichen Anteilen auf beide Zweige gelangen. Mit entsprechenden Pegelanhebungen und -absenkungen im Decoder lassen sich die Anteile wieder separieren. Eine einfache Addition beider Signale hingegen lässt die Surround-Signale vollständig auslöschen im Center-Kanal

und decodieren. Es ist auch in der Lage, digitale Formate (also Dolby Surround) zu decodieren.

Pro Logic IIz Gegenüber dem IIx-Verfahren wird hier die Kanalzahl weiter erhöht auf 9.1. Im Detail zeigt sich dies durch zusätzliche zwei Kanäle im vorderen Bereich, nämlich einen linken und einen rechten Kanal, die höher angeordnet sind als die Standard-Stereo-Lautsprecher.

1.9.8 DTS

Auf dieses Verfahren, dessen Name ein Kürzel von *Digital Theater Systems* ist, soll nur ganz kurz eingegangen werden. Prinzipiell liegen wie bei Dolby Digital Daten in digitaler Form vor, und zwar für jeden der Kanäle separat. Die Aufbereitung erfolgt mittels verlustbehafteter Kompression. Die Standardversion des Verfahrens verarbeitet wie Dolby Digital 5.1 Kanäle, also fünf diskrete Lautsprecherpositionen sowie einen Subwoofer-Kanal. Abwandlungen allerdings können auch mehr Kanäle wiedergeben (z. B. DTS-HD für High-Definition-Bildformate). Es gibt auch DTS-Verfahren – DTS-ES Matrix 6.1 und DTS NEO:6 –, die ähnlich Dolby Pro Logic mit einer Matrix die Kanäle verkoppeln.

Die höherwertigen Verfahren von DTS kommen meist ausschließlich in der Kinotechnik zur Anwendung. Hier kommt teilweise auch ein sog. Effektkanal zum Tragen, mit dem man bei der Vorführung spezielle Lichteffekte oder dergleichen steuern kann.

DTS- und Dolby-Systeme sind grundsätzlich untereinander nicht kompatibel. Natürlich gibt es im Consumer-Bereich Wiedergabegeräte, die beide Systeme korrekt verarbeiten können.

Neben autarken Geräten für DTS gibt es natürlich auch zahlreiche Programme und Plug-ins für PCs, um die Formate importieren und das Material weiterverarbeiten zu können, so z. B. AC3filter.

1.9.9 Wellenfeld-Synthese

Ein relativ aufwendiges Verfahren ist die Wellenfeld-Synthese. Bei der Wiedergabe sind viele Einzellautsprecher mit separaten Ansteuersystemen nötig, sodass ein solches Verfahren für den Heimbereich kaum praktikabel ist. Das Prinzip ist relativ einfach und ähnelt dem der Holografie: Die Klangquellen, die an verschiedenen Stellen des (virtuellen) Raumes platziert sind, senden eine Wellenfront aus, deren Zentrum die Quellen selbst darstellen. Eine linear in einer Ebene platzierte Lautsprecherreihe muss nun möglichst exakt diese Wellenfront nachbilden. Hierzu muss jeder Lautsprecher ein Signal definierter Phasenlage abgeben. Richtig angesteuert, können diese Lautsprecher jede x-beliebige Klangquelle und deren Wellenfront abbilden, wie schematisch in Abb. 1.17 zu erkennen ist.

Das Huygenssche Prinzip besagt, dass jeder Punkt einer Wellenfront zu einem Ausgangspunkt einer neuen Wellenfront führt.

Solange Lautsprecher lediglich auf einer Ebene angeordnet sind, wird die Wiedergabe auch nur originalgetreu in einer Ebene erfolgen können. Allerdings kann der Zuhörer sich frei auf dieser Ebene bewegen, ohne dass Beeinträchtigungen zu erwarten sind. Vom Idealzustand abweichend ist natürlich auch die Lückenhaftigkeit der Lautsprecher, was sich als geringfügige Pegeleinbrüche äußert. Dies ist jedoch vom menschlichen Ohr kaum zu bemerken. Ein weiteres Handicap ist die begrenzte Ausdehnung der Lautsprecherfront.

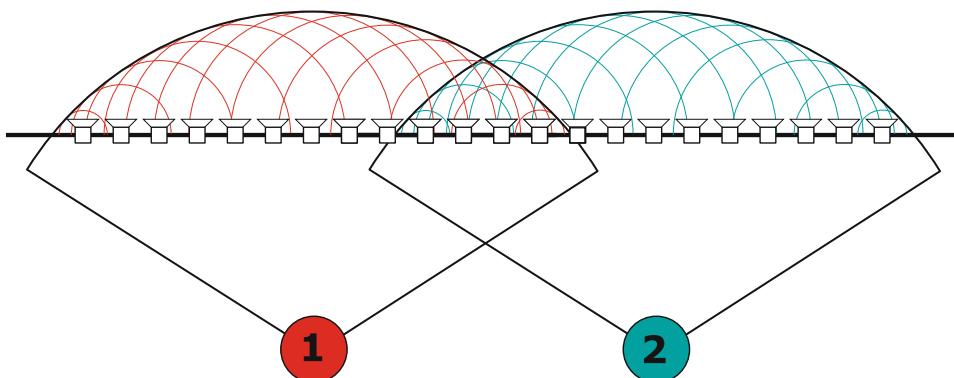


Abb. 1.17 Oberhalb der Lautsprecherwand entstehen Wellenfronten der beiden Phantomschallquellen (1) und (2) so, als ob diese hinter der Wand positioniert stünden. Die Lautsprecher verarbeiten pro Schallquelle ein Signal mit individueller Phasenlage. Würden alle dasselbe Signal abgeben, entstünde eine geradlinige Wellenfront mit einem virtuellen Zentrum im Unendlichen

Als Folge davon entsteht im Außenbereich ein Abbruch der Wellenfront, die sich als eine Art „Schattenwelle“ äußert.

Bei der Aufnahme sind die einzelnen Instrumente bzw. Stimmen normalerweise möglichst steril zu halten, also ohne jegliche Raumakustik. Diese sowie die gewünschte Position im virtuellen Raum besorgt ein Programm (z. B. *Wonder* oder *WFSCollider*), welches außerdem alle Signale für die einzelnen Lautsprecher generiert.

Bei der Wiedergabe muss der Raum möglichst neutral und ohne Reflexionen sein, da andernfalls störende künstliche Wellenfronten entstünden, die das gesamte Schallbild zerstören würden.

1.9.10 Weitere Raumklang-Systeme – Raumklang mit nur zwei Lautsprechern

Obwohl es sich beim *SRS TruSurround XT* nicht wirklich um ein Verfahren mit Rauminformationen handelt, ist es dennoch geeignet, eine Raumimpression zu erzeugen. Die Wirkungsweise beruht auf psychoakustischen Effekten, die wiederum in der Tatsache begründet sind, dass Signale mit bestimmten spektralen Bereichsanhebungen durch das Gehirn an spezifische Richtungen im Raum zugewiesen werden.

Die spektrale Raumzuordnung nach Abb. 1.18 geschieht nach den *Blauertschen Bändern*, da die Grundlagen hierzu bereits in den 70er-Jahren von *Jens Blauert* erforscht wurden. Die „Fähigkeit“ des Gehirns, bestimmte Spektralgebiete einer Raumrichtung zuzuordnen, entsteht durch Erfahrung. Alle von außen kommenden Klänge durchlaufen bis zum Mittelohr eine richtungsabhängige Filterung, die man sich durch den geometrischen Aufbau des Außenohres und des Gehörgangs erklären kann. S. 42 [9].

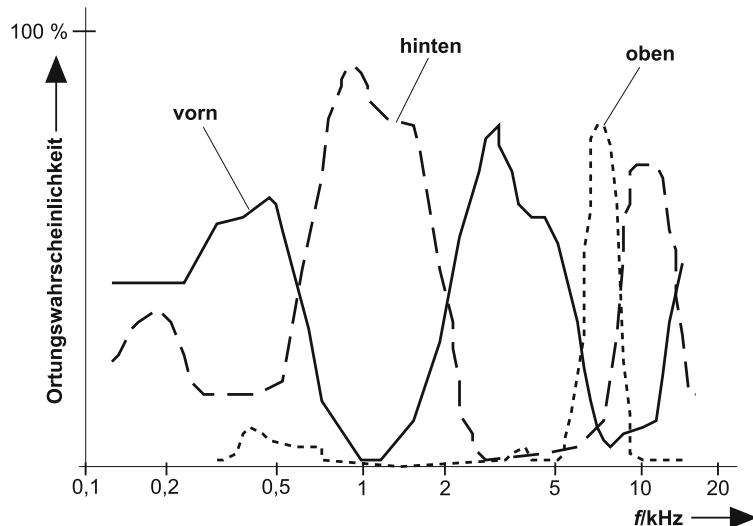


Abb. 1.18 Blauertsche Bänder. Audioversuche mit Personen führen zu dem Ergebnis, dass bestimmte Frequenzbereiche an verschiedenen Stellen des Raum-Meridians bevorzugt geortet werden. Eine über 50-%-ige Trefferquote bedeutet dabei bereits eine hohe Sicherheit, dass die Ortung bei der Mehrheit der Probanden übereinstimmend ist. Eine seitliche Ortung ist aus Symmetriegründen verständlicherweise nicht möglich

Am eindrucksvollsten arbeitet dieses Verfahren, wenn man an einem Rauschsignal die besagten Spektralbereiche um maximal 6 dB anhebt. So kann man also Rauschklänge erzeugen, die der Hörer als von vorne, von hinten oder von oben kommend einordnet. Man kann es auch auf ein allgemeines Audiosignal anwenden, allerdings muss man dabei beachten, dass die ursprünglichen spektralen Anteile auch zu den Anhebungen passen.

1.10 Speicherung von Audiosignalen

Wir wollen hier noch ganz kurz verschiedene, derzeit aktuelle Speicher- bzw. Aufzeichnungsverfahren von Audiosignalen erwähnen. Da sich in diesem Bereich in den letzten Jahren sehr viel getan hat und künftig noch entwickeln wird, sollte man bei der Tonverarbeitung am Computer einige grundlegende Dinge kennen. Es sollen hier die wichtigsten Systeme und deren Vor- und Nachteile genannt werden. Das physikalische Prinzip soll dabei nur ganz oberflächlich zur Sprache kommen.

1.10.1 Analoge Magnetbandaufzeichnung

Dieses Verfahren der *Tonspeicherung* ist auch heute noch sehr verbreitet und wird auch in Studios noch viel eingesetzt.

Arbeitsweise Der Tonträger ist eine ferromagnetische Schicht, die sich auf einem Kunststoffband befindet. Magnetköpfe, die mittels feinem Spalt Magnetfeldlinien austreten lassen können, magnetisieren diese Schicht nach den analogen Kriterien der Signalspannung. Auf umgekehrte Weise „liest“ ein Magnetkopf die magnetisch gespeicherte Signalinformation. Es existieren zwei verschiedene Ausführungen: der Kassettenrecorder und das Spulengerät. Während ersterer fast ausschließlich im Consumer-Bereich zu finden ist, sind die hochwertigen Spulengeräte hauptsächlich für professionelle Studioanwendungen gedacht.

Grenzdaten Bedingt durch das Medium „Band“ lässt sich der Klirrfaktor lediglich auf ca. 0,3 % drosseln, und der Dynamikbereich liegt mit etwa 60 dB auch nicht besonders hoch (bedingt durch das Rauschen des Tonträgers; durch *Rauschunterdrückungsverfahren* wie z. B. *Dolby C* kann der Dynamikbereich noch um ca. 20 dB gesteigert werden). Bei genügend hohen Bandgeschwindigkeiten und optimalem Bandmaterial ist ein Frequenzgang bis 20 kHz bei Abweichungen von weniger als 2 dB zu schaffen. Bei guten Geräten beträgt der Gleichlauf weniger als 0,1 %. Die Kanaltrennung liegt bei Werten >45 dB. Diese Daten (vor allem Frequenzgang, Dynamik und Klirrfaktor) jedoch sind nur bei abgeglichenem und eingemessenem Gerät erreichbar; eine regelmäßige Wartung ist im Studiobetrieb daher unumgänglich.

Eigenschaften des Tonträgermediums Die Verwendung von *Kassetten* ist recht komfortabel, dagegen ist ein Umgang mit Spulen ziemlich mühsam – ein Aufsuchen bestimmter Stellen ist umständlich und zeitaufwendig. Das Tonträgermaterial ist empfindlich gegenüber Wärme, magnetischen Feldern und mechanischer Beanspruchung. Das Abspielen bedeutet *Abrieb* und somit eine kontinuierliche Abnahme der Tonqualität. Letzteres empfiehlt daher das Anlegen von Master-Aufzeichnungen und mehreren Arbeitskopien, wenn das Material lange seine Tonqualität aufrechterhalten soll. Daneben treten auch *Kopiereffekte* auf, d. h. das Durchdringen der Magnetisierung auf naheliegende Bandwindungen, was besonders bei warmer Lagerung fortschreitet. Nach einiger Zeit ist generell ein Nachlassen der Magnetisierung festzustellen. Die Spieldauer beträgt für Kassetten 60 oder 90 min (C60 oder C90); es sind jedoch Sonderausführungen bis 120 min erhältlich. Spulentonbänder sind als Wickel bis 26,5 cm erhältlich. 18er-Spulen mit Doppelspielband haben eine Durchlaufzeit (einseitig) von ca. 60 min bei einer Bandgeschwindigkeit von 19 cm/s.

1.10.2 Digitale Magnetbandaufzeichnung

Grundsätzlich stehen bei der digitalen Aufzeichnung auf ein Band zwei Methoden zur Verfügung, nämlich das *R-DAT-System* und das DCC-System. In Studios sind vor allem DAT-Geräte zu finden.

Arbeitsweise R-DAT (Rotary Digital Audio Tape) ist von der Mechanik und der Aufzeichnungsart mit Videoaufzeichnung zu vergleichen. Eine schräg stehende, rotierende Kopftrommel sorgt bei relativ langsamem Bandvorschub für eine hohe effektive Schreibgeschwindigkeit. Diese ist auch notwendig, denn der zu verarbeitende digitale Datenstrom beträgt 7,5 Mbit/s. Es kommen jedoch nicht nur die reinen PCM-Daten der Audiokanäle zur Aufzeichnung, sondern auch Redundanzinformation für die *Fehlerkorrektur* (wichtig bei Dropouts des Datenträgers) sowie Zeitcode- und User-Daten. Das Auslesen der Daten erfolgt mit einem Quarztakt, was durch Servosteuerung des Bandvorschubs erreicht wird. S. 42 [7].

DCC (*Digital Compact Cassette*) setzt – wie der Name bereits andeutet – eine Compact-Kassette ein, welche jedoch spezielles Bandmaterial beinhaltet. Im Gegensatz zu DAT handelt es sich hier um feststehende Köpfe, und der Bandvorschub beträgt wie beim Analog-Kassettendeck 4,76 cm/s. Diese Köpfe verfügen über 9 Spalte, wovon 8 für die Audiodaten genutzt werden. Dadurch wird die mögliche Datenrate erhöht. Zur Aufzeichnung kommt hier jedoch nicht der direkte PCM-Audio-Code, sondern Daten, die durch das System *PASC* (*Precision Adaptive Sub-Coding*) umgewandelt wurden. Dieses Verfahren vermindert die anfallende Datenmenge auf den Wert 384 kbit/s. Die Datenreduktion basiert auf der Tatsache, dass das menschliche Ohr viele Signalkomponenten überhaupt nicht wahrnehmen kann (*Überdeckung*, Trägheit usw.). Nach der Digitalisierung des Audiosignals wird über den PASC-Encoder eine Bandaufteilung vorgenommen und die genannte Datenschrumpfung durchgeführt. Leise und laute Stellen können dabei mit derselben Auflösung verarbeitet werden, weil von einzelnen PASC-Rahmen ein Skalenfaktor gewonnen und mitaufgezeichnet wird. In Bezug auf Datenredundanz, Zeitcode und User-Daten ist DCC ebenfalls ausgestattet. Das Auslesen der Daten erfolgt im Quarztakt durch eine Servosteuerung des Bandvorschubs.

Grenzdaten Beide Systeme zeigen etwa dieselben erreichbaren Qualitätsstandards. Der Klirrfaktor liegt unter 0,01 % und die Dynamik sowie die Kanaltrennung über 90 dB. Beide Werte sind sozusagen systembedingt und hängen auch nicht von Einstellungen ab (Justagen dienen hier nur der Systemsicherheit); Eigenschaften analoger Einheiten können i. Allg. nur geringen Einfluss auf die Eckdaten nehmen. Der Frequenzgang entspricht dem Bereich 20 ... 20.000 Hz innerhalb einer Toleranz von $\pm 0,5$ dB. Die Gleichlaufschwankungen sind bei digitalen Systemen generell unterhalb der Messgrenze, da die Wiedergabe im Quarztakt erfolgt.

Eigenschaften des Tonträgermediums Grundsätzlich sind dieselben Abhängigkeiten und Einflüsse wie beim Analogband zu erwarten. Jedoch werden Effekte wie Kopiermagnetisierung oder allmähliche Entmagnetisierung i. Allg. keine Auswirkungen haben, weil messbare Veränderungen normalerweise unterhalb der Codeschwelle liegen, d. h. der Digitalcode kann nach wie vor richtig identifiziert werden. Aus dem gleichen Grunde sind digitale Überspielungen auch verlustfrei. Bandbeschädigungen und die daraus resultierenden Datenverluste können bis zu einem gewissen Grade durch eine Fehlerkorrektur kompensiert werden. Sowohl DAT als auch DCC verwenden ausschließlich Kassetten, deren Handhabung einfach und sicher ist. Trotzdem sollte das Material natürlich vor Wärme, mechanischen und magnetischen Einflüssen geschützt werden. DAT-Kassetten gibt es für eine maximale Spieldauer von 120 min (R 120), und DCC-Kassetten sind mit einer Spieldauer bis 90 min (D 90) erhältlich.

1.10.3 Digitale Aufzeichnung auf MiniDisc (MD)

Die MiniDisc stellt eine Neuerung der Aufzeichnungsmethoden dar. Sie ist eine Weiterentwicklung von Speichermedien für Computer. Der Einsatz reicht von Consumer-Geräten bis hin zum Studio.

Arbeitsweise Das Trägermaterial befindet sich auf einer Scheibe, die durch ein Gehäuse geschützt ist. Der Aufbau ähnelt damit stark den Computer-Disketten mit 3,5 Zoll Größe. Das Aufzeichnungsverfahren bedient sich magneto-optischen Vorgängen; mit einem Laser wird eine diskrete Stelle der sich drehenden Disc auf ca. 220 °C erhitzt, was die Koerzitivkraft so stark herabsetzt, dass der Magnetkopf genau hier eine Magnetisierung des Filmes erzwingt. Die Schreibdichte hängt also vor allem von der punktuellen Erhitzung des Lasers und nicht durch ein konzentriertes Magnetfeld ab.

Die Magnetisierung bewirkt Veränderungen der Reflexionseigenschaften im Trägermaterial. Bei der Wiedergabe erfolgt nun eine Abtastung durch den Laserstrahl, dessen zurückgeworfenes Licht auf Fotosensoren gelangt und danach verstärkt wird.

Grenzdaten Aufgrund der ähnlichen Datenaufbereitung zum DAT- bzw. CD-System entspricht die Audioqualität der MD den gleichen Grenzwerten.

Eigenschaften des Tonträgermediums Die MiniDisc ist recht unempfindlich und zugleich komfortabel in der Handhabung. Die geometrische Form (Scheibe) lässt ein schnelles Auffinden einer bestimmten Stelle zu. Die Suchzeit ist dabei vergleichbar mit der bei CD-Playern und beträgt ca. 2 s. Eine Empfindlichkeit gegenüber Magnetfeldern besteht im Normalfall nicht, somit ist auch nicht mit einer Entmagnetisierung bzw. Veränderung der magneto-optischen Schicht zu rechnen. Vor Hitze ist das Speichermedium jedoch zu schützen. In Bezug auf die Spielzeit (max. ca. 80 min) erreicht die MD eine sehr kompakte

Bauweise (ca. 70 mm Seitenmaß). Die Abtastung erfolgt optisch, also berührungsfrei und bewirkt somit keinerlei Abrieb.

1.10.4 Hard-Disk-Recording

An dieser Stelle soll die Möglichkeit der Aufzeichnung von Audiosignalen per Computer bereits erwähnt werden. Wir werden jedoch noch genauer in den nächsten Kapiteln darauf zu sprechen kommen.

Arbeitsweise Eine im Computer eingebaute Audiokarte (oder eine auf dem Mainboard integrierte Einheit) ermöglicht das Digitalisieren des Audiosignals. Die PCM-Daten, die i. Allg. ohne Fehlerschutzcode zur Aufzeichnung kommen, können mit einem geeigneten Programm editiert werden. Welches genaue Format gespeichert wird, hängt von der Hardware und insbesondere von der Software ab, die das Hard-Disk-Recording durchführen (siehe auch Kap. 4). Auf diese Weise ist Schnittbetrieb mit allen Komforts möglich. Als Speichermedium dient die Festplatte, die meist als permanent rotierende Mehrfach-Scheibe mit magnetisch beschreibbarer Beschichtung ausgeführt ist.

Grenzdaten Die Werte für die Audioqualität ergeben sich aus der Qualität der A/D-Wandlung auf der Karte. Erreichbar sind grundsätzlich die gleichen Standardgrenzwerte wie bei den Systemen DAT, MD und DCC.

Eigenschaften des Speichermediums Da eine Festplatte ständig rotiert, ist die Zugriffszeit relativ kurz (Größenordnung 10 ms). Ein Auffinden eines Punktes innerhalb einer Audiodatei jedoch hängt auch von der Geschwindigkeit des Programms ab. Zur Archivierung werden die Audiodateien meist auf wechselbare Platten oder kleinere Passagen auch teilweise auf Disketten kopiert. Eine Alterung ist unter normalen Bedingungen kaum festzustellen; mitunter kommt es jedoch zu Totalausfällen der Festplatte, was meistens bedeutet, dass die elektronische Steuerung Funktionsstörungen zeigt. Hitze, mechanische und magnetische Einwirkungen sind auch bei diesem Medium zu vermeiden. Die Spieldauer hängt von der Speichergröße der Festplatte ab; diese kann mehrere Gigabyte betragen.

1.10.5 Compact Disc (CD)

Obwohl das Medium Compact Disc standardmäßig nicht wiederbeschreibbar ist (außer beim CD-RW-Träger), wollen wir es hier kurz erwähnen, weil die Herstellung von Kleinserien heute schon recht billig angeboten wird.

Arbeitsweise Die Herstellung des Datenträgers hängt u. a. von der Stückzahl ab. Bei der Massenfertigung handelt es sich um einen recht komplizierten Prozess. S. 42 [3, 6].

Die Datenaufbereitung erfolgt digital, wobei bei der Aufzeichnung wie beim DAT Redundanz-Daten integriert werden, damit bei der Wiedergabe eine Fehlerkorrektur möglich ist. Ferner sind die Informationen Zeitcode, Spurnummer und Indexnummer enthalten. Die Abtastung erfolgt berührungslos mit einem Laser; gelangt der Strahl auf die spiegelnde Oberfläche der Informationsseite der Disc (Unterseite), so erfolgt Reflexion, beim Auftreffen auf sog. Pits dagegen wird der Laserstrahl gestreut. Die auf diese Weise generierte Helligkeitsmodulation wird mittels Fotosensor in Datenströme gewandelt, welche verstärkt und digital aufbereitet werden (Synchronisation, Fehlerkorrektur, Rückwandlung zum Analogformat).

Grenzdaten Aufgrund der ähnlichen Datenaufbereitung zum DAT- bzw. CD-System entspricht die Audioqualität der CD den gleichen Grenzwerten.

Eigenschaften des Tonträgermediums Die CD ist relativ unempfindlich und zugleich einfach in der Handhabung. Das Auffinden einer bestimmten Stelle dauert i. Allg. höchstens ca. 2 s. Da eine berührungslose (optische) Abtastung erfolgt, gibt es keinen Verschleiß. Vor Hitze ist das Speichermedium jedoch zu schützen. Die Spielzeit beträgt maximal 80 min.

Betriebsarten als CD-ROM In der Betriebsart 1 liegt eine hauptsächliche Nutzung der CD-ROM als Programmsspeicher vor. Dies setzt voraus, dass eine gewisse Datenredundanz vorliegen muss, damit Fehler im Datenträger (Pressfehler, Kratzer, Schmutz usw.) sich nicht durch Absturz beim späteren Programmlauf äußern, sondern diese durch Zusatzdaten über einen Fehlerschutzalgorithmus aufgefangen werden. Pro Block fallen für diese Hilfsdaten 288 Bytes an, und für die Nutzdaten verbleiben somit 2048 Bytes.

In der Betriebsart 2 verzichtet man auf die Hilfsdaten zur Fehlerkompensation; das kann geschehen, wenn es sich hauptsächlich um Audio- oder Videodaten handelt, bei denen eventuelle Fehler keine derart fatalen Auswirkungen zeigen. Hier gibt es also einen unbeschränkten Umfang der Nutzdaten von 2336 Bytes pro Block.

Elektrische und geometrische Daten Aus Tab. 1.1 gehen die Daten des Speichermediums hervor. Die geometrischen Daten entsprechen exakt denen von CD-Audio und sind in Tab. 1.2 zu finden.

Die Codierung erfolgt nach EFM (Eight-to-Fourteen-Modulation) mit der Fehlerschutzcodierung nach CIRC (Cross Interleaved Reed Solomon Code), jedoch wird bei CD-ROM auf eine Fehlerverdeckung durch Interpolation (letzte Stufe der Fehlerkompensation bei CD-Audio) verzichtet. S. 42 [1, 6, 8].

Datenformate Zur Standardisierung der Datenformate sind einige Festlegungen erfolgt, die man in den sog. Color Books niedergelegt hat. Für CD-ROM-Anwendungen spielen dabei vor allem zwei Kategorien eine wichtige Rolle, nämlich das Red Book und das Yellow Book.

Tab. 1.1 Daten des Speichermediums CD-ROM

Größe	Wert/Modus 1	Wert/Modus 2	Einheit
Speicherkapazität gesamt	≈635	≈635	MByte
Speicherkapazität der Nutzdaten	≈553	≈631	MByte
Blockgröße	2352	2352	Byte
Frames pro Block	98	98	–
Datenrate	176400		Byte/s
Fehlerrate	<10 ⁻¹²	–	–
Übertragungsrate der Nutzdaten	153,600	175,200	Byte/s
Spieldauer	≈80	≈80	min
Zugriffszeit Spur-Spur	≈1	≈1	ms
Zugriffszeit Innenspur-Außenspur	≈1000	≈1000	ms

Tab. 1.2 Geometrische und mechanische Daten von CD-ROMs

Größe	Wert	Einheit
Plattendurchmesser	120	mm
Plattendicke	1,2	mm
Mittelloch-Durchmesser	15	mm
Innenspur-Durchmesser (Beginn)	50	mm
Außenspur-Durchmesser (Ende)	116	mm
Spurabstand	1,6	µm
Pit-Breite	0,4	µm
Pit-Tiefe	0,12	µm
Pit-Länge	0,83 … 3,1	µm
Drehsinn (betrachtet von der Leseseite)	Gegenuhrzeigersinn	–
Abtastgeschwindigkeit	CLV	–

Red Book Unter dieser Kategorie sind alle Festlegungen in einem Pflichtenheft vereinigt, die Aufzeichnungen für Audio (CD-DA Compact Disc Digital Audio oder auch CDA) beschreiben. Diesem Standard entsprechen alle üblichen CDs für den Audiobereich, die jedoch keinerlei Programm- oder Dokumentdaten beinhalten. Die Normung, die daraus abgeleitet wurde, heißt IEC 908 und sieht die Aufzeichnung unter der Fehlerschutzcodierung nach Reed Solomon (CIRC) vor.

Die Daten für Audio gehorchen der linearen, unkomprimierten PCM-Form. Eine gewöhnliche CD-DA lässt sich normalerweise in jedem CD-ROM-Laufwerk abspielen, allerdings sind damit die Fähigkeiten einer solchen Einrichtung längst nicht ausgeschöpft.

Yellow Book In diesem Pflichtenheft befinden sich Festlegungen für die Nutzung der Compact Disc als allgemeinen Datenträger. Dies beinhaltet sowohl die Möglichkeit der Aufnahme von Programm- und Dokumentdaten, als auch die Integration von PCM-Daten

nach IEC 908. Allerdings kommt letzteres nie exklusiv vor, wie dies etwa bei reinen Audio-CDs der Fall ist. Außerdem ist bei Audio- und Videodaten die Betriebsart 2 zulässig, die zugunsten einer größeren Speicherausnutzung auf den Fehlerschutzcode verzichtet. Normen sind ECMA-119 und ECMA-130 bzw. ISO 9660 und ISO 10149.

Sektorenaufbau Wie bereits erwähnt, kommen grundsätzlich zwei Betriebsarten in Frage. Während Betriebsart 1 wegen des Fehlerschutzcodes für Programm- und Dokumentdaten zwingend erforderlich ist, kann für reine Audio- und Videodaten wegen der Platzersparnis auch Modus 2 zur Anwendung kommen, wo kein Fehlerschutz besteht.

XA-Format Innerhalb des Yellow Book gibt es ein Datenformat mit erweiterten Eigenschaften, das sog. XA-Format (Extended Architecture). Die Sektoren umfassen hier insgesamt eine Größe von 3234 Byte. Im verwendeten Modus 2 existieren zwei Formate (Form 1 und Form 2), die sich im Umfang der Fehlerkorrektur unterscheiden.

Die XA-Betriebsart kann für CD-ROM-Anwendungen vorkommen; allerdings ist sie dann durchgängig für die gesamte CD gültig. Moderne CD-ROM-Geräte können diese Daten lesen. Bei der sog. CD-I (Compact Disc Interactive) kommt das Sektorenformat ebenfalls vor. Audiodaten können im beschriebenen Format nach dem ADPCM-Verfahren aufbereitet und komprimiert sein.

Sektorenplan Für Programmdaten muss der Datenträger auch einem bestimmten Plan entsprechen, nach dem eine Zuordnung der Sektoren geschieht. Die innersten Sektoren der Nummern 0 bis 15 gehören dem Systembereich an, während die Nummern ab 16 Datenträger-, Datei-, Verzeichnisbeschreibungen und Pfadtabellen beinhalten. Eine Datei muss nicht aus zusammenhängenden Sektoren bestehen, sondern kann durch Sektoren anderer Dateien unterbrochen sein. Dies ist vor allem bei untermalenden Audiodateien wichtig; bei dieser Art von Codespreizung entfällt ein dauerndes Positionieren zwischen Programm- und Audiodatei, was ja u. U. zu Diskontinuitäten beim Ablauf führen würde.

In den Datenträgerbeschreibungen kann ein Boot-Record definiert sein; somit ist der Boot-Vorgang von der CD-ROM möglich.

Neben der CD-ROM kommt heute auch die **DVD (Digital Versatile Disc)** als Speichermedium für Audiosignale in Betracht; sie ist ähnlich aufgebaut, bietet jedoch etwa sechs mal so viel Speicherkapazität. Allerdings ist ein spezielles Schreibgerät nötig, da DVD-Spuren eine höhere Auflösung benötigen.

Die DVD wird häufig für Videoaufzeichnungen eingesetzt, daneben spielt sie aber auch für die immer größer werdenden Speichervolumina zu installierender Programme eine bedeutsame Rolle. Im zweiten Teil des Buches in Abschn. 7.7.1 kommen wir noch auf das Datenformat der DVD zu sprechen.

1.10.6 Unbewegte Speichermedien

Speicher, die als Karte in entsprechende Slots in den PC einzuführen sind, werden immer wichtiger und sind nicht mehr wegzudenken. Die meisten PCs haben sogar universell verwendbare Slots. Natürlich gibt es auch extern anzuschließende Lese-/Schreibgeräte für Speicherkarten.

Der prinzipielle Aufbau der meisten Speicherkarten ist Flash, also ein beschreibbarer Halbleiterspeicher, der nichtflüchtig ist (also keine Batterie benötigt). Obwohl er bei weitem nicht so schnell beschreibbar ist wie RAM-Speicher, sind doch viele von den Flash-Medien heute in der Lage, sogar einen Videostream zu bewältigen (Abb. 1.19).

Es sind verschiedene Untergruppen auf dem Markt, die sich je nach Applikation besser oder schlechter eignen. Hier eine kleine Auflistung:

- Compact Flash (CF)
- Multimedia Card (MMC)
- Secure Digital Memory Card (SD)
- Smart Media (SM)
- xD-Picture Card (xD)

Die beiden Typen *SM* und *xD* arbeiten ohne internen Controller – die Daten müssen also extern entsprechend aufbereitet sein. Die restlichen drei Typen besitzen einen Controller, so dass der Schreib-/Lesezugriff sehr viel einfacher zu gestalten ist.



Abb. 1.19 Verschiedene Formen von Speicherkarten. Neben den Karten ist auch ein USB-Speicherstick zu sehen

Für Audiodaten sollte es bei den meisten Arten dieser Solid-State-Speicher keine wesentliche Begrenzung geben, es sei denn, es handelt sich um ein Multitrack-Audiosignal. Der Vorteil einer Speicherung auf Karten liegt sicher darin, dass sie keine beweglichen Teile haben und somit verschleißfreier und zuverlässiger sind. Außerdem ist die Wechselmöglichkeit bezüglich Archivierung komfortabler ist. Allgemein ist jedoch der Schreibvorgang wesentlich langsamer als der Lesevorgang. Für hohen Speicherbedarf dürften die Kosten von Halbleiterspeichern wesentlich höher ausfallen als für bewegliche Speicher, vornehmlich Festplattenspeicher. Möglich, dass sich künftig noch eine Verschiebung zugunsten der Halbleiterspeicher ergibt.

Tab. 1.3 Datenanordnung bei MP3-Tags der Version ID3v1

Offset (Hex)	Länge (Hex)	Bedeutung
00	03	Kennung ID3v1-Blocks
03	1E	Songtitel
21	1E	Künstler/Interpret
3F	1E	Album
5D	04	Erscheinungsjahr
61	1E	Beliebiger Kommentar
7F	01	Genre

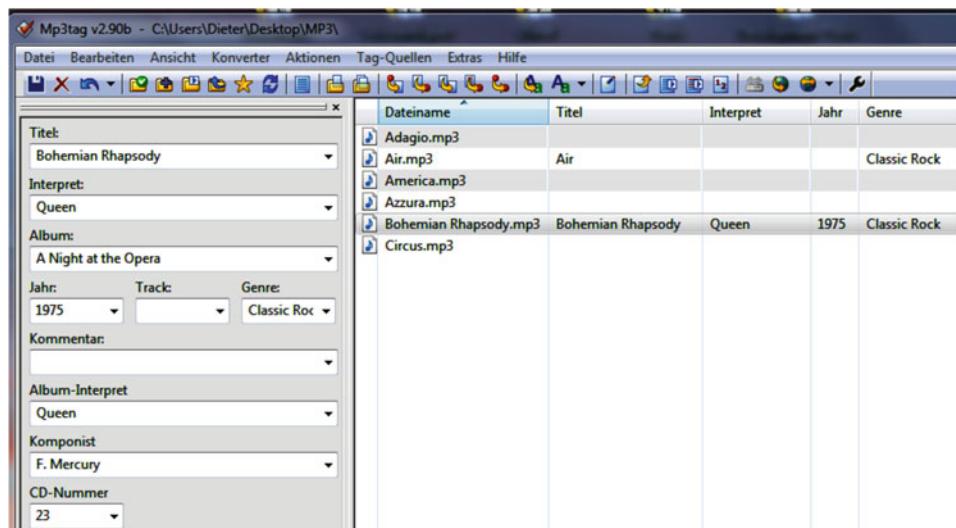


Abb. 1.20 Das Freeware-Programm *MP3tag* zur Verarbeitung, Editierung und Filterung von MP3-Daten. Im linken Bereich sind Filtereinstellungen möglich, während rechts zeilenweise die einzelnen Titel bzw. Dateien gelistet sind. Welche der Tags zur Anzeige kommen, ist einstellbar (Software *MP3tag* von Florian Heidenreich)

1.10.7 Archivierung und Datenbanken

Zur Archivierung von Audio- und Musikdateien im MP3-Format ist am Ende der Datei ein Block von 128 Bytes vorgesehen, der für Metadaten vorgesehen ist. Die Anordnung der Daten ist in Tab. 1.3 zu sehen. Leider gibt es viele unterschiedliche Versionen, die auch untereinander nicht kompatibel sind.

Viele Abspielprogramme für Musikdateien sind gleichzeitig in der Lage, die Tags getrennt abzuspeichern, damit diese in einer Datenbank übernommen werden können. Daneben sind Programme erhältlich, welche nur die MP3-Tags verwalten. Daneben sind natürlich auch Abspiellisten erstellbar, die sich gerade nach spezifischen Kriterien zusammenstellen lassen. Ein Beispiel eines komfortablen Verwaltungsprogramms ist in Abb. 1.20 zu sehen.

Literatur

1. Skritek, P.: Handbuch der Audio-Schaltungstechnik. München: Franzis-Verlag 1988.
2. Webers, J.: Tonstudientechnik. München: Franzis-Verlag 1989.
3. Stotz, D.: Audio- und Videogeräte richtig einmessen und justieren.
München: Franzis-Verlag 1994.
4. Jecklin, J.: Musikaufnahmen. München: Franzis-Verlag 1980.
5. Philippow, E.: Taschenbuch Elektrotechnik, Band 2.
Berlin: VEB Verlag Technik Berlin 1987.
6. Krieg, B.: Praxis der digitalen Audiotechnik. München: Franzis-Verlag 1989.
7. Biaesch-Wiebke, C.: CD-Player und R-DAT-Recorder. Würzburg: Vogel-Verlag 1992.
8. Stotz, D.: Erfolgreiche Fehlersuche an CD-Playern. Poing: Franzis-Verlag 1995.
9. Jens Blauert and Ning Xiang: Acoustics for engineers.
Berlin, Heidelberg, New York: Springer-Verlag 2009.



Abtastung und Digitalisierung

2

2.1 Einleitung

In vielen Publikationen erscheinen Begriffe wie Quantisierungsfehler, Abtastspektren usw. immer im Zusammenhang eines Gesamtsystems, wo alle Effekte vermischt auftreten. Für das Verständnis ist jedoch entscheidend besser, alle Vorgänge für sich isoliert zu untersuchen, soweit dies möglich ist.

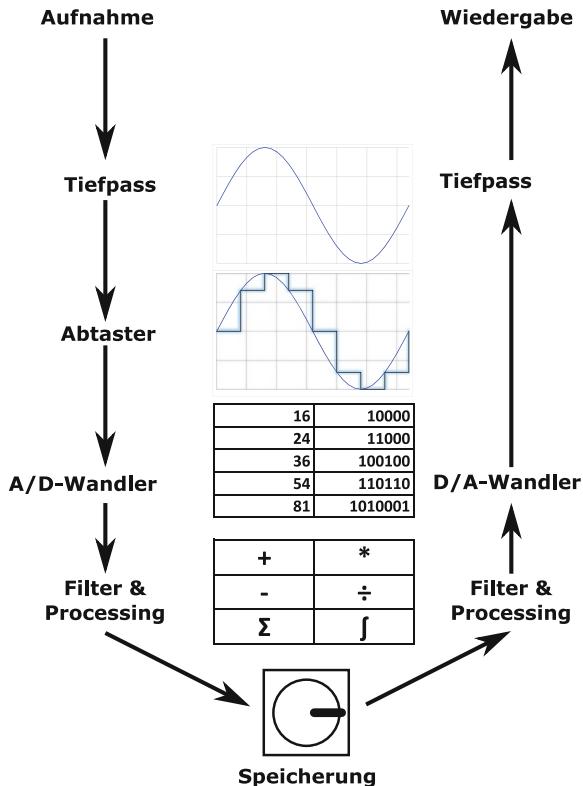
Bei der Reproduktion, also Aufzeichnung, Speicherung und Wiedergabe eines Audiosignals sind nämlich zwei Vorgänge von zentraler Bedeutung: Die Abtastung und die Analog-Digital-Wandlung, also die Quantisierung. Beides bringt spezifische Effekte zutage, die die Qualität entscheidend beeinflussen können. Diese Effekte können separat behandelt werden, d. h. wir betrachten die reine Abtastung und lassen die Wandlung in Digitalwerte, die Speicherung und die Rückwandlung dabei vorläufig außer Acht. Die Abtastung liefert bereits Abtastspektren, die im originalen Analogsignal noch nicht vorhanden sind.

Weiterhin ist bei der Betrachtung der Quantisierung und die damit zusammenhängenden Effekte die Abtastung irrelevant, wenn man sich gedanklich vorstellt, das Analogsignal unendlich fein zeitlich aufzulösen. Die Wandlung in Digitalwerte unterliegt jedoch Beschränkungen, die sich als Quantisierungsfehler äußern.

Am Ende sind alle Vorgänge wieder zu verschmelzen, wenn es um die Reproduktion eines Audiosignals geht. Die Einzeleffekte addieren sich dann verständlicherweise (Abb. 2.1).

Speziell tauchen Themen wie Abtastung und Überabtastung in beiden Kategorien auf, sodass man sich hier auf das Wesentliche beschränken kann.

Abb. 2.1 Übersicht zur Signalwandlung. Die Darstellung spiegelt eine gewisse Symmetrie wider, weil einige Prozesse in ähnlicher oder inverser Art auf beiden Seiten, der Aufnahme und der Wiedergabe, existieren



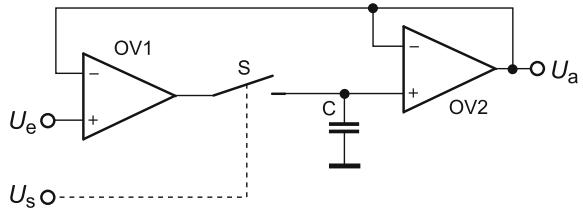
2.2 Abtastung

2.2.1 Notwendigkeit und technische Ausführung

Ein bestehendes Analogsignal ist in seinem herkömmlichen Sinne bereits ein Signal. Wenn man jedoch ein Signal als Werte abspeichern möchte, ist es mit einer dafür speziell ausgelegten Schaltungseinheit zu wandeln. Dies übernimmt ein Analog-Digital-Wandler, auf den wir weiter unten noch genauer zu sprechen kommen. Siehe hierzu Abschn. 2.3. Doch dieser Vorgang erfordert die Entnahme zeitdiskreter Proben, das ein Abtast-Halte-Glied übernimmt (Sample&Hold oder S&H). Abb. 2.2 zeigt den schematischen Aufbau eines derartigen Gliedes.

Die beiden Operationsverstärker $OV1$ und $OV2$ dienen der Entkopplung vom eigentlichen Abtast-Halte-Glied, das aus dem elektronischen Schalter S und der Kapazität C besteht. Bei jedem Sample-Takt am Schalteingang U_S schließt der Schalter; nach kurzer Zeit hat sich C auf die momentane Signalspannung aufgeladen. Nach dem Öffnen des Schalters bleibt diese Ladespannung erhalten und gelangt über den Operationsverstärker $OV2$ als belastbare Spannung zum Ausgang des Gliedes. Die Qualität des Sampling-Vorgangs

Abb. 2.2 Abtast-Halte-Schaltung (Sample&Hold)



hängt von verschiedenen Eigenschaften ab, die u. a. durch die Kapazität C und auch durch den Schalter S gegeben sind. S. 97 [1]

Heutige moderne A/D-Wandler besitzen bereits eine geeignete S&H-Schaltung. Die gesonderte Betrachtung ermöglicht jedoch, den Abtastvorgang völlig lösgelöst von allen weiteren Bearbeitungsschritten zu behandeln.

2.2.2 Abtastspektren

Beim Abtastvorgang entstehen Spektren um Vielfache der Abtastfrequenz herum. Im Folgenden schauen wir uns die Entstehung und die Eigenschaften der Abtastspektren etwas genauer an.

Bei einer idealen Abtastung werden unendlich kurze Proben entnommen, die Werte werden multipliziert mit dem *Dirac-Kamm* (Abb. 2.3). Zwischen den Abtastpunkten befinden sich Nullwerte. Bei Erhöhung der Probendichte, also bei häufigerer Abtastung entstehen Abtastspektren bei entsprechend höheren Frequenzen (Abb. 2.4).

Intuitiv würde man meinen, die Höhe der Abtastspektren müsste nach oben hin abnehmen. Bei dieser idealen Abtastung entstehen durch Fourier-Analyse jedoch Spektren, welche alle Einheitshöhe aufweisen, was bei einer endlichen Abtastlänge (also z. B. bei Entnahme zweier aufeinanderfolgender Werte) nicht der Fall wäre. Dann nämlich entstünden Spektren abfallender Höhe.

Bei der realen Abtastung durch das weiter oben beschriebene S&H-Glied entstehen Stufensignale gemäß Abb. 2.5. Die Grobheit der Stufen bzw. deren Höhe hängt u. a. ab von der Frequenz des Nutzsignals bzw. von deren Verhältnis zur Abtastfrequenz. Die Stufung selbst führt dazu, dass sich auch hier Abtastspektren ausbilden, die jedoch im Gegensatz zur idealen Abtastung sehr viel kleinere Amplituden aufweisen. Auch hier ist naheliegend, dass große Stufen auch zu großen Spektrumsamplituden führen. Der Vergleich bei verschiedenen Nutzsignalfrequenzen ist in Abb. 2.6 zu sehen.

Natürlich ist es auch möglich, das Stufensignal mit einem (digitalen) Hochpass zu filtern, damit nur noch die Abtastspektren vorhanden sind. Siehe hierzu Abb. 2.7. Es verbleiben Schwingungszüge von abfallenden e-Funktionen. Beim Berechnen der zu erwartenden Spektrumsamplituden ist dies zu berücksichtigen. (Die Annahme von Sägezahnschwingungen mit einer Amplitude aus gemittelter Stufenhöhe führt zu falschen Ergebnissen.)

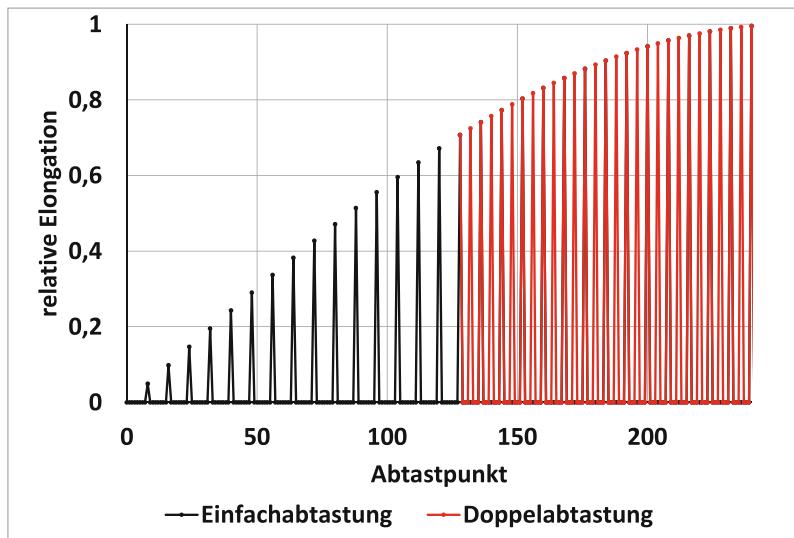


Abb. 2.3 Ausschnitt eines Sinussignals mit Einfachabtastung und Doppelabtastung. Nur jeweils ein einzelner Punkt nimmt den Funktionswert an, dazwischen sind Nullwerte. Die Länge der Abtastwerte ist somit unendlich klein und die Abtastung entspricht idealen Verhältnissen

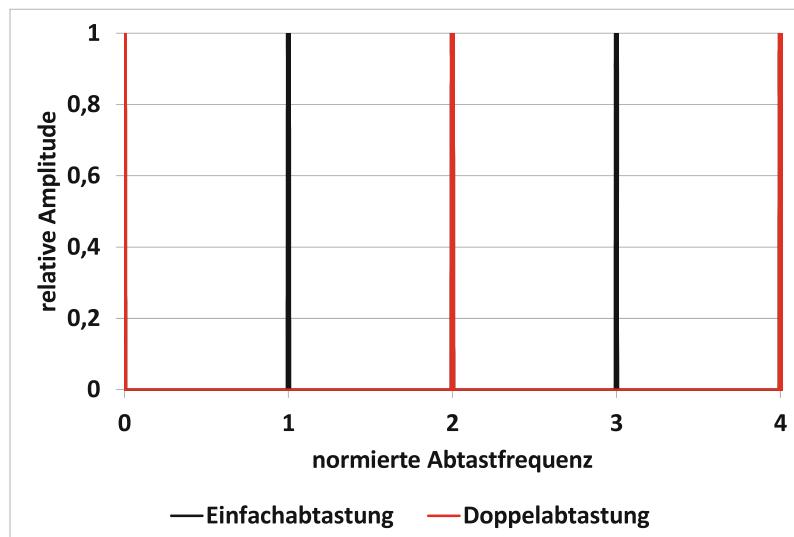


Abb. 2.4 Spektren bei Einfach- und Doppelabtastung aus dem Signal nach Abb. 2.3. Doppelte Abtastung bedeutet Verschiebung des ersten Abtastspektrums entsprechend in der Frequenz nach oben. Die Amplituden sind alle vom Betrag 1. Bei nichtidealer Abtastung – also bei mehr als einem Funktionswert in Folge – würden die Amplituden mit steigender Ordnung abfallen

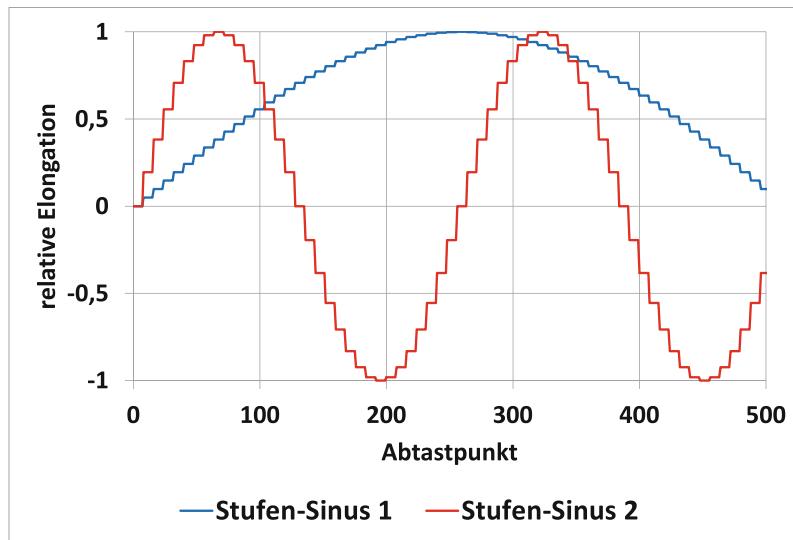


Abb. 2.5 Stufensignale unterschiedlicher Frequenz: Es leuchtet ein, dass die Stufungshöhe proportional zur Frequenz des Nutzsignals ist

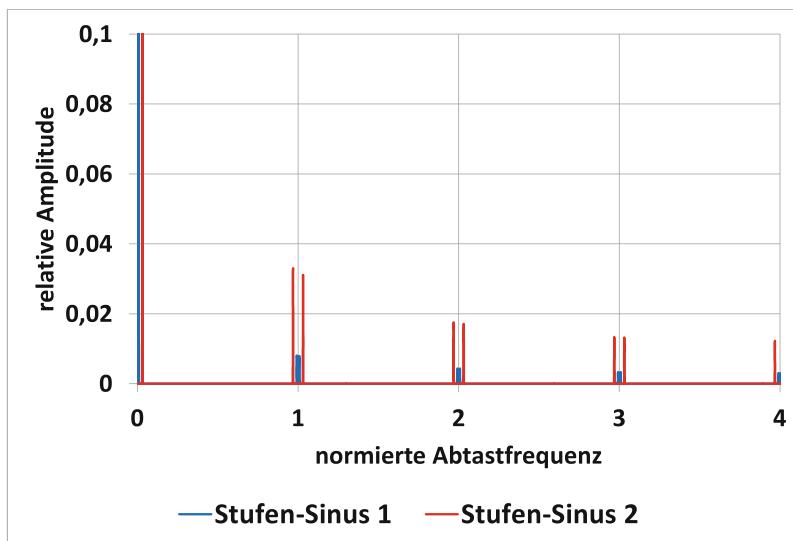


Abb. 2.6 Spektren der Stufensignale unterschiedlicher Frequenz. Aufgrund der unterschiedlichen Stufenhöhe im Zeitbereich resultieren in den Spektren auch entsprechend unterschiedliche Höhen der Frequenzanteile

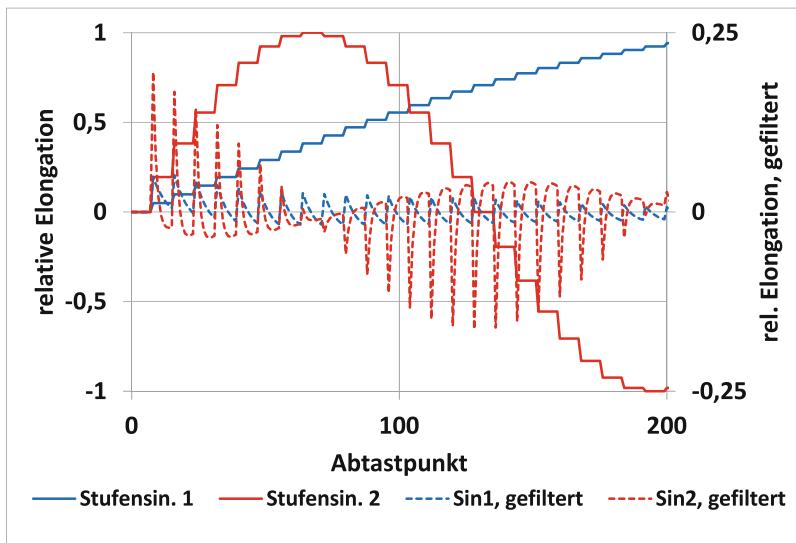


Abb. 2.7 Stufensignale und ihr Filterprodukt. Zur Verdeutlichung seien hier die Stufensignale gefiltert, daraus können dann die unterschiedlichen Amplituden der Abtastspektren abgeleitet werden

2.2.3 Aliasing

Nach dem Theorem von Shannon ist für die höchste Signalfrequenz pro Schwingung mindestens eine zweimalige Abtastung erforderlich, damit erstens die Schwingung noch übertragen wird und zweitens keine Störfrequenzen entstehen. Dieser Extremfall ist in Abb. 2.8 dargestellt.

Der dargestellte Fall ist zwar ein Sonderfall; er ist deswegen jedoch zur Veranschaulichung besonders gut geeignet. Die allgemeine Situation wird durch folgende Gleichungen beschrieben:

$$F_A \geq 2F_S \quad (\text{Abtastbedingung}) \quad (2.1)$$

$$F_X = F_A - F_S \quad (\text{Störfrequenz}) \quad (2.2)$$

Die Abtastbedingung ist zwar notwendig, für eine ausreichende Rekonstruktion des Originalsignals sollte die Abtastfrequenz jedoch besser deutlich größer sein. Denn je nach Phasenlage liefert die Abtastung für viele Takte nur Signal-Nulldurchgänge.

Bei zu geringer Abtastfrequenz jedoch kommt es zur Spektrenüberlappung nach Abb. 2.9.

Da die entstehenden Störfrequenzen in keinem harmonischen Zusammenhang zum Nutzsignal stehen, werden sie als sehr auffallend und somit störend empfunden. Außerdem tritt der Aliasing-Fall meist nur bei sehr hohen Signalfrequenzen auf, und diese werden

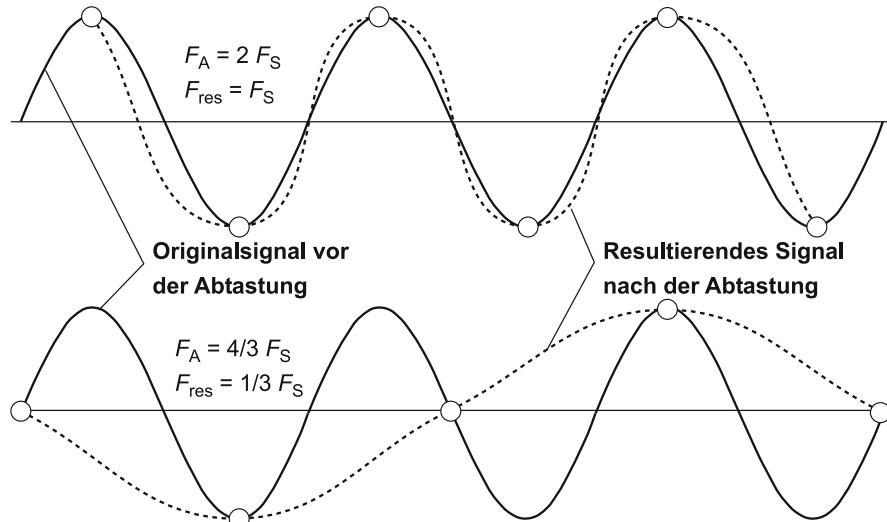


Abb. 2.8 Unterschiedliche Abtastfrequenzen und ihre Auswirkungen. Es ist zu sehen, dass hier durchaus noch eine etwaige Reproduktion möglich ist. Die resultierenden Signale entstanden durch eine stillschweigend vorausgesetzte Interpolation von Zwischenwerten – sonst müsste man an dieser Stelle noch Stufensignale annehmen. Darunter ist eine Situation einer Abtastung dargestellt, bei der die Abtastrate F_A nur das $4/3$ -fache der Signalfrequenz F_S beträgt. Hier entsteht eine Frequenz, die sehr viel niedriger ist als die des zu übertragenden Signals, nämlich $1/3$ der Signalfrequenz

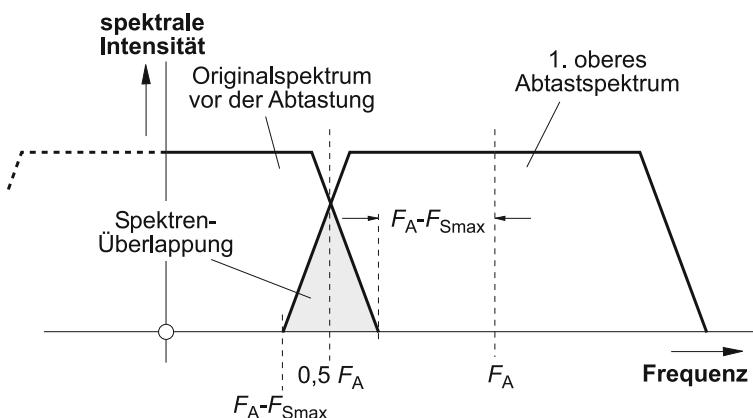


Abb. 2.9 Entstehung von Aliasing-Störungen durch Überlappung der Spektren. Die höchstmögliche Signalfrequenz F_{Smax} bildet zusammen mit der Abtastfrequenz F_A eine Differenzfrequenz $F_A - F_{Smax}$, die als Störsignal ins Nutzspektrum zurückfällt

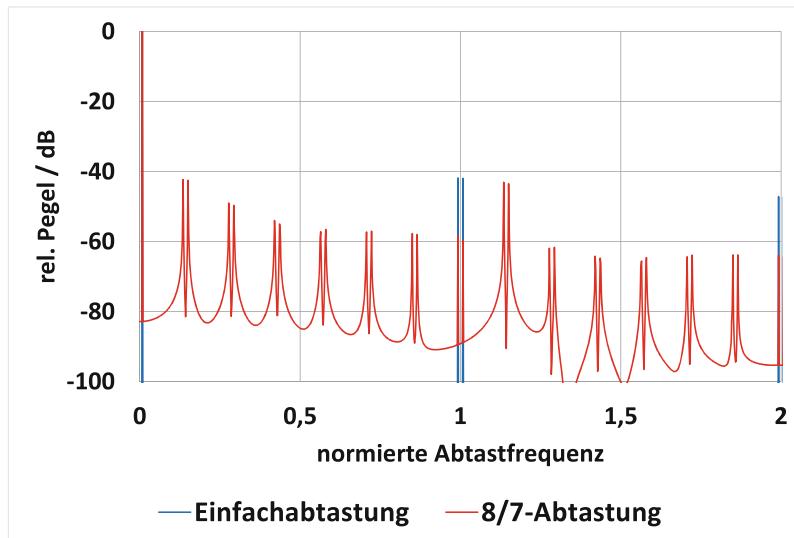


Abb. 2.10 Nochmalige Abtastung mit varierter Abtastfrequenz: Ein Stufensignal wurde mit einer mit Faktor 8/7 höheren Frequenz abgetastet. Als Resultat ergeben sich viele Aliasing-Produkte auch auf sehr niederfrequenten Bereichen

ohnehin schwächer wahrgenommen. Dagegen treten die Aliasing-Störungen selbst mitunter als recht tieffrequente Signale in Erscheinung.

Selbst wenn ein Stufensignal mit einer höheren Frequenz erneut abgetastet wird, können Aliasing-Störungen auftreten, da ja in diesem Fall die Abtastbedingung nach Shannon und Nyquist verletzt ist. Siehe hierzu Abb. 2.10.

Beim Versuch, mit der Audiosoftware *Audacity* ein Sinussignal der Frequenz 1 kHz bei einer Abtastfrequenz von 44,1 kHz zu erzeugen und dieses dann erneut mit 48 kHz abzutasten, ergeben sich neben dem normalen Sinussignal auch noch Aliasing-Produkte mit 2,9 kHz und 4,9 kHz. Die Ursache ist leicht zu verstehen: Die Differenz der beiden Abtastfrequenzen beträgt 3,9 kHz, diese Frequenz subtrahiert und addiert sich zur Nutzsignalfrequenz von 1 kHz. Ersteres wird an der Nullfrequenz-Achse gespiegelt, dadurch entsteht 2,9 kHz.

Audiokarten für Computer sollten daher mit steilflankigen Tiefpassfiltern ausgestattet sein, die eine Grenzfrequenz von deutlich unter der halben Abtastrate aufweisen. Meist ist über Software eine Umschaltung der Abtastfrequenz möglich; das impliziert für das Antialiasing-Filter am Eingang ebenfalls eine Umschaltung der Grenzfrequenz. Probehalber kann man ein Signal mit der Karte aufzeichnen, dessen Frequenz z. B. genau 1 kHz unter der Sampling-Frequenz liegt, das entspricht also bei $F_A = 44,1$ kHz einer Signalfrequenz von 43,1 kHz. Zur Untersuchung des zu erwartenden Aliasing-Signals mit 1 kHz kann eine reine Software-Messung (Frequenzanalyse) oder eine Signalmessung bei der Wiedergabe zur Anwendung kommen, wobei am besten ein Selektivfilter mit 1 kHz Mittenfrequenz zur Isolation des relevanten Störsignals dient. Zum Vergleich dient ein bei gleicher Aussteuerung aufgezeichnetes 1-kHz-Sinussignal als Referenz. Das Spannungsverhältnis des Referenzsignals zum Aliasing-Signal kann dann als lineares Maß für die Aliasing-Unterdrückung herangezogen

werden. Audiokarten mit guten Filtern können durchaus Unterdrückungsfaktoren von mehr als 10 000, also 80 dB erreichen. Es kommt leider mitunter bei Audiokarten auch vor, dass der Anwender selbst dafür Sorge tragen muss, dass das Signalspektrum sicher unterhalb der halben Abtastfrequenz liegt; diese Forderung ist jedoch selbst bei Mikrofonaufnahmen keinesfalls so einfach zu erfüllen, besonders wenn die Sampling-Frequenz einmal etwas niedriger gewählt wird. Die Forderung an das Tiefpassfilter entschärft sich, wenn man durch Überabtastung die Abtastfrequenz wesentlich höher setzt, als dies für übliche Nutzfrequenzbereiche notwendig wäre. Siehe auch entsprechendes Unterkapitel 2.2.4 weiter unten.

Die Aliasing-Unterdrückung hängt natürlich entsprechend vom Frequenzgang des Tiefpassfilters ab; besonders wichtig ist deshalb das Durchlassverhalten direkt unterhalb der Grenzfrequenz und das Sperrverhalten bei halber Abtastfrequenz. Einen solchen Frequenzgang zeigt schematisch Abb. 2.11.

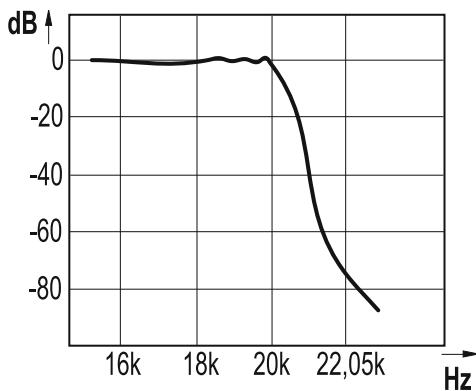
Im Zeitkontinuum muss diese Tiefpassfilterung natürlich auf analoger Ebene stattfinden. Zur Anwendung kommen oft Cauer-Tiefpässe oder FDNR-Tiefpässe (Frequency Dependent Negative Resistance), die eine sehr hohe Sperrdämpfung bei kleinen Abweichungen innerhalb des Frequenzgangs aufweisen und sich zudem leicht in puncto Eckfrequenz umschalten lassen. S. 97 [2]

2.2.4 Überabtastung

Das oben Geschriebene stimmt zwar formal, allerdings müssen sich die Eigenschaften des analogen Tiefpasses nur auf die Abtastfrequenz beziehen. Wenn man jetzt letztere einfach drastisch hochsetzt, so entfernen sich die Abtastspektren von einer für Audioverhältnisse hinlänglichen Bandbreite um den Faktor der höher gewählten Abtastfrequenz (Abb. 2.12).

Überabtastung – oder auch Oversampling genannt – kann sowohl auf Aufnahme- wie auch auf Wiedergabeseite erfolgen. Grundsätzlich ist ein Vielfaches an Abtastpunkten vorhanden, als es alleine zur Erfüllung der Abtastbedingung notwendig wäre.

Abb. 2.11 Frequenzgang eines Anti-Aliasing-Filters zwischen Grenzfrequenz und halber Abtastfrequenz (hier 22,05 kHz)



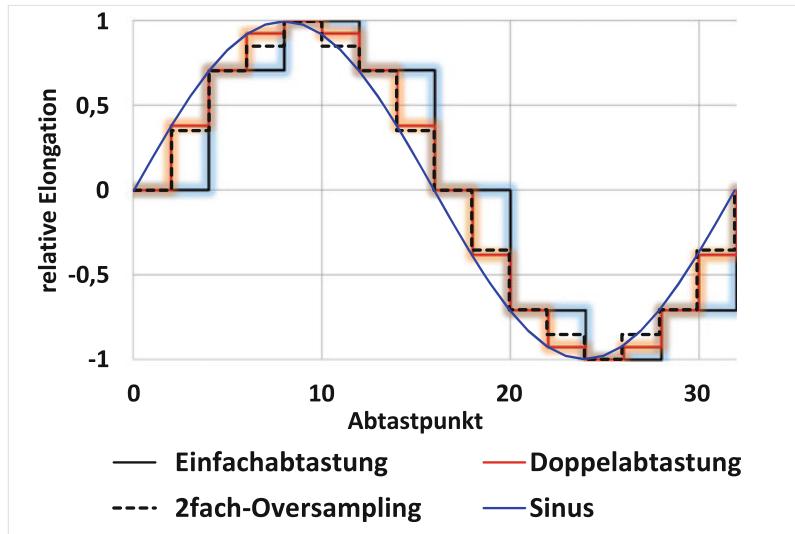


Abb. 2.12 Stufensignale bei Einfach- und Doppelabtastung. Zur Verdeutlichung werden Stufensignale unterschiedlicher Abtastung dargestellt, auch das 2fach-Oversampling, dessen Zwischenpunkte durch Interpolation künstlich erzeugt wurden

Aufnahmeseite Findet Überabtastung auf Aufnahmeseite statt, stehen normalerweise lauter echte Abtastwerte vor. Durch die hohe Abtastfrequenz jedoch kann ein „echter“ Hardware-Tiefpass einfach ausfallen, da dieser erst vom Hörbereich weit entfernte Frequenzanteile unterdrücken muss (und deshalb nicht steilflankig sein muss).

Erst während der weiteren Verarbeitung wird das Datenvolumen wieder verringert, damit kein unnötiger Speicherbedarf entsteht. Man nennt diesen Vorgang auch Dezimation.

Eine einfache Möglichkeit zur Reduzierung der Abtastrate besteht darin, alle Daten mit einem digitalen Tiefpass zu behandeln. Dabei entstehen Daten, die oberhalb einer Grenzfrequenz eine starke Dämpfung erfahren. Die Ordnung des Filters bestimmt dabei die Steilheit bei dieser Grenzfrequenz. Die Reduzierung der Abtastrate besteht nun darin, dass mehrere Stützstellen wieder als Mittelwerte zusammengefasst werden (Abb. 2.13).

Wie derartige *FIR-Filter* arbeiten, ist in Abschn. 3.2.1 genauer behandelt.

Als Zahlenbeispiel sei folgende Situation genannt: Die nominelle Abtastrate eines Audiosignals liege bei 44,1 kHz, es soll jedoch 8fache Überabtastung vorliegen. Der reale Tiefpass am Eingang muss lediglich bei 176 kHz eine hinreichend gute Dämpfung aufweisen, um Aliasing zu vermeiden. Die sich der A/D-Wandlung anschließende Digitalfilterung und Dezimation sorgt dafür, dass der Speicherbedarf dem entspricht, der bei 44,1 kHz vorliegt.

Die Abb. 2.14 und 2.15 zeigen anschaulich, wie Aliasing auf der Aufnahmeseite durch Überabtastung wirksam unterdrückt wird.

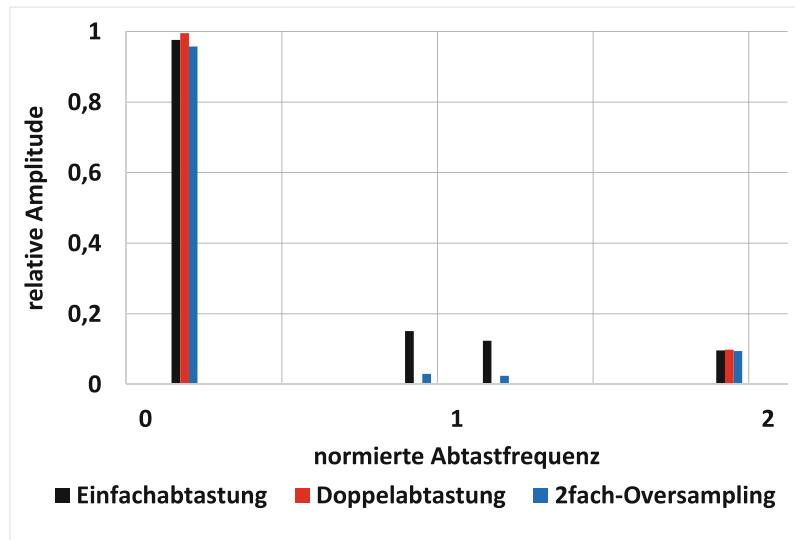


Abb. 2.13 Spektren der Stufensignale bei Einfach-, Doppelabtastung und 2fach-Oversampling. Die Frequenzverschiebung verhält sich wie gewohnt, allerdings hat das 2fach-Oversampling durch die „unechten“ Punkte auch Residuen bei der einfachen Abtastfrequenz

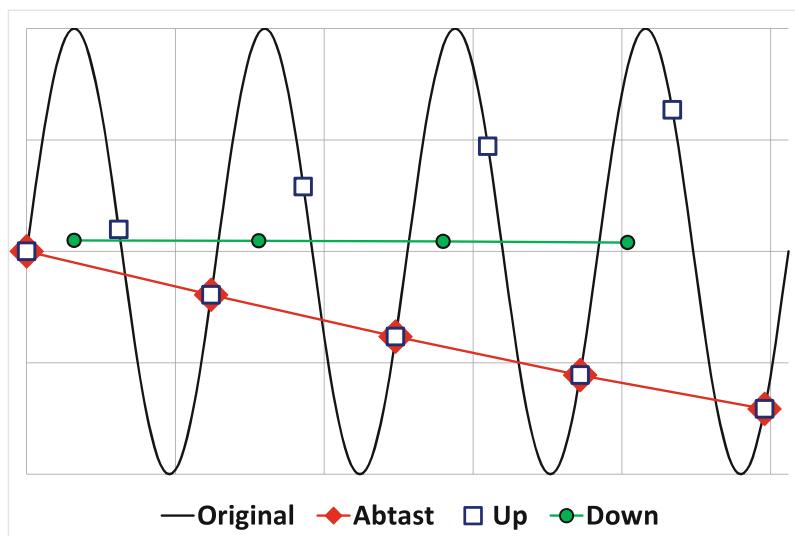


Abb. 2.14 Aliasing durch Nutzsignal mit Frequenz nahe der Abtastfrequenz. Das störende Aliasing-Artefakt entsteht zwar (rote Linie), wird jedoch durch Down-Sampling fast restlos unterdrückt, siehe grüne Linie. Die Stützstellen hierzu entstehen durch Interpolation der quadratischen Oversampling-Punkte

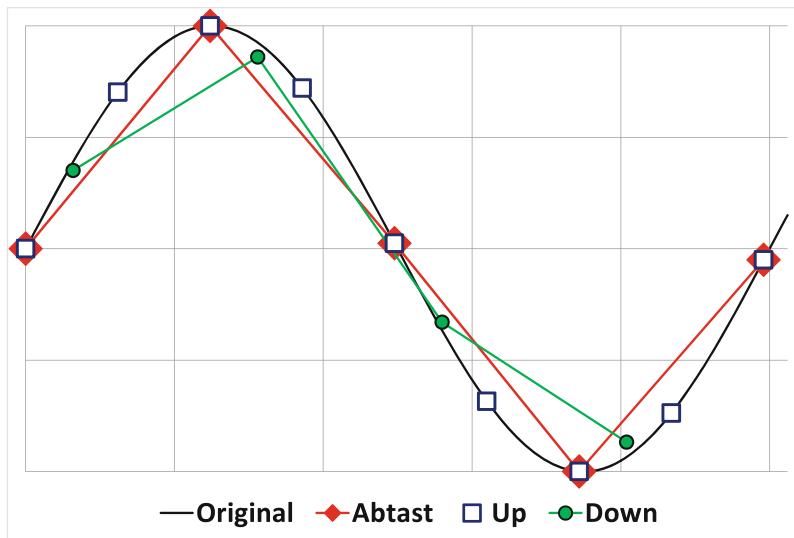


Abb. 2.15 Überabtastung auf A/D-Seite mit anschließender Tiefpassfilterung und Down-Sampling. Niedrigere Frequenzen des Nutzsignals erscheinen wieder ordnungsgemäß (vgl. Abb. 2.14)

Wiedergabeseite Bei der Wiedergabe muss Überabtastung anders funktionieren, denn das Mehr an Signalpunkten liegt ja de facto nicht vor. Wie kann man hier agieren? Nun, wir haben es bereits in Abb. 2.12 und 2.13 gesehen, dass Zwischenpunkte einfach durch Interpolation bzw. Mittelwertbildung entstehen können.

Warum ist das sinnvoll? Nun, eine einfache Audio-CD hat normalerweise Daten, die Stützpunkte im Takte von 44,1 kHz repräsentieren. Würden diese einfach gelesen, dem D/A-Wandler zugeführt und das Resultat nach draußen geführt, läge bei genanntem Takt ein unerwünschtes Abtastspektrum vor und bei Vielfachen dieser Frequenz würden ebenfalls Störspektren auftreten.

Die zeitliche Auflösung steigt also auch hier durch Überabtastung an. Da diese zusätzlichen Daten jedoch keinen erhöhten Speicherraum benötigen (ad sie ja nur temporär vorhanden sind), muss keine Dezimation stattfinden. Stattdessen wird dem Wandlerausgang einfach ein einfaches Tiefpassfilter nachgeschaltet. Da dieses wie bei der Aufnahmeseite eine um Faktoren erhöhte Störfrequenz unterdrücken muss, kann es wiederum relativ einfach ausfallen.

2.3 Quantisierung – Wandlung in Digitalwerte

Unter Quantisierung ist die Wandlung der Analogspannung in Digitalwerte zu verstehen.

Nach der Abtastung des analogen Signals erfolgt die Wandlung in das Digitalformat. Man kann hier ebenfalls von Spannungen sprechen, jedoch handelt es sich ja bei der

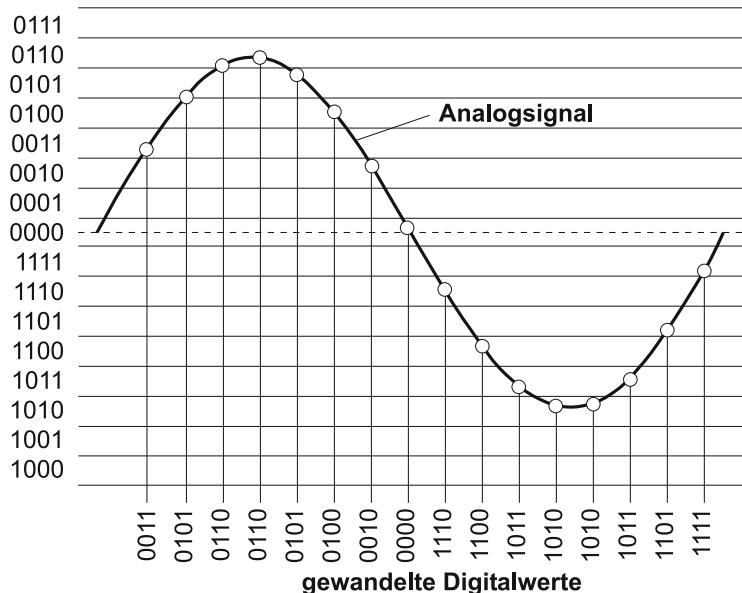


Abb. 2.16 Zuordnung digitaler 4-Bit-Werte einer abgetasteten analogen Signalspannung. Die PCM-Methode ordnet eine Analogspannung innerhalb eines vorgegebenen Intervalls einem Raster zu, wobei jedes Rasterfeld einem Digitalwert entspricht. Man spricht in diesem Zusammenhang auch von Quantisierungsstufen, weil das Spannungsintervall in definierte Elemente unterteilt wird

Pulsecode-Modulation (PCM) um mehrere, zeitlich parallel nebeneinander vorhandene Digitalspannungen, so dass die Betrachtung als Binärmuster bzw. Digitalwertes passender ist, zumal die Software im Endeffekt ja mit Dualzahlen operiert. Abb. 2.16 verdeutlicht die Zusammenhänge an einem Modell, das eine Signalspannung in 4-Bit-Werte umzuwandeln vermag.

2.3.1 Digitalformate nach der Wandlung

Im abgebildeten Beispiel wird stillschweigend vorausgesetzt, dass das Nullpotential der Analogspannung (gestrichelte Linie) auch tatsächlich dem Nullwert des digitalen Wortes entspricht und die Codierung über das Zweierkomplementformat festgelegt ist. Eine Alternative hierzu stellt die sog. Offset-Dual-Codierung dar, bei der nur positive Binärwerte zugelassen sind und daher die digitale Null dem kleinsten Analogwert entspricht und die analoge Null mit dem halben Maximum des Digitalbereichs korrespondiert. Mathematisch gilt dabei:

$$N_{D+} = N_A + \frac{N_{\max}}{2} \quad (\text{Offset-Dual}) \quad (2.3)$$

$$N_{D2} = N_A \quad \text{für } N_A \geq 0 \quad (\text{Zweierkomplement}) \quad (2.4)$$

$$N_{D2} = -|N_A| \quad \text{für } N_A < 0$$

N_A der der Analogspannung äquivalente Analogzahlenwert, $N_{\max}/2$ der Halbbereichswert (die Hälfte aller möglichen Werte).

Welches Zahlenformat gewählt wird, ist reine Festlegungssache. Es ist nicht einmal abhängig vom A/D-Wandler, da der Digitalwert einer Audiodatei beliebig konvertiert sein kann. Bei WAV-Dateien (Microsoft-Audioformat) liegt bei 8-Bit-Auflösung das Format nach Gl. 2.3 vor, während die 16-Bit-Auflösung das Zweierkomplement verwendet.

Die Auflösung der Wandlung bestimmt, mit wie vielen Quantisierungsstufen die momentane Analogspannung erfasst werden kann. Die Anzahl dieser Stufen ist definiert durch:

$$n_{\max} = 2^D \quad (2.5)$$

Diese Stufenzahl wächst also exponentiell mit der Auflösung D , welche der Stellenzahl des Binärcodes entspricht. Gebräuchliche Auflösungen sind $D = 8$ (256 Stufen) und $D = 16$ (65,536 Stufen). Vereinzelt setzen Audiokarten jedoch auch eine Auflösung von 12 Bit (4096 Stufen) ein. Hochqualitative Studioaufnahmen setzen einen Standard von 16 Bit voraus; eine solche Auflösung wird z. B. auch bei CDs oder DATs eingesetzt. Alle darunter liegenden Auflösungen kommen nur für Demonstrations- oder Testzwecke in Frage. Die Auflösung 12 Bit stellt einen gewissen Kompromiss zwischen Qualität und geringem Aufwand dar; viele Anwendungen können jedoch eine entsprechende Audiodatei nicht konvertieren.

Bei der bisher besprochenen Wandlungsproblematik wurde immer von einer linearen Wandlungskennlinie ausgegangen. Linear heißt in diesem Falle, dass die Binärcodes mit gleichem Abstand auf den wandlungsfähigen Analogspannungsbereich verteilt sind.

Wenn man sich ein Sinussignal vorstellt, dessen Amplitude sich ständig verkleinert, so wird klar, dass die auf diese Amplitude bezogene Quantisierung immer größer wird. Abb. 2.17 zeigt die Gegenüberstellung zweier Signale mit unterschiedlichen Amplituden, die mit linearer Kennlinie gewandelt wurden. Als Auflösung wurde für dieses Beispiel 8 Bit verwendet – prinzipiell zeigt sich quantitativ der gleiche Effekt natürlich auch bei höheren Auflösungen, es ist dann lediglich ein größerer Dynamiksprung notwendig.

Der zutage tretende Effekt ist mit den bereits besprochenen Quantisierungsfehlern vergleichbar. Man muss sich vergegenwärtigen, dass bei einer Aussteuerung von -40 dB nur noch maximal vier Quantisierungsstufen zur Verfügung stehen – also besteht eigentlich die Situation einer extrem eingeschränkten Auflösung. In diesem Zusammenhang fällt auch häufig der Begriff Granular-Rauschen.

Verteilt man das Quantisierungsraster hingegen logarithmisch auf den Analogspannungsbereich, so entsteht eine Situation nach Abb. 2.18. Durch die höhere Steigung in der Nähe des Analog-Nullpunkts fallen hier mehr Quantisierungsstufen an, sodass die digitale Beschreibung bei kleinen Amplituden wesentlich genauer ausfällt.

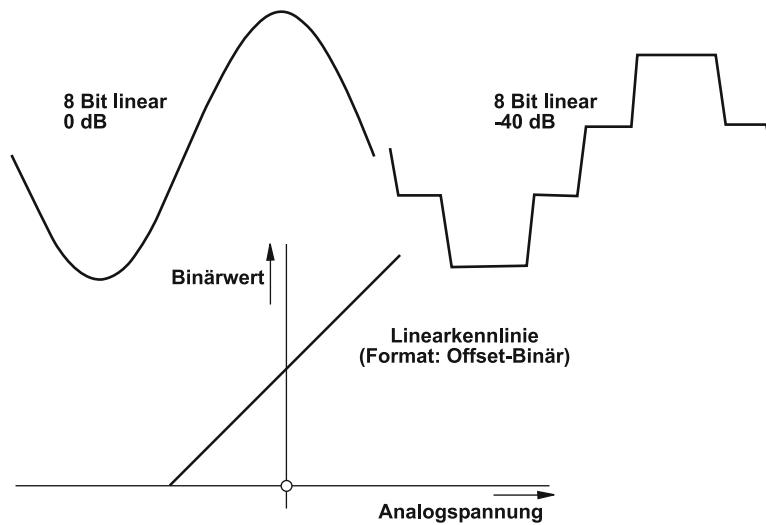


Abb. 2.17 Lineare Wandlungskennlinie und Sinussignale verschiedener Aussteuerung

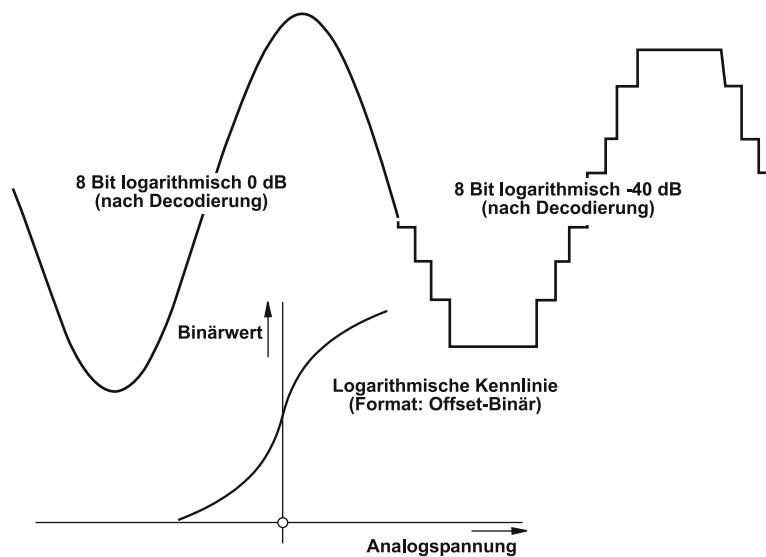


Abb. 2.18 Logarithmische Wandlungskennlinie und Sinussignale unterschiedlicher Aussteuerung. Dargestellt sind hier die Signale nach der Rückwandlung. Nach dem Encoding erscheint das Sinussignal mit Vollaussteuerung sehr verzerrt – eben entsprechend der Kennlinie; bei kleiner Aussteuerung zeigt sich dagegen kaum ein Unterschied zwischen endodierter und decodierter Signal, da innerhalb des kleinen Intervalls die Steigung der Kennlinie wenig Änderungen aufweist

Eine solche logarithmische Wandlungskennlinie hat sich bei der Encodierung von Musik nicht bewährt. Der Grund ist relativ einfach, denn es ist einzusehen, dass Mischsignale aus verschiedenen Frequenzen und Amplituden mitunter sehr mangelhaft umgesetzt werden können. Man stelle sich nur vor, eine kleine Amplitude eines hohen Tones ist der großen Amplitude eines niederfrequenten Tones überlagert. Gelangt nun letzterer zu einer großen Elongation, so ist ja hier systembedingt eine kleinere Dichte an Quantisierungsstufen. Dadurch gerät hier jedoch die Umsetzung des hohen Tones mit der kleinen Amplitude in ungünstige Bedingungen. Das Verfahren ist daher vor allem für Codierung von Sprach- und anderen, monofonen Signalen beschränkt, wo es bei gleichbleibender Auflösung eine Qualitätsverbesserung gegenüber der linearen Codierung darstellt, bzw. bei gleicher Qualität die Datenmenge zu reduzieren vermag.

2.3.2 Methoden der A/D-Wandlung im Überblick

Für die Technik und somit die Technologie der A/D-Wandlung stehen grundsätzlich vier Methoden zur Verfügung:

- Zählverfahren (Dual-Slope)
- Wägeverfahren (Sukzessive Approximation)
- Direktverfahren (Flash-Wandlung)
- Delta-Sigma-Verfahren

Neben diesen reinen Methoden gibt es natürlich Varianten, die eine Mischung der Verfahren anwenden.

Das Zählverfahren ist unter kleinstem Aufwand zu realisieren. Die Genauigkeit bzw. Auflösung kann sehr hoch sein, die Geschwindigkeit der Wandlung ist jedoch für den Einsatz im Audiobereich zu gering, so dass dieses Verfahren hier nicht in Frage kommt. Das Grundprinzip beruht darauf, dass ein Binärzähler den gesamten Zahlenbereich durchläuft. Sein Ausgangsmuster gelangt an den Eingang eines D/A-Wandlers, dessen Ausgangsspannung wiederum auf einen Eingang eines Komparators führt. Der andere Eingang des Komparators erhält die zu wandelnde Analogspannung. Sobald nun der Komparator umkippt, ist dies ein Zeichen dafür, dass das Binärmuster soeben den entsprechenden Wert durchlaufen hat, der dann z. B. mit dem Umschaltimpuls des Komparators in ein Register übernommen wird.

Das Direktverfahren setzt viele Komparatoren ein, wobei jeder eine andere Referenzspannung erhält. Das Ausgangsmuster aller Komparatoren gelangt auf einen Coder, der das endgültige Digitalwort ausgibt. Da die Komparatoren praktisch simultan ihre Ausgangszustände annehmen, spricht man von paralleler Wandlung. Die Wandlungsgeschwindigkeit ist dadurch sehr hoch (bis über 100 MHz), jedoch sind genauso viel Komparatoren nötig wie Quantisierungsstufen vorgesehen sind, was den Aufwand mit steigender Auflösung enorm erhöht (bei 8 Bit 256 Komparatoren, bei 10 Bit 1024

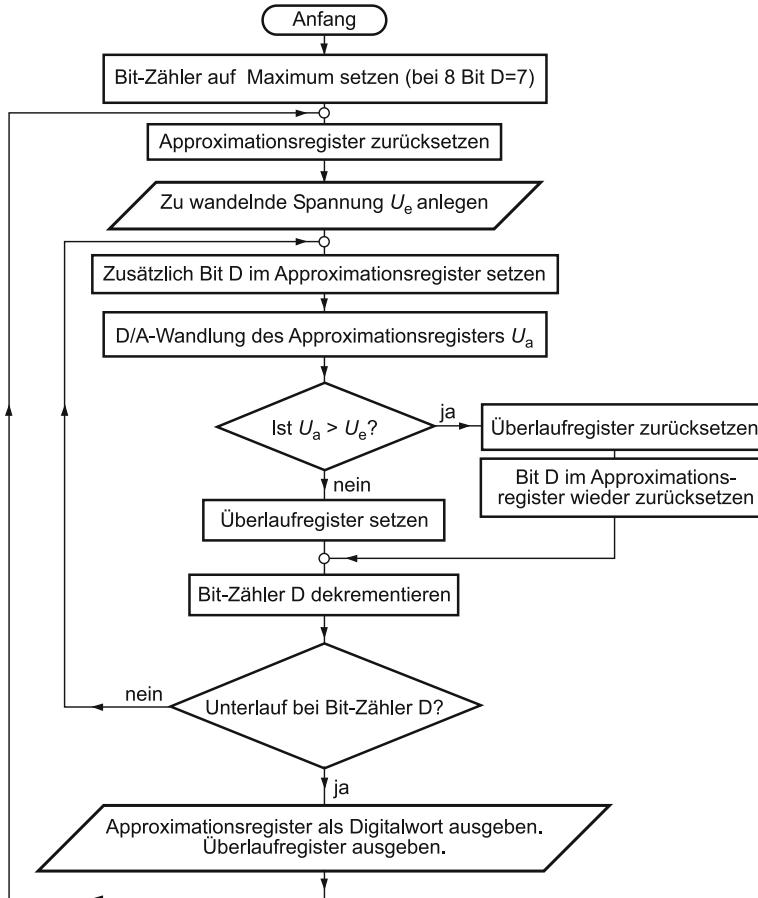


Abb. 2.19 Ablaufdiagramm der sukzessiven Approximation (Wägeverfahren)

Komparatoren). Die bei einer Auflösung von $D = 16$ Bit notwendigen $2D = 65536$ Komparatoren und Referenzspannungen macht das Direktverfahren für hochqualitative Aufzeichnungen im Audiobereich unwirtschaftlich.

In Bezug auf Wandlungsgeschwindigkeit und Aufwand (bzw. Auflösung) bleiben nur das Wägeverfahren und das Delta-Sigma-Verfahren als annehmbar übrig. Die sog. sukzessive Approximation (allmähliche Annäherung) arbeitet nach dem Flussdiagramm in Abb. 2.19.

Es werden hier also der Reihe nach die Bits eines Registers gesetzt (angefangen beim MSB), bis der nach analog rückgewandelte Wert größer als die Eingangsspannung ist. Ist dies der Fall, wird das soeben verwendete Bit im Register wieder zurückgesetzt. Der Vorgang dauert solange, bis man bei der feinsten Stufe, also dem LSB angelangt ist – ein optimaler, zeitsparender Vorgang, der auch beim Wiegen mit Gewichten angewandt wurde.

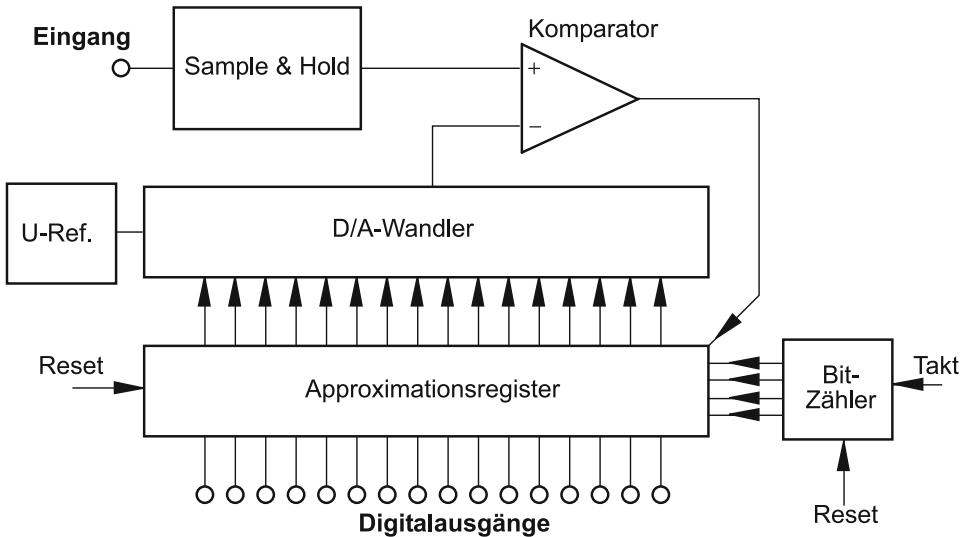


Abb. 2.20 Blockdiagramm der A/D-Wandlung nach der sukzessiven Approximation

Abb. 2.20 zeigt das Blockfunktionsschema eines solchen Wandlers. Die Konversionszeit steigt linear mit der Auflösung, d. h. 16 Bit benötigen die doppelte Zeit als 8 Bit bei gleichem Aufbau des Konverters. Die Genauigkeit hängt vom integrierten D/A-Wandler ab, der dem Komparator den Istwert liefert. Im Detail gibt es hier natürlich einige Schaltungsvarianten.

Die Alternative zur sukzessiven Approximation sind die heute üblicheren Verfahren der Delta-Sigma-Wandlung. Grundsätzlich entstehen hierbei aus Analogwerten ein serieller Bitstream, also eigentlich eine Folge aus 1-Bit-Werten. Ein Beispiel, bei dem Zahlenwerte von $-4 \dots +4$ in eine Binärfolge gewandelt wird, ist in Abb. 2.21 dargestellt. Der Wandler selbst arbeitet nach dem entsprechenden Blockdiagramm.

Als Ergänzung diene Tab. 2.1, die die ersten zwei Zeilen einer Tabellenkalkulation darstellt, welche zur Erzeugung der Grafik in Abb. 2.21 diente.

Der Delta-Sigma-Wandler besticht nicht nur durch seinen einfachen Aufbau – auch seine Linearität, seine Resistenz gegenüber Störungen und keine Missing Codes sind als Vorteile zu nennen. Außerdem kommt das Prinzip ohne Sample&Hold-Glied aus. Das dem Modulator nachgeschaltete Digitalfilter ist der aufwändigste Teil des Wandlers. Wie wir auch später sehen werden, hat es die Aufgabe, den Bitstrom in Digitalwerte bestimmter Breite umzuformen.

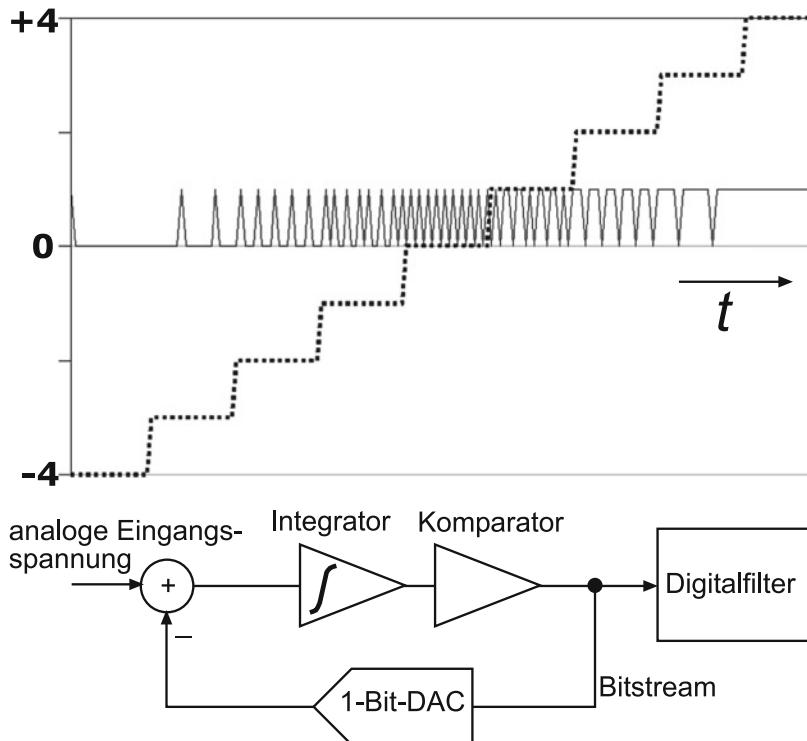


Abb. 2.21 Delta-Sigma-Wandlung verschiedener Analogwerte in ein Bitstream-Muster. Der analoge Spannungswert – der zwischen dem negativen und positiven Bereichswert liegen muss – gelangt über einen Addierer zum Integrator. Dieser summiert zu seinem letzten Wert den neuen Wert aus der Addierstufe. Der Komparator vergleicht den Integrationswert mit einem Schwellwert – im einfachsten Fall mit null. Bei Überschreiten gibt er 1 aus, sonst 0. Der DA-Wandler formt daraus einen Analogwert, und zwar den positiven bzw. den negativen Bereichswert. Dieser Wert wird negativiert und der Addierstufe zugeführt

Tab. 2.1 Tabellenkalkulation zur Simulation von Delta-Sigma. Der Bereich reicht von $-4 \dots +4$, die Eingangswerte sollten sich mindestens für 20 Zeilen wiederholen, damit sich auch ein Bit-Anteil einstellen kann

	A	B	C	D	E
1	Eingang	Add.-Ausg.	Integr.-Ausg.	Komparator-Ausg.	DAC-Ausg.
2	-4	=A2-E1	0	=WENN(C2<0;0;1)	=WENN(D2=1;4;-4)
3	-4	=A3-E2	=C2+B3	=WENN(C3<0;0;1)	=WENN(D3=1;4;-4)

2.3.3 Quantisierungsrauschen und Linearitätsfehler

Betrachten wir nochmals Abb. 2.16, so ist festzustellen, dass sich die ursprüngliche Analogspannung ja innerhalb eines Quantisierungsintervalls verändern könnte, ohne dass der gewandelte Binärwert eine Änderung zeigt. Das heißt mit anderen Worten, dass eine kontinuierliche Änderung der Analogspannung ein sprunghafte Digitalsignal hervorruft. Die Fehler bei der Quantisierung – und um solche handelt es sich letztlich – sind natürlich um so kleiner, je feiner die Abstufung, also je größer die Auflösung (in Bit bzw. Anzahl der möglichen Werte) bei der Wandlung ist.

Natürlich liegt auch der Einwand nahe, dass der Abtast-Halte-Vorgang ebenfalls eine Stufung ohne Zwischenwerte hervorruft. Die abgetasteten Werte sind jedoch (theoretisch) genau und bewirken somit keine den Quantisierungsfehlern ebenbürtige Effekte. Es sind also Quantisierungsfehler zu unterscheiden von der durch reine Abtastung entstehende Stufung eines Signals (die ja noch auf Analogseite entsteht). Stellt man sich vor, die Abtastrate sei sehr groß gegenüber der Frequenz des Nutzsignals, so sind Abtast-Artefakte weitab vom Nutzsignal, das Quantisierungsrauschen ist jedoch im hörbaren Bereich.

Die auftretenden Fehler bei der A/D-Wandlung kommen Verzerrungen gleich, jedoch treten sie bei einer sich zeitlich verändernden Eingangsspannung (z. B. Sinusspannung) statistisch auf. Eine solche statistische Fehlerspannung jedoch kennzeichnet sich immer als Rauschen und wird als solches auch empfunden. Rein zahlenmäßig lässt sich der Quantisierungs-Störabstand ausdrücken durch:

$$a_Q = 20 \cdot \log(2^D) + a_L \quad [\text{dB}] \quad (2.6)$$

D Auflösung (z. B. 8 oder 16 Bit), a_L Aussteuerungsmaß (bei Vollaussteuerung 0 dB).

Diese Bestimmungsgleichung beruht auf dem Verhältnis des Gesamtbereichs 2^D bezogen auf einen maximalen Quantisierungsfehler von $\pm 1/2$ LSB. Sie ist nicht exakt, da das Quantisierungsrauschen auch von der Signalform abhängig ist. Für Sinussignale erhöht sich a_Q etwa um 1,7 dB, hingegen kann sich bei allgemeinen Audiosignalen auch eine Verschlechterung (Abnahme) um bis zu -15 dB ergeben. Liegt das Aussteuerungsmaß unterhalb der Vollaussteuerung, verkleinert sich der Störabstand ebenfalls genau um diesen Wert.

Abb. 2.22 veranschaulicht das Entstehen des Quantisierungsrauschens, welches unabhängig von der Abtastrate ist (selbst bei unendlich hoher Abtastrate bliebe das Störsignal vorhanden). Abb. 2.23 zeigt zwei Sinusperioden eines 1-kHz-Signals, welches mit 20 dB unterhalb der Vollaussteuerung mittels Audiokarte aufgezeichnet wurde. Beide Perioden zeigen die Gegenüberstellung von 8- und 16-Bit-Auflösung.

Audiokarten mit einer Auflösung von 8 Bit lassen bei der Wiedergabe bereits bei Vollaussteuerung deutlich Rauschen vernehmen. Um das Quantisierungsrauschen möglichst gering zu halten, müsste hier der Aussteuerungsgrad ständig knapp unterhalb des Maximums eingestellt sein – die Gefahr einer Übersteuerung ist dann jedoch sehr groß. Bei Audiosignalen, die stark

Abb. 2.22 Statistische Quantisierungsfehler und sich daraus ergebendes Quantisierungsrauschen. Der Fehler entsteht hier lediglich durch Diskretisierung der Werte, wobei das Zeitkontinuum bei dieser Betrachtung bestehen bleibt. Das Fehlersignal entsteht durch Differenzbildung von Quantisierungssignal und Originalsignal. Man kann auch sagen, man erhält letztendlich die Summe aus Originalsignal und Störsignal. Letzteres hat stochastischen Charakter, obwohl auch eine Korrelation zum Originalsignal besteht

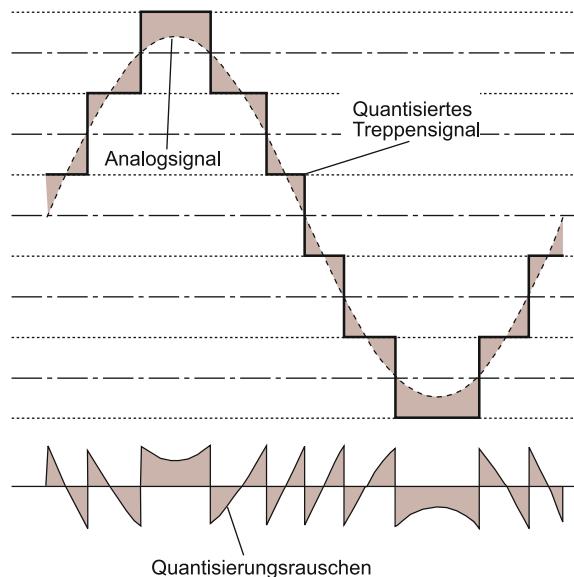
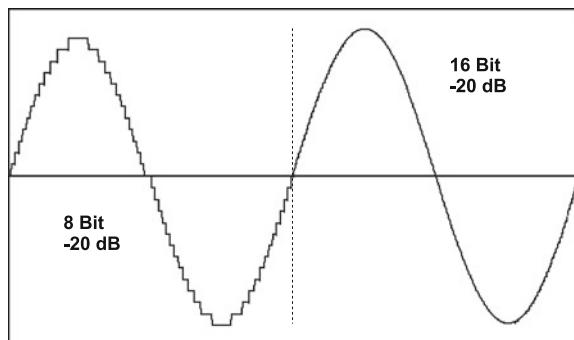


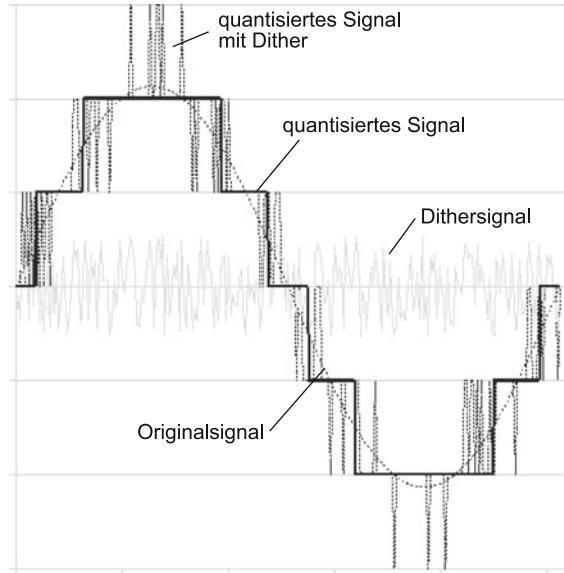
Abb. 2.23 Aufzeichnung je einer Periode mit einer Auflösung von 8 bzw. 16 Bit bei jeweils -20 dB Aussteuerung. Bei dieser geringen Aussteuerung gibt es bei 8 Bit nur noch ca. 25 Stufen, bei 16 Bit ca. 6500 Stufen



im Pegel schwanken und teilweise wenig überdeckende Höhenanteile beinhalten, kann dieses Quantisierungsrauschen noch wesentlich stärker hervortreten. Qualitätsaufnahmen sollten deshalb auf keinen Fall mit einer PCM-Auflösung von nur 8 Bit hergestellt werden. Der auflösungsbedingte Störabstand beträgt hier maximal 48 dB, bei 16 Bit sind es etwa 96 dB.

Methoden zur Reduzierung von Quantisierungsrauschen Man kann Quantisierungsrauschen durch Zugabe eines geringen echten Rauschanteils physiologisch – also in Bezug auf den Klangeindruck – abmildern, obwohl der rechnerische Rauschabstand dadurch geringer wird. Man spricht bei dieser Methode vom *Dithering*. Da die Diskretisierung beim Wandeln in Digitalwerte fehlerbehaftet ist, überdeckt das Originalsignal bei sehr kleinen Signalamplituden nur noch eine geringe Anzahl von Quantisierungsstufen (siehe Gl. 2.6). Das Quantisierungsrauschen nimmt einen großen Anteil des Gesamtsignals ein.

Abb. 2.24 Durch Dithering reduziertes Quantisierungsrauschen. Das Rauschsignal (Dither = zittern) sollte den Bereich von 1 LSB überstreichen. Es addiert sich dem Originalsignal, dann erfolgt Quantisierung



Wird nun *vor* der Quantisierung ein kleines Rauschsignal addiert, so verbessert sich die Werte-Auflösung, weil sich die Wahrscheinlichkeit der Niveaus nach dem Wert des Originalsignals ausrichtet. Im Extremfall – wenn das Originalsignal kleiner als eine Quantisierungsstufe ist, wird sich ohne Dithering keine Änderung des Digitalsignals ergeben – als quantisiertes Signal käme nur noch ein waagerechter Verlauf, also kein Signal. Mit Dithering jedoch ergeben sich mehrere Stufen, deren Häufigkeit mit dem Originalsignal einher geht. Obwohl das Rauschen *de facto* größer ist, tritt das Quantisierungsrauschen in den Hintergrund. Dither erhöht also den Rauschanteil, mindert jedoch den mittleren Quantisierungsfehler (Abb. 2.24).

Eine weitere weit verbreitete Methode zur Reduzierung des Quantisierungsgeräusches ist der Einsatz von *Noise Shaping*. Dieses Verfahren benötigt den Vergleich zwischen Wandlungswert und Originalwert. Nach jedem Abtasttakt wird diese Differenz invertiert auf den Eingang rückgeführt (auch *Error Feedback* genannt). Der Fehler wird also nach jedem Wandeltakt kompensiert, und die Mittelung aus den zwei Werten ergibt nur noch die Hälfte des ursprünglichen Fehlers. Damit diese Kompensationssprünge in möglichst hohe Frequenzbereiche verlagert sind, ist die Verwendung von *Überabtastung* angezeigt (siehe Abschn. 2.4 weiter oben).

Beides – Dithering und Noise Shaping – findet seinen Einsatz jedoch nicht nur generell bei der A/D-Wandlung, sondern vor allem bei der Reduzierung der Quantisierungstiefe eines Signals. So ist z. B. bei der Reduzierung von einer Auflösung von 18 Bit auf 16 Bit ohne die genannten Maßnahmen mit einem Zuwachs an Quantisierungsrauschen im Verhältnis zum Gesamtignal zu rechnen.

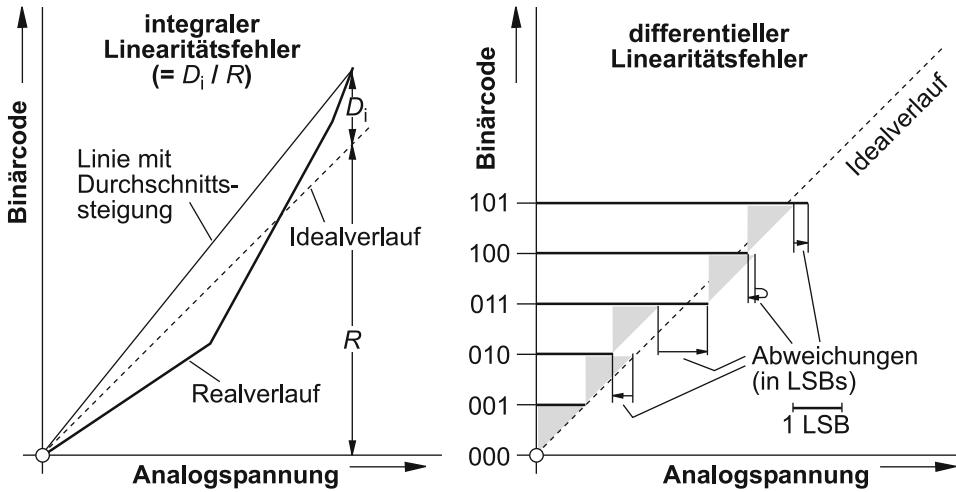


Abb. 2.25 Linearitätsfehler bei der Wandler. A/D-Wandler arbeiten nicht ganz ohne Fehler. Man kann hier prinzipiell in zwei Gütekategorien unterteilen, nämlich in absolute und in relative Linearität. Die absolute oder auch integrale Linearität sagt aus, wie groß die Abweichung der mittleren Steigung der Kennlinie vom Soll ist. Die Abweichung wird sodann relativiert, also bezogen auf die Sollsteigung. Hier kann die Kennlinie durchaus ohne Krümmungen sein, jedoch weicht sie durch Verstärkungsabweichung und Offset vom idealen Verlauf ab

Relative oder differentielle Linearitätsfehler sind Merkmale der Kennlinie eines A/D-Wandlers, siehe Abb. 2.25. Sind hier irgendwelche Krümmungen zu verzeichnen, so äußert sich dies bei der späteren Wiedergabe durch harmonische Verzerrungen, d. h. durch Entstehung von zusätzlichen Obertönen eines reinen Sinussignals. Bei der Bestimmung dieses Fehlers wird die Differenz zweier Analogwerte, die durch zwei genau benachbarte Binärkodes entstehen, gebildet. Diese Differenz gilt nun als individuelle Steigung an der betreffenden Stelle der Kennlinie. Der relative Fehler dieser Steigung bezüglich der sog. Full-Scale-Steigung entspricht dem differentiellen Linearitätsfehler. Diese Bestimmung wird an allen benachbarten Codes vorgenommen; der maximale Wert der differentiellen Nichtlinearität dient dann als Spezifikation, welche in %, ppm oder sehr häufig auch in LSBs (Abweichung des niederwertigsten Bit) Wandler mit der Angabe *no missing codes* definieren eine Fehlerangabe, die unter einem LSB bleibt. Es leuchtet ein, dass Abweichungen der Kennlinie von mehr als 1 LSB für eine linear ansteigende Analogspannung bei der Wandlung das Überspringen von Binärkodes verursachen. Man spricht hier auch von einer gestörten *Monotonie*.

Wie schon erwähnt, arbeiten Wandler nach dem Prinzip Delta-Sigma ohne systematische Linearitätsfehler.

Im Gegensatz zu analog arbeitenden Systemen wie Verstärkern, Kassettenrecordern usw. haben wir bei Digitalsystemen (wie z. B. einer Audiokarte) eine Kennlinie, die einen scharfen Knick oberhalb der Aussteuerungsgrenze aufweist (Abb. 2.26). Die Beschneidung (Clipping) setzt hier also abrupt ein, und sie ist außerdem normalerweise absolut

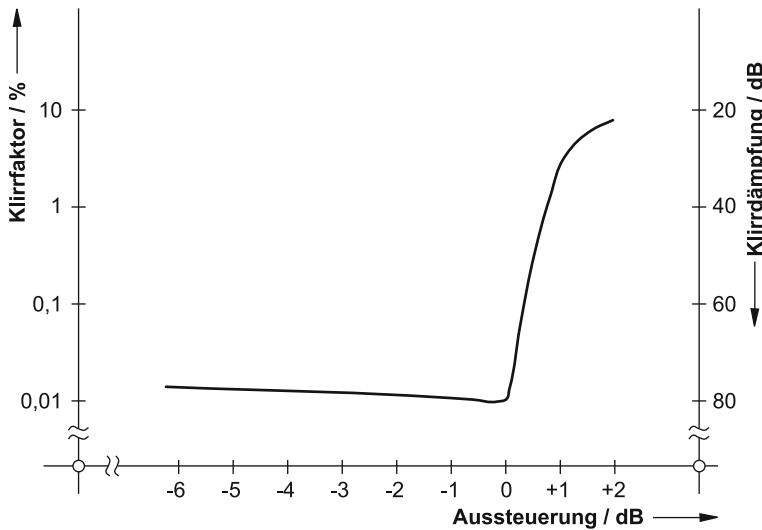


Abb. 2.26 Klirrfaktor in Abhängigkeit von der Aussteuerung bei digitalen Systemen. Am Ende der Kennlinie befindet sich zwangsläufig ein Knick – hier ist die absolute Grenze der Aussteuerung. Wird diese Grenze überschritten, wächst der *Klirrfaktor* drastisch an

symmetrisch zum Signalnullpunkt. Ein zeitweise übersteuerndes Musik- oder Sprachsignal erzeugt bei der Wiedergabe kurze Knistergeräusche (ähnlich wie beim Abspielen verkratzter Analogplatten), während Analogsysteme im Falle von ebenfalls kurzen Übersteuerungen meist noch überhaupt nichts Auffälliges erkennen lassen. Erst bei chronischer, starker Übersteuerung ähneln sich die Symptome bei Analog- und Digitalsystemen.

2.4 Verfahren zur Datenreduktion

Beim der digitalen Audioaufzeichnung fällt eine ansehnliche Datenmenge an. Als Bestimmungsgleichung für den Datendurchsatz gilt:

$$S = K \cdot F_A \cdot D / 8 \quad [\text{Byte/s}] \quad (2.7)$$

K Anzahl der Kanäle, F_A Abtastrate (Sampling-Frequenz), D Signalauflösung in Bit.

Das würde für den konkreten Fall bedeuten, dass ein Stereosignal ($K = 2$) bei einer Auflösung von 16 Bit und einer Sampling-Frequenz von 44,1 kHz pro Sekunde 176,400 Bytes geschrieben bzw. gelesen werden. Das würde für ein Musikstück mit einer Länge von 3 min immerhin einem Speicherplatz von ca. 30 MByte entsprechen.

In Anbetracht dieser Situation sind natürlich Bestrebungen aller Art im Gang, die Datenmenge reduzieren zu wollen. Abhängig davon, wie groß der Qualitätsverlust sein darf, sind hier durchaus konkrete Wege gangbar. Allerdings sind hier im Gegensatz zur

Kompression bei der Videotechnik keine so drastischen Ergebnisse zu erwarten. Das liegt hauptsächlich daran, dass im Audiobereich keine diskreten Datenpakete mit viel Redundanz und Ähnlichkeiten vorliegen, wie das ja bei Videosequenzen in Form von Einzelbildern der Fall ist. Infolgedessen sind die Kompressionsfaktoren für Audiodateien i. Allg. sehr viel ungünstiger aus. Wir werden dennoch weiter unten sehen, dass unter Berücksichtigung der physiologischen Höreigenschaften viel Information enthalten ist, was gar nicht wahrgenommen wird.

2.4.1 Datenreduzierung durch Parameteränderung

Wie wir im letzten Abschnitt erfahren und auch aus Gl. 2.7 entnehmen konnten, ist eine direkte Reduzierung der Datenmenge bzw. der Dateiengröße möglich, indem die Sampling-Frequenz oder auch die Auflösung herabgesetzt wird. Natürlich ist auch eine Beschränkung auf einen Kanal u. U. zu diskutieren.

Eine Sampling-Datei im Mono-Modus zu generieren ist dann empfehlenswert, wenn es sich tatsächlich um inhaltlich identische Signale auf beiden Kanälen handelt. Eine Panoramaeinstellung kann beim Abmischen auch software-technisch realisiert werden. Allerdings sind dann lediglich für die abzumischenden Stimmen Mono-Dateien hinreichend – die endgültige Abmischung muss wieder ein Stereoformat besitzen. Der allgemeine Fall schreibt jedoch wegen der aufzuzeichnenden Raumakustik auch für die Ressourcen den Stereomodus vor.

Eine nachträgliche Konvertierung einer Audiodatei ins Monoformat ist natürlich ohne weiteres möglich (was jedoch irreversibel ist). Das gleiche gilt für Änderungen der Sampling-Frequenz und der Auflösung.

2.4.2 Datenreduzierung durch Kompression

Wir sprachen weiter oben bereits einige Fakten an, die die Kompression für Audiodateien betreffen. Dieses Thema berührt natürlich auch Belange der Dateiformate, d. h. der Art und Weise der Anordnung von Audiodaten. Die am meisten praktizierte Verarbeitung und Speicherung geschieht im üblichen PCM-Modus. Das entspricht einer unveränderten Aneinanderreihung der Sampling- bzw. Quantisierungsdaten. Abhängig vom verwendeten Dateiformat (siehe Anhang A) ist jedoch jeder Datei ein sog. *Header* vorangestellt, der das genaue Datenprotokoll mit den Parametern und dem Datenarrangement repräsentiert.

Verschiedene Wege einer Kompression sind wiederum nicht für alle Anwendungsfälle geeignet. Beispielsweise sind komplexe Verfahren oftmals ungeeignet, sie in Echtzeit anzuwenden, d. h. während des Sampling-Vorgangs. Die Leistung des Prozessors im Rechner ist damit meist überfordert. Bessere Chancen bieten natürlich spezielle Signalprozessoren, die in Bezug auf spezielle Instruktionen besonders gut ausgerichtet sind und daher höhere Geschwindigkeiten aufweisen. In jedem Falle ist das Datenaufkommen nach Gl. 2.7 zu

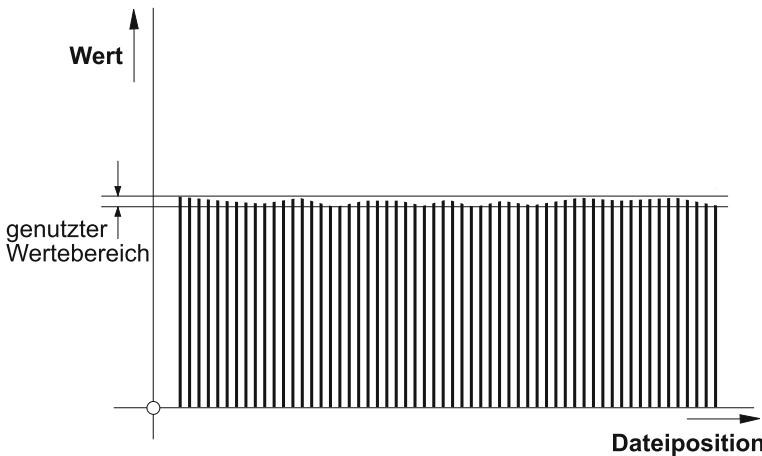


Abb. 2.27 Verminderte Nutzung des zur Verfügung stehenden Zahlenbereiches

bewältigen, sonst kommt es zu digitalen Dropouts, was besonders unangenehm und schwer zu ändern ist.

Jegliche Art von Kompression beruht auf der Tatsache, dass sich Datenmuster wiederholen oder aufeinanderfolgende Bytes in Bezug auf Wert nur wenig ändern. Zur Verdeutlichung diene Abb. 2.27.

Die hier sichtbaren Datenwerte (als Säulen dargestellt) mögen einer Audiodatei angehören, und zwar einer etwas leiseren Passage.

Verlustlose Kompression Gelingt es, die Daten in Abb. 2.27 einer kleineren Grundmenge zuzuordnen, so ist dies auch gleichzeitig mit einer Kompression verbunden. Nehmen wir an, es handele sich um eine Datei mit einer Auflösung von 8 Bit. Damit sind 256 Mengenwerte möglich. Bei sehr kleinen Änderungen von einem Dateiwert zum nächsten könnten z. B. auch 16 Mengenwerte ausreichen. Somit genügt eine Breite von 4 Bit (Nibble) zur Übertragung oder Speicherung aller hier vorliegenden Informationen, und wir kommen in diesem Fall auf einen Kompressionsfaktor von immerhin 0,5.

Das Audiosignal, welches im Profil der Säulen erkennbar ist, kann natürlich auch längerfristige Änderungen aufweisen. Abb. 2.28 zeigt ein Beispiel für solche „langsam“ Schwankungen, die jedoch immerhin den gesamten Mengenbereich ausfüllen.

Das untere Signal wurde aus oberem gewonnen, und zwar durch Differenzbildung zum vorangehenden Wert. Wie man sieht, besteht auch hier eine Möglichkeit, trotz Ausnutzung des gesamten Wertebereichs beim Original eine Reduzierung wie im ersten Fall vorzunehmen.

Unter bestimmten Situationen lässt sich der Wertebereich noch weiter verkleinern, indem die Bildung der Differenzwerte ein zweites Mal vorgenommen wird. Ob dadurch tatsächlich etwas gewonnen ist, hängt jedoch stark vom Inhalt des Audiosignals ab.

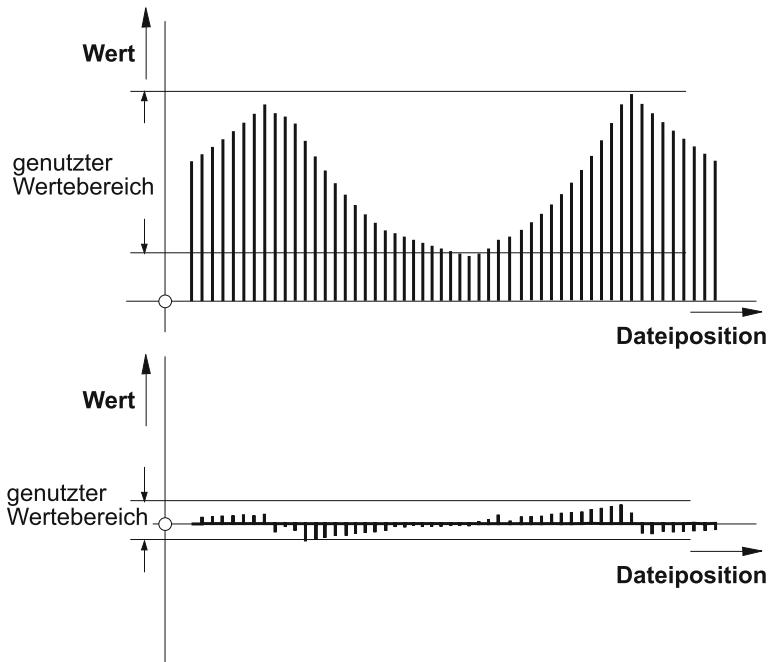


Abb. 2.28 Niederfrequente Auslenkungen und Reduzierung durch Differenzbildung

Ein Datenformat, das auf die oben beschriebene Weise aus dem PCM-Format gewonnen wird, heißt ADPCM. Teilweise erfolgt nach der Differenzbildung auch bereits die Kompression, wobei der Kompressionsfaktor vom Verlauf des Audiosignals abhängt. Leider fallen die durchschnittlichen Ergebnisse solcher Kompressionen nicht besonders günstig aus. Der Kompressionsfaktor beim ADPCM-Format hängt u. a. auch von der primären Quantisierungstiefe ab. Angewandten ADPCM-Verfahren haften teilweise auch Verluste an, die von der Quantisierung abhängen.

Verlustbehaftete Kompression Lässt man bei der Differenzbildung das LSB des Ergebnisses außer Acht, so fallen sehr viele Daten-Bytes auf einen Wert von null zurück, was sich sehr günstig auf den Kompressionsfaktor auswirkt. Dies kann mit höherwertigen Bits ebenfalls durchgeführt werden. Verständlicherweise schrumpft damit natürlich die Auflösung, was sich wiederum durch Quantisierungsrauschen und Verzerrungen äußert.

Weitere Möglichkeiten einer Kompression bestehen durch das sog. PASC-Verfahren (Precision Adaptive Sub-Coding), welches ein Stereosignal mit einem Datenstrom von ca. 50 kByte/s (anstatt ca. 170 kByte/s bei normaler PCM-Speicherung) verarbeiten kann. Das Prinzip beruht auf einem Splitting in verschiedene Frequenzbänder, die dann mittels PASC-Algorithmus auf Signalkomponenten hin untersucht werden, die aufgrund des Überdeckungseffektes und der Lautstärke-Trägheit für das menschliche Gehör nicht von Bedeutung sind. Trotzdem ist natürlich klar, dass solche Verfahren die Originaldaten nicht

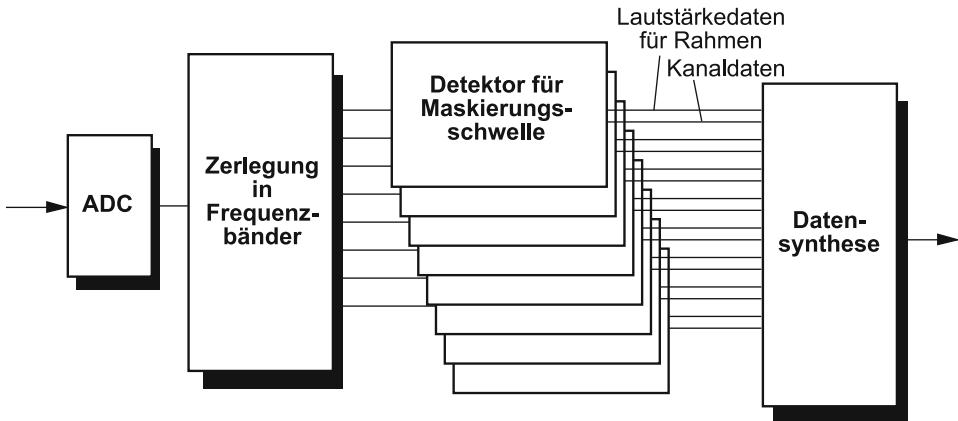


Abb. 2.29 Datenreduktion nach der PASC-Methode

erhalten können, obschon die Veränderungen wohl in den seltensten Fällen wahrgenommen werden können. Das Prinzip des PASC-Encoding und -Decoding ist Abb. 2.29 zu entnehmen.

Methoden, die dynamisch und in Abhängigkeit vom Obertongehalt die Abtastfrequenz verändern, sind ebenfalls möglicherweise eine vielversprechende Sache, vor allem auch in der Kombination mit der ADPCM-Codierung.

2.4.3 MP3-Verfahren

Ein weiteres Verfahren zur Datenreduzierung ist das MP3-Verfahren (eigentlich MPEG Layer 3), welches in weiten Bereichen Einzug gehalten hat. Den Hauptanteil dieser Erfindung trägt das Fraunhofer-Institut. Es handelt sich ebenfalls um ein Verfahren, welches verlustbehaftet ist. Der Anwender ist jedoch in der Lage, den Parameter für die Kompression selbst zu wählen. Eigentlich handelt es sich um die Angabe der Bitrate, die aufrecht erhalten bleiben soll. Dies bedeutet für ein Signal mit schwankender Komplexität eine variierende Audioqualität. Alternativ gibt es bei manchen MP3-Codern auch die Möglichkeit, die Qualität vorzugeben und damit den Datenstrom flexibel zu gestalten.

Eine genaue Beschreibung zur Funktionsweise wäre zu umfangreich – hier sei auf spezielle Literatur verwiesen. Die einzelnen Schritte im Verfahren sollen jedoch erläutert werden. S. 97 [3]

Die Bezeichnung *Layer 3* deutet darauf hin, dass es auch vorangehende Verfahren gab, die erst später ihre Perfektionierung erlangten. Vor allem das Hinzufügen einzelner Verarbeitungsblöcke (wie z. B. das MDCT=Modified Diskrete Cosinus Transformation) brachte einen entscheidenden Qualitätsschritt bei gleichem Kompressionsgrad.

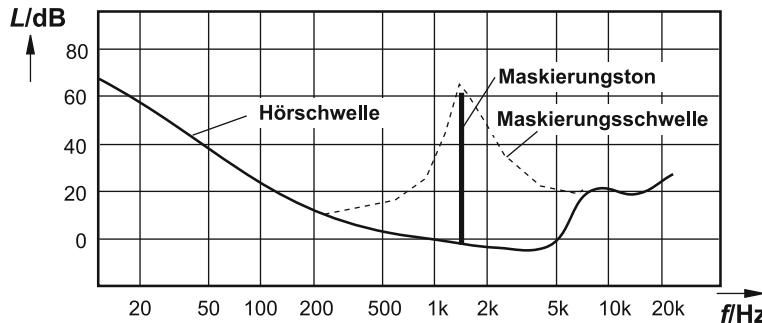


Abb. 2.30 Hörschwelle und Veränderung durch Maskierung. Die untere Linie zeigt den Verlauf der Hörschwelle – nur Töne mit oberhalb liegendem Schalldruck werden überhaupt wahrgenommen. Sobald ein starker Ton erscheint (dicke Linie), verändert sich die Hörschwelle, und zwar hauptsächlich in unmittelbarer Nachbarschaft zur Frequenz des starken Tones. Dann werden nur noch Töne wahrgenommen, die einen Pegel oberhalb der Maskierungslinie (gestrichelt) aufweisen

Hauptidee Der grundlegende Schritt für die MPEG-Kompression besteht darin, dass man nicht die zeitlichen Abtastwerte zur Speicherung heranzieht, sondern die im Frequenzbereich. Das eröffnet mehr Möglichkeiten zur Beurteilung und Behandlung der Daten zur Datenreduktion. Außerdem ist diese Methode – nicht die in der Zeitdomäne anfallenden Daten zu speichern, sondern die in Blöcken aufgeteilten Daten in der Frequenzdomäne – entscheidend für einen immens großen Schritt der Kompressionsgrads. Eine Verschiebung der Tonhöhe (Pitch) ist nach der Wandlung in Spektraldata ebenfalls sehr einfach möglich, warum viele MP3-Spieler die Tonhöhenverschiebung ebenfalls als Option anbieten.

Verdeckungseffekte Ähnlich wie beim PASC-Verfahren (siehe 2.4.2), nutzt auch MP3 gewisse Unzulänglichkeiten des menschlichen Ohres aus. Dabei spielen sog. Verdeckungseffekte (auch *Maskierung* genannt) die größte Rolle. Man stelle sich vor, es liegen zwei Sinussignale zeitgleich vor, die sich nur wenig in ihrer Frequenz, aber stark in ihrem Pegel unterscheiden. Es ist zu erwarten, dass der schwächere Ton nicht mehr zu vernehmen ist. Es gibt einen Verlauf für die Mindestlautstärke, die ein Ton bestimmter Frequenz aufweisen muss, um wahrgenommen werden zu können. Diese Kurve ist die sog. Hörschwelle und ist frequenzabhängig, wie Abb. 2.30 zu entnehmen ist. Ist ein starker Ton in einem Frequenzgebiet vorhanden, so verläuft zufolge Verdeckungseffekts die Hörschwelle in der Nachbarzone stark angehoben, d. h. frequenzähnliche Töne benötigen einen starken Pegel, um die Schwelle zu übersteigen und damit vernehmbar zu werden.

Analog hierzu gibt es einen zeitlichen Verdeckungseffekt, bei dem ein abrupt abbrechender, lauter Ton von einer Pause von wenigen Millisekunden gefolgt wird, in denen das Ohr sehr leise Ereignisse nicht registriert. Eine solche „Pause“ spart verständlicherweise dann Speicherplatz.

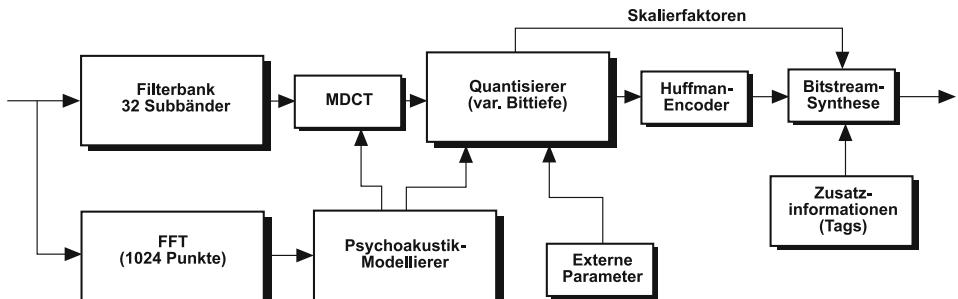


Abb. 2.31 Blockschema des MP3-Encoders

Schritte beim Encodieren Die einzelnen Phasen während der MP3-Codierung sind aus dem prinzipiellen Aufbau des Coders in Abb. 2.31 zu entnehmen.

Erster Schritt: Aufteilung des Datenstroms in Frames und Frequenzbänder Jedes Frame enthält entweder 1152 (long frame) oder 384 Samples (short frame). Daraus werden zunächst spektrale Daten für 32 Subbänder generiert. Für die untersten beiden Subbänder können die längeren Blöcke zur Anwendung kommen, damit die Frequenzauflösung bei tiefen Frequenzen noch hinreichend ist, während die höheren Frequenzen mit 384 Samples pro Block noch eine genügend hohe Auflösung aufweisen. Zu den 32 Subbändern ist außerdem hinzuzufügen, dass diese keine gleichmäßige Aufteilung des Gesamthörbereichs darstellen, sondern eine gehörrichtige. Das heißt, im Bereich der höchsten physiologischen Empfindlichkeit und Auflösung von etwa 1 bis 4 kHz befinden sich dichtere Bänder als darunter oder darüber. Der mögliche Bereich der Abtastraten der Ein- und Ausgangsdaten ist von 16 bis 48 kHz.

Zweiter Schritt: Eindimensionale Kosinustransformation (MDCT) Dieser Vorgang bewirkt im Endeffekt eine weitere, weitaus feinere Unterteilung in endgültig 576 Frequenzbänder. Dies ist ein spezieller Vorgang, bei dem die sonst entstehenden starken Artefakte durch Überlappung – wie sie bei der Mehrphasenfiltertechnik entstehen – unterdrückt werden. Stattdessen kann es durch die verbesserte Auflösung im Frequenzbereich und folglich verschlechterte Auflösung im Zeitbereich bei starken Pegelsprüngen zu sog. *Vorechos* kommen. Die oben erwähnte duale Blocklänge reduziert die Entstehung dieser Artefakte etwas. Grundsätzlich handelt es sich eigentlich nicht um regelrechte Vorechos, sondern um Quantisierungsrauschen, welches vor dem Einsetzen eines Signalpegelsprungs auftritt. S. 97 [4]

Zeitgleich werden vom Eingangssignal durch eine getrennte FFT Spektrumsdaten gewonnen und für den Bereich des psychoakustischen Modells bereitgestellt. Die Vorgänge dort sind recht komplex, denn das spektrale Profil muss allerhand Kriterienprüfungen durchlaufen. Danach wird entschieden, wie der Verlauf der Hörschwellenkurve zu wählen

ist – alles, was bezüglich Pegel darunter liegt, kann ignoriert werden und ist demnach nicht relevant für zu speichernden Code.

Dritter Schritt: Die eigentliche Codierung Die weitere Verarbeitung erfolgt im Quantisierer, der berücksichtigt, welche Frequenzdaten nach der Hörschwellen-Bearbeitung noch bestehen bleiben, die dann letztlich auch tatsächlich zu codieren sind. Die Bittiefe wird dynamisch so gewählt, dass das durch Bitreduktion entstehende Quantisierungsrauschen gerade noch verdeckt bleibt. Gleichzeitig werden entsprechende Skalierungsfaktoren der späteren Datensynthese übermittelt.

Der Quantisierer reduziert seine Bittiefe natürlich auch nach externen Vorgaben des Anwenders. Dabei sind nicht nur feste Werte einstellbar, sondern auch die variable Codierung, welche die Auflösung je nach Signaldaten ständig anpasst.

Vierter Schritt: Huffman-Codierung Diese Art der Codierung bringt eine weitere, teilweise beträchtliche Datenreduktion. Nun werden die Häufigkeiten der Sampling-Werte (der Frequenzen) untersucht. Zahlreich vorkommende Werte erhalten einen Ersatzwert in einer Tabelle, der klein ist und somit wenige Bits benötigt. Selten auftretende Samplingwerte dagegen erhalten einen großen Ersatzwert. Diese Huffman-Tabelle muss als Code ebenfalls im endgültigen Datenstrom mit eingebaut werden. Sie hat nur Gültig für das jeweils gerade zu bearbeitende Frame.

Fünfter Schritt: Bitstream-Ausgabe Im letzten Schritt sind alle vorhandenen Daten zusammenzusetzen. Natürlich handelt es sich dabei nicht nur um die Huffman-Tabelle, die Skalierfaktoren und die eigentlichen Signaldaten, sondern auch die sog. MP3-Tags, der Fehlercode und die sonstigen Beschreibungsdaten für das spezielle MP3-File.

Joint Stereo Der Anwender kann nicht nur die Qualität der Kompression über die Bitrate definieren, sondern kann auch festlegen, welche Art der Stereo-Verarbeitung zu wählen ist. Im allgemeinen Fall müssten beide Kanäle getrennt behandelt werden, denn sie könnten ja völlig unabhängigen Inhalt (also einen Korrelationsgrad von null) haben. In der Praxis sind jedoch häufiger beide Stereokanäle ähnlich, d. h. ein gewisser Anteil ist bei beiden Seitensignalen gleich. Sind im Extremfall sogar die Unterschiede nur gering, so ist die Generierung eines MS-Stereosignals vorteilhaft. Im Gegensatz zum normalen Stereosignal sind also nicht die Signale L und R weiter zu verarbeiten, sondern das Mittensignal durch $M = L/2 + R/2$ und das Seitensignal durch $S = L/2 - R/2$ definiert. Siehe auch Abschn. 1.9.1. Durch Addition und Subtraktion von M und S entstehen dann wieder die ursprünglichen Signale L und R .

Ob die Wahl von *Joint Stereo* zur Verkleinerung der Datenmenge führt, hängt allein vom Signal ab. Bei großer Signalähnlichkeit von L und R ist J. S. auf jeden Fall zu empfehlen, während es bei geringer Ähnlichkeit auch zu Nachteilen kommen kann. Auch Qualitätseinbußen sind denkbar (Abb. 2.32).

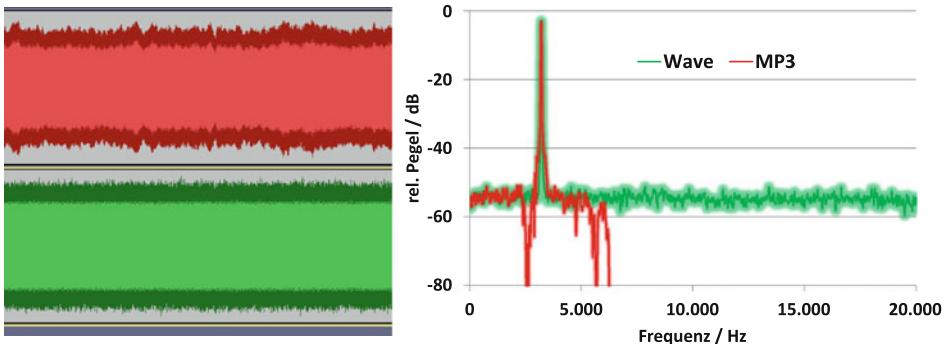


Abb. 2.32 Frequenz-Verdeckungseffekt. Einem starken Sinuston von 3200 Hz wird Weißes Rauschen mit geringerem Pegel überlagert. Danach wird daraus eine MP3-Datei erzeugt (32 kBit/s). Links sind die Signale im Zeitbereich dargestellt, unverändert in Grün und MP3 in Rot. Rechts sind die Spektren dargestellt. Man erkennt deutlich die Veränderungen nach der MP3-Codierung, sowohl die etwas holprig anmutende Amplitudensprünge als auch die Einschnitte im Frequenzspektrum. Allerdings ist auch zu bemerken, dass die Speicherbedarfe von Original und MP3 etwa im Verhältnis von 20:1 stehen

Metadaten und Tags Neben den eigentlichen, für das Audiosignal relevanten Daten, gibt es noch einige Anwenderdaten, die mit in die MP3-Datei aufgenommen werden können. Es handelt sich dabei um Daten wie Kennung des Tags, Songtitel, Künstler/Interpret, Album, Erscheinungsjahr, beliebiger Kommentar und Genre. Es gibt unterschiedliche Versionen von Tags, die mitunter nicht kompatibel zueinander sind und sich auch darin unterscheiden, ob sie am Anfang oder am Ende der MP3-Datei befinden. Auch gibt es teilweise Probleme, die diverse MP3-Player haben, wenn sie auf Tag-Daten stoßen.

Rückwandlung/Decodierung Der umgekehrte Weg zurück zum normalen PCM-Signal geschieht entsprechend dem Encoder-Blockschaltbild. Allerdings entfallen aufwändige Funktionsbereiche wie Psychoakustik-Modellierer, der im Encoder ja umfangreiche Berechnung zur Generierung der Hörschwellenkurve durchführen soll.

MP3 Surround Das Standard-MP3-Format kann eigentlich nur ein einfaches Stereosignal verarbeiten. Ein um den Raumklang 5.1 erweiterter Codec ist der MP3 Surround. Es handelt sich eigentlich um die Gewichtung der Kanäle ergänzten MP3-Daten. Der Schlüssel hierzu ist *Binaural Cue Coding (BBC)*. Aus einem Multiple-Kanal-Signal werden demnach Zusatzdaten gewonnen, welche die Raumverteilung spezifizieren. Siehe auch Abschn. 1.9, S. 97 [5,6]

Simultan durchlaufen die Kanaldaten ein Downmixing-Modul, welches das kompatible binaurale Stereosignal generiert. Dieses gelangt dann in einen MP3-Coder. Die oben genannten Surround-Daten werden in den restlichen Bitstream des Coders integriert.

Bei der Rückwandlung in die einzelnen Kanaldaten muss der Decoder lediglich die Multikanal-Zusatzdaten wieder extrahieren und dem BCC-Decoder zuführen, der dann seinerseits aus dem normalen Binaural-Stereosignal das mehrkanalige Signal formt.

2.4.4 Sonstige Verfahren

Advanced Audio Coding AAC Das Verfahren ist eine Weiterentwicklung aus dem MP3-Standard und den Kriterien von MPEG-4. Auch hier handelt es sich um verlustbehaftete Kompression. Die einzelnen Punkte, mit denen sich AAC gegenüber MP3 heraushebt, sind:

- Weiterer Bereich der möglichen Abtastraten von 8 bis 96 kHz (lediglich 16 bis 48 kHz bei MP3).
- Polyphonie bis 48 Kanäle (herkömmliches MP3 bietet nur 2 Kanäle).
- Predictorale Kompression
- Joint Stereo: Schnellere Bitrate-Anpassung
- Verbessertes, flexibleres Huffman-Coding
- Reines MDCT-Filter statt eines Hybridfilters

Vorbis Ein recht populäres Format der Firma *Xiph.Org Foundation* mit dem Namen Vorbis. Oft wird auch vom OGG-Format gesprochen, was jedoch nicht ganz korrekt ist, denn OGG bezeichnet lediglich ein Containerformat.

Zunehmender Zustimmung erfreut sich dieses Format bzw. Kompressionsverfahren nicht zuletzt deswegen, weil man hier keinerlei Patentrechte beachten muss – jedermann kann es zu jedwedem Zweck anwenden. Encoder und Decoder sind nicht nur als fertige Software erhältlich, sondern liegen auch als Quellcode vor.

Natürlich werden bei alternativen Verfahren immer Vergleiche angestellt – so auch hier: Es wird dabei oft behauptet, dass die Vorbis-Audiodateien bei gleicher Größe bessere Audioqualität liefern würden als MP3-Coder. Diese Vergleiche setzen klar definierte Randbedingungen voraus, damit reproduzierbare Ergebnisse entstehen. Außerdem ist zu beachten, dass es verschiedene Encoder bei beiden Verfahren gibt.

Auch der Autor hat einen solchen Test vollzogen: Das oben genannte Beispiel von Abb. 2.33 macht bei ähnlicher Dateigröße beim Vorbisformat nach dem Wiedereinführen ein Signal, welches vom Original kaum zu unterscheiden ist. Die Pegelschwankungen wie beim MP3-Format sind hier nicht vorhanden, folglich ist auch keinerlei Rauheit des 440-Hz-Tones hörbar. Ein solches Testsignal stellt jedoch auch einen Sonderfall dar – übliche Signale aus Musik enthalten wesentlich mehr unterschiedliche Situationen, bei denen mehrere Kompressionsmechanismen greifen würden.

Was ebenfalls bei diesem Test auffiel, sind fehlende Vorechos und ein Ausbleiben des zeitlichen Versatzes.

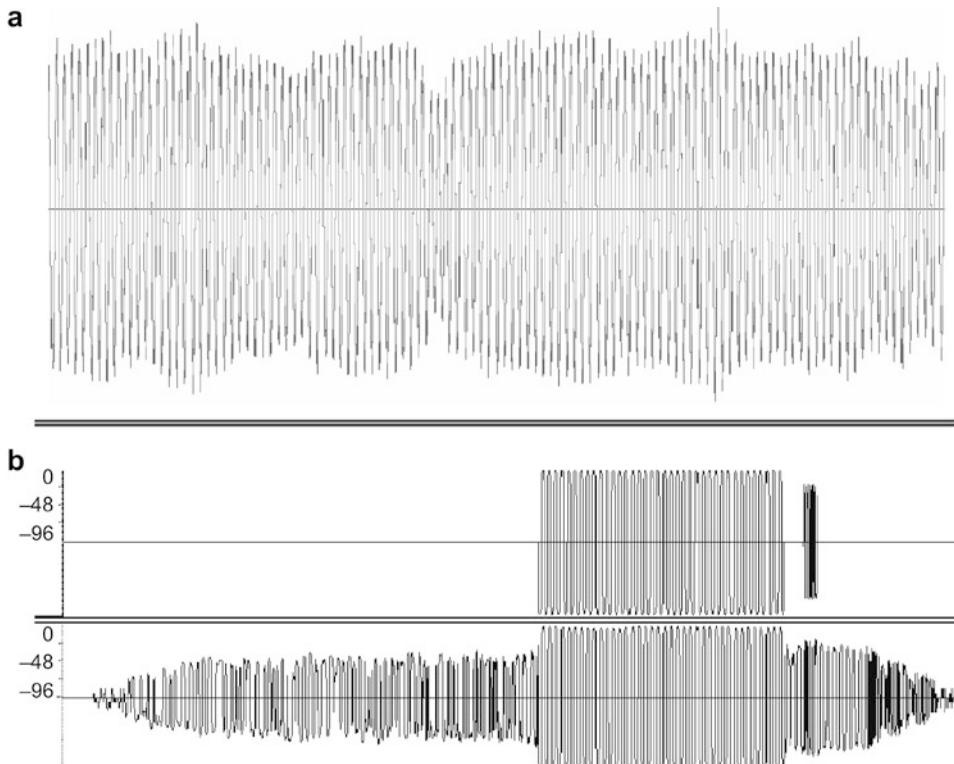


Abb. 2.33 Signale, die per MP3 deutlich verändert wurden. Beide Signale entstanden durch einen Lame-Encoder, 64 kBit/s. Das obere Signal (a) ist eine Mischung aus einem 440-Hz-Sinuston mit -2 dB und einem Rauschsignal vom Pegel -22 dB. Man sieht deutlich die enormen Pegelschwankungen, die durch Rauschüberlagerung hervorgerufen werden. Diese Artefakte sind auch sehr deutlich zu hören – der Ton wird rau empfunden. Beim unteren Signal (b), welches zur deutlicheren Darstellung logarithmisch skaliert ist, handelt es sich um einen abrupt einsetzenden Sinuston von 3000 Hz, gefolgt von einer kurzen Pause, danach ein sehr kurzer hoher, aber leiser Ton. Das Original ist in der oberen Hälfte dargestellt. Nach dem Encodieren (untere Hälfte) zeigen sich deutliche Vorechos sowie ein sehr unsauberer Einsatz des nachfolgenden hohen Tones

Vorbis setzt stets auf variable Bitrate, um sich den laufenden Erfordernissen flexibel anzupassen. Dabei gibt es keine minimalen oder maximalen Blockgrößen. Auf die Verarbeitung mithilfe des psychoakustischen Modells wird für die Kompression Wert gelegt, ebenso wie auf die Verwendung eines Hybridfilters zur Verbesserung der Reaktion auf signalmäßige Transienten bzw. zeitlich kritische Vorkommnisse. S. 97 [7]

Square-Polar-Mapping Die Eigenheit von Vorbis gegenüber MP3 ist jedoch nicht die Art der Maskierung – also Berücksichtigung der Verdeckungseffekte –, sondern vielmehr das Handling bei mehr als einem Kanal (der oben angeführte Test wurde anhand eines

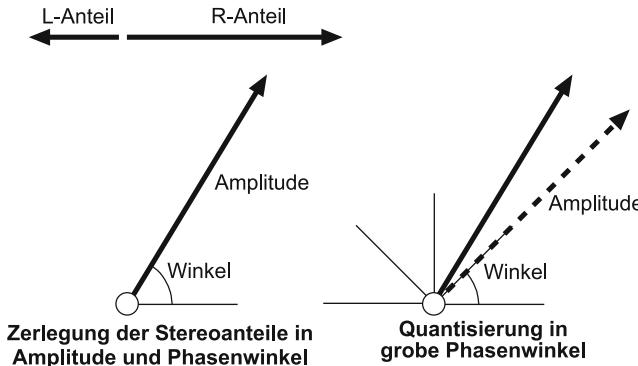


Abb. 2.34 Zerlegung eines Abtastwertes eines einzigen Spektrumintervalls in Amplitude und Phase beim Square-Polar-Mapping. Im rechten Teil der Abb. wird zwar derselbe Amplitudenwert verwendet, jedoch einer der fünf Phasenwinkel, der dem Original am besten entspricht

Monosignals durchgef\u00fchrt). W\u00e4hrend MP3 ja entweder unabh\u00e4ngige binaurale Stereo-kan\u00e4le oder aber nur ein Monosignal und ein Seitensignal verarbeitet (bei Joint-Stereo), wird bei Vorbis ein anderer Weg gegangen.

Nach der Aufteilung in die kleinen Frequenzb\u00e4nder, also noch vor dem Sampling der Spektralwerte, werden f\u00fcr beide Kan\u00e4le (bei Stereo) die Intensit\u00e4tswerte individuell ermittelt und diese gespeichert. Es handelt sich also f\u00fcr jeden Spektralwert um zwei zu quantisierende Parameter, die Amplitude und den Phasenwinkel. W\u00e4hrend ersteres mit der \u00fcblichen Genauigkeit erfolgen sollte, ist die Aufl\u00f6sung des Phasenwinkels wesentlich geringer und erf\u00fcllt dennoch die Anforderung an die r\u00e4umliche Genauigkeit hinreichend. Wie dies schematisch aussieht, ist in Abb. 2.34 dargestellt.

Mit dieser Methode gelingt nat\u00fcrlich nicht nur die Abbildung binauraler Signale, sondern auch von Raumsignalen mit sechs (5.1) oder mehr Kan\u00e4len, ohne dass der Speicherbedarf proportional mit der Anzahl der Kan\u00e4le ansteigt. Allein die Tatsache, dass nat\u00fcrliche Musiksignale mit Raumklangeigenschaften einer Verteilungsbeschreibung folgen, erm\u00f6glicht eine hohe Kompression, weil die Kan\u00e4le einen hohen Redundanzanteil besitzen.

Das verwendete Kanal-Modell kann variieren – die in Abb. 2.34 dargestellte Methode hat lediglich schematischen Charakter. Ebenso ist es m\u00f6glich, sogar innerhalb einer Audiodatei das Kanal-Modell zu wechseln.

Noise-Normalization Vorbis beinhaltet auch eine Methode, die sich nicht Reduzierung des Quantisierungsrauschen zum Ziel macht, sondern die Bewahrung der spektralen Energie. Bei sehr niedrigem Signal-/Rauschabstand kann es bei anderen Verfahren zu metallisch klingenden Artefakten kommen, die man bei Hervortreten des Quantisierungsrauschen beobachtet (siehe Abschn. 2.3.3). Ferner kann es bei minimiertem Quantisierungsrauschen auch zu zeitweise Unterdr\u00fcckung des Nutzsignals kommen, was bei Noise-Normalization verhindert wird. \u00c4hnliche Zielf\u00fchrung besteht bei Einsatz von

Dither-Methoden (siehe Abschn. 2.3.3), auch dort ist eine Verbesserung der Nutzsignal-Verständlichkeit zu verbuchen, obwohl ja die Rauschleistung vergrößert wird. S. 97 [7]

Die Werte der spektralen Energien, die jedem zu codierenden Block zugeordnet sind, werden quantisiert und bei nachfolgenden Blöcken nur noch der Unterschied zu Vorhersagewerten erfasst (Delta-Codierung). Die Abweichung zur Spektrumenergie des Original-PCM-Signals führt man einer speziellen Quantisierung zu (*Pyramid Vector Quantisation*). S. 97 [8]

AC3-Codec Während MP3 sich auf den Audiobereich beschränkt, handelt es sich bei AC-3 um einen Kompressionscodec für den Video- und Filmbereich, entwickelt von *Dolby Laboratories*. Das Verfahren bearbeitet Kanäle von Mono bis 5.1. Das unveränderte Datenformat muss dabei zunächst in den Abtastraten 32 kHz, 44,1 kHz oder 48 kHz vorliegen, die Datentiefe kann bis 20 Bit gehen. Auf das eigentliche Kompressionsverfahren soll hier nicht näher eingegangen werden, die Arbeitsweise ist generell dem von MP3 ähnlich.

2.5 Rückwandlung und Wiedergabe

Ausgehend von gespeicherten und gelesenen Digitaldaten kommt bei der Wiedergabe die Prozedur der Rückwandlung in Spannungswerte zum Zuge, um möglichst genau das ursprüngliche Analogsignal zu gewinnen. Wiederum sind einige verschiedene Wandlungsmethoden zu nennen sowie ein paar ihrer Schwächen.

2.5.1 D/A-Wandlungsarten

Zählverfahren Ein in der Technik der Mikrocontroller häufig anzutreffendes Verfahren ist das Zählverfahren. Ein Taktgenerator inkrementiert einen Zähler laufend von null bis 2^n , und dessen Stand wird dabei laufend mit dem momentan übergebenen Digitalwert verglichen. Je nachdem, wie dieser Vergleich ausfällt, erfolgt H- oder L-Ausgabe. Das ausgegebene Signal entspricht einer Pulsweitenmodulation (PWM), und die Rückwandlung kann durch einen einfachen Tiefpass geschehen. Dieses Verfahren ist für die anspruchsvolle Audiotechnik weniger geeignet, da man für eine Auflösung von 16 Bit 65,536 Zählerstufen benötigt. Daraus folgt für eine maximal zu übertragende Frequenz von 20 kHz zufolge Shannon-Theorem eine Abtastfrequenz von mindestens 40 kHz, und während einer Periode sind die 65,536 Stufen zu durchlaufen. Die erforderliche Taktfrequenz für den Zähler wäre somit etwa 2,6 GHz, was technisch unpraktikabel bzw. zu aufwändig ist.

Direktverfahren Bei der direkten Wandlung wird der Digitalcode zunächst in ein 1-aus-n-Signal gewandelt. Das bedeutet z. B. bei einem 8-Bit-Wert, dass von 256 Leitungen genau eine H-Zustand hat, während alle anderen 255 L-Zustand aufweisen. Die Leitungen

sind mit Schaltern verbunden, welche wiederum an einer Serienschaltung von Widerständen eine entsprechende proportionale Spannung abgreifen. Für Audiotechnik ist auch diese Methode nicht praktikabel, da für eine Auflösung von 16 Bit ein aufwändiger 1-aus-65536-Decoder und 65536 Schalter notwendig wären.

Parallelverfahren Bei der parallelen Wandlung benötigt man im Gegensatz zum Direktverfahren nicht 2^n , sondern nur n Schalter. Alle Schalter führen zu in 2er-Potenzschritten gewichteten Widerständen, so dass letztlich pro Bit Ströme in ebensolcher Gewichtung zur Verfügung stehen. Siehe Abb. 2.35.

Alternativ hierzu kommt ein geschaltetes R - $2R$ -Netzwerk in Frage, welches mit nur zwei Widerstandswerten auskommt und deshalb mit höherer Genauigkeit und einfacher zu realisieren ist. Siehe Abb. 2.36. S. 97 [1]

Abb. 2.35 D/A-Wandlung nach dem Wägeverfahren mit gewichteten Widerständen

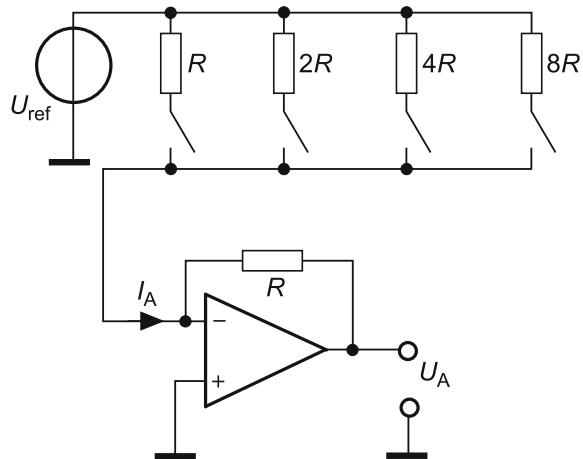
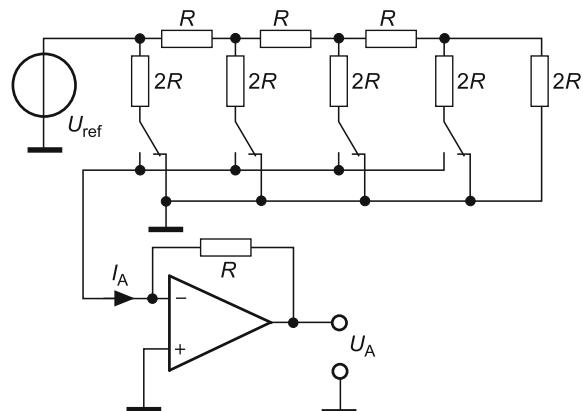


Abb. 2.36 D/A-Wandlung nach dem Wägeverfahren mit R – $2R$ -Widerstandsnetzwerk



Delta-Sigma Das Verfahren wurde bereits bei der A/D-Wandlung erörtert (siehe Abschn. 2.3.2). Es erlaubt auch den umgekehrten Weg zurück ins Analoge. Liegen auf Speicherseite keine echten Digitaldaten, sondern ein Bitstream der Breite 1 Bit vor – was ja ein Delta-Sigma-A/D-Wandler ohne Digitalfilter direkt liefern würde –, fiele die Rückwandlung recht einfach aus, denn ein analoger Tiefpass ohne besondere Eigenschaften würde genügen. Eine solche Methode kommt übrigens bei der *Super Audio Compact Disc* (SA-CD) zur Anwendung.

2.5.2 D/A-Quantisierungsrauschen

Die Ursachen für das Entstehen von Quantisierungsrauschen bei der A/D-Wandlung hatten wir bereits am Beginn dieses Kapitels (Abschn. 2.3.3) kennengelernt: Die diskreten Wandlungsstufen kommunizieren mit jeweils einem ganzen Bereich von analogen Spannungswerten, und der bei der Abtastung tatsächlich vorliegende Spannungswert ist innerhalb dieser Grenzen rein zufällig. Die fluktuierende Diskrepanz bedeutet Rauschen.

Der Sachverhalt bei der Rückwandlung ist eine andere. Wir gehen bei dieser Betrachtung davon aus, die zeitdiskreten Digitalwerte, die der D/A-Wandlung zugeführt werden, seien exakt. Je nach Wandlungsart ist mit statistisch verteilten Linearitätsfehlern zu rechnen. Allerdings werden heute häufig Delta-Sigma-Wandler eingesetzt, die systembedingt keine Linearitätsfehler aufweisen. Andere Wandler (z. B. solche mit R-2R-Widerstandnetzwerken) unterliegen einer gewissen Bauteilstreuung, die Fehler im Ausgangssignal bewirken.

2.5.3 Überabtastung auf der D/A-Seite

Das Thema *Überabtastung* wurde bereits in Abschn. 2.2.4 behandelt. Grundsätzlich findet vor der Wandlung in ein Analogsignal ebenfalls eine Überabtastung statt, damit die Unterdrückung von Abtastspektren gelingt.

2.6 Signalaufbereitung auf der Analogseite

2.6.1 Aussteuerung

Die Aussteuerung bei einem digitalen System ist mindestens ebenso kritisch wie bei einem analogen. Bei falscher Aussteuerung sind jedoch andersartete Phänomene zu erwarten.

Während das analoge System bei zu kleiner Aussteuerung ein verrausches Wiedergabesignal zeigt, fällt beim digitalen System die effektive Auflösung ab. Gleichermassen steigt relativ zum Gesamtsignal das Quantisierungsrauschen an, welches jedoch ein anderes Verhalten hat als normales Verstärker- oder Bandrauschen. Es erscheint nämlich

nur mit einem Nutzsignal. Bei zu großer Aussteuerung (Übersteuerung) setzt beim digitalen System eine abrupte Beschneidung ein, was sich durch einen steil anwachsenden Klirrfaktor äußert (siehe Abschn. 2.3).

Bei der Aufzeichnung von Live-Ereignissen kann man nie mit absoluter Sicherheit sagen, welchen Dynamikumfang das Signal noch erreichen wird. Hier ist es lediglich möglich, einen angemessenen Sicherheitsabstand zur Vollaussteuerung einzuhalten. Soll um jeden Preis eine optimale Aussteuerung erzielt werden, kommt man nicht umhin, mehrere Sampling-Systeme mit unterschiedlicher Aussteuerung simultan laufen zu lassen. Nach der Aufzeichnung ist eine Angleichung der lautesten Stelle an das Bereichsmaximum durch einen Software-Algorithmus möglich.

Damit das aufgezeichnete Signal gewissen Qualitätsansprüchen gerecht wird, müssen bestimmte Richtlinien beachtet werden. Das Signal und seine Eigenschaften erfordern jedoch auf der gesamten Verarbeitungsstrecke die ganze Aufmerksamkeit des Aufnahmetechnikers. Das beginnt bei Mikrofonaufnahmen bereits bei der Wahl und der Anwendung von Mikrofonen. Doch auch Kriterien der Auswahl von Zusatzgeräten sollen zur Sprache kommen.

2.6.2 Mikrofonwahl und -aufstellung

Signale, die nur akustisch vorliegen, müssen mit Mikrofonen in elektrische Signale gewandelt werden. Bei der Wahl, Aufstellung und Zusammenschaltung von Mikrofonen sind einige Richtlinien zu beachten, weil sonst die Qualität einer Aufzeichnung stark leiden kann und dies in den meisten Fällen auch nicht mehr rückgängig zu machen ist. S. 97 [9–11]

Aufnahme von Einzelinstrumenten Zunächst ist zu entscheiden, ob Stereofonie gewünscht wird oder nicht. Falls ja, sind Stereomikrofone oder entsprechende Kombinationen (siehe Kap. 1) zu verwenden. Jedoch auch bei Monoaufnahmen kann die Raumakustik eine wichtige Rolle spielen. Sollen Wandreflexionen möglichst wenig in Erscheinung treten, ist ein Mikrofon mit Nierencharakteristik zu wählen, das in einem Abstand von 0,5 ... 1 m Entfernung auf das Instrument gerichtet wird. Die Höhe des Mikrofons sollte dabei etwa derjenigen des Instruments entsprechen.

Grenzflächenmikrofon Bei Hartböden ist zur Vermeidung von Interferenzen mit dem Reflexionsschall das sog. Grenzflächenmikrofon zu empfehlen. Diese Mikrofonart ist meist als flacher Körper ausgeführt, der am Boden kaum eine Änderung der Schalleneigenschaften bewirkt. Aus diesem Grunde integriert sich das Mikrofon fast störungsfrei in die Grenzfläche des Bodens. Der Schall kommt nur einmal an und unterliegt somit keinen Interferenzerscheinungen. Der Frequenzgang dieser Anordnung ist demzufolge ohne Einbrüche bzw. entspricht weitgehend dem des Mikrofons selbst, wie dies auch aus Abb. 2.37 hervorgeht.

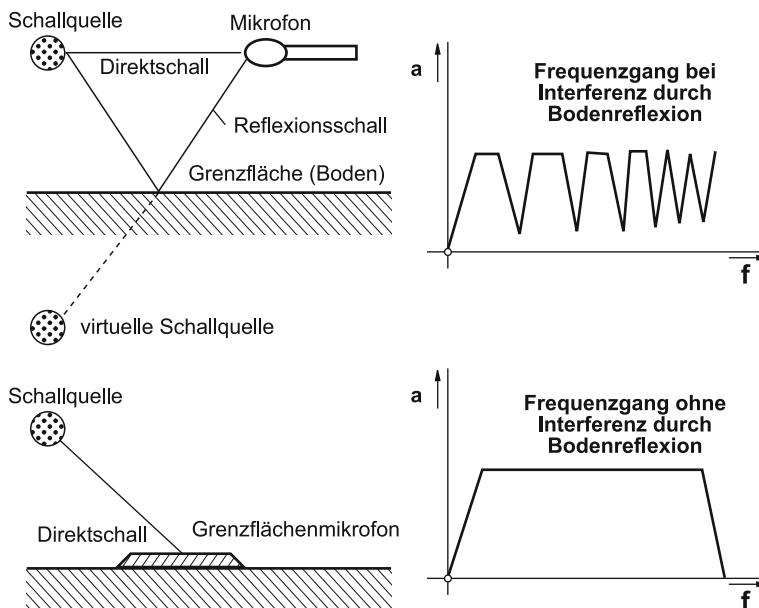


Abb. 2.37 Prinzipielle Einflüsse der Grenzflächen-Reflexion auf den Frequenzgang und Abhilfe durch Verwendung des Grenzflächenmikrofons

Der kammförmige Verlauf des Frequenzgangs bei Bodenschallreflexionen entsteht durch Auslöschung bzw. Verstärkung für Schall entsprechender Wellenlängen (vgl. Abschn. 1.8.1).

Richtmikrofon Bei starker Richtwirkung des Mikrofons (z. B. Superniere oder Keule) lassen sich ebenfalls störende Reflexionen von der Aufzeichnung fernhalten. Beim Einsatz von zwei oder mehr Mikrofonen lässt sich auch auf einfache Weise eine Laufzeitstereofonie verwirklichen. Die Mikrofone sollten hierbei einen Abstand von mehr als 0,5 m voneinander haben, wobei sie in einer Linie mit dem Instrument aufgestellt sein sollten. Der klangliche Eindruck hängt neben dem gewählten Abstand auch noch von der Pegelbalance der Mikrofone ab.

Aufnahme mehrerer Instrumente Handelt es sich um mehrere Instrumente, welche simultan zur Aufzeichnung kommen sollen, wird man wohl immer zur Stereofonie greifen. Grundsätzlich lässt sich dies auf zwei unterschiedliche Verfahren verwirklichen, die sich natürlich auch im Aufwand unterscheiden.

Aufzeichnung mit Stereomikrofon Um ein stereofones Raumsignal zu erhalten, sind verschiedene Methoden denkbar. Einerseits können ausgesprochene Stereomikrofone gewählt werden, die entweder im MS- oder XY-Modus arbeiten. Andererseits kommt auch eine Aufzeichnung mit dem Kunstkopf in Frage. Als Alternative zum Kunstkopf bietet sich

die sog. OSS-Technik an, bei der zwei Druckempfänger einen Abstand von ca. 15–20 cm voneinander aufweisen; zwischen ihnen ist eine Trennscheibe mit einem Durchmesser von ca. 30 cm befestigt, an deren Außenflächen Dämmmaterial befestigt ist. Im Gegensatz zur Kunstkopfstereofonie ist die OSS-Technik auch für Lautsprecherwiedergabe geeignet. Die Aufstellung bzw. Platzierung des Stereomikrofons entspricht etwa dem Klangindruck an der jeweiligen Stelle; deshalb sollte dieser Platz möglichst sorgfältig mit Hörproben ausgewählt werden.

Ferner lässt sich Stereofonie natürlich auch mit zwei räumlich separaten Mikrofonen realisieren, die einen gewissen Abstand voneinander aufweisen (siehe auch Abschn. 1.9.2).

Aufzeichnung mit diskreten Mikrofonen Der Stereoeindruck kann auch dadurch erzeugt werden, dass man die einzelnen Instrumente durch zugeordnete Mikrofone „abnimmt“. Eine Platzierung im Klangraum geschieht dann im Mischer, der die Einstellung der Einzelpegel und des Panoramas gestattet. Aufzeichnungen auf Mehrspursysteme erlauben sogar ein nachträgliches Abmischen (Mixdown) der einzelnen Instrumentstimmen; das fertige (meist statische) Mischen während der Aufnahme entfällt somit. Bei guter Trennung der Einzelsignale können sogar nachträglich Einzeleffekte zugemischt werden, die diskret für ein Instrument wirken.

Die einzelnen Mikrofone sollten wie bei der monauralen Aufnahme eines Instruments relativ dicht an letzterem aufgestellt sein.

Aufnahme der menschlichen Stimme Bei Gesangsaufnahmen werden meist direkt postierte Mikrofone eingesetzt. Zur Vermeidung von Popp- und Blasgeräuschen dient ein Poppfilter, welches als Schaumstoffhülse oder als Dämpfungswand ausgeführt ist. All diese Filter zeigen jedoch auch leicht höhenabsenkenden Charakter. Poppgeräusche können auch durch Erhöhung des Mikrofonabstandes abgeschwächt werden.

Aufnahme von Geräuschen Die Technik bei der Aufnahme von Geräuschen hängt stark von letzteren selbst ab. Einerseits ist die Separation von Geräuschen und andererseits deren Aufzeichnungsmethode teilweise problematisch.

Leise Geräusche Sollen Geräusche mit kleinem Schallpegel zur Aufzeichnung kommen, besteht oft das Problem, dass andere, unerwünschte Schallquellen ebenfalls im Signal enthalten sind. Permanente Störgeräusche können nur relativ zum Nutzsignal abgeschwächt werden, indem das Mikrofon möglichst nahe am aufzuzeichnenden Objekt aufzustellen ist. Ist eine Verringerung der Entfernung nicht ohne weiteres möglich, sind Mikrofone mit extremen Richteigenschaften (Richtrohr oder Hohlspiegel) zu verwenden. Bei schwachem Mikrofonsignal ist zu befürchten, dass der Rauschabstand nicht mehr ausreichend ist. Optimale Voraussetzungen schafft hier ein besonders rauscharmer Mikrofonvorverstärker.

Laute Geräusche Bei kräftigem Schall bestehen die oben genannten Probleme nicht; dafür kann es zu Übersteuerungen kommen. Nehmen wir an, das Geräusch des Donners

ist aufzuzeichnen. Der Schallpegel ist hier so gut wie überhaupt nicht vorhersagbar. Ist keine durchgehende Aufzeichnung notwendig, so können ausschließlich die Passagen mit ordnungsgemäßer Aussteuerung Verwendung finden. Protokollierte Werte und Aufnahmedingungen können in diesem Zusammenhang ebenfalls sehr nützlich sein. Bei Aufzeichnungen mit einem digitalen System (wie der Audiokarte) kann durchaus eine Untersteuerung von ca. -20 dB in Kauf genommen werden, wenn man dafür das Risiko einer Übersteuerung (und damit Begrenzung) sicher vermeiden kann.

Vermeidung von Trittschall Falls Mikrofone ungedämpft an Stativen usw. befestigt sind, kann es zur Übertragung von unerwünschtem Bodenschall kommen. Es handelt sich hierbei um tiefe Frequenzen, die sich durch Vibrationen über Boden, Wand oder Decke fortpflanzen. Die Ursachen dieser Vibrationen können weit gestreut sein; im einfachsten Falle handelt es sich tatsächlich um Bodenschwingungen, die sich durch Schritte entstehen.

Da es sich um besonders niederfrequente Erschütterungen handelt, sind diese auch relativ schwer vom Mikrofon fernzuhalten. Eine Verbesserung kann erzielt werden, indem das Mikrofon an einer weichen (langen) Schnur oder einem Gummifaden an der Decke befestigt wird. Professionelle Lösungen setzen ein Trittschallelement ein, das prinzipiell nach Abb. 2.38 aufgebaut ist.

Natürlich verfügen gute Mischpulte über zuschaltbare Trittschallfilter; dennoch ist ein grundsätzliches Vermeiden der Störungen immer besser, weil elektronische Filter immer zwangsläufig auch das Nutzsignal beeinflussen.

Mikrofon-Wandlertypen Grundsätzlich sind unterschiedliche Wandlerarten der Mikrofonkapseln nicht unbedingt auch verschiedenen Anwendungszwecken vorbehalten. Das Wandlungsprinzip (z. B. dynamisch, Elektret- oder Kondensatorwandler) bestimmt jedoch teilweise die akustische Arbeitsweise und somit auch das Richtungsverhalten mit. Dadurch ergeben sich natürlich zwangsläufig spezifische Anwendungen.

Abb. 2.38 Aufbau eines Trittschallelements. Der dehbare Aufnehmer wird durch zwei Strangsysteme aus Gummi gefedert gehalten



Funktionsprinzip Grundsätzlich gibt es zwei Arten: den elektrodynamischen und den elektrostatischen Wandler. Das elektrodynamische Prinzip beruht auf der Tatsache, dass senkrecht zum Magnetfeld bewegte Leiter eine Spannung induzieren, die proportional von der Bewegungsgeschwindigkeit abhängt. Dies wird beim Schwingspulenmikrofon und beim Bändchenmikrofon ausgenutzt. Beide Male taucht in den Luftspalt eines Permanentmagneten ein Leiter, der durch den Schall in Bewegungen versetzt wird und dadurch Spannungen induziert. Die Technologie der elektrostatischen Wandler heißt Kondensatormikrofon und Elektret-Mikrofon. Die prinzipiellen Ausführungen dieser Mikrofone sehen wir in Abb. 2.39.

Störabstand Unterschiede bestehen ferner in Bezug auf Empfindlichkeit bzw. Rauschabstand. Während dynamische Schwingspulenmikrofone keine Spannungsversorgung

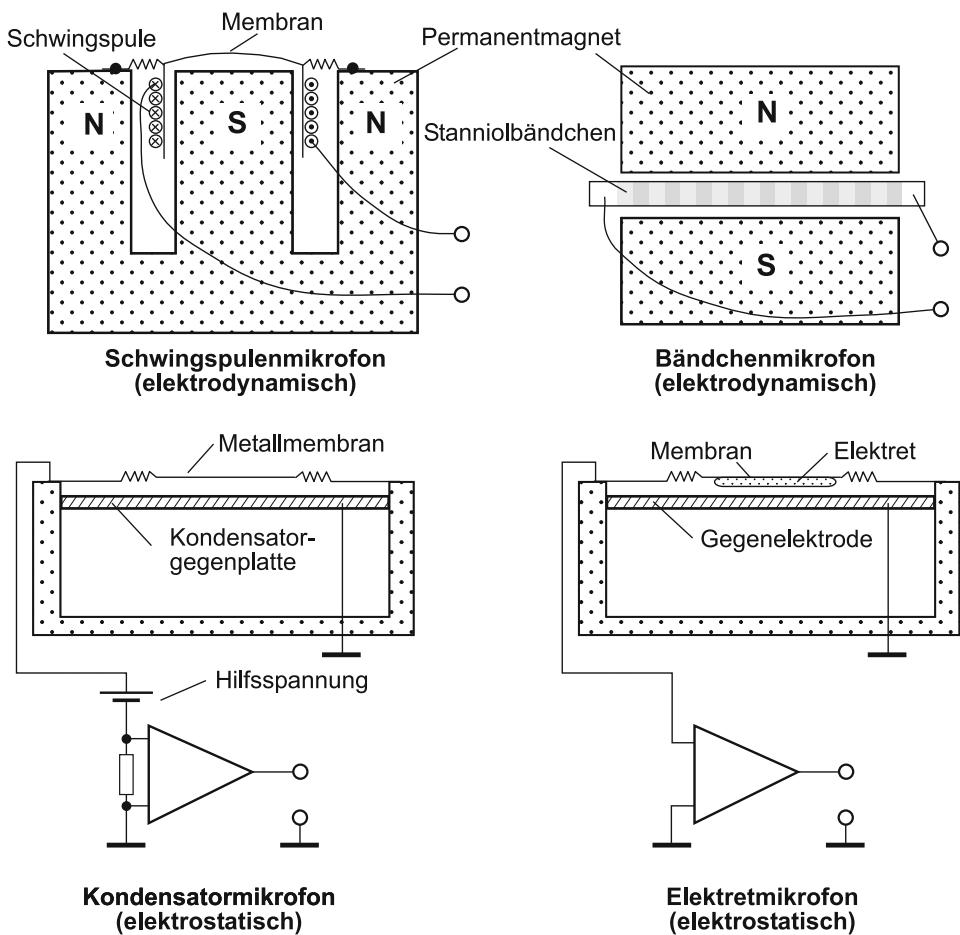


Abb. 2.39 Schematischer Aufbau der wichtigsten Mikrofonkapseln

benötigen, ist diese bei Elektret- und Kondensatormikrofonen durchaus erforderlich; außerdem werden bei beiden Bauarten Vorverstärker für eine Ankopplung mit niedriger Lastimpedanz eingesetzt. Das Halbleiterrauschen ist hierbei deutlich stärker als das Leiterrauschen beim Schwingspulenmikrofon, weshalb bei Kondensatormikrofonen der Geräuschspannungsabstand merklich herabgesetzt ist (ca. 50 ... 70 dB). Der effektive Geräuschspannungsabstand hängt jedoch vor allem bei dynamischen Mikrofonen natürlich nicht nur von den entsprechenden Mikrofondaten ab, sondern auch vom Rauschen des Mikrofonvorverstärkers des Mischpults; im Anhang Rechenbeispiele sind diesbezüglich Rechenbeispiele zur Bestimmung von Störabständen zu finden.

Empfindlichkeit der Mikrofone In einem Mikrofondatenblatt ist ein Bezug zwischen abgegebener Spannung (ohne Last) und dem dabei herrschenden Schalldruck angegeben. Dieser Wert wird in mV/Pa (= 0,1 mV/ μ bar) ausgedrückt und wird mit korrektem Terminus als Feld-Leerlauf-Übertragungsfaktor bezeichnet. Dieser Wert gilt natürlich nur für die Bezugsfrequenz (meist 1000 Hz). Als ungefähre Richtwerte für die meist verwendeten Wandlertypen können gelten: dynamisches Schwingspulenmikrofon: 2 mV/Pa, Elektret-Mikrofon: 5 mV/Pa und Kondensatormikrofon: 20 mV/Pa. Werden für verschiedene Frequenzen die Empfindlichkeiten logarithmisch abgetragen, so ergibt sich das Feld-Leerlauf-Übertragungsmaß oder auch kurz der Frequenzgang.

Akustische Arbeitsweisen Die akustischen Arbeitsweisen wurden bereits in Kap. 1 kurz erwähnt. Sie hängen vom geometrischen Aufbau und damit teilweise auch vom physikalischen Arbeitsprinzip ab. Man kann hauptsächlich eine Einteilung in die Kategorien Druck-, Druckgradienten- und Schnellempfänger vornehmen. Die hier aufgeführten drei Typen stellen physikalische Einheiten dar, die nur die genannten Eigenschaften aufweisen. In der Praxis hingegen handelt es sich sehr oft auch um gemischte Arbeitsweisen, die durch besondere Gestaltung der Schallleitung und Membranen bzw. deren Befestigungsstücke erzielt werden. Auf diese Weise erst lassen sich z. B. Kapseln mit Nieren-Richtcharakteristik herstellen. Andererseits entsteht natürlich die Richtcharakteristik letztlich auch durch den Einbau der Kapsel in das Mikrofongehäuse.

Druckempfänger Kapseln, bei denen der Schall nur auf eine wirksame Membranseite einfallen kann, arbeiten meist als Druckempfänger, d. h. sie geben eine Momentanspannung ab, die dem momentanen Druck der Schallwelle entspricht. Eine solche Abhängigkeit bedeutet für die Funktionseinheit Mikrofonkapsel (ohne Zusatzeinrichtungen) immer notwendigerweise, dass Kugelrichtcharakteristik besteht, also die Empfindlichkeit nach allen Richtungen in etwa gleich ist.

Druckgradientenempfänger Ist die Membran von beiden Seiten für den Schall zugänglich, so reagiert die Kapsel auf den Druckunterschied zwischen Vorder- und Hinterseite und stellt somit einen Druckgradientenempfänger dar. Dadurch entsteht beim reinen Druckgradientenempfänger zwangsläufig eine Acht-Richtcharakteristik, weil ja auch ein

Phasenunterschied von 180° zwischen beiden Seiten zum Tragen kommt. Da der Druckgradient bei konstantem Schalldruck mit der Frequenz ansteigt, ist hier auch eine entsprechende Anhebung bei hohen Frequenzen zu erwarten, was natürlich elektronische Entzerrungsmaßnahmen erfordert.

Schnelleempfänger Wird die Membranbewegung beim Gradientenempfänger mechanisch gedämpft, so reagiert die Kapsel proportional zur Schallschnelle, welche nicht mit der Frequenz ansteigt. Demzufolge ist auch die abgegebene Mikrofonspannung (theoretisch) nicht von der Frequenz abhängig. Doch auch hier entsteht Acht-Richtcharakteristik. Bei der Anwendung von Schnelleempfängern ist zu beachten, dass die proportionalen Abhängigkeiten näherungsweise in der Nähe der Schallquelle $1/r^2$, bei größeren Entfernung dagegen $1/r$ lautet. Dabei hängt der Übergang zwischen diesen beiden Abhängigkeiten zudem frequenzabhängig. Das bedeutet im Klartext, dass bei einer Annäherung des Schnelleempfängers an die Schallquelle die tiefen Frequenzen stärker angehoben werden als die hohen. Das liegt daran, dass Kugelwellencharakteristik für tiefe Frequenzen noch bei größerer Entfernung besteht. S. 97 [9]

Zusammenschaltung mehrerer Mikrofone Sollen mehrere Mikrofone zusammenge schaltet oder einem Mischpult zugeführt werden, besteht vor allem bei Billigfabrikaten die Gefahr, dass keine einheitliche Phasenlage besteht. Selbst wenn die einzelnen Mikrofone durchaus akzeptable Eigenschaften aufweisen, so wirkt der Klangeindruck äußerst „dünn“, wenn es aufgrund von Verpolungen einzelner Mikrofone zu Auslöschen diskreter Frequenzen kommt.

Ob sich zwei Mikrofone in Bezug auf Phase äquivalent verhalten, kann mit geringem technischen Aufwand überprüft werden. Es wird ein Prüfton (Sinus 1000 Hz) über einen Lautsprecher wiedergegeben. Beide Mikrofone befinden sich direkt nebeneinander und in einem Abstand von ca. 1 m vom Lautsprecher entfernt. Ein Zweikanal-Oszilloskop, an dessen Eingänge die vorverstärkten Mikrofonsignale gelangen, zeigt anhand der Signalzüge direkt eine eventuelle Verpolung an. Doch auch mit einem Einkanal-Oszilloskop ist solche eine Prüfung durchführbar. Am besten betreibt man es im XY-Modus, also mit abgeschalteter Zeitablenkung. Die jeweils vorverstärkten Mikrofonsignale gelangen an den Y- bzw. X-Eingang und erzeugen auf dem Schirm eine undeutliche Figur, die bei einheitlicher Phasenlage der Mikrofone in die Quadranten I und III (rechts oben und links unten) reicht, im anderen Falle in die Quadranten II und IV (links oben und rechts unten).

Professionelle Technik setzt die symmetrische Beschaltung ein. In diesem Falle liegen zwei Anschlüsse vor, an denen die Signalspannung anliegt, wobei jedoch keinerlei Masse- bzw. Erdbezug vorhanden ist. Der Vorteil liegt in einer wesentlich kleineren Störempfindlichkeit gegenüber Brummeinstreuungen, da ja auf beide Adern das Störfeld gleichphasig einwirkt und sich somit kompensiert. Dagegen nimmt die einzige Ader der asymmetrischen Beschaltung bei längeren Leitungen die Störung auf, die sich ja weitgehend auf Masse bezieht. Im Anhang sind Beispiele für verschiedene Beschaltungen und deren Konvertierungen angegeben.

2.6.3 Echtzeitbeeinflussung des Signals

Zur Aufbereitung des Signals auf der Analogseite gibt es viele unterschiedliche Geräte, die u. U. hilfreich sind und in Echtzeit wirken. Grundsätzlich ist es angebracht, irgendwelche Spezialeffekte oder auch Störunterdrückungen durch solche Geräte zu bewerkstelligen – eine entsprechende zeitaufwendige Nachbearbeitung entfällt somit.

Vorverstärker Zur Pegelanhebung ist es insbesondere bei Verwendung von Mikrofonen als Signalquellen erforderlich, Vorverstärker einzusetzen.

Störabstand und Übersteuerungsfestigkeit Diese müssen für die niedrigen Signalspannungen geeignet sein, d. h. ein gewisser Mindeststörabstand sollte gewährleistet sein. Andererseits muss man natürlich auch Übersteuerungsfestigkeit für große Mikrofonpegel voraussetzen. Beides erreichen Geräte mit bestimmten Operationsverstärkern. Für hochqualitative Aufnahmen sollte bei einem Nennschalldruck von 1 Pa (94 dB, bei 1000 Hz ca. 94 Phon) ein Geräuschspannungsabstand von >60 dB zu erzielen sein. Das entspricht einer Ersatzlautstärke von <34 dB. Im Anhang Rechenbeispiele ist die Ermittlung von Störabstandbilanzen zu finden. Beziiglich dieses Nennschalldrucks ist meist die Mikrofonspannung angegeben (als Maß für die Empfindlichkeit); die Übersteuerungsfestigkeit des Vorverstärkers sollte mindestens 20 dB darüber liegen, also dürfen bei einem Schallpegel von 114 dB in Verbindung mit dem betreffenden Mikrofon keine Übersteuerungen auftreten. Die Mikrofonspannung ist in diesem Falle das 10fache der Nennspannung. In diesem Zusammenhang ist auch zu erwähnen, dass Kondensatormikrofone systembedingt Verzerrungen ab einem gewissen Schallpegel zeigen, welcher aus dem Datenblatt hervorgeht.

Symmetrischer Eingang Handelt es sich um Mikrofone mit symmetrischem Ausgang, so sollte ein Vorverstärker mit symmetrischem Eingang zum Einsatz kommen; eine alternative wäre das Zwischenschalten eines Übertragers, dessen Ausgang dann auch an einen Vorverstärker mit asymmetrischem Eingang angeschlossen werden kann. Im Anhang Beschaltungen sind entsprechende Informationen aufgeführt.

Eingangsimpedanz Bei Verwendung moderner Operationsverstärker ist die Eingangsimpedanz (Scheinwiderstand am Eingang) meist größer als $50\text{ k}\Omega$, so dass ein Mikrofonübertrager eine Übersetzung aufweisen kann, wodurch oft ein Gewinn des Störabstands zu verzeichnen ist. Dabei ist jedoch zu beachten, dass mit dem Übertrager die effektive Eingangsimpedanz des Verstärkers um das Quadrat des Übersetzungsverhältnisses herabgesetzt wird. Die empfohlene Abschlussimpedanz für das Mikrofon (ca. das 10-fache der Nennimpedanz) darf dabei nicht unterschritten werden.

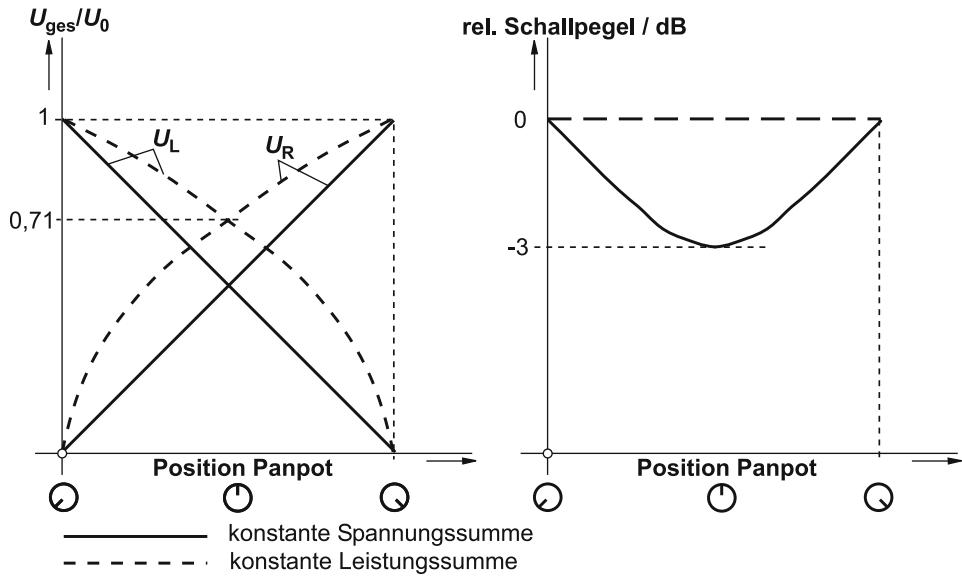


Abb. 2.40 Lautstärkeverhalten von PanPots

Mischpulte Ein Mischpult erlaubt die Beeinflussung der Pegel aller an ihm angeschlossenen Signalquellen. Das Mischsignal ist für unterschiedliche Zwecke bestimmt – sehr häufig gelangt es zu einem Aufzeichnungsgerät.

Panorama Sollen alle monauralen Quellen zu einem Stereosignal abgemischt werden, so muss für jeden Kanal eine Einrichtung zur Positionierung im Stereohorizont vorhanden sein, nämlich das sog. Panorama. Dabei wird einfach eine pegelabhängige Verteilung des Signals auf beide Stereokanäle erreicht. Für das Stellverhalten ist hierbei wünschenswert, dass sich der ergebende Schallpegel theoretisch nicht ändert. Viele Panpots aber verteilen lediglich die Signalspannung auf die entsprechenden Kanäle, so dass die Summe der Einzelspannungen konstant gehalten wird. Das bringt jedoch einen deutlichen Lautstärkeeinbruch in Mittelstellung. Abb. 2.40 verdeutlicht die Situationen in der Gegenüberstellung.

Pegeleinstellung Die Pegeleinstellung ist die wichtigste Einrichtung. Mit ihrer Ausführung steht und fällt die Qualität eines Mischpults. Die Skala der Schieberegler sollte dB-Teilung aufweisen, und ein entsprechend langer Schiebeweg sowie die Qualität der logarithmischen Widerstandsaufteilung ermöglicht die Nutzung eines Pegelintervalls von mindestens 40 dB. Zur Pegelanpassung der einzelnen Quellen sind separate Voreinsteller sehr nützlich.

Pegelanzeige Zur Kontrolle der eingestellten bzw. voreingestellten Pegel (am besten umschaltbar) sind Anzeigen mit LED-Ketten und Spitzenwertspeicherung sinnvoll. Zeigerinstrumente, die lediglich in VU geeicht sind, sind Mittelwertanzeiger und daher zu träge, um kurzzeitige Übersteuerungsspitzen anzuzeigen.

Klangregelung Es ist nicht für jeden Mischpultkanal eine komplette parametrische Filtergruppe notwendig, um den Klang zu optimieren bzw. Störgeräusche zu unterdrücken. Jedoch ist eine Klangeinstellung der Tiefen, Mitten und Höhen sehr empfehlenswert, zumal der Aufwand hierfür nicht besonders groß ist.

Effektzumischung Besonders häufig benötigte Effekte wie Hall und Echo sollten sich über ein entsprechendes Zusatzgerät realisieren lassen. Jeder Kanal des Mischpults wird hierbei über einen besonderen Steller ausgekoppelt und über einen speziellen Anschluss dem Effektgerät zugeführt. Das Einschleifen des Effektsignals erfolgt dann meist wieder im Summenbaustein.

Kanäle Die Anzahl der Kanäle richtet sich natürlich nach der geplanten Anwendung. Optimal ist die Ausführung in Modultechnik, welche nicht nur den Ausbau auf die momentan benötigten Kanäle, sondern auch eine individuelle Konstellation von Mono- und Stereomodulen erlaubt. Oftmals werden nämlich beiderlei Arten gebraucht – eine Realisierung einer Stereogruppe durch jeweils zwei Monogruppen ist nicht akzeptabel.

Eingänge Von Vorteil sind in jedem Falle symmetrische Eingänge, zumindest was die Mikrofonkanäle anbelangt. Doch auch beim Verbinden von Geräten, die eine erdbezogene Masse aufweisen, kommt es ohne Potentialtrennung oder Symmetrierung sehr oft zu Störungen aufgrund einer Brummschleife. Wenn man die prinzipiellen Ursachen solcher Phänomene kennt, sind effektive Vorsorgemaßnahmen ergreifbar. Zur Vermeidung von Brummschleifen sei an dieser Stelle auf den Anhang Beschaltungen verwiesen.

Stereoeingänge Neben Stereomikrofonen machen auch Stereogeräte wie DAT, CD, andere Audiokarte usw. Stereoeingänge am Mischpult erforderlich. Obwohl ein Tandem-Schieberegler zur Pegeleinstellung sinnvoll ist, wird das Panpot durch einen Balance-Steller und gegebenenfalls auch durch einen Basisbreitensteller ersetzt. Die restlichen Bedienelemente entsprechen den Einrichtungen beim Monomodul und wirken auf beide Stereokanäle gleichermaßen.

Einstellbare Verzögerungszeit Zur Erzeugung eines Raumklanges kann bei hochwertigen Mischpulten jeder Kanal zeitlich verzögert werden. Man erspart sich bei Mikrofonaufnahmen das entsprechende Positionieren der Mikrofone.

Filter Zur Erzeugung von Spezialeffekten und Unterdrückung von Störfrequenzen ist der Einsatz von Filtern zwischen Signalquelle und Audiokarte zweckmäßig.

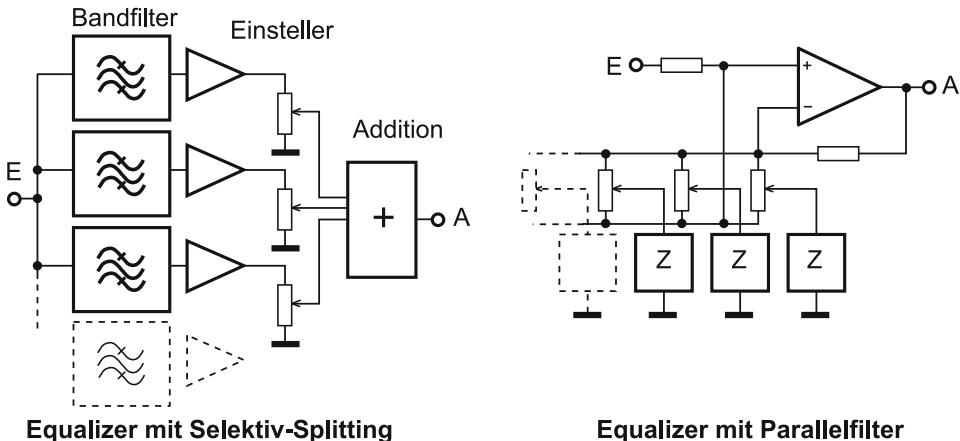


Abb. 2.41 Funktionsprinzipien verschiedener Equalizer

Grafischer Equalizer Eine Anordnung, bei der der eingestellte Frequenzgang durch die Stellung der Schiebereglern sehr übersichtlich hervorgeht, ist der Grafische Equalizer. Das Gesamtsignal wird durch Bandfilter (meist Oktav- oder Terzfilter) aufgeteilt, deren Ausgangssignale wiederum über Lautstärkesteller zusammengemischt werden. Alternativ zum Frequenz-Splitting findet man auch häufig Mehrfachfilter, die mittels Potentiometer eine Abschwächung oder Verstärkung bewirken und die als Gyratorschaltung (Substitution einer Induktivität durch aktiven Vierpol, Kondensator und Widerstände) ausgeführt sind. Das Funktionsprinzip beider Verfahren ist in Abb. 2.41 zu sehen.

Parametrischer Equalizer Im Gegensatz zum Grafischen Equalizer kann man beim Parametrischen Equalizer den eingestellten Frequenzgang nicht auf Anhieb anhand der Potentiometerstellungen ablesen. Jedoch sind hiermit bestimmte Einstellungen gezielt mit wenigen Reglern zu bewerkstelligen. Bei den Parametern handelt es sich um Frequenz, Bandbreite und Verstärkung. Darüber hinaus ist auch oft zwischen den Arbeitsweisen Tiefpass, Hochpass, Bandpass und Bandsperre zu wählen. Die Wirkungen eines solchen Filters ist in Abb. 2.42 zu sehen.

Zur Unterdrückung einer bestimmten Störfrequenz (z. B. 50 Hz Netzbrumm) ist ein Parametrisches Filter eher geeignet als ein Grafischer Equalizer, weil es bei optimaler Einstellung die Störung selektiver und zugleich wirksamer unterdrückt. Die Synthese mehrerer solcher Filter ergibt den Equalizer, wobei natürlich lediglich Bandsperren in Serie zu schalten sind, alle anderen Filtertypen verlangen Parallelschaltung.

In Mischpultmodulen sind häufig Parametrische Filter zu finden.

Effektgeräte An Mischpulte können Effektgeräte angeschlossen werden, sofern sie über eine entsprechende Möglichkeit der Effekteinblendung verfügen. Doch auch das Signal eines einzelnen Instruments oder Mikrofons ist mit Zusatzgeräten manipulierbar.

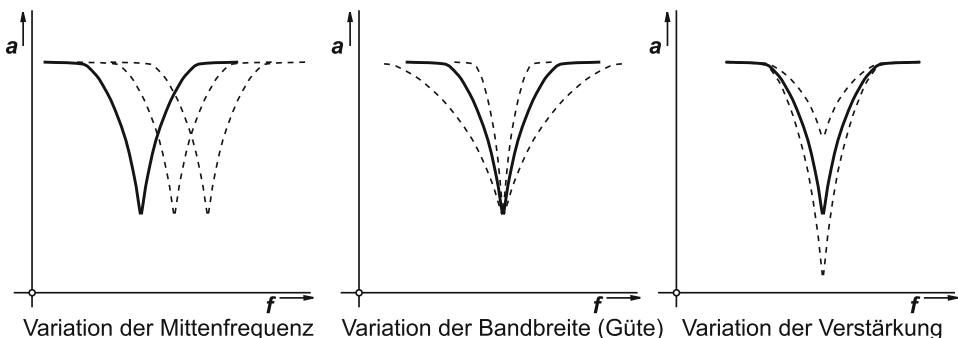


Abb. 2.42 Einstellparameter einer Bandsperre

Hall- und Echogeräte Diese Einrichtungen zählen zum Standard der Ausrüstung. Man kann sie eigentlich nicht mehr nur als Effektgeräte bezeichnen, weil bei sehr dosierter Anwendung keine echte Verfremdung, sondern sozusagen nur eine Klangergänzung bzw. sogar ein Schritt zum Erreichen eines natürlicheren Klanges erzielt wird. Ältere Echogeräte arbeiten mit Tonbändern, während moderne Geräte fast ausschließlich vollelektronisch aufgebaut sind. Hierbei ist wiederum eine Aufteilung in Digital- und Analoggeräte möglich. Für hochwertige Echoeffekte sind sicherlich Digitalgeräte vorzuziehen, weil die Verzögerungszeit nicht die Tonqualität beeinträchtigt, was bei analog arbeitenden Eimerkettenspeichern der Fall ist. Generell hängt die Tonqualität bei Digitalechogeräten jedoch von den gleichen Parametern ab wie beim Sampling von Audiosignalen, nämlich von der Sampling-Frequenz, der Auflösung und dem Oversampling-Faktor. Herkömmliche Hallgeräte, welche ja bekanntlich einen diffusen Echocharakter erzeugen sollen, setzen im semiprofessionellen Bereich Hallspiralen und im professionellen Bereich Hallplatten und -räume ein. In allen Fällen erfolgt eine Anregung des Mediums durch elektromagnetische oder akustische Wandler, und es erfolgt eine Reflexion an den Unstetigkeiten des Mediums (Aufhängung bzw. Begrenzung des Mediums). Auch für Hallgeräte gibt es bereits eine vollelektronische Simulation, die in puncto Klangqualität durchaus einem Hallraum nahekommt. Darüber hinaus jedoch verfügt das elektronische Hallgerät über weitaus mehr Einstellmöglichkeiten. Hallspiralen hingegen zeigen in Bezug auf Klirrfaktor, Frequenzgang und Fremdspannungsabstand recht ungünstige Werte; außerdem neigen sie bei bestimmten Frequenzen zu „Blubbergeräuschen“ und eignen sich deshalb nur für semiprofessionelle Anwendungen. Neben solchen recht kompakt gebauten Hallspiralen (für Transversalschwingungen) gibt es jedoch auch hochwertige Ausführungen, die die Federn zu Torsionsschwingungen anregen. Mit dieser Anordnung lassen sich sehr gute Ergebnisse erzielen.

Harmonizer Zur Transponierung der Tonhöhenlage sind (meist digital arbeitende) Geräte konstruiert worden, die mitunter sehr kostspielig sind. Die Verschiebung (Pitch Ratio) kann mehrere Oktaven betragen und sowohl nach unten als auch nach oben erfolgen. Somit

sind starke Verfremdungen möglich (z. B. unkenntliche Stimme); aber auch sehr feine Verschiebungen, die lediglich der Vermeidung einer akustischen Rückkopplung dienen, sind möglich (für diesen Zweck sind jedoch auch einfachere Geräte ausreichend). Der Harmonizer wird i. Allg. vor eine einzelne Quelle geschaltet – eine Effektzumischung über ein Mischpult ist wenig sinnvoll.

Phaser und Flanger Sehr effektvoll sind Geräte, die zum Eingangssignal ein phasenverschobenes generieren und ersterem wieder zumischen. Die Phasenverschiebung verändert sich noch dazu automatisch mit Hilfe eines Generators niedriger Frequenz. Zusätzlich sind einstellbar die Frequenz, der Hub der Phasenverschiebung und teilweise auch die Schwingungsform des Steuergenerators. Da die Phasenverschiebung von der Eingangsfrequenz annähernd linear abhängt, kann man sich dafür auch eine Zeitverschiebung vorstellen, deren Verzögerungszeit sich jedoch langsam durch den Steuergenerator ändert. Im Frequenzbereich äußert sich dieser Effekt dergestalt, dass an sehr vielen Stellen Nullstellen ausgebildet sind, welche jedoch durch den Steuergenerator eine Verschiebung erfahren. Durch diese ständige Veränderung des kammartigen Frequenzgangs entsteht ein Klangeindruck, bei dem man ein in der Luft schwebendes Orchester assoziiert. Realisiert wird die frequenzabhängige Phasenverschiebung durch sog. Allpässe, die alle Frequenzen gleichermaßen passieren lassen. Der Flanger ist sozusagen eine Weiterentwicklung; hier verwendet man regelrecht Zeitverzögerungsglieder (Eimerketten-Speicher), um die frequenzabhängige Phasenverschiebung zu erzielen. Die Wirkung kann hierbei gegenüber dem Phaser gesteigert werden, weil die äquivalente Zeitverzögerung beim Flanger größer ist und dadurch wesentlich mehr Nullstellen im Frequenzbereich entstehen können. Auch diese beiden Effektgeräte kommen vorzugsweise nur für einzelne Instrumente in Betracht.

Leslie Das Leslie nutzt den Doppler-Effekt aus, der ja bekanntlich bei Bewegungen des Schallempfängers gegenüber dem Sender entsteht. Die mathematischen Zusammenhänge sind in Abschn. 1.8.4 zu finden. Die relative Bewegung wird meist realisiert durch rotierende Lautsprecher oder Schalltrichter. Der Doppler-Effekt impliziert mit der Frequenzverschiebung natürlich zwangsläufig auch eine Phasenverschiebung, wobei sich durch die Rotation mithin beides quantitativ ändert. Eine sehr effektvolle Nachbildung mit geringem Aufwand ist bei Mikrofonaufnahmen möglich, wobei die Mikrofone durch Armbewegungen zu schwenken sind. Wenn man voraussetzt, dass der Schall durch Reflexion verschiedene Weglängen zurückzulegen hat, hat der Leslie-Effekt sehr viel Gemeinsames mit dem Phasing-Effekt, wo ja auch eine variable Phasenverschiebung stattfindet.

Vocoder Ein sehr komplexes Effektgerät stellt der Vocoder dar. Für die Verfremdung von Stimmen wird er sehr häufig eingesetzt. Zur Veranschaulichung der Funktionsweise betrachten wir Abb. 2.43. Zunächst erfolgt für das Steuersignal (z. B. Stimme) eine Aufteilung in viele verschiedene Frequenzbänder. Die gewonnenen Einzelsignale gelangen jeweils über einen Gleichrichter und einen Tiefpass und stellen jetzt sehr niederfrequente

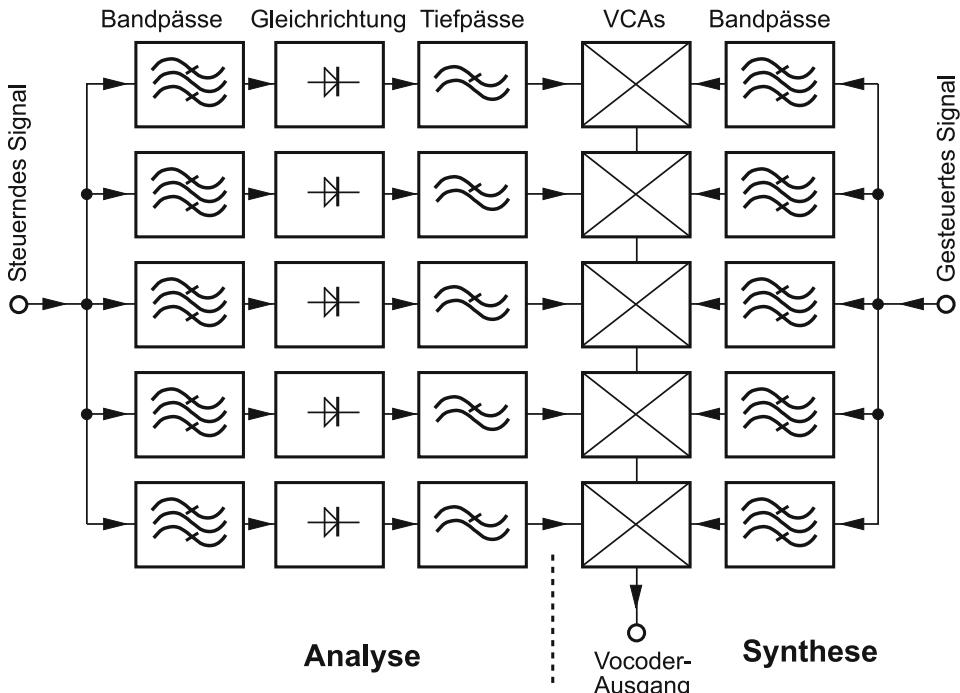


Abb. 2.43 Schematische Arbeitsweise eines Vocoder

Steuersignale dar, mit denen man z. B. auch einen Spektrum-Analyzer realisieren könnte. Der beschriebene Komplex stellt den Analyseteil dar. Der Syntheseteil ist komplementär aufgebaut und steuert nun mit den im Analyseteil gewonnenen Spannungen VCAs bzw. Modulatoren an, die entsprechend ihrer Steuerspannung das jeweilige Synthesesignal mehr oder weniger passieren lassen. Beim Synthesesignal handelt es sich wiederum um ein in Frequenzbänder gesplittetes, gesteuertes Signal, wobei jedoch keine Gleichrichtung und Tiefpassfilterung erfolgt. Dieses gesteuerte Signal kann z. B. von einer elektronischen Orgel kommen. Als Ziel soll aus der Stimme ein neues Signal geschaffen werden, das aus den gleichen Spektralanteilen besteht wie die Stimme. Es ergibt sich sozusagen eine artikulierende Orgel.

Natürlich lassen sich auch andere Signale zur Synthese heranziehen, z. B. *Weißes Rauschen*, welches ja bekanntlich alle Frequenzen mit gleicher Amplitude aufweist. In der Praxis sind jedoch mehr spektrale Kanäle wie in obiger Abbildung vorhanden – zehn und mehr sind hier keine Seltenheit. Nur auf diese Weise lässt sich der Klang einer Stimme als solche wieder erkennen. Es liegt natürlich auf der Hand, dass das gesteuerte Signal die spektralen Bereiche des steuernden Signal genügend abdecken muss.

Bekanntester Effekt ist die „Computerstimme“, jedoch lassen sich durchaus auch nuancenreiche Effekte erzielen, wenn dem Vocoder etwas die Penetranz genommen wird.

Ringmodulator Dieses Gerät wurde in den Anfängen elektronischer Musik häufiger verwendet als heutzutage. Ähnlich wie beim Vocoder ist auch mit dem Ringmodulator der Effekt meist sehr aufdringlich und prinzipiell wenig variabel; auch ihm haftet ein wenig die Assoziation „Computerstimme“ an, obwohl natürlich in puncto Signalquellen sehr wohl Variationen möglich sind. Das Funktionsprinzip des Ringmodulators ist sehr viel einfacher als beim Vocoder. Es wird sozusagen nur ein einziger VCA in Form eines Vierquadranten-Multiplizierers eingesetzt, an dessen gleichwertige Eingänge zwei Signalquellen anzuschließen sind. Als Produkt dieser multiplikativen Mischung bzw. Amplitudenmodulation entstehen Seitenfrequenzen, die sonst nur bei nichtlinearer Verzerrung entstehen. Handelt es sich tatsächlich um völlig äquivalente Signaleingänge eines Multiplizierers, so ergibt sich bei der Mischung zweier Signale unterschiedlicher Frequenz w_1 und w_2 folgendes Produkt:

$$\text{Einzelschwingungen: } f_1(t) = A_1 \cos(\omega_1 t), \quad f_2(t) = A_2 \cos(\omega_2 t)$$

$$f_{\text{ges}}(t) = f_1(t) \cdot f_2(t) = \frac{A_1 A_2}{2} \cdot (\cos(\omega_1 t - \omega_2 t) + \cos(\omega_1 t + \omega_2 t)) \quad (2.8)$$

Die beiden Amplituden A_1 und A_2 erscheinen in Gl. 2.8 ebenfalls als Produkt; bei der elektronischen Realisierung gibt es natürlich immer einen definierten Spannungsbezug von Produkten (z. B. 10 V). Wie leicht zu erkennen ist, besteht das Produkt aus der Summen- und der Differenzfrequenz der Ausgangskomponenten. Sobald eine der beiden Schwingungen mit einer Offset-Spannung versehen ist, erscheint die jeweils andere Schwingung nach der Modulation zusätzlich zu den beiden Seitenfrequenzen. In Abb. 2.44 ist der Vorgang im Frequenzbereich zu sehen.

Im Idealfall (ohne Offset-Spannung) ist ein Signal nicht zu hören, solange das andere fehlt. In der Urform des Ringmodulators arbeitete man mit zwei Übertragern und vier Dioden.

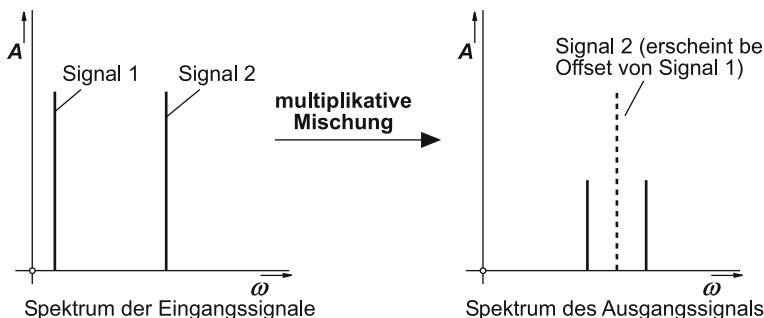


Abb. 2.44 Spektrale Frequenzen bei der multiplikativen Mischung

Rauschunterdrückung Diverse Rauschunterdrückungsverfahren finden vornehmlich Verwendung bei herkömmlichen Bandaufzeichnungen in Analogtechnik. Ziel ist es dabei, das Rauschen des Speichermediums gegenüber dem Nutzsignal möglichst klein zu halten. Dies geschieht bei dynamischen Systemen durch Reduktion der Dynamik (Kompression) während der Aufnahme und Spreizung der Dynamik (Expansion) während der Wiedergabe. Bekannte und weit verbreitete Verfahren sind Dolby, HighCom und DBX.

Die oben genannten Rauschunterdrückungsverfahren kommen bei Sampling-Vorgängen nicht in Betracht, es sei denn, das Quellmaterial des Zuspielers wurde mit einem dieser NR-Systeme aufgezeichnet.

Da Audiokarten digitale Systeme sind, ist das Rauschen des Speichermediums nicht relevant. Bei kleinen Pegeln könnte sich lediglich das Quantisierungsrauschen bemerkbar machen, was jedoch erst unterhalb -80 dB bei 16-Bit-Auflösung auffällig ist. Bei 8-Bit-Auflösung könnte eine Dynamikkompression während des Sampling sinnvoll sein; dadurch wäre ein Dynamikgewinn (abhängig vom System) zu erzielen.

Emphasis Eine weitere Möglichkeit, ein Rauschen bei der Wiedergabe zu reduzieren, besteht in der Anhebung höherer Frequenzen bei der Aufnahme (Pre-Emphasis oder Vorentzerrung) und einer entsprechenden Absenkung dieser Frequenzbereiche während der Wiedergabe (De-Emphasis oder Entzerrung). Auf diese Weise kommt es nicht nur zu einer Reduzierung des Quantisierungsrauschen, sondern auch des Störrauschens im Analogteil des Übertragungssystems.

Zur Festlegung einer Emphasis genügt die Angabe einer Zeitkonstante, welche dem Reziprokwert der Grenzkreisfrequenz eines Hoch- bzw. Tiefpasses 1. Ordnung entspricht, also:

$$R \cdot C = \frac{1}{\omega_g} \quad [\mu\text{s}] \quad (2.9)$$

Die prinzipielle Arbeitsweise einer De-Emphasis würde während der Wiedergabe auch mit Hilfe eines Tiefpasses 1. Ordnung bewerkstelligt werden können, wie dies in Abb. 2.45 zu sehen ist, wohingegen der Hochpass für die Pre-Emphasis anders wirken muss, damit eine entsprechende Knickstelle im Frequenzgang erreicht wird (z. B. über ein RC-Glied im Gegenkopplungszweig eines Verstärkers). Die Spezifikation der Emphasis erfolgt wie erwähnt über eine dem RC-Glied entsprechende Zeitkonstante, welche i. Allg. in μs anzugeben ist.

Das Verfahren kann vor allem dann zur Anwendung kommen, wenn die Höhenanteile des Signals in ihrer Amplitude schwächer ausfallen als das restliche Spektrum. Dieser Eigenschaft werden die meisten Audiosignale gerecht. Füllen dagegen die Höhen den gesamten Aussteuerbereich aus, so erfordert eine Pre-Emphasis eine grundsätzliche Reduzierung des Gesamtpegels, um Übersteuerungen zu vermeiden. Eine solche Reduzierung beeinträchtigt jedoch wieder die Verbesserung des Geräuschspannungsabstands.

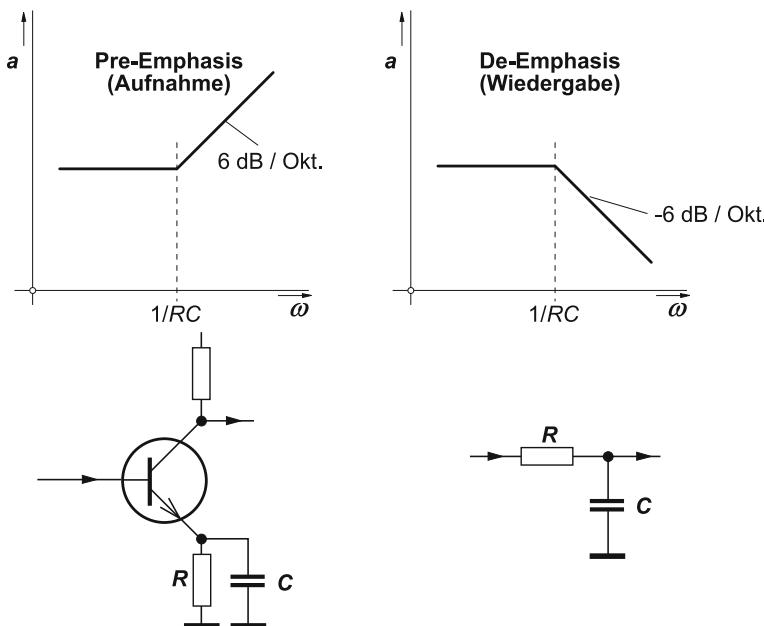


Abb. 2.45 Prinzip der Emphasis

Das Emphasis-Verfahren ist recht störsicher, da es im Gegensatz zu den DNR-Verfahren statisch arbeitet, d. h. ohne Reaktion auf die Hüllkurve des Signals auskommt. Teilweise kommt die Emphasis auch bei Digitalgeräten wie CD-Player zum Einsatz, wobei spezielle Subcodes beim wiedergebenden Gerät automatisch eine eventuell erforderliche De-Emphasis aktivieren.

Literatur

1. Tietze, U., Schenk, C.: Halbleiter-Schaltungstechnik.
Berlin, Heidelberg: Springer-Verlag 1991.
2. Skritek, P.: Handbuch der Audio-Schaltungstechnik. München: Franzis-Verlag 1988.
3. K. Brandenburg, G. Stoll: ISO-MPEG-1 Audio:
A Generic Standard for Coding of High Quality Digital Audio, 1996.
4. J. P. Princen and A. W. Johnson and A. B. Bradley:
Subband/transform coding using filter bank designs based on time domain
aliasing cancellation (P. 2161 – P. 2164) 1987.
5. J. Herre, J. Hilpert, Ch. Ertel, A. Hoelzer, C. Spenger, S. Disch, and K. Linzmeier:
An Introduction to MP3 Surround.
6. J. Breebaart, Ch. Faller: Spatial audio processing:
MPEG surround and other applications. Wiley-Interscience, 2007.
7. Xiph.org Foundation: Vorbis I specification.
8. Xiph.org Foundation: CELT.

9. Webers, J.: Tonstudientechnik. München: Franzis-Verlag 1989.
10. Jecklin, J.: Musikaufnahmen. München: Franzis-Verlag 1980.
11. Maurer, R., Taucher, G., Brantner, G.: Das Vierspur-Aufnahmestudio. München: Franzis-Verlag 1989.



Bearbeitung von Sampling-Dateien

3

3.1 Schneidevorgänge

Entsprechend der Schneidetechnik beim herkömmlichen Magnettonband sind auch Sampling-Dateien dem Schnitt unterziehbar. Es handelt sich hier natürlich nur sinnbildlich um ein „Schneiden“, da die Manipulationen ja nur virtuell vorgenommen werden, d. h. das Einfügen, Kopieren usw. erfolgt lediglich in Bezug auf eine Speicherkontinuität, nicht zwangsläufig auf eine physische Kontinuität.

3.1.1 Bearbeitungsschritte

Alle Schneidevorgänge besitzen Äquivalente zur Text- und Grafikbearbeitung. Es handelt also auch hier in der Hauptsache um die Arbeitsfunktionen Kopieren, Einfügen, Ersetzen, Ausschneiden und Löschen.

Kopieren Das Duplizieren eines Dateibereichs in einen Speicherbereich dient dem Zwecke der späteren Wiederverwendung dieses Teils, z. B. zum Einfügen an anderer Stelle oder in eine andere Datei. Was als Speicher hierfür in Frage kommt, hängt in erster Linie vom Umfang des zu kopierenden Stücks und auch vom Anwendungsprogramm und Betriebssystem ab.

Rechntypen wie Apple-Macintosh und auch IBM (mit Betriebssystem MS Windows) setzen häufig die Zwischenablage ein, die als Programm einen RAM-Bereich verwaltet. Je nach Ausstattung der RAM-Größe fällt auch der verfügbare Raum der Zwischenablage aus. Übersteigt die notwendige Speichermenge die Maximalgröße der Zwischenablage, erfolgt ein Zugriff auf die Festplatte, was natürlich den gesamten Vorgang erheblich

verlangsamt. Das Kopieren eines bestimmten Teilstücks einer Sampling-Datei geschieht i. Allg. durch Markieren des entsprechenden Bereichs, wie dies in Abb. 3.1 zu sehen ist.

Es handelt sich hier übrigens um eine Stereodatei mit einer Sampling-Frequenz von 22,05 kHz und einer Auflösung von 16 Bit. Das verwendete Anwendungsprogramm zeigt die Signale beider Kanäle in einem Fenster mit horizontalen und vertikalen Bildlaufleisten. Auf diese Weise sind wahlfreie Verschiebungen möglich. Am unteren Bildrand erscheinen die Zeitwerte der Datei. Der allgemein verwendete Befehl Kopieren legt das markierte Teilstück als Binärdaten-Paket in die Zwischenablage oder in einen anderen Speicher.

Einfügen Dieser Befehl verwendet nun das im Hilfsspeicher (Zwischenablage o. ä.) befindliche Teilstück einer Audiodatei und fügt dieses normalerweise an die Stelle ein, wo sich vorher der Cursor befunden hat. Diese Nahtstelle öffnet sich somit, und das einzufügende Stück gliedert sich zwischen die entstandenen Enden (Abb. 3.2).

Die zeitliche Länge der Audiodatei verlängert sich somit zwangsläufig um die Länge des einzufügenden Teils, wenn das Programm hier keine besonderen Schritte veranlasst. Der Inhalt des Zwischenspeichers bleibt nach dem Einfügevorgang jedoch normalerweise

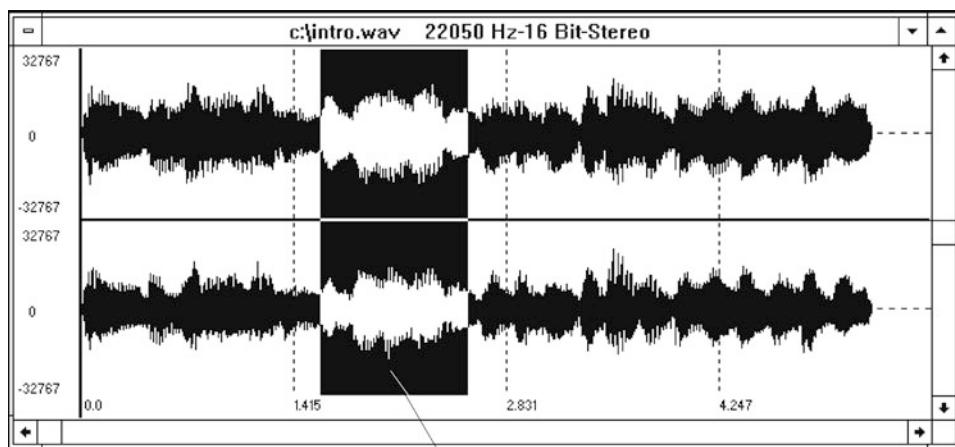
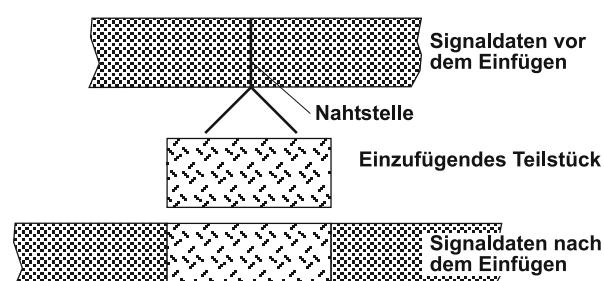


Abb. 3.1 Kopieren eines Bereichs einer Sampling-Datei

Abb. 3.2 Schema des Einfügevorgangs



bestehen, so dass der Insert-Schnitt theoretisch beliebig oft und an beliebigen Stellen wiederholt werden kann. Das ist für Spezialeffekte sicherlich sehr interessant.

Ersetzen Der Arbeitsschritt Ersetzen rückt die Nahtstelle nicht auseinander, sondern überschreibt einen Bereich der Datei. Die Anfangsposition dieses Bereichs wird durch die Cursor-Position definiert, während die Länge des Bereichs automatisch mit dem im Hilfsspeicher liegenden Teilstücks übereinstimmt. Den gesamten Vorgang spiegelt Abb. 3.3 schematisch wider.

Die Gesamtlänge der Audiodatei bleibt bei diesem Vorgang erhalten, weil hier der Überschreibemodus und nicht der Einfügemodus angewandt wird. So können ganze Passagen durch ausgewählte, bessere Versionen ersetzt werden. Ausschneiden. Soll ein markierter Bereich in den Hilfsspeicher kopiert und gleichzeitig aus der Chronologie entfernt werden, so ist der Arbeitsschritt Ausschneiden zu wählen. Die Lücke wird hierbei wieder geschlossen. Dieser Arbeitsgang erlaubt ein Bewegen bestimmter Portionen an andere Stellen der Datei. Man kann diese Funktion als Kombination der beiden Schritte Kopieren und Löschen ansehen. Abb. 3.4 zeigt schematisch die einzelnen Schritte.

Bei jedem weiteren Ausschneiden geht der alte Inhalt des Hilfsspeichers normalerweise verloren. Bei manchen Programmen ist jedoch eine bestimmte Anzahl von Bearbeitungsschritten wieder rückgängig zu machen, so dass man auf diese Weise wieder zu einer der letzten Versionen kommt.

Abb. 3.3 Schema beim Ersetzen

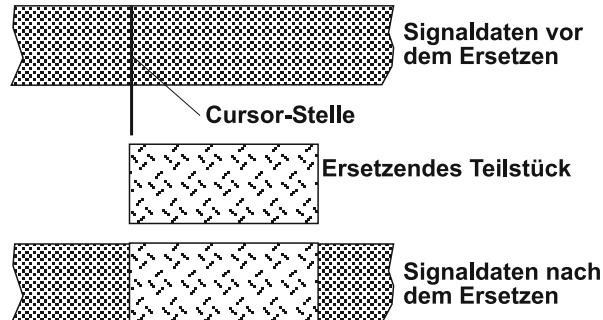


Abb. 3.4 Schema beim Ausschneiden

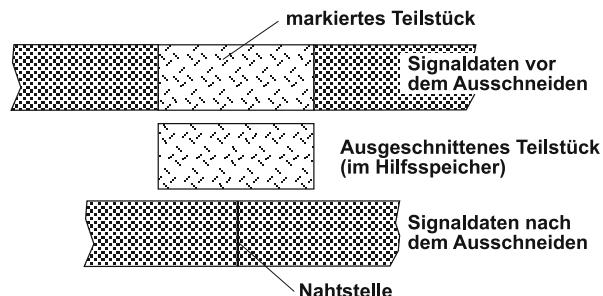




Abb. 3.5 Prinzip der Mute-Operation

Löschen Das Löschen entfernt einen markierten Bereich, wobei jedoch im Gegensatz zum Ausschneiden das herausgetrennte Stück nicht in den Hilfsspeicher gelangt (in ihm sind die Daten des letzten Kopier- oder Ausschneidevorgangs, ggf. ist er auch leer).

Mute Das Stummschalten eines markierten Teilstücks entspricht dem Löschen, es wird hier jedoch die Lücke nicht geschlossen, sondern sie bleibt als „stumme“ Sequenz erhalten (Abb. 3.5).

Das „stumme“ Teilstück der Audiodatei kann jetzt mit speziellen Signalen überschrieben werden, z. B. auch durch Werteingabe oder durch das Zeichnen eines Signalverlaufs.

3.1.2 Nahtstellen nach dem Schneidevorgang

Treffen die Enden zweier Teilstücke einer Audiodatei zusammen, die unterschiedlichen Zeitpunkten angehörten, so ergibt sich i. Allg. eine Unstetigkeit, d. h. es erfolgt möglicherweise ein abrupter Spannungssprung zwischen den Grenztastungen der Teilstücke. Abb. 3.6 zeigt eine solche Situation bei Verwendung eines 1-kHz-Sinussignals, welches an ungünstigen Stellen verschmolzen wird (bei entgegengesetzter Maximalelongation).

Um diesen Effekt beim Schnitt zu vermeiden, führen viele Anwenderprogramme für die Bearbeitung von Sampling-Dateien Interpolationen an den Schnittpunkten (bzw. Nahtstellen) durch. Diese Interpolationen betreffen jedoch nicht wie beim Oversampling zusätzliche Werte zwischen den vorgegebenen, sondern die benachbarten Werte der Nahtstelle. Abb. 3.7 zeigt den interpolierten Verlauf sowie das sich ergebende Frequenzspektrum über der Zeit.

Natürlich lässt sich eine kleine Welle im zeitlichen Spektrum nicht ganz unterdrücken. Der Algorithmus der Interpolation bestimmt natürlich entscheidend die Hörbarkeit der Schneidestelle. Außerdem ist meist die Zeit bzw. die Anzahl der Sampling-Zyklen anzugeben, über die sich die Interpolation erstrecken soll.

Anwenderprogramme ohne Schnittinterpolation machen eine „Nachbehandlung“ der Schneidestellen von Hand erforderlich; alternativ hierzu könnte man beim Schneiden bereits

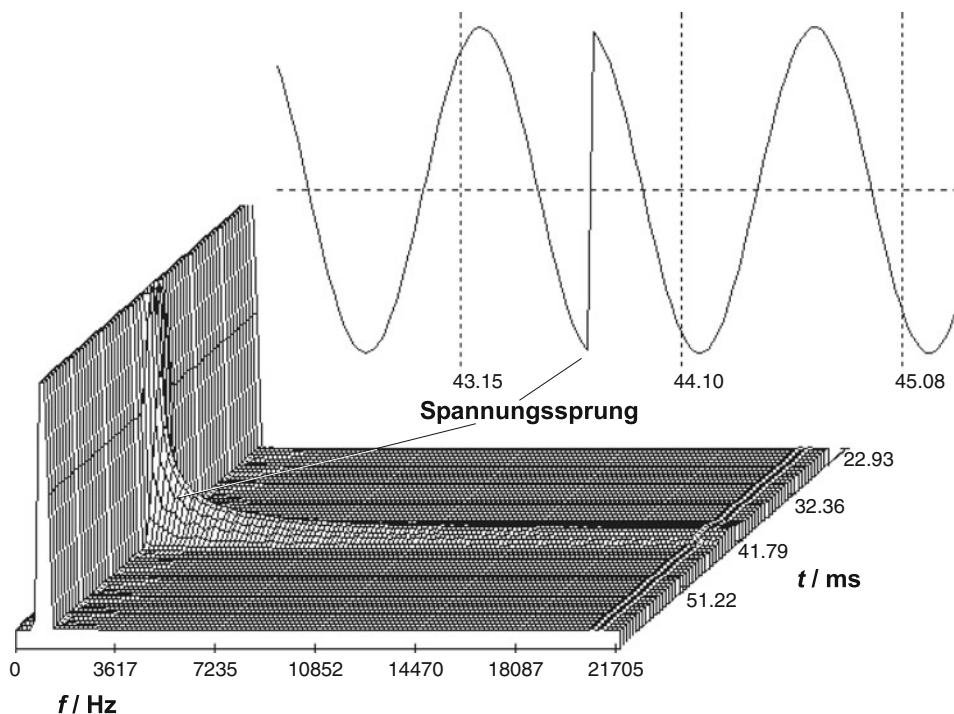


Abb. 3.6 Sprungstelle als Folge eines Schneidenvorgangs. Der Spannungssprung im Zeitbereich hat eine Welle im Frequenzbereich (genauer gesagt im Zeit-Frequenzbereich durch FFT) zur Folge. Die spektralen Anteile reichen erwartungsgemäß bis in die höchsten Bereiche. Akustisch äußert sich dies durch einen kurzen Knackgeräusch

geeignete Stellen suchen, die sich besser aneinanderfügen lassen. Eine andere Möglichkeit besteht durch Cross-Mix, d. h. einem Überblenden zwischen beiden Bereichen, die sich vor diesem Vorgang jedoch überlappen müssen. Dieses kurze Überblenden entspricht eigentlich auch einer Interpolation, wenn auch einer relativ weiträumigen.

Bei einem natürlichen Signal ist die Hörbarkeit an Schnittübergängen ohnehin auch subjektiv bedingt, d. h. eine gewisse Änderung ist unvermeidbar – einen Knacks empfindet das Ohr jedoch aufgrund der Erfahrung, die uns sagt, dass dieses Geräusch nicht von einem Instrument gekommen sein kann.

3.2 Filter

Falls der Frequenzgang während der Aufzeichnung keiner Korrektur unterzogen wurde, kann eine Filterung von Audiodateien erforderlich werden. Grundsätzlich sind auf digitalem Wege unterschiedliche Filterprinzipien realisierbar, die auch Differenzen im Verhalten und Ergebnis zeigen.

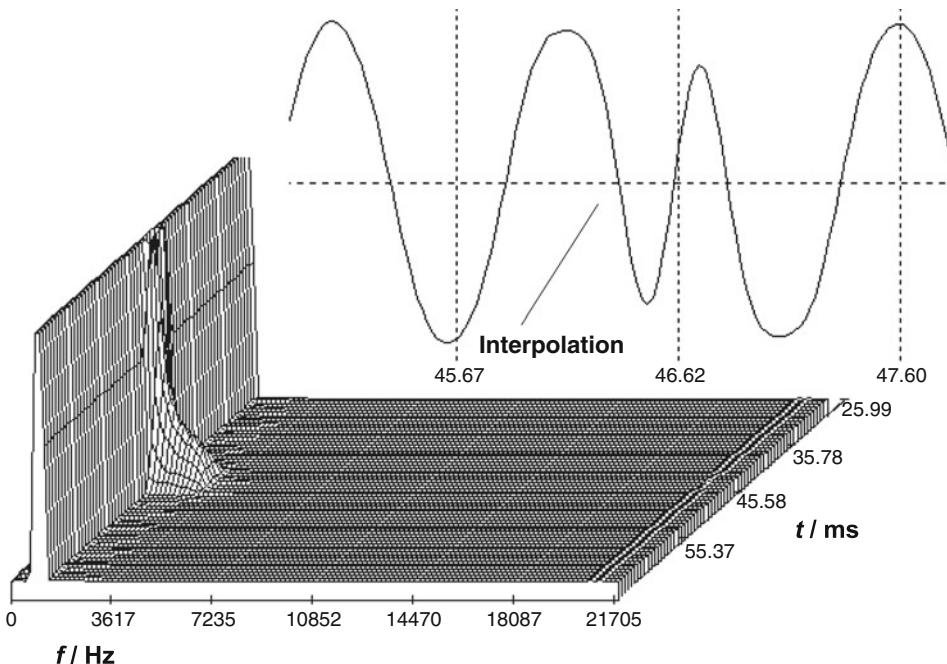


Abb. 3.7 Dämpfung des Knackgeräusches durch Grenzwert-Interpolation. Die Interpolation wurde hier aus Demonstrationszwecken durch eine FIR-Filterung erreicht, die sich über die gesamte Datei erstreckt; deshalb erscheint hier die kritische Stelle auch zeitlich verschoben gegenüber Abb. 3.6. In der Praxis darf natürlich keine solche Filterung stattfinden, da die Interpolation ja örtlich begrenzt stattfinden muss

3.2.1 FIR-Filter

Filter mit einer endlichen Impulsantwort nennt man FIR-Filter (Finite Impulse Response). Der schematische Aufbau ist in Abb. 3.8 zu sehen (für jeden Audiokanal sind natürlich separate Filter notwendig).

Nach jedem Taktzyklus erfolgt ein Weiterschieben der Audiodaten in der Kette der Speicherglieder. Gleichzeitig gelangen alle Ausgangswerte der Speicherkette auf einen Multiplikationskomplex, dessen Ausgänge wiederum bei jedem Schiebetakt akkumuliert, also aufsummiert werden. Vor jeder Zusammenfassung muss der Akkuinhalt jedoch geleert bzw. initialisiert werden. Das Ergebnis des Akkus entspricht dem Filterausgangswert. Mathematisch ergibt sich folgender Ausgangswert:

$$Y_N = \alpha_0 X_N + \alpha_1 X_{N-1} + \dots + \alpha_{N-1} X_1 + \alpha_N X_0 \quad (3.1)$$

In der Praxis trifft man FIR-Filter sowohl „fest verdrahtet“ innerhalb von Signalprozessoren, als auch Software-Simulationen zur Nachbearbeitung von Audiodateien in

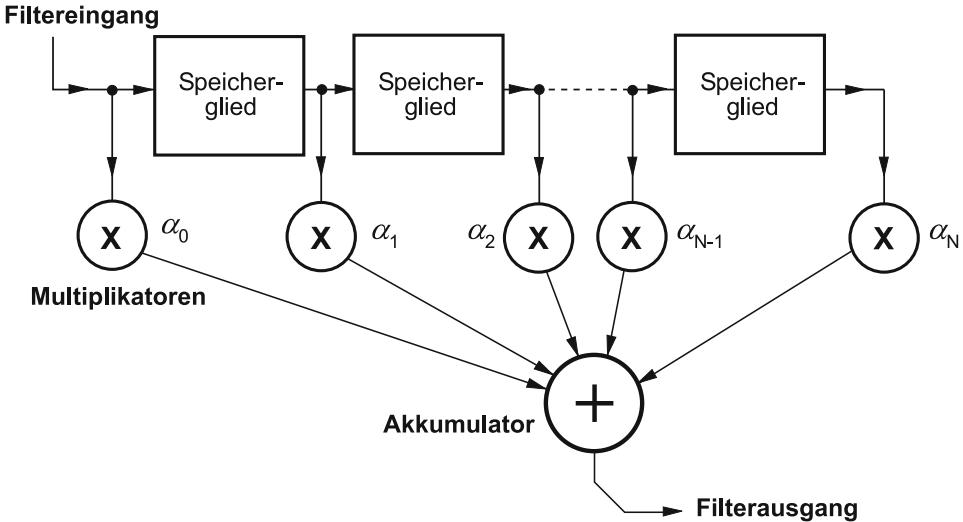


Abb. 3.8 Funktionsweise von FIR-Filters

Non-Realtime. Hardware-FIR-Filter kamen in Abschn. 2.2.4 bereits zur Sprache, und zwar im Zusammenhang mit dem Oversampling-Verfahren. S. 132 [2,3]

Prinzipielle Berechnung der Filterkoeffizienten Die a -Werte sind feste Filterkoeffizienten, die für jede Filterfunktion vorgegeben sind, also in erster Linie vom Filtertyp abhängen. Je nach ihrer Dimensionierung ergibt sich Tiefpass-, Hochpass-, Bandpass- oder Bandsperrencharakteristik. Jedoch hängen die Koeffizienten auch stark von der Filterordnung, also von der Anzahl der Stufen ab. S. 132 [2]

Die Filterkoeffizienten ergeben sich beim Tiefpass zu:

$$a_k = \frac{\sin[(2k - N)\pi \cdot F_G^*]}{(2k - N)\pi \cdot F_G^*} \cdot [0, 54 - 0, 46 \cdot \cos(2\pi k/N)] \quad (3.2)$$

k Index des Koeffizienten, N Ordnung des Filters, F_G^* normierte Grenzfrequenz (Verhältnis Grenzfrequenz zu Abtastfrequenz).

In Gl. 3.2 werden die ursprünglichen Koeffizienten (linker Teil des Terms) mit einem Bewertungsfaktor (rechter Teil des Terms) behandelt, welcher mit der Hamming-Fensterfunktion zu ermitteln ist. Dieses Verfahren ist recht einfach und liefert bereits gute Ergebnisse, was die tatsächlich erreichte Grenzfrequenz anbelangt. Vor allem für niedrige Grenzfrequenzen ($F_G^* < 0,1$) und kleine Filterordnungen sind dennoch Modifikationen der Koeffizienten erforderlich, damit die gewünschte Grenzfrequenz auch hinreichend genau erreicht wird.

In der obigen Gleichung ist ferner zu beachten, dass die Koeffizientenwerte sich spiegelbildlich nach dem Erreichen der mittleren Ordnung wiederholen; es ist demnach

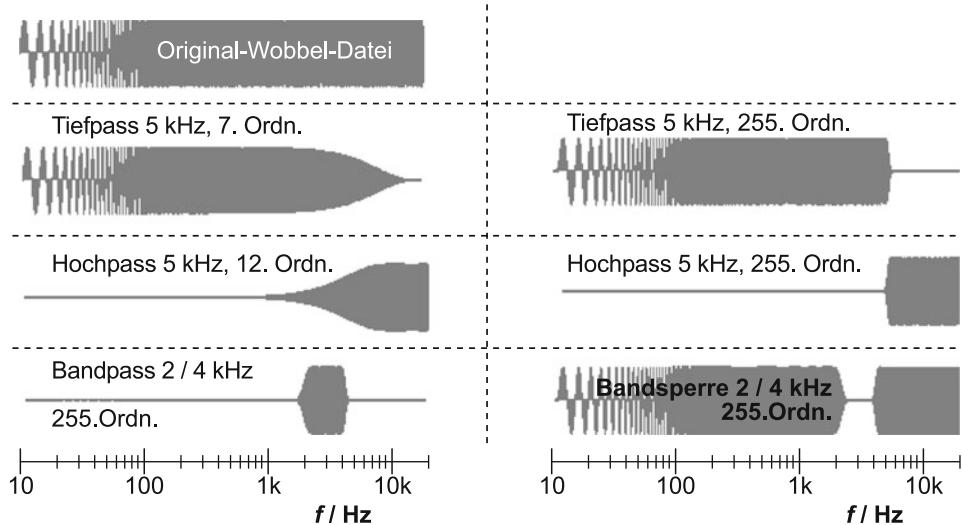


Abb. 3.9 Filterung einer Audiodatei mit einem Software-FIR-Filter

nicht nötig, alle Koeffizienten zu berechnen. Bei ungeraden Ordnungen existiert jedoch nur ein mittlerer Koeffizient. Für diesen Fall gilt außerdem $k = N/2$, und der Wert des Bruches ist auf 1 zu setzen.

In Abb. 3.9 ist das Ergebnis einer Bearbeitung mittels FIR-Filter zu sehen. Dieses ist in der Lage, Filterarten Tiefpass, Hochpass, Bandpass und Bandsperre mit Filterordnung bis 255 zu realisieren. Zu filtern sind damit WAV-Dateien im Monoformat und verschiedenen Auflösungen und Sampling-Frequenzen, z. B. eine Sweep-Datei, die Sinusfrequenzen von 10 ... 20000 Hz durchläuft. Der Frequenzverlauf ist bei dieser Datei logarithmisch zur Zeitachse, was die Beurteilung der Filterung erleichtert.

Tiefpass und Hochpass Grundsätzlich ergeben sich nach Gl. 3.2 die Koeffizienten für einen Tiefpass. Die Koeffizienten für einen Hochpass entstehen, wenn man die Koeffizienten des äquivalenten Tiefpasses (also mit gleicher Grenzfrequenz) von denen des Allpasses subtrahiert. Letzterer ist bekanntlich ein Glied, das alle Frequenzen gleichermaßen passieren lässt, wobei jedoch eine frequenzabhängige Phasenverschiebung entsteht. Die Koeffizienten für Allpassverhalten gewinnt man mit Gl. 3.2, indem man für die Grenzfrequenz die Hälfte der Sampling-Frequenz einsetzt (also $F_G^* = 0,5$).

Bandpass und Bandsperre Die Koeffizienten eines Bandpasses resultieren aus der Subtraktion der Koeffizienten zweier Tiefpässe, welche die Grenzfrequenzen des Bandpasses aufweisen. Ferner ergibt sich eine Bandsperre durch Subtrahieren der Bandpass-Koeffizienten von denen des Allpasses.

Die Realisierung mehrerer Bandpässe bzw. -sperren erfordert den mehrmaligen Durchlauf der Daten durch das Filter mit jeweils entsprechender Anpassung der Koeffizienten.

Auf diese Weise lassen sich theoretisch beliebige Frequenzverläufe erzielen. Allerdings steigt damit auch der Rechenaufwand und deswegen die erforderliche Rechenzeit.

Besondere Eigenschaften von FIR-Filtern Die Gruppenlaufzeit bei FIR-Filtern ist konstant. Das bedeutet mithin, dass das Signal für den Filterdurchlauf bei einer Gruppenordnung N und der Abtastfrequenz F_A folgende maximale Durchlaufzeit benötigt:

$$T_g = \frac{N}{F_A} \quad (3.3)$$

Bei jeder Filterung ist also grundsätzlich mit einer Verzögerung zu rechnen. Dies ist natürlich zu bedenken, wenn Daten verschiedenen Ursprungs zu mischen sind. Mitunter muss dann eine zeitliche Kompensation vorgenommen werden.

Das oben Gesagte lässt natürlich sofort den Schluss zu, dass FIR-Filter auch für reine Verzögerungsverfahren (wie z. B. Echo usw.) geeignet sind. Hierzu sind alle Koeffizienten auf 0 und nur der letzte auf 1 zu setzen. Die Grundverstärkung des Filters lässt sich proportional verändern, wenn man die Koeffizienten mit einem konstanten Faktor multipliziert.

3.2.2 IIR-Filter

Im Gegensatz zu den FIR-Filtern zeigen IIR-Filter theoretisch eine unendlich lange Impulsantwort. Weiterhin ergeben sich folgende Unterscheidungsmerkmale gegenüber FIR-Filtern:

- hohe Selektivität bei niedriger Ordnung
- geringer Speicherbedarf und wenig MAC-Operationen
- keine konstante Gruppenlaufzeit, somit Phasenverzerrungen möglich
- Stabilität nicht grundsätzlich gewährleistet, d. h. Schwingneigung vorhanden
- Koeffizienten müssen genau angegeben werden

IIR-Filter setzen zusätzlich einen Rückkopplungspfad ein, um den Filtereffekt bei kleinem Aufwand zu erreichen. Abb. 3.10 zeigt den schematischen Aufbau. Hierdurch gewinnt man zunächst den Eindruck, dass hier ein Mehraufwand nötig ist, jedoch kommen IIR-Filter mit wesentlich weniger Stufen aus als FIR-Filter. Trotz des Rückkopplungspfades sinkt damit der Schaltungsaufwand und somit die Rechenzeit erheblich.

Manche Hardware-Filter erlauben beide Funktionsarten – FIR- und IIR-Filter. Die Umschaltung erfolgt dann extern über den Status eines Anschlusses.

$$\begin{aligned} Y_N &= \alpha_0 X_N + \alpha_1 X_{N-1} + \dots + \alpha_{N-1} X_1 + \alpha_N X_0 \\ &\quad - \beta_1 Y_{N-1} - \beta_2 Y_{N-2} - \dots - \beta_{N-1} Y_1 - \beta_N Y_0 \end{aligned} \quad (3.4)$$

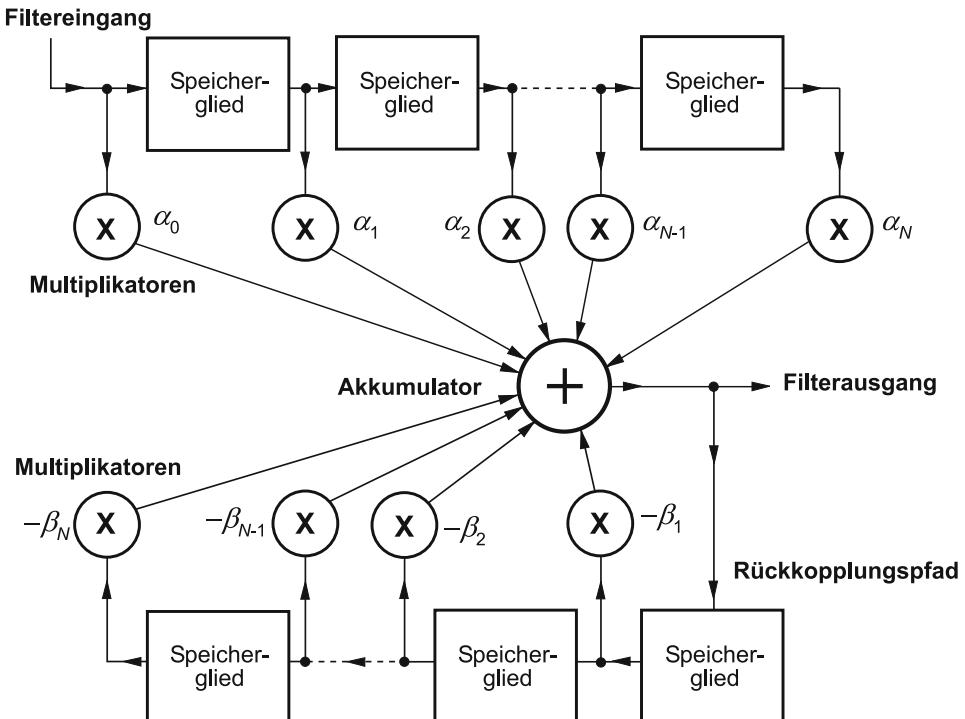


Abb. 3.10 Funktionsweise von IIR-Filters

An dieser Gleichung ist wegen den Y-Gliedern sofort erkennbar, dass es sich um rekursive Vorgänge handelt, d. h. die sich ergebenden Ausgangswerte erscheinen bei neuen Berechnungen wieder im rechten Teil der Gl. 3.4. S. 132 [2]

Auch dieses Filter ist leicht mittels Software zu simulieren und es lassen sich beliebige Filtertypen verwirklichen, die Tschebyscheff-, Butterworth- und Bessel-Charakteristik aufweisen. Hierzu sind lediglich wieder die Koeffizienten einzustellen. Wir wollen hier jedoch auf eine detailliertere Mathematik verzichten, die zur Berechnung der Filterkoeffizienten notwendig ist; an dieser Stelle sei auf spezielle Fachliteratur verwiesen. S. 132 [2]

Die Flankensteilheit von FIR-Filters hängt unmittelbar mit deren Ordnung zusammen; daraus ergibt sich eine Vergrößerung der Rechenzeit, wenn die Selektivität des Filters ansteigen soll. Bei IIR-Filters hingegen kann wie bei Analog-Aktivfiltern durch geeignete Wahl der Koeffizienten auch die Filtergüte manipuliert werden, ohne dass sich die Geschwindigkeit der Operationen ändert. Es sind jedoch – wie oben bereits erwähnt – auch durchaus nicht unerhebliche Nachteile zu verbuchen, weshalb dem IIR-Filter nicht immer der Vorzug gegeben werden kann. Die Entscheidung muss im Einzelfall getroffen werden.

3.3 Effekte

Für die Bearbeitung von Audiodateien stellen Anwendungsprogramme die unterschiedlichsten Effekte zur Verfügung, wobei die wohl wichtigsten Echo, Hall und diverse Phasing- und Flanging-Effekte darstellen. Wir wollen hier deshalb auf deren Funktionsprinzip und Besonderheiten eingehen. Bereits im Abschn. 2.6.3 sind einige Erläuterungen dazu gemacht worden – jedoch im Zusammenhang mit der Echtzeitbearbeitung mittels besonderer Geräte.

Die Simulation auf Software-Basis zeigt viele Ähnlichkeiten, nur sind hier meist wesentlich mehr Einstellmöglichkeiten verfügbar.

3.3.1 Echo

Die Erzeugung eines Echos erfordert eine Verzögerung des Originalsignals; danach erfolgt eine Beimischung zum Originalsignal. Abb. 3.11 verdeutlicht dieses Prinzip.

Einstellbar sind bei diesem Vorgang vor allem zwei Parameter, nämlich die Verzögerungszeit sowie die Intensität, mit der die verzögerten Signale dem Originalsignal beigemischt werden. Der Abschwächer lässt sich oftmals so einstellen, dass er als Verstärker wirkt. In diesem Falle führt die Rückkopplung zu einem Aufschaukeln des Signals, bis es zur Begrenzung kommt.

Ferner sind teilweise auch Sondereffekte möglich, z. B. das Einmal-Echo. Das Originalsignal initiiert in diesem Falle lediglich ein einziges Echo. Der Ausgang der Addierstufe liegt dann nicht mehr am Eingang der Verzögerungsstufe, somit ist der Weg für verzögerte Signale zum Verzögerungsglied versperrt. Eine Rückkopplung entfällt also, was zum Schema nach Abb. 3.12 führt.

Realisierung von Echos auf Software-Ebene Der Datenfluss nach Abb. 3.11 kann mittels Software nachgebildet werden. Als Verzögerungsglied dient ein als Puffer wirkender Speicherbereich, durch den die Audiodaten hindurch zu schieben sind. Die aus der letzten

Abb. 3.11 Prinzip einer künstlichen Echoerzeugung

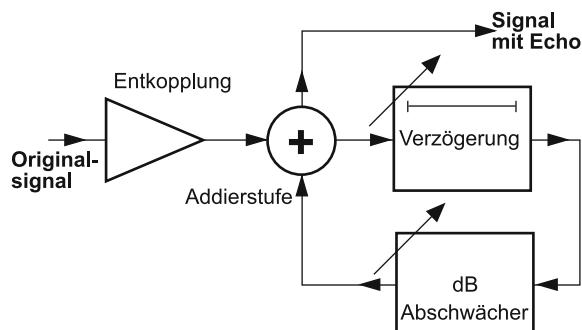
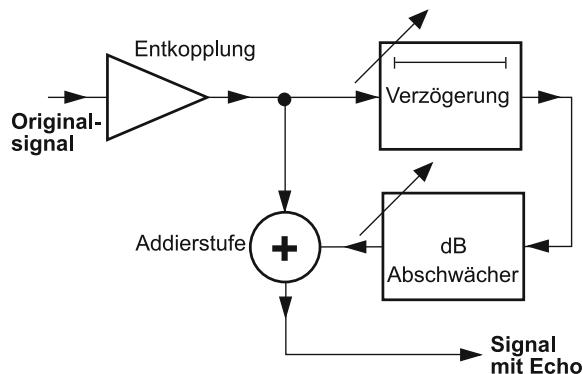


Abb. 3.12 Erzeugung von Einmal-Echos



Speicherzelle hinausgeschobene Information stellt die zeitlich verzögerten Daten dar. Eine additive Mischung, die Originaldaten und verzögerte Daten zusammenführt, erfolgt auf digitaler Ebene natürlich rein rechnerisch.

Bei allen Mischvorgängen muss darauf geachtet werden, dass das Summensignal nicht übersteuert ist, was bei einem optimal ausgesteuerten Originalsignal zu befürchten ist. Zur Mischung bzw. Addition von zwei Signalen mit optimaler Aussteuerung sollte daher zunächst eine Reduzierung der Pegel von jeweils 6 dB vorgenommen werden. Komfortable Anwenderprogramme lassen eine solche Option simultan mit dem Mischvorgang zu, so dass kein weiterer Arbeitsschritt notwendig ist.

Zwei Signale, die z. B. nur Sinussignale gleicher Frequenz repräsentieren, erzeugen eine Summenamplitude, die nur von der Phasenlage beider Komponenten zueinander abhängt. Bei beliebigen Signalen ist die Sachlage wesentlich anders; natürlich hängen hier eventuell auftretende Peaks ebenfalls vom Phasenunterschied beider Signale ab, jedoch lässt sich dies nicht vorhersagen, wenn der Signalverlauf nicht genau bekannt ist. Die Situation ähnelt der Addition zweier Rauschsignale, deren Summenamplitude im Durchschnitt auf den Wert

$$\bar{U} = \frac{\bar{U}_0}{\sqrt{n}} \quad (3.5)$$

abfällt, was bei zwei Signalen mit $n = 2$ einem Rückgang auf ca. 71% entspricht. Das gilt jedoch nur für den Durchschnitt – nach der Wahrscheinlichkeit ist nicht auszuschließen, dass an diskreten Stellen auch eine Zunahme der Elongation stattfindet. Komfortable Programme sind in der Lage, eine Untersuchung auf Peaks sozusagen im Voraus durchzuführen, um sodann für die Mischung eine angepasste Reduzierung vorzunehmen.

Bei Programmen, die eine Mischung mit $1/n$ der Amplituden vornehmen, ist eine Übersteuerung nicht zu befürchten.

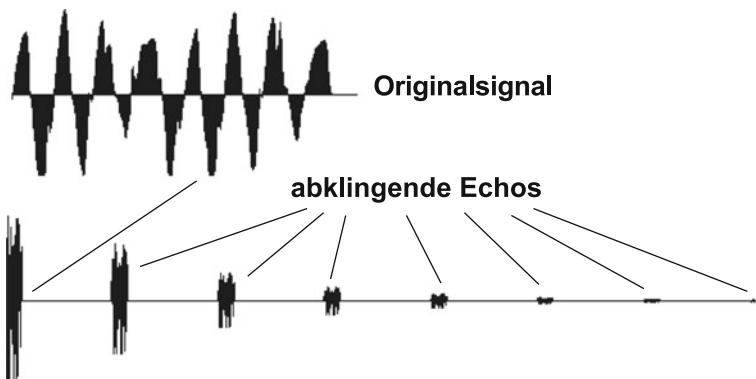


Abb. 3.13 Erzeugung abklingender Echos bei einem Echopegel von 50 % (−6 dB)

Abb. 3.13 zeigt ein Beispiel für die Erzeugung eines Echos, wobei das Originalsignal ein kurzes Wellenpaket ist, damit die Echos deutlich sichtbar werden (in der Praxis überlagern sich die Echos i. Allg. dem Originalsignal).

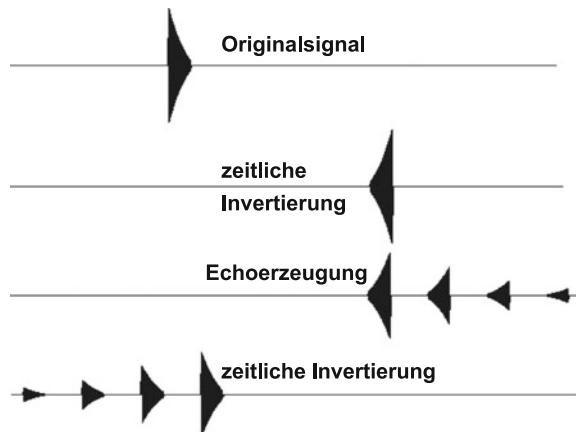
Dargestellt sind hier natürlich nur wenige Nachechos, obwohl bei dieser Einstellung durchaus bis zu ca. 15 Echos zu erwarten sind.

Invers-Echos Echos, die vor dem Originalsignal mit ansteigendem Pegel erscheinen, nennt man Invers- oder Rückwärts-Echos. Dieser Effekt kommt unter natürlichen Umständen nicht vor – man muss ihn künstlich erzeugen. Die Methode ist relativ einfach, und bereits mit üblichen Magnetbandmaschinen waren Rückwärtsechos machbar, wenn das Gerät über eine normale Echoeinrichtung verfügte. Das Verfahren erfolgt nach folgenden Schritten:

- Das Originalsignal ist rückwärts abzutasten; dies geschieht entweder durch vollständige zeitliche Umkehrung und Normalabtastung oder durch direktes Rückwärtslesen einzelner Dateiblöcke, welche vom Ende bis zum Anfang der Ursprungsdatei zu laden sind.
- Zum diesem rückwärts abgetasteten Digitalsignal wird ein Echo nach bekannter Methode erzeugt.
- Jetzt ist die Datei wieder in der ursprünglichen Richtung zu lesen.

In Abb. 3.14 ist der gesamte Vorgang zu sehen. Es wurde dabei angenommen, dass das Programm zunächst eine zeitliche Umkehr durchführen muss, die am Ende wieder rückgängig zu machen ist. Das ist natürlich nicht bei jedem Programm der Fall; manche mögen in der oben beschriebenen Weise die Daten einer Datei rückwärts lesen können, und somit entfallen die Umkehrschrifte am Anfang und Ende des Verfahrens.

Abb. 3.14 Erzeugung eines Rückwärts-Echos



Bei der Echobildung gibt es natürlich meist eine längere Wirkung, als dies aus der Abbildung hervorgeht; soll das Originalsignal bis zu seinem Ende (nach der zeitlichen Drehung wird der Anfang zum Ende) mit dem Echoeffekt bearbeitet werden, so entsteht eine zeitliche Verlängerung durch die Echoerzeugung bis über den ursprünglichen Anfang hinaus, was nur künstlich unterbunden werden kann (was jedoch den Echoeffekt am Anfang der Datei verstümmelt).

3.3.2 Hall

Zu den etwas aufwändigeren Effekten zählt der Hall; er entsteht durch die Generierung (theoretisch) unendlich vieler Echos, die wiederum unendlich viele unterschiedliche Laufzeiten aufweisen.

In der Praxis muss man sich auf eine gewisse Anzahl von Reflexionspunkten (sinnbildlich gesprochen) beschränken; die Rechenzeit für eine Hallerzeugung hängt unmittelbar mit dieser Anzahl zusammen. Abb. 3.15 zeigt eine Hallsimulation, wie sie mit relativ kleinem Rechenaufwand möglich ist. Sichtbar (und natürlich zwangsläufig für empfindliche Ohren auch hörbar) sind die einzelnen diskreten Echos, die mit einer scheinbaren zeitlichen Zufälligkeit auftreten – ein Zeichen für die Unzulänglichkeit der Nachbildung. Jedoch für bestimmte Anwendungen mag der Verdeckungseffekt durch andere Klangquellen so groß sein, dass die diskreten Echos hier nicht mehr über die Identifikationsschwelle gelangen.

Bei jeder Hallentwicklung ist eine Grundgröße die erste Verzögerungszeit, welche die Laufzeit beschreibt, bis das primäre Hallsignal erscheint. Danach jedoch können diskrete Echos zu unregelmäßigen Zeitpunkten auftreten. Man kann sich die Vorgänge in etwa so vorstellen, wie wenn man in einem mit Wasser gefüllten, flachen Gefäß an einer Stelle eine Welle initiiert; nach kurzer Zeit kommen dort die an den Wänden gespiegelten Wellen an,

Abb. 3.15 Hallerzeugung durch Addition von Echosignalen mit pseudo-zufälligem zeitlichen Charakter

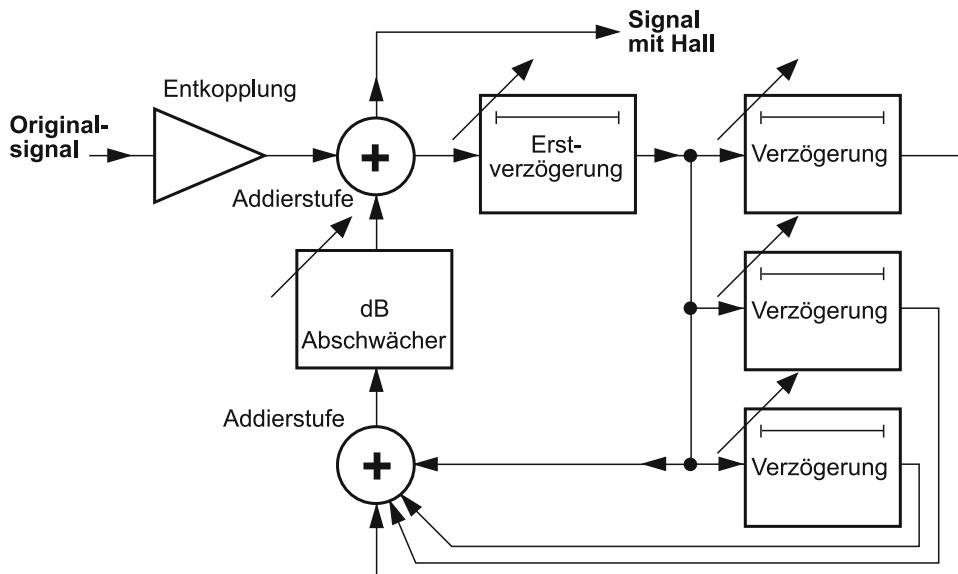


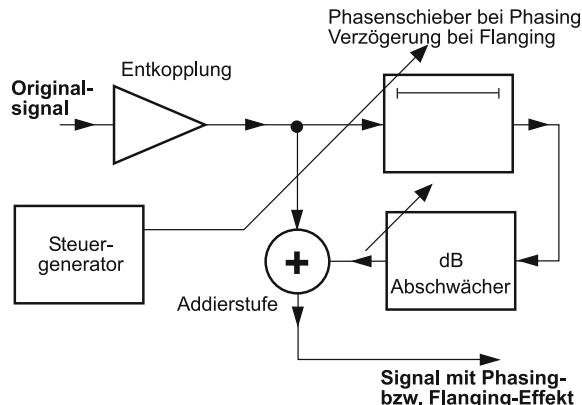
Abb. 3.16 Prinzip einer künstlichen Hallerzeugung

und danach erfolgen unzählige Reflexionen, die durch unterschiedliche Laufzeiten schnell zu einem chaotischen System werden. Ähnliche Vorgänge spielen sich in einem Raum oder Saal mit schwacher Grenzflächendämpfung für den Schall ab.

Simulation des Halls In Abb. 3.16 sehen wir die Umsetzung einer künstlichen Hallerzeugung. Auf diese Weise erfolgt prinzipiell auch die Verwirklichung auf Software-Ebene. Die Anzahl der Verzögerungsstrecken macht die Qualität des Halls und natürlich auch die Rechenzeit aus. Die Halldauer kann mit dem Abschwächer eingestellt werden; er definiert sozusagen die Dämpfung an den reflektierenden Grenzflächen.

Oft ist lediglich die Zeit der Erstverzögerung einstellbar; die Zeiten für die restlichen Verzögerungsglieder sind auf spezifische Größenordnungen für Hall eingestellt, wobei sie sich natürlich unterscheiden und auch keine ganzzahligen Vielfache voneinander sind. Die Anzahl der Verzögerungsglieder beträgt hier vier – hochqualitative Hallsysteme arbeiten jedoch mit zehn und mehr Gliedern.

Abb. 3.17 Schematische Arbeitsweise des Phasing bzw. Flanging (vgl. Abb. 3.12)



3.3.3 Phasing und Flanging

Phaser und Flanger kommen in der Regel mit einem Phasenglied bzw. Zeitglied pro Kanal aus. Damit hat der prinzipielle Aufbau Ähnlichkeiten mit dem des Einmal-Echos, mit der Ausnahme, dass die Phase bzw. Verzögerungszeit über einen Generator langsam verändert wird. Außerdem betragen die Verzögerungszeiten hier maximal ca. 100 ms.

Arbeitsweise Wie bereits in Abschn. 2.6.3 beschrieben, setzt das Phasing einfache Phasenschieberketten (Allpässe) ein, während beim Flanging Zeitverzögerungsglieder zur Anwendung kommen (Abb. 3.17). Im Prinzip handelt es sich dabei zwar um das-selbe, jedoch erreicht man mit Zeitgliedern bei entsprechender Dimensionierung eine wesentlich steilere lineare Abhängigkeit zwischen Phase und Frequenz. Als Folge stellen sich wesentlich mehr Nullstellen im Frequenzgang ein als beim Phaser-System. Diese Nullstellen (bzw. Auslösungen durch gegensätzliche Phasenlagen) verschieben sich im Frequenzbereich, wenn sich der Parameter des Phasen- bzw. Zeitglieds verändert.

Objektive und subjektive Wirkung Wir hatten bereits die sich ergebenden Nullstellen im Frequenzgang erwähnt; oft wird auch von einem kammartigen Frequenzgang gesprochen. Abb. 3.18 zeigt ein Wobbelsignal, das mit den Effekten modifiziert wurde. Dargestellt wurden hier Ergebnisse für drei verschiedene Verzögerungszeiten (innerhalb eines Wobbeldurchgangs blieb die Verzögerungszeit jeweils konstant). Zur Demonstration des Effektes wurden übrigens einfach zwei identische, unveränderte Wobbelsignale mit einer entsprechenden Zeitverzögerung „übereinandergelegt“, also eine Mischung aus beiden generiert.

Werden Klänge von Instrumenten mit reichem Spektrum (z. B. Orgel) oder ganze Orchester mit dem Phasing oder Flanging verändert, so entsteht der Eindruck, die Musik schwebt förmlich am Zuhörer vorbei und erhebt sich in die Lüfte.

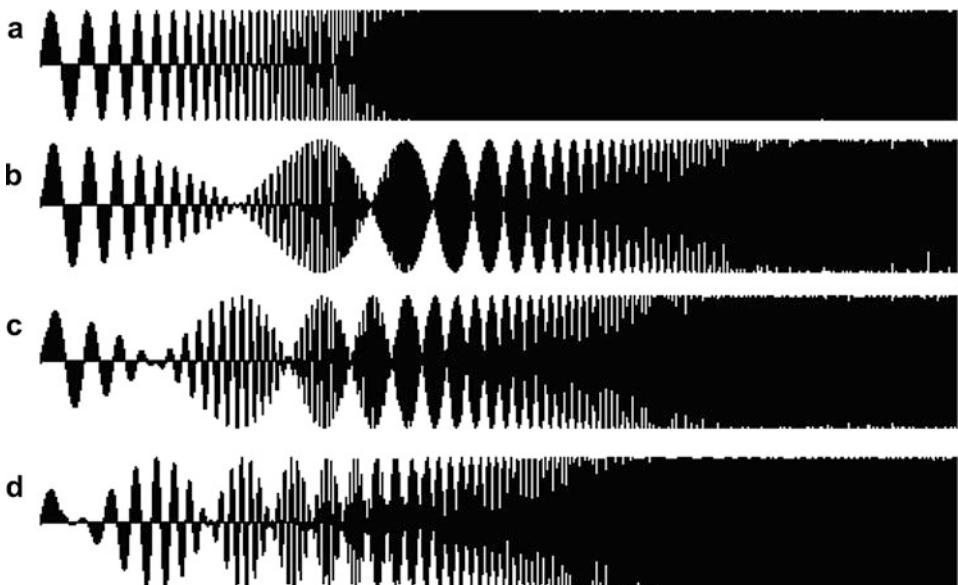


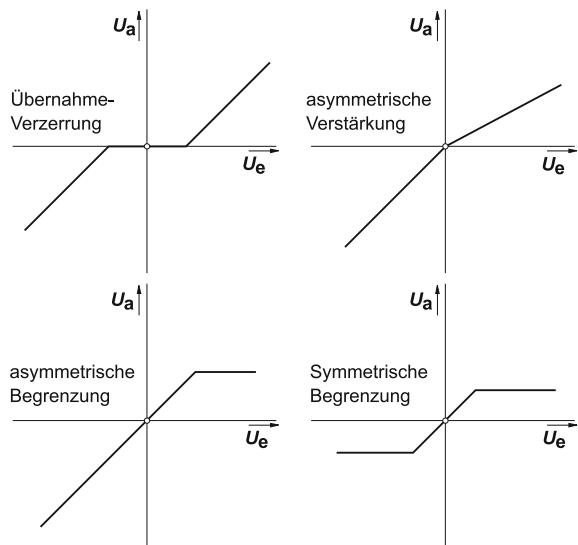
Abb. 3.18 Wirkungen des Flanging im Frequenzbereich. (a) unverändertes Wobbelsignal, (b) Flanging mit einer momentanen Verzögerungszeit von 10 ms, (c) bei 20 ms und (d) bei 40 ms

Realisierung auf Software-Ebene Die Beschreibung der Umsetzung ist recht einfach. Das Prinzip orientiert sich weitgehend am Echoeffekt und die Realisierung erfolgt auf die gleiche Weise. Der Steuergenerator zur Variation der Verzögerungszeit ist über eine „wandernde“ Adressierung realisierbar; die Puffergröße kann in diesem Fall konstant bleiben. Einstellbar sind normalerweise auch verschiedene Kurvenformen und natürlich die Amplitude des Steuersignals, damit die Tiefe des Effekts beeinflussbar ist. Der „feststehende“ Effekt, wo also keine Verschiebung der Nullstellen stattfindet, kann wie erwähnt durch das Mischen identischer, zeitversetzter Signale erzielt werden.

3.3.4 Verzerrung

Der Verzerrungseffekt ist anfänglich vor allem für elektrische Gitarren eingesetzt worden. Ein Verstärker wird dabei außerhalb seiner linearen Kennlinie betrieben, wo ja bekanntlich Begrenzungseigenschaften einsetzen. Ist das ursprüngliche Signal ein reiner Sinus (mit einer einzigen Linie im Frequenzbereich), so zeigt eine gekrümmte oder abgeflachte Degenerierung starke Obertöne, d. h. zur Grundfrequenz kommen noch Harmonische, welche ein ganzzahliges Vielfaches der Grundfrequenz aufweisen (vgl. hierzu auch Abschn. 1.3).

Abb. 3.19 Kennlinien einiger typischer Verzerrungsarten



Verzerrungsarten Je nachdem, welchen Charakter die statische Kennlinie des Übertragungssystems aufweist, ergeben sich unterschiedliche spektrale Zusammensetzungen der Obertöne. Abb. 3.19 zeigt eine Auswahl einiger Verzerrungscharakteristika, welche im Frequenzbereich große Unterschiede mit sich bringen.

Übernahmeverzerrungen sind meist unerwünscht, weil sie klanglich relativ uninteressant sind und weil eben unterhalb einer gewissen Schwelle alles verschluckt wird.

In der Praxis finden vor allem Begrenzungsmethoden Anwendung, obwohl auch die asymmetrische Verstärkung sehr interessante Merkmale aufweist. Sie generiert nämlich als einzige lediglich Harmonische gerader Ordnung, also mit $2f, 4f, 6f$ usw. Außerdem tritt kein Sustain-Effekt auf, d. h. die Ausgangsspannung steht immer in einem festen Verhältnis zur Eingangsspannung.

Die symmetrische Begrenzung lässt Harmonische ungeradzahliger Ordnung entstehen, während die asymmetrische Begrenzung alle Harmonischen aufweist. Abb. 3.20 verdeutlicht die Effekte in Bezug auf Verhalten im Zeit- und Frequenzbereich am Beispiel eines Sinussignals mit einer Grundfrequenz von 1 kHz. Die starken 0-Hz-Komponenten bei den beiden asymmetrischen Methoden röhren von der natürlichen Mittelwertsverschiebung und dem daraus resultierenden DC-Offset. Dieser Effekt kann jedoch unterbunden werden, unabhängig davon, ob die Verzerrung auf Hardware- oder Software-Ebene entsteht.

Realisierung auf Software-Ebene Im Gegensatz zu den Filter-, Hall- und Echovorgängen kann eine Verzerrung ohne Beachtung benachbarter Abtastpunkte erfolgen, d. h. bei der Manipulation eines Abtastwerts sind die benachbarten Stützstellen nicht von Belang.

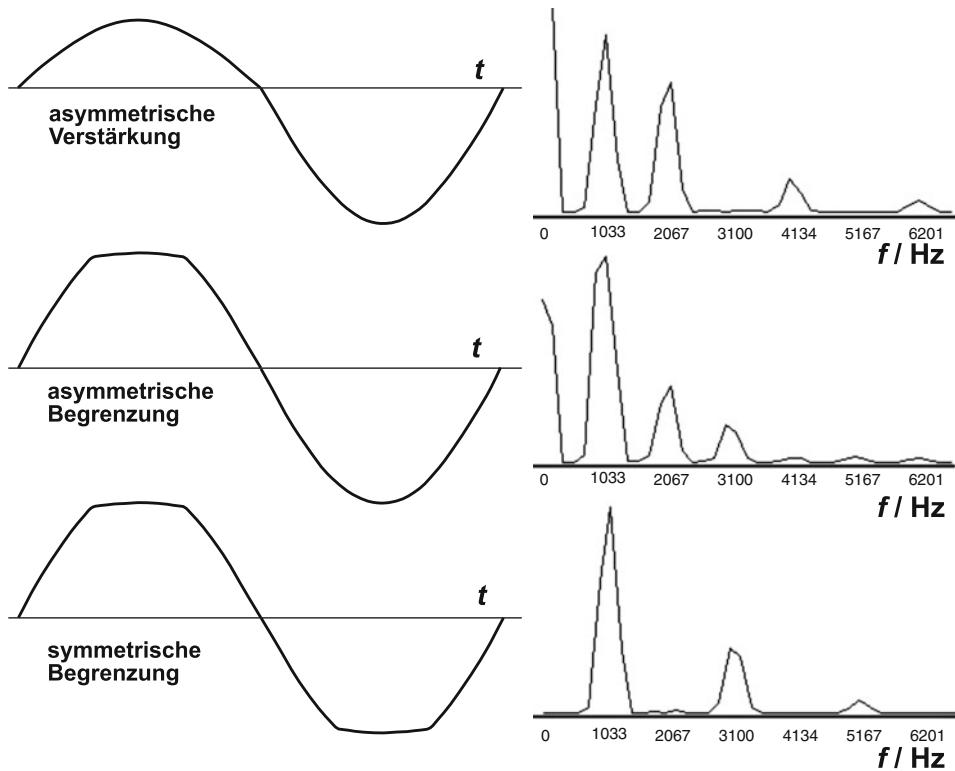


Abb. 3.20 Verzerrungsmerkmale im Zeit- und Frequenzbereich

Liegen die Werte im 2er-Komplement vor, so ergeben sich folgende Verzerrungsalgorithmen:

$$\text{Asymmetrische Verstärkung: } Y = \begin{cases} X, & \text{für } X \leq 0 \\ c \cdot X, & \text{für } X > 0 \text{ und } 0 \leq c \leq 1 \end{cases} \quad (3.6)$$

$$\text{Asymmetrische Begrenzung: } Y = \begin{cases} X, & \text{für } X \leq 0 \\ C_0, & \text{für } X > C_0 \end{cases} \quad (3.7)$$

$$\text{Symmetrische Begrenzung: } Y = \begin{cases} X, & \text{für } |X| \leq C_0 \\ C_0, & \text{für } X > C_0 \\ -C_0, & \text{für } X < -C_0 \end{cases} \quad (3.8)$$

X ursprünglicher Abtastwert, Y neuer Abtastwert, c Stauchkonstante, C_0 Schwellenwert für Begrenzung.

Grundsätzlich ist es möglich, auch das Format Offset-Dual zu behandeln, wenn zunächst übergangsweise ins 2er-Komplement konvertiert wird.

3.3.5 Veränderung der Abspielgeschwindigkeit

Stellt man sich die Sampling-Datei als Gummiband vor, so erfolgt eine Reduzierung der Abspielgeschwindigkeit, indem man dieses Gummiband streckt, während eine Stauchung einer Erhöhung der Geschwindigkeit gleichkommt. Wir betrachten hierzu in Abb. 3.21 eine kurze Sampling-Datei, die aus zwei Sinusschwingungen besteht.

Der **Zeitstreifen** spiegelt die Veränderung wider, d. h. die zeitliche Länge der Datei folgt dem Dehnungsfaktor. Gleichzeitig ist natürlich eine Frequenzänderung zu verbuchen. Nennen wir die normierte Abspielgeschwindigkeit V (bei Normalgeschwindigkeit ist $V = 1$), so ergeben sich für Dateilänge t und Wiedergabefrequenz f :

$$t = \frac{t_0}{V} \quad \text{Dateilänge} \quad (3.9)$$

$$f = f_0 \cdot V \quad \text{Wiedergabefrequenz} \quad (3.10)$$

Mit der zeitlichen Änderung der Datei variiert natürlich auch ihr Speicherbedarf.

Realisierung auf Software-Ebene Grundsätzlich sind für die Umsetzung zwei Mechanismen denkbar, und zwar eine hardware-nahe und eine reine Software-Lösung.

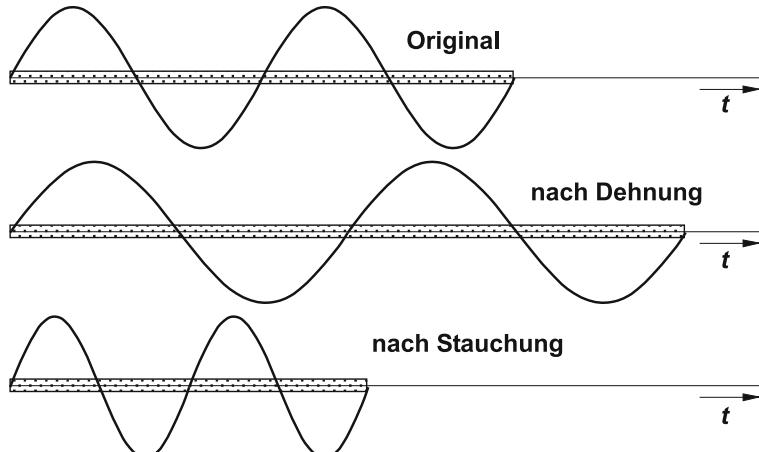


Abb. 3.21 Dateidehnung und -stauchung und die damit verbundene Änderung der Abspielgeschwindigkeit

Sampling-Raten-Manipulation Die erste Methode beeinflusst das Verhältnis zwischen eingestellter Sampling-Frequenz der Audiokarte und der vorgegebenen Rate der Datei. Im Header der Datei ist eine Eintragung, aus der die vorgesehene Abtastfrequenz hervorgeht. Eine Manipulation dieses Werts ergibt (theoretisch) den gewünschten Effekt einer Änderung der Wiedergabegeschwindigkeit. Obwohl diese Methode in Echtzeit vollziehbar ist, ergeben sich einige Handicaps, die diesen Weg vereiteln können. Eine solche Änderung kann eine Datei nur über die gesamte Länge interpretieren, weil im Header nur einmal die Sampling-Rate spezifiziert ist. Aus dem gleichen Grunde können natürlich auch keine Dateien gemischt oder verbunden werden, die in diesem Punkte inkompatibel sind (zumindest kann dieser Arbeitsschritt nicht ohne speziellen Algorithmus erledigt werden). Beachtenswert ist außerdem, dass nicht alle Audiokarten eine Abweichung von den Standard-Sampling-Raten ohne Qualitätsverlust hinnehmen. Manche reagieren darauf mit unangenehmen Nebengeräuschen. Vermeidbar sind solche Effekte nur bei Wahl der doppelten oder halben Frequenz, sofern die Audiokarte dafür ausgelegt ist.

Berechnung von Interpolationswerten Die gängigste Methode arbeitet vollständig auf Software-Ebene und erlaubt somit i. Allg. keine Echtzeitverarbeitung, es sei denn, hierfür stünde ein separater Signalprozessor zur Verfügung. Bei der Änderung der Geschwindigkeit ergibt sich eine Umverteilung der Stützstellen bzw. Abtastpunkte. Bei einer Vergrößerung der Wiedergabegeschwindigkeit erfolgt ein Auseinanderziehen der Abtastpunkte auf der angenommenen Signallinie, eine Verlangsamung erfordert hingegen eine Stauchung der Abtastpunkte. Ist daher der Verlauf durch Stützstellen nach Abb. 3.22 vorgegeben, so können auch Werte zwischen den diskreten Stützstellen durch Interpolation gewonnen werden. Sei die duale Form der Spannungsfunktion $U(t_n)$, so kann also eine Berechnung der neuen Stützstellenwerte $U(t_m)$ an beliebigen Zeitpunkten erfolgen. Die konstanten Zeitintervalle müssen nun nicht mehr dem ursprünglichen Abtastabstand Δt_n entsprechen, sondern können mit Δt_m auch eine Abweichung nach unten oder oben aufweisen.

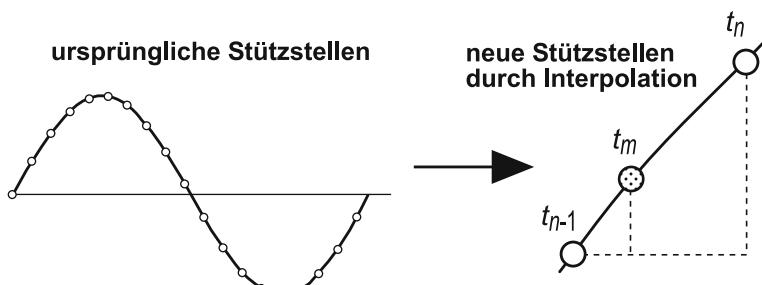


Abb. 3.22 Berechnung neuer Abtastwerte durch lineare Interpolation

Der Algorithmus für eine lineare Interpolation gestaltet sich also wie folgt:

$$t_m \leftarrow t_m + \Delta t_m \quad \text{Zeiterhöhung} \quad (3.11)$$

$$t_{n-1} = \Delta t_n (t_m \text{div} \Delta t_n) \quad \text{und} \quad t_n = t_{n-1} + \Delta t_n \quad \text{Nachbarstützstellen} \quad (3.12)$$

$$U(t_m) = U(t_{n-1}) + \frac{t_m - t_{n-1}}{t_n - t_{n-1}} \cdot (U(t_n) - U(t_{n-1})) \quad \text{Linearinterpolation} \quad (3.13)$$

Die Berechnung der Nachbarstützstellen kann auch mittels sog. Überlaufregister einfacher erfolgen, man erhält dadurch sofort die Adresse der Nachbarstützstellen, die man ja zur Ermittlung der zugehörigen Werte benötigt, und das Zeitintervall $t_m - t_{n-1}$ ergibt sich dann als Überlaufwert.

Nachdem die neue Datei auf die oben beschriebene Weise aus der alten entstanden ist, muss man sich vor Augen halten, dass das Abspielen der neuen Abtastwerte natürlich wieder mit den gleichen Zeitintervallen geschieht. Die t_m -Werte wurden lediglich für die Interpolation benötigt.

Die Steigerung der Abspielgeschwindigkeit birgt möglicherweise Gefahren, wenn dadurch die Abtastbedingung nicht mehr erfüllt ist. Besteht auf der Software-Seite die Option einer Aliasing-Unterdrückung, so sollte diese zur Anwendung kommen. Grundsätzlich ist eine Geschwindigkeitserhöhung um den Faktor V (vgl. Gl. 3.10) auch mit einer entsprechenden Frequenzerhöhung verbunden. Dies verpflichtet zur Herabsetzung der oberen Grenzfrequenz (vor der Veränderung) auf:

$$f_g \leq \frac{F_A}{2 \cdot V} \quad (3.14)$$

Mit Hilfe einer FIR-Filterung hoher Ordnung gelingt eine wirksame Unterdrückung von Frequenzen, die nach der Erhöhung der Geschwindigkeit die Grenze der halben Abtastfrequenz überschreiten würden.

3.3.6 Veränderung der Tonlage (Pitch)

Während die Abspielgeschwindigkeit sowohl die Gesamtzeit als auch die Tonlage bestimmt, beeinflusst die alleinige Veränderung der Tonlage die Abspielzeit nicht. Soll die Tonlage beispielsweise um eine Quinte nach oben transponiert werden, so ergibt sich die Situation nach Abb. 3.23. Die Frequenz steigt um den Faktor 1,5, und die Gesamtabspielzeit bleibt jedoch konstant.

Wie man sich leicht vorstellen kann, bedeutet eine solche Bearbeitung der Audiodaten einen weitaus größeren Aufwand, als das bei einer reinen Änderung der

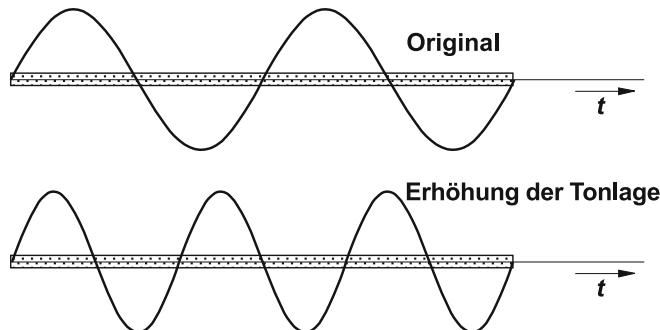


Abb. 3.23 Tonlagentransponierung um den Faktor 1,5 (Quinte nach oben)

Abspielgeschwindigkeit der Fall ist. Schließlich handelt es sich nicht mehr um eine bloße Umstrukturierung der vorhandenen Daten, sondern es sind Teildaten zu entfernen oder hinzuzufügen.

Die Transponierung selbst wird entweder in Prozent (Pitch Ratio p_r) oder bei musikalischer Orientierung in n Halbtönen angegeben. Eine Umrechnung erfolgt nach der Beziehung:

$$p_r = 2^{n/12} \quad (3.15)$$

Als elektronisch ausgeführtes Gerät trägt eine solche Einrichtung die Bezeichnung Harmonizer – ein meist digital arbeitendes Effektgerät.

Realisierung auf Software-Ebene Einfachere Algorithmen vergrößern bei einer Transponierung nach oben mit z. B. $p_r = 1,5$ die Abspielgeschwindigkeit um den gleichen Faktor durch das bereits beschriebene Verfahren; die sich ergebende Zeitschrumpfung wird durch das Einfügen ganzer Perioden der Grundschwingung (je eine Periode nach zwei bestehenden Perioden) wieder kompensiert. Diese Methode ist natürlich recht ungenau, außerdem können je nach Ausgangssignal unangenehme Nebengeräusche entstehen.

Ein besseres Verfahren transformiert feste Intervalle vom Zeitbereich in den Frequenzbereich. Das verwendete Analyseverfahren hierzu ist meist *Fast Fourier Transformation*, bei dem Datenblöcke der Größe einer 2er-Potenz zur Umrechnung herangezogen werden. Die so ermittelten Spektralanteile für jeweils wenige Millisekunden erfahren nun die eigentliche Transponierung um den gewünschten Wert. Anschließend erfolgt wieder eine Rück-Transformation in den Zeitbereich. Abb. 3.24 zeigt schematisch die Vorgänge, auf die wegen ihrer mathematischen Komplexität hier nicht genauer eingegangen werden kann. S. 132 [4,5]

Auch bei dieser Operation treten zwangsläufig Ungenauigkeiten auf, die mit der Unschärfe der Fensterfunktionen zu tun haben.

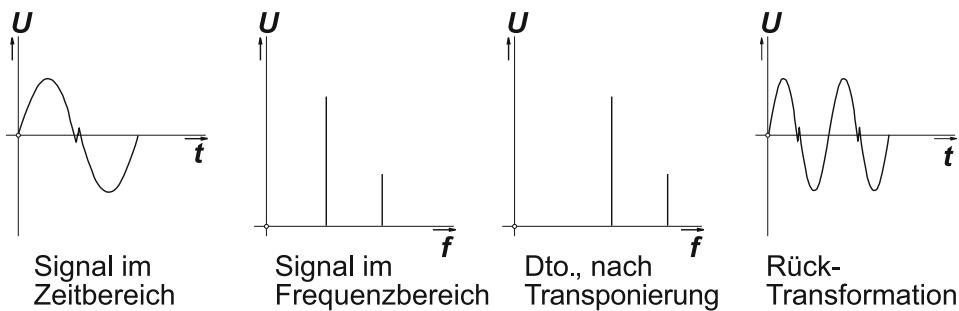


Abb. 3.24 Vorgänge bei der Tonlagen-Transponierung

3.3.7 Zeitdehnung

Eine Zeitdehnung soll die Tonhöhenlage belassen und trotzdem die Gesamtdauer der behandelten Audiodatei verändern. Dieser Vorgang kommt somit einer Kombination aus Tonlagen- und Geschwindigkeitsänderung gleich. Der umgekehrte Weg wäre bei Vorhandensein einer eigenständigen Zeitdehnungsmethode natürlich auch möglich: Die Tonlagenverschiebung wäre dann eine Kombination aus Zeitdehnung und Geschwindigkeitsänderung.

Bei manchen besonderen Formen des Signals ist auch ein Hinzufügen oder Entfernen von Perioden der Hauptschwingungen – falls eine solche existiert – oder ganzen Paketen davon, möglich. Das gelingt jedoch i. Allg. nur bei Signalen, deren Profil sich nicht laufend innerhalb weniger Millisekunden ändert. Leider ist natürlich im Normalfall eine solche Signalauswahl nicht möglich, daher scheidet dieses Verfahren zum Erreichen vernünftiger Ergebnisse aus.

Abgesehen von Sound-Effekten eröffnet eine Zeittrimmung bei Klangcollagen die Möglichkeit von exakten Einpassungen in zeitliche Freiräume. Darüber hinaus lassen sich unabhängige Stimmen oder Tracks auf anderes Material synchronisieren.

3.4 Pegelbearbeitung

3.4.1 Einstellung von Einzelbereichen

Ein wichtiger Bearbeitungsschritt ist das Anpassen von Einzelpegeln zueinander und auch das Ein- und Ausblenden innerhalb einer Spur. Die drei wichtigsten Parameter sind dabei Pegelveränderung, Form sowie Zeitintervall der Änderung. Abb. 3.25 verdeutlicht die Parameter.

Logarithmisches Aus- oder Einblenden entspricht eher einer linearen Blende im Sinne einer physiologischen Lautstärkeänderung. Der zeitliche Wirkungsbereich ist oftmals

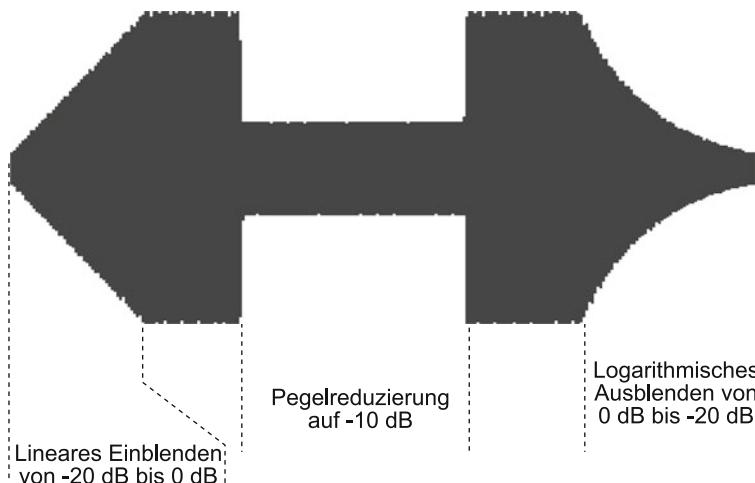


Abb. 3.25 Parameter bei Ein- und Ausblendungen

einfach visuell zu kennzeichnen, und die Pegelgrenzen sind entweder numerisch oder über Schiebesteller zu definieren.

3.4.2 Hüllkurvenkontrolle

Eine ausgesprochen komfortable Methode zur Manipulation von Pegeln (vor allem über einen längeren Zeitraum hinaus) ist die Hüllkurvenkontrolle. Grundsätzlich handelt es sich um die multiplikative Mischung zweier Signale, wobei das erste das zu steuernde Audiosignal und das zweite die Kurve zur Steuerung der Amplitude des ersten und somit dessen Hüllkurve ist. Ein Beispiel einer solchen Steuerung veranschaulicht Abb. 3.26.

Das Ursprungssignal zeigt hier konstante Amplitude, was natürlich ein Sonderfall ist, bei dem nämlich das manipulierte Signal dasselbe Amplitudenprofil aufweist wie die Steuerkurve selbst. Im allgemeinen Fall trifft dies meist nicht zu; das endgültige Signal folgt jedoch in Bezug auf die Amplitudenveränderungen entsprechend dem Steuersignal. Für letzteres ist auch eine Kurve denkbar, die sich im negativen Areal aufhält und deshalb die Phasenlage des gesteuerten Signals um 180° dreht.

3.4.3 Dynamikkompression

Eine automatische Veränderung der Amplitude eines endgültigen Stereosignals erfolgt bei der Dynamikkompression, welche bei der Aufzeichnung auf ein analog arbeitendes Magnetband oft zur Unterdrückung von Mediumrauschen angewandt wird. S. 132 [5–7]

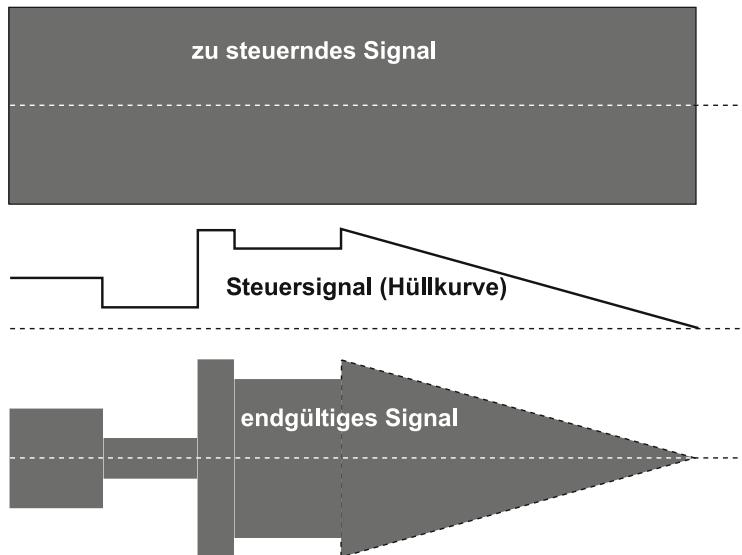


Abb. 3.26 Amplitudensteuerung eines Signals über eine Hüllkurve

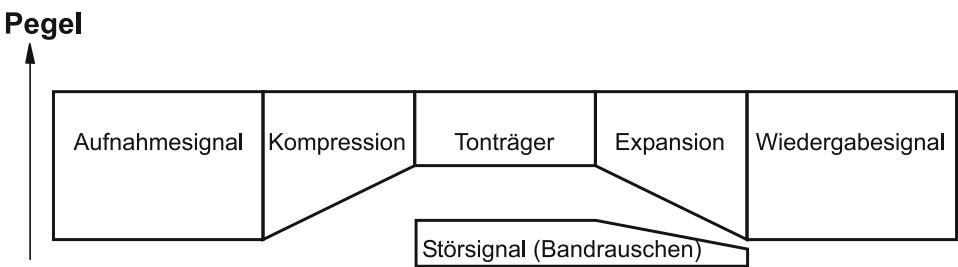


Abb. 3.27 Schema der Pegelbeeinflussung bei Comander-Systemen

Die Dynamikkompression erfolgt je nach System (z. B. Dolby A, Dolby B, Dolby C, HighCom, DBX usw.) auf unterschiedliche Weise, was beim Abspielen streng zu berücksichtigen ist. Dient die Bandmaschine als Standalone-Wiedergabegerät, so muss sie mit dem entsprechenden Expander für die Wiedergabe ausgerüstet sein. Für den Fall, dass das Band jedoch nur als Zwischenspeicher dient, um das Material ggf. wieder mit der Audiokarte auf die Festplatte zu spielen, kann sowohl für die Kompression als auch für die Expansion eine Bearbeitung auf Software-Ebene erfolgen. Doch auch hier müssen sich natürlich Kompressor und Expander komplementär verhalten, sonst entstehen möglicherweise drastische Signalveränderungen.

Da eine Realisierung von Comander-Systemen mittels Software nur selten benötigt wird und es zudem viele unterschiedliche Systeme gibt, wollen wir hier aus Platzgründen auf eine genaue Beschreibung verzichten. Zur Veranschaulichung der allgemeinen Funktionsweise diene Abb. 3.27.

Die Kompression bzw. Expansion erfolgt bei Dolby-Systemen getrennt für einzelne Frequenzbereiche, wobei die einfacheren Systeme (B und C) mit einem sog. Gleitband arbeiten, um auf die spektrale Gegebenheit dynamisch reagieren zu können. Dagegen setzen Breitband-Compander meist zusätzlich eine (feste) Emphasis ein, um bereits dadurch eine Erhöhung des Rauschabstands zu erreichen.

3.5 Unterdrückung von Knackgeräuschen

Bei der Verarbeitung von älteren oder beschädigten Tonaufzeichnungen können oftmals unangenehme Knack- und Knistergeräusche vorkommen, wie man sie auch vom Abspielen von Schallplatten her kennt. Bei der digitalen Tonbearbeitung gibt es eine Reihe von Möglichkeiten, mit denen eine Reduzierung der Störungen möglich ist. Betrachten wir zunächst Abb. 3.28, wo die Signallinie eines Geräusches als Nadelimpuls sichtbar wird, welcher auch mittels geeigneter Software stark reduziert wurde.

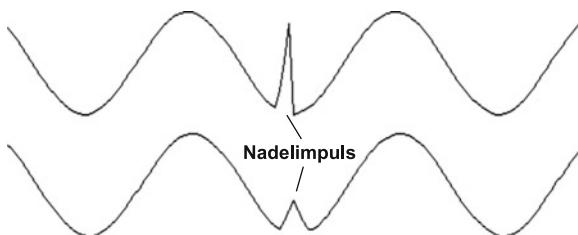
Es ist natürlich denkbar, dass die restliche Ausbuchtung ebenfalls beseitigt werden kann; dennoch ist eine solche Operation nicht ohne weiteres möglich, ohne das restliche Signal zu beeinflussen.

3.5.1 Slewrate-Filterung

Eine einfache Möglichkeit zur Unterbindung von Nadelimpulsen besteht in der Behandlung durch ein Slewrate-Filter. Peaks haben die eindeutige Eigenschaft, dass zwei darin enthaltene, benachbarte Abtastwerte eine große Differenz aufweisen, oder anders ausgedrückt, die Elongationsschnelle ist besonders groß. Andererseits zeigen jedoch die höheren Frequenzanteile des Nutzsignals normalerweise keine solch hohe Amplitude, so dass deren Schnelle nicht sehr groß ist.

Ein Filter, das lediglich eine maximale Differenz zwischen zwei Abtastwerten zulässt bzw. größere Differenzen auf diese Obergrenze beschränkt, vermag genau diese Eigenschaften von unerwünschten Peaks zu drosseln. Natürlich kann eine derartige Slewrate-Begrenzung nicht unterscheiden zwischen beabsichtigten und unbeabsichtigten Nadelimpulsen. Aber erfahrungsgemäß gilt die oben genannte Situation, dass nämlich hohe Frequenzen des Nutzsignals nur mit schwacher Amplitude vorkommen.

Abb. 3.28 Nadelimpuls und teilweise Unterdrückung durch Slewrate-Kontrolle



Die Werteberechnung erfolgt mit der vorgegebenen Slewrate s nach dem Algorithmus:

$$\Delta X = X_n - X_{n-1} \quad (\text{vorgegebener Sprungwert}) \quad (3.16)$$

$$X_n = \begin{cases} X_{n-1} + s & \text{für } \Delta X > s \quad (\text{Sprungbegrenzung nach oben}) \\ X_{n-1} - s & \text{für } \Delta X < s \quad (\text{Sprungbegrenzung nach unten}) \\ X_{n-1} + \Delta X & \text{für } \Delta X < |s| \quad (\text{keine Sprungbegrenzung}) \end{cases} \quad (3.17)$$

Der Wert für die Slewrate muss für diesen Zweck sorgfältig ausgewählt werden; bei zu großem Wert zeigt sich kaum eine Wirkung, und ein zu kleiner Wert vermag die Höhen zu bedämpfen.

3.5.2 Einzelwert-Bearbeitung

Handelt es sich lediglich um sehr vereinzelte Stellen, an denen Knackgeräusche entfernt werden sollen, kann man auch die manuelle Bearbeitung in Betracht ziehen. Die verschiedenen Bearbeitungsprogramme sind hier mit diversen Möglichkeiten ausgestattet. Die einfachste und vielleicht naheliegende Lösung liegt in der visuellen Bearbeitung: die Signallinie wird einfach wie gewünscht mit dem Zeichengerät (Maus o. ä.) nachgefahren. Auf diese Weise lassen sich Peaks wie im oben beschriebenen Beispiel einfach „glätten“. Eine weitere Möglichkeit besteht darin, die zu bewegenden Abtastpunkte zu markieren und sie dann z. B. mit Cursor-Steuertasten in gewünschter Weise zu bewegen.

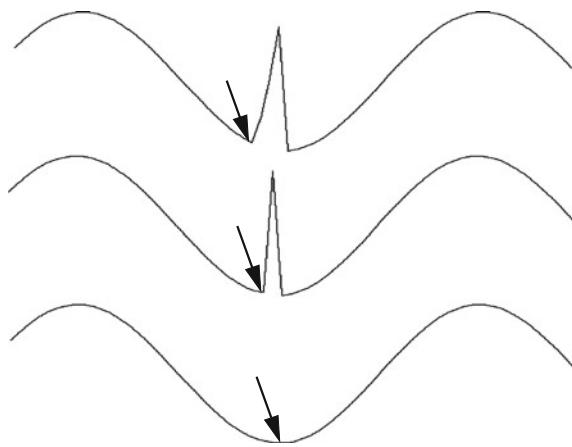
Beides erfordert natürlich die notwendige Vergrößerung (Zoom) auf den betreffenden Signalausschnitt, und dies wiederum kann erst nach dem Orten des Geräusches in Bezug auf den Zeitpunkt geschehen. Abb. 3.29 zeigt das vollständige Ausmerzen des Störimpulses mit der Zeichenfunktion.

Es ist einzusehen, dass dieses Verfahren nur für diskrete Stellen geeignet ist, weil der Zeitaufwand doch nicht unerheblich ist. Bei eindeutiger Identifikation allerdings kann mit einer fast vollständigen Entfernung der Störstelle gerechnet werden. Als Zwischenlösung von automatischer und manueller Entfernung von Peaks ist eine Software, die Sprungstellen bestimmten Grades innerhalb der Datei aufsuchen kann und sie dann mit einem gewissen Umgebungssignal auf dem Bildschirm zeigt, so dass sowohl eine visuelle als auch klangliche Beurteilung des Impulses möglich ist.

Im Anhang Computerprogramme ist der Pascal-Quellcode für ein Slewrate-Filter zu finden.

Eine solche Filterung lebt von der Tatsache, dass laute Störgeräusche eine sehr große Steigung des Signals hervorrufen, während normale Signale noch nicht in den Filterbereich fallen.

Abb. 3.29 Verschiedene Stufen der Entfernung eines Nadelimpulses mithilfe des Nachfahrens der Signallinie



3.6 Echtzeitbearbeitung per DSP

Obwohl heutige Prozessoren bereits so schnell sind, dass damit Algorithmen in Echtzeit möglich sind, gibt es doch höhere Anforderungen und komplexere Aufgaben, denen nur noch *Digitale Signal-Prozessoren* gewachsen sind.

Obwohl es für die DSP-Programmierung viele Einsteiger-Tools gibt, ist eine in die Tiefe gehende Implementierung von Programmen sehr komplex. Hier sei auf Spezialliteratur bzw. auf die Websites der DSP-Herstellerfirmen verwiesen.

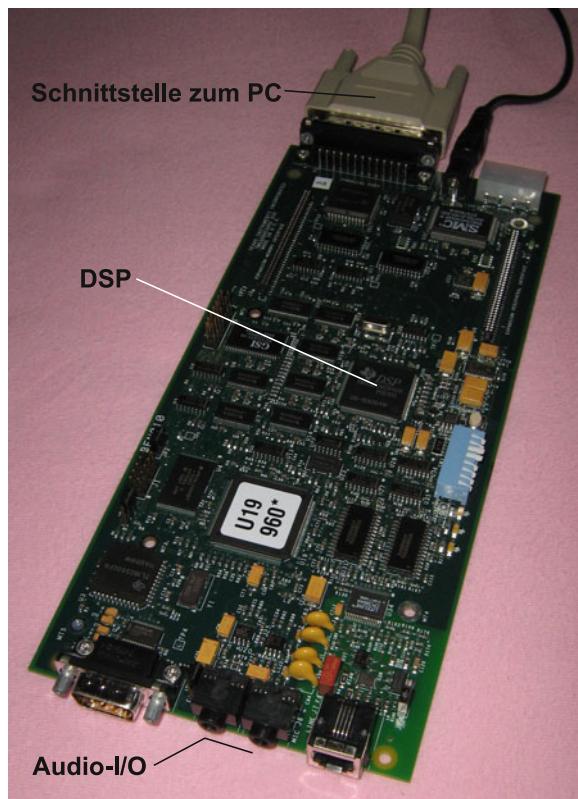
3.6.1 Grundlagen zum DSP

Der *Digitale Signalprozessor* (DSP) hat gegenüber herkömmlichen Prozessoren einen dahingehend geänderten Aufbau, als dass spezielle, immer wiederkehrende Operationen der Signaltechnik schon in ihrer Hardware verdrahtet sind. Das bedeutet im Einzelnen, dass sonst recht komplex aufzubauende Unterprogramme oder Funktionen durch eine einzige Befehlszeile *und* durch einen einzigen Maschinenzyklus erledigt werden (eine einzige Befehlszeile bedeutet im Allgemeinen nicht zwingend eine kurze Maschinencode-Struktur bzw. eine schnelle Abarbeitung).

DSPs besitzen i. Allg. Harvard-Architektur. Das heißt, Daten und Programmcode liegen auf getrennten Speichern und sind somit gleichzeitig in die Rechenwerke ladbar. Dies beschleunigt die Arbeitsgeschwindigkeit.

In Abschn. 3.2.1 ist das Prinzip eines FIR-Filters zu sehen. Auffällig ist hier die Verwendung der Operation in Gl. 3.1, die man auch als MAC-Operation (Multiply & Accumulate) bezeichnet. Da dieser Vorgang sehr häufig im Ablauf eines DSP-Programms vorkommt, ist es naheliegend, den Prozessor so aufzubauen, dass er mit wenigen Aktionen

Abb. 3.30 Kleines Entwicklungstool für den DSP TMS320 von Texas Instruments. Dazu gehören ein (meist abgespeckter) C-Compiler sowie einige Beispielprogramme als Software-Unterstützung



auskommt. Im Wesentlichen handelt es sich dabei um das Holen der Daten und der Instruktion, den Rest erledigt die Hardware.

Ein DSP-System zum externen Anschluss an einen PC ist in Abb. 3.30 zu sehen.

3.6.2 Bearbeitungs-Mechanismen

Um das Hauptziel *Echtzeitbearbeitung* zu erreichen, haben DSPs einen speziellen Aufbau. Dieser soll eine optimierte Abarbeitungsgeschwindigkeit für häufig wiederkehrende Operationen garantieren. Man kann eigentlich alle Operationen auch auf Software-Ebene mit jedem PC anwenden, wenn es jedoch auf Schnelligkeit ankommt und gleichzeitig eine hohe Datenrate vorliegt, wird ein PC-Prozessor schnell an seine Grenzen gelangen.

Der DSP eröffnet viele Anwendungsbereiche, wie z. B.:

- Echtzeit-FFT
 - Digitale Filter
 - Dynamik-Kompression

- Rauschunterdrückung und Störungsbeseitigung
- Echo, Hall und andere Effekte
- Audiokompression (MP3 usw.)
- Spracherkennung und -generierung
- Messen und Beurteilen von Signalen

3.7 Praxisbeispiel einer Audio-Bearbeitung

Wir wollen ein paar Tonbearbeitungsschritte anhand der Bearbeitungs-Software *Audacity*® durchführen. S. 132 [1]

Dieses Programm – als Freeware erhältlich – dient den Schritten, Aufzeichnung, Schnitt, Effektbearbeitung, Abmischen und Formatkonvertierung. Nach dem Start eröffnet sich ein Fenster wie in Abb. 3.31. Es ist nicht nur frei erhältlich und nutzbar, sondern es liegt auch der Quellcode vor (Open Source).

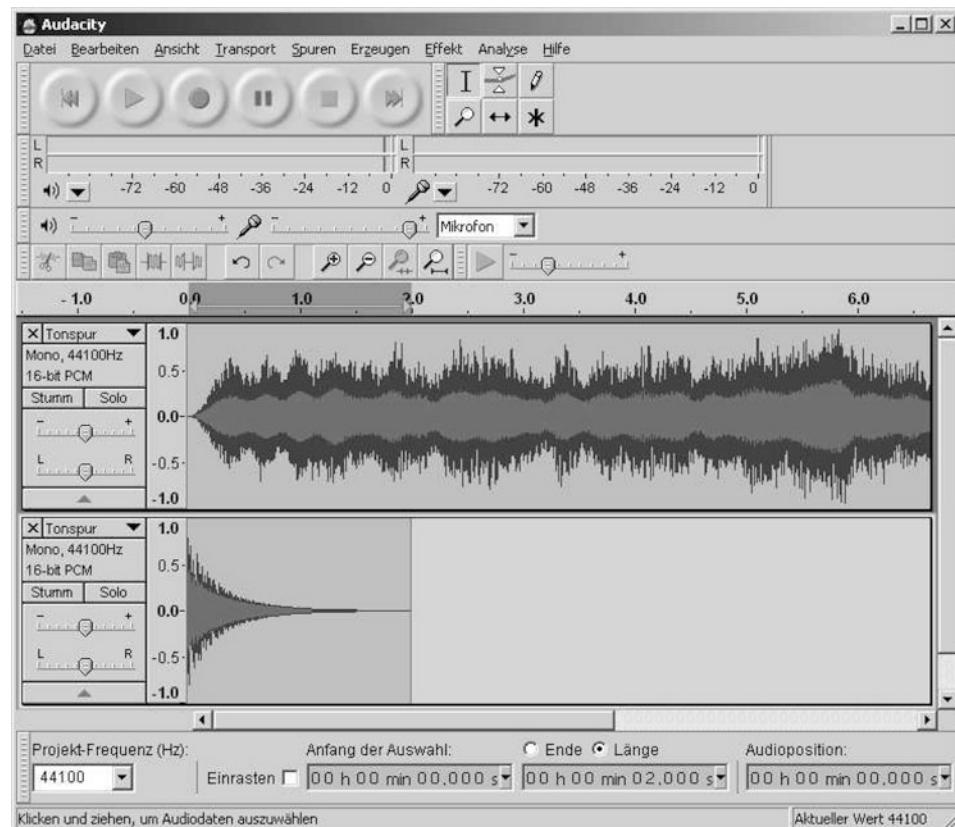


Abb. 3.31 Bearbeitungs-Software Audacity®(S. 132 [1])

3.7.1 Inversecho

Ein eindrucksvoller Effekt ist der des Inversechos. Es handelt sich dabei um vorauseilende Echos, bevor das Originalsignal erscheint. Anstelle von Echos kommt auch häufig der Halleffekt zum Einsatz. Vor allem im Zusammenhang mit Stimmen kommt ein merkwürdiger Atmungseffekt hinzu, der vor allem der Verfremdung dient. Wie aus dem Originalklang ein mit Inversecho behaftetes entsteht, das zeigt Abb. 3.32.

Analoge Studiotechnik war bereits lange zuvor in der Lage, eine derartige Effektbearbeitung zu leisten. Der Arbeitsaufwand war jedoch ungleich höher. Die notwendigen Schritte hierzu beinhalteten vor allem bei Bandmaschinen die Vollspurtechnik, die es erlaubte, Tonmaterial zeitlich rückwärts abspielen zu können. Die Zugabe von Echos war bei Mehrkopftechnik kein Problem, wobei der Zeitparameter – also die Echodistanz – nur bei speziellen Maschinen variabel ist (hier sind entweder Kopfabstand oder Bandgeschwindigkeit einstellbar).

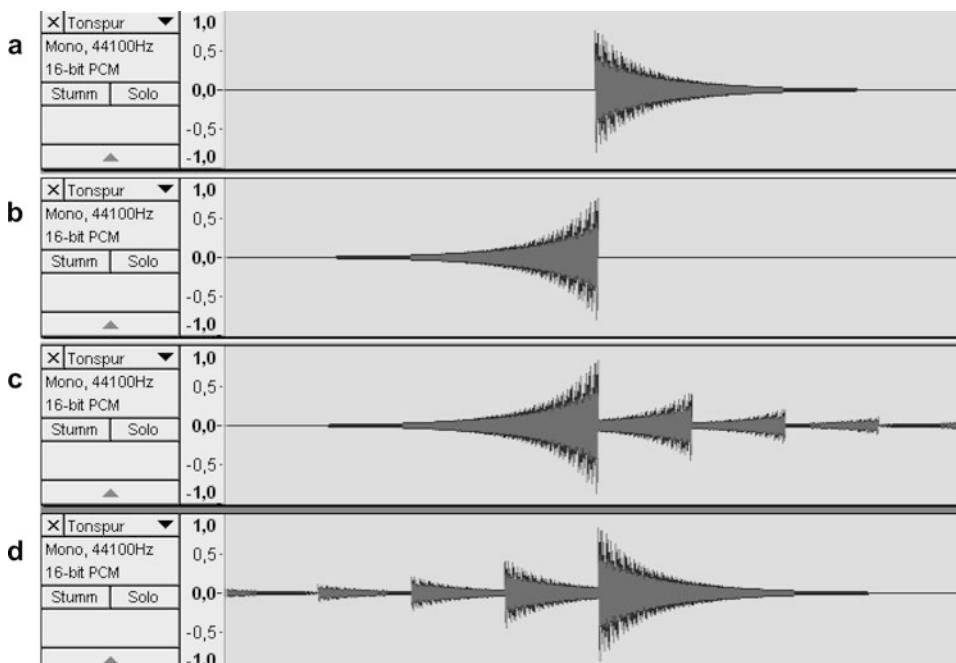


Abb. 3.32 Effektbearbeitung mit der Software Audacity® an einer Audiospur. Es wird ein sog. Inversecho hinzugefügt. Das Originalsignal (a) – ein Gongschlag – wird zeitlich gespiegelt (Effekt *Rückwärts*). Das entstandene Signal (b) erhält ein Echo (Effekt *Echo*), Signal (c), dann erfolgt aufs Neue eine Zeitumkehr (Effekt *Rückwärts*) und das endgültige Signal (d) hat deutliche vorauseilende Echos (S. 132 [1])

3.7.2 Erzeugung von Dolby Pro Logic

Im zweiten Beispiel soll eine Synthese von Pro Logic entstehen. Zunächst stehen die vier separaten Signale L, R, C und S zur Verfügung. Die letzteren beiden erfahren eine Pegeländerung bzw. einen 90° -Versatz per Hilbert-Transformation. Beides kann Audacity® problemlos bewältigen, so dass man am Ende die beiden modifizierten Signale C-3dB und die Seitensignale addieren kann zum endgültigen Produkt Lt und Rt. Die Hilbert-Transformation muss nur einmal erfolgen, denn von $+90^\circ$ auf -90° kommt man durch eine einfache Inversion. In Abb. 3.33 sind die einzelnen Vorgänge ersichtlich.

3.7.3 Hüllkurven-Bearbeitung

Das letzte Beispiel demonstriert, wie man mit Audacity® eigene Klänge generieren und danach über eine *Hüllkurve* manuell bearbeiten kann. Die Tonerzeugung beruht auf schmalbandiger Filterung von Weißem Rauschen. Die Töne, die der Klang beinhalten soll, werden alphanumerisch in MIDI-Manier eingegeben. Weitere Parameter sind Bandbreite zu den Grundtönen sowie Anzahl und Art der Obertöne. Man kann ähnlich wie bei einer Orgel spezielle Registraturn vornehmen. Demzufolge wird der resultierende Klang auch ähnlich dem einer Orgel sein. Der Effekt heißt hier *Harmonic Noise*.

Im zweiten Schritt kommt ein Werkzeug zum Einsatz, mit dem man Hüllkurven setzen und auch später noch nachbearbeiten kann. In Abb. 3.34 sehen wir oben das generierte

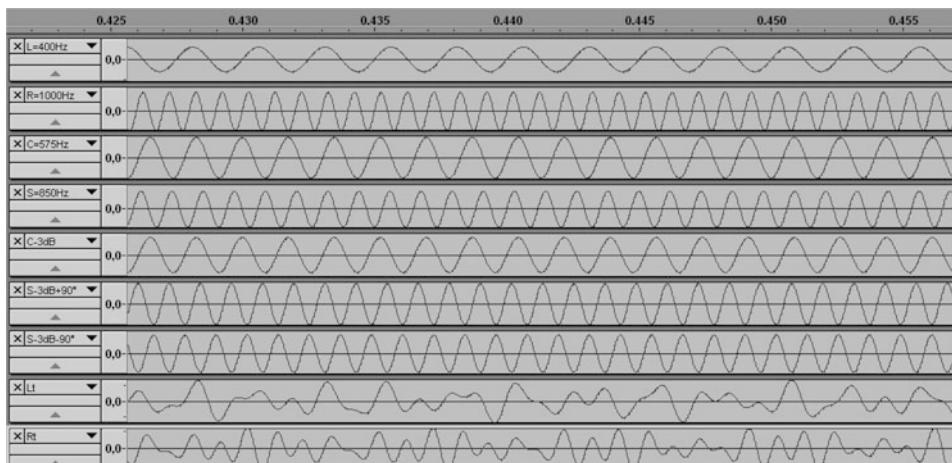


Abb. 3.33 Generierung von Pro Logic per Audacity® aus vier Einzelsignalen. Die oberen vier Spuren repräsentieren das Originalmaterial, die vier Spuren der Signale L, R, C und S. Die Namen der Signale sind im linken Bereich zu sehen. Die Signale C-3dB, S-3dB+90° sowie S-3dB-90° sind Zwischenprodukte, die nach dem Verfahren in Abb. 1.15 per Addition mit den Signalen L und R dann zum Endprodukt, den Signalen Lt und Rt, führen (S. 132 [1])

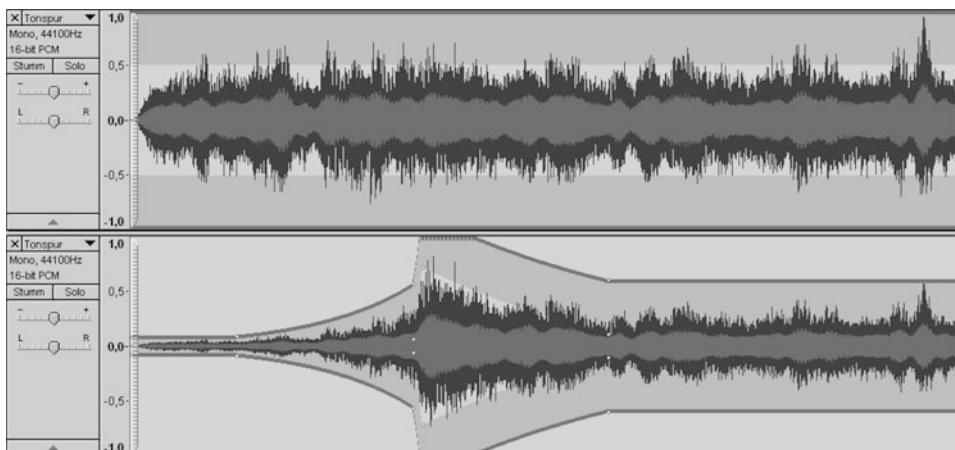


Abb. 3.34 Einfache und schnelle Hüllkurven-Bearbeitung mit Audacity®(S. 132 [1])

Signal und unten die Formung durch eine Hüllkurve. Selbstverständlich ist auch ein Ein-, Aus- und Überblenden auf halbautomatischem Wege möglich.

Literatur

1. Audacity®software is copyright (c) 1999–2011 Audacity Team. Internet: SourceForge.
2. Tietze, U., Schenk, C.: Halbleiter-Schaltungstechnik.
Berlin, Heidelberg: Springer-Verlag 1991.
3. Prochaska, E.: Digitale Signalprozessoren. Ehningen: Expert-Verlag 1988.
4. Brigham, E. O.: FFT, Schnelle Fourier-Transformation.
München: Oldenbourg-Verlag 1992.
5. Skritek, P.: Handbuch der Audio-Schaltungstechnik. München: Franzis-Verlag 1988.
6. Webers, J.: Tonstudientechnik. München: Franzis-Verlag 1989.
7. Stotz, D.: Audio- und Videogeräte richtig einmessen und justieren.
München: Franzis-Verlag 1994.



4.1 MIDI-Hardware

4.1.1 Anschlüsse bei MIDI-Instrumenten

Drei verschiedene Anschlüsse sind beim MIDI-Standard verfügbar: MIDI-Out, MIDI-In und MIDI-Thru. Nicht alle Instrumente oder Geräte, die sich für MIDI-Applikationen eignen, sind auch immer mit allen drei Anschlussarten ausgerüstet.

MIDI-Out Daten, die ein MIDI-Instrument oder ein Steuercomputer ausgibt, liegen an der Buchse MIDI-Out an. Nicht alle MIDI-Instrumente weisen diesen Anschluss auf, mitunter lassen sich solche MIDI-Instrumente auch nur steuern. Ein Steuercomputer allerdings verfügt grundsätzlich über MIDI-Out.

MIDI-In An den MIDI-In-Anschluss eines MIDI-Instruments gelangen die Daten eines steuernden Organs (also z. B. auch ein MIDI-Instrument oder ein Steuercomputer). Die verschlüsselten Daten gelangen vom MIDI-In-Anschluss zu einem Prozessor innerhalb des MIDI-Instruments und werden dort interpretiert und als Befehle und Daten bei der Klangerzeugung ausgeführt.

MIDI-Thru Um die Daten eines Gerätes, die aus seiner MIDI-Out-Buchse kommen, für mehr als nur ein MIDI-Instrument verwenden zu können, müssen letztere einen durchschleifenden Datenausgang besitzen. Ein solcher MIDI-Thru-Ausgang leitet also sozusagen nur die an MIDI-In anstehenden Daten weiter; allerdings muss der Anschluss gegenüber dem Eingang entkoppelt sein, d. h. er muss einen eigenen Treiberverstärker verwenden. Der Grund dafür liegt in der Tatsache, dass die MIDI-In-Eingänge eine Belastung darstellen, welche normalerweise nur jeweils einmal für jeden Ausgang auftreten darf.

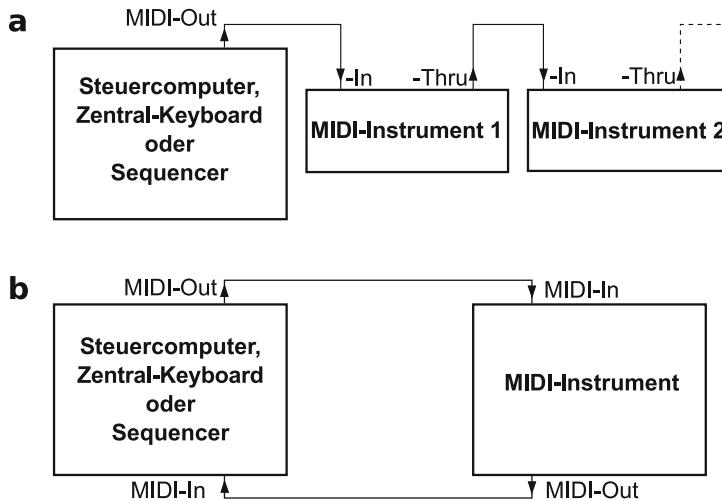


Abb. 4.1 MIDI-Verschaltungen: (a): Lineare Verschaltung, Geräte erhalten die Daten über die Thru-Buchse des vorangehenden Instruments oder Geräts. (b): Zirkuläre Verschaltung, bei der das gesteuerte Instrument oder Gerät auch seinerseits wieder Daten zurückgibt

4.1.2 Verschaltung von MIDI-Instrumenten und -Geräten

Je nach dem MIDI-Ablaufprogramm kommen unterschiedliche Verschaltungen von MIDI-Instrumenten bzw. -Geräten in Frage. Abb. 4.1 zeigt zwei prinzipiell unterschiedliche Verbindungsschemata.

Beide Arten der Verschaltungen können u. U. auch gemischt werden.

Bei der linearen Verkettung gibt es in der Praxis selbstverständlich eine Grenze in Bezug auf die Anzahl der verbundenen Geräte. Dieses Limit hängt mit der Laufzeit zusammen, die die Schnittstelle zwischen MIDI-In und MIDI-Thru erzeugt. Daraus resultierende Timing-Fehler addieren sich und ergeben eine hörbare Beeinträchtigung in der Steuerung, sobald eine gewisse Anzahl von in Serie geschalteten Geräten überschritten ist. Zur Verbesserung dieses Verhaltens ist es ratsam, MIDI-Geräte mit schnellen Optokopplern einzusetzen, was bei den heutigen Bauelementen keine Sorge mehr bereiten sollte. Anstiegs- und Abfallzeit sollten jeweils unter einem Wert von $2 \mu\text{s}$ liegen, was z. B. vom Typ 6N138 oder PC900 erfüllt wird.

4.1.3 Ausführung der Schnittstelle

Steuergrößen Es handelt sich bei MIDI um eine serielle Schnittstelle, die im Stromschleifen-Modus arbeitet. Diese Betriebsart ist nicht so störungsanfällig wie eine Steuerung mit Spannungssignalen (der Nachteil ist jedoch, dass die Ankopplung nicht

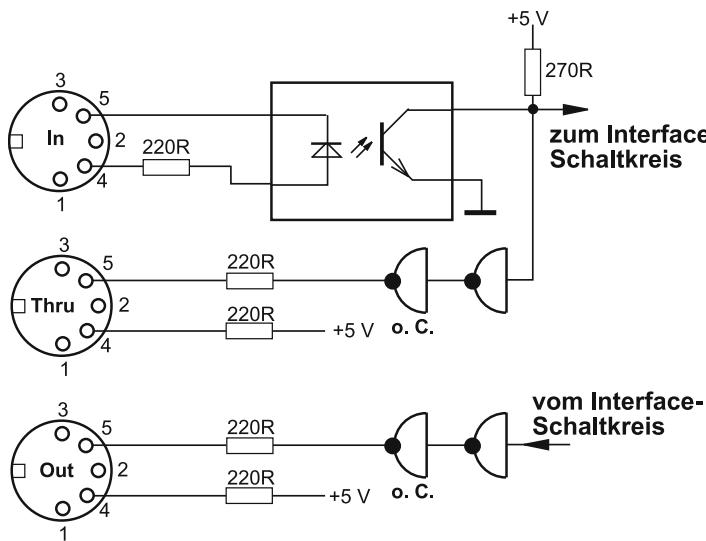


Abb. 4.2 Elektronische Ankopplung der Interface-Schaltung an die Buchsen. Die zu den Anschlüssen Thru und Out führenden Gatter sind normalerweise als Open-Collector-Device oder als Transistor ausgeführt – ein Strom kann somit nur fließen, wenn eine Verbindung mit der entsprechenden Plusleitung besteht; außerdem besteht so keine Gefahr eines Kurzschlusses, falls einmal zwei Ausgangsanschlüsse miteinander verbunden werden. Der Standard schreibt ebenfalls eine galvanische Trennung am MIDI-In-Anschluss vor, was mittels Optokoppler bewerkstelligt wird. Auf diese Weise sind Störungen, die aufgrund von Masseschleifen immer zu befürchten sind, vermeidbar

lastfrei erfolgen kann). Der Strom beträgt 5 mA, wobei Stromfluss eine logische 0 bedeutet (Abb. 4.2).

Datenprotokoll Die Datenrate beträgt 31,25 kBaud ($\pm 1\%$), der Datenfluss ist asynchron, und ein Zeichen benötigt ein Start-Bit, 8 Daten-Bits und ein Stop-Bit. Aufgrund der Datenrate besteht keinerlei Verwandtschaft mit der seriellen RS232-Schnittstelle; letztere lässt sich daher nicht für MIDI-Zwecke heranziehen. Zur Übertragung eines Zeichens sind daher für je 10 Bit $320\ \mu\text{s}$ notwendig. Das Start-Bit ist eine logische 0 (Strom ein) und das Stop-Bit eine logische 1 (Strom aus). Bei der Übertragung von Mehr-Byte-Daten erscheint zuerst das LSB und zuletzt das MSB (Datenformat nach Intel). Außerdem erfolgt die Übertragung eines Bytes ebenfalls mit dem niederwertigsten Bit und endet mit dem höchstwertigen. Abb. 4.3 verdeutlicht nochmals dieses Datenprotokoll anhand eines zu übertragenden Zeichens, Kommandos oder Wertes.

Verbindungen Als Anschlüsse dienen Steckverbindungen nach DIN 41524 (Diodenverbindung 5-polig, 180°), wobei die Geräte immer mit Buchsen ausgestattet sind, während die Leitungen jeweils an beiden Enden Stecker aufweisen. Siehe hierzu auch Abb. 4.5. Das positive Potential liegt jeweils an den Kontakten 4, und die Kontakte 1 und 3 sind

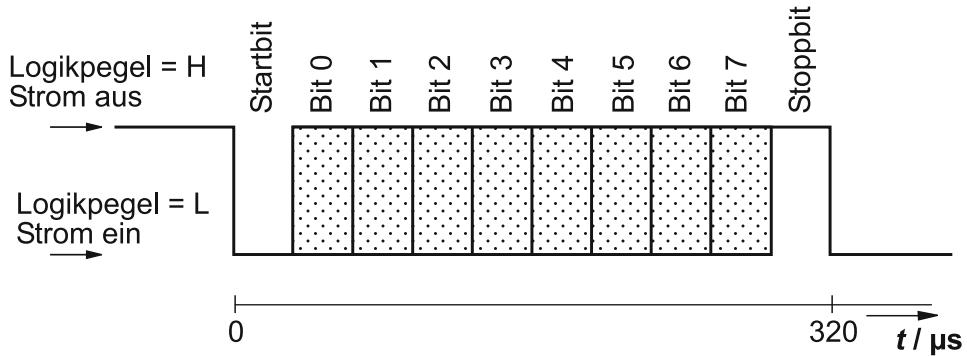


Abb. 4.3 Schema der Datenübertragung bei MIDI

unbelegt und dürfen auch nicht verwendet werden. Kontakt 2 hat Erdpotential und dient lediglich Abschirmzwecken; als Bezugspotential darf dieser Kontakt nicht zur Anwendung kommen und sollte auch nicht mit dem Masseanschluss verbunden sein (diese Forderung lässt sich jedoch in der Praxis nicht immer erfüllen). Einzelne Leitungen sollten eine Maximallänge von ca. 15 m nicht überschreiten, und beide Adern sind vorzugsweise verdrillt (mit gemeinsamer Abschirmung) zu wählen.

Audiokarten, die für MIDI-Applikationen geeignet sind, setzen häufig 9polige Sub-D-Buchsen ein, um alle drei MIDI-Anschlüsse darin zu vereinigen. Da jedoch diesbezüglich noch keine allgemeingültigen Festlegungen gelten, muss hier auf eine Wiedergabe von Anschluss schemata verzichtet werden. Adapterkabel splitten diesen Anschluss üblicherweise auf dem kürzesten Wege in die drei gewünschten Diodenkupplungen.

4.2 MIDI-Software-Standard

Zur Steuerung von MIDI-Instrumenten und -Geräten müssen bezüglich des Datenformats genaue Richtlinien befolgt werden. Es gibt in der Struktur der MIDI-Daten eine Reihe von unterschiedlichen Botschaftstypen¹, welche wir im folgenden etwas näher betrachten wollen.

4.2.1 Begriffserläuterungen

In den folgenden Abschnitten werden immer wieder die gleichen Begriffe auftauchen, die jedoch zum Teil zu Missverständnissen führen können. Leider haben sich bislang terminologisch keine Vereinheitlichungen ergeben, so dass hier keine Anpassung an eine bestehende Terminologie vorgenommen werden kann. Deshalb sollen an dieser Stelle zunächst einige Erklärungen abgegeben werden, die allerdings nur innerhalb dieses Werkes gelten.

Instrument In diesem Zusammenhang muss eine klare Trennung zwischen physischem und virtuellem Instrument stattfinden. Während ein physisches Instrument immer eine komplette Einheit darstellt und in sich abgeschlossen ist, kann sein Klang nicht immer eindeutig definiert werden. Betrachten wir z. B. elektronische Tasteninstrumente, welche sehr oft auf vielerlei Instrumentenkänge einstellbar sind. Ein solches physisches Instrument kann also mehrere virtuelle Instrumente sozusagen beinhalten, von denen wir jedoch definitiv den spezifischen Klang vorhersagen können (also z. B. Violine, Flöte, Gitarre usw.). Wenn in den folgenden Ausführungen der Ausdruck *Instrument* fällt, so ist dort immer ein virtuelles Instrument gemeint. Bei MIDI-Geräten oder -Instrumenten fällt in diesem Zusammenhang immer wieder der Ausdruck *multitimbral*, was lediglich ein werbewirksamer Ausdruck für die Tatsache ist, dass mehrere Instrumentenklangfarben gleichzeitig verfügbar sind, die jedoch auf einzelne Kanäle zu verteilen sind (siehe auch Programm).

MIDI-Instrument Einheiten, die einem physischen Instrument entsprechen, also durchaus verschiedene spezifische Klänge erzeugen und somit mehrere virtuelle Instrumente nachbilden können, nennen wir nachfolgend MIDI-Instrument. Voraussetzungen hierfür sind jedoch, dass eine Ansteuerung durch MIDI-Steuerimpulse möglich ist und dass eine Klaviatur oder ein Griffbrett zur manuellen Bedienung vorliegt. Keyboard-Synthesizer, die sowohl eine unabhängige Klangerzeugung als auch eine MIDI-Schnittstelle aufweisen, sind Beispiele für MIDI-Instrumente.

MIDI-Gerät Dies entspricht einem MIDI-Instrument mit der Einschränkung, dass keine Möglichkeit der manuellen Bedienung besteht, d. h. es existiert keine Klaviatur. In den nachfolgenden Ausführungen dieses Kapitels taucht häufiger der Begriff MIDI-Gerät häufiger auf als der Begriff MIDI-Instrument. Jedoch kann grundsätzlich davon ausgegangen werden, dass die jeweils angesprochenen Belange auch MIDI-Instrumente betreffen. Umgekehrt ist dies natürlich nicht immer der Fall, vor allem wenn es explizit um die Klaviatur oder das Griffbrett geht. Als Beispiel für ein MIDI-Gerät kann eine Audiokarte gelten.

Manuelles MIDI-Steuergerät Es handelt sich hierbei im Gegensatz zum MIDI-Gerät um eine Klaviatur oder ein Griffbrett, das lediglich MIDI-Steuerimpulse ausgeben kann. Zusammen mit einem MIDI-Gerät ergibt sich ein MIDI-Instrument. MIDI-Keyboards, die über keine interne Klangerzeugung verfügen bzw. sie in der entsprechenden Situation nicht benötigen, stellen Beispiele dar.

Automatisches MIDI-Steuergerät Die MIDI-Steuerimpulse entstehen hier nach dem Start eines Ablaufprogramms. Eine Klaviatur kann, muss aber zur Programmierung nicht vorhanden sein. Sequenzer sind ein Beispiel hierfür.

Stimme In einem MIDI-Gerät ist eine Klangerzeugung integriert, die es je nach Anzahl verfügbarer Stimmen erlaubt, auch simultan umzusetzende Noten wiederzugeben. Zu verwechseln ist der Ausdruck Stimme jedoch nicht mit Instrument, und eine mehrstimmige Abfolge von Noten bedeutet nicht hinreichend, dass auch mehrere (virtuelle) Instrumente daran beteiligt sind. Zur Veranschaulichung möge die Vorstellung dienen, an einer Gitarre würden alle sechs Saiten gleichzeitig angeschlagen. In diesem Falle handelt es sich um ein 6stimmiges Instrument. Beim Anschlag weniger Saiten wird dieser Stimmenumfang jedoch nicht vollständig ausgeschöpft.

Programm Einem einzelnen Programm ist ein spezifischer Klang zugeordnet, der von einem (virtuellen) Instrument stammt. Der MIDI-Standard lässt 128 verschiedene Programme zu, die nur noch über eine Nummerierung (der sog. Programm-Nummer) angesprochen werden. Die mit einer bestimmten Programm-Nummer verbundenen Klangfarbe wird auch oft mit Timbre bezeichnet. Während die Programm-Nummern im Sprachgebrauch einen Bereich von 1–128 abdecken, liegen die Datenwerte als 7-Bit-Format um 1 niedriger, reichen also von 0–127. Zu unterscheiden sind Programme, die einen Steuerungsprozess bewirken (also eine Aneinanderreihung von MIDI-Anweisungen und -Daten darstellt); diese werden Ablaufprogramm genannt.

Aftertouch Dieser feststehende englische Begriff steht für den Anschlagsdruck einer Taste nach dem ersten Betätigen. Dieser Parameter kann unterschiedliche Auswirkungen zeigen; im einfachsten Falle reguliert er die Lautstärke während dem Halten einer Taste. Aftertouch ist jedoch nicht nur Kriterium für physische Vorgänge, sondern stellt auch eine virtuelle Größe dar, die sich mittels MIDI-Datenübertragung verarbeiten lässt.

Velocity Dieser ebenfalls aus dem Englischen stammende Begriff ist eigentlich wörtlich übersetzt gleichbedeutend mit *Geschwindigkeit*. Gemeint ist hier jedoch die Anschlagsgeschwindigkeit einer Taste und die damit erreichte Lautstärke bzw. Anschlagsdynamik beim Anklingen einer Note. Ähnlich dem Aftertouch kann die Velocity Datenwerte im Bereich 0–127 annehmen und als MIDI-Information übermittelt werden.

Master und Slave Innerhalb eines MIDI-Netzes gibt es verschiedene Kategorien der Priorität. Ein Gerät, das als Master fungiert, gibt Daten an seiner MIDI-Out-Buchse aus und steuert damit Slaves. Ein Master ist zunächst einmal unabhängig und gibt nur Daten aus, ohne auf welche von außen zu reagieren (i. Allg. ist seine MIDI-In-Buchse nicht belegt). Ein Sequenzer übernimmt beispielsweise meist die Stelle des Master.

Notennummer Sie korrespondiert mit der Tonhöhe einer Note und spezifiziert somit die vertikale Position in der Notation. Die Note C4 (mittleres C) repräsentiert eine Notennummer von 60 innerhalb des erlaubten Bereichs von 0–127.

4.2.2 Überblick zu den MIDI-Messages

Jegliche Vorgänge, die über MIDI initiiert werden sollen, erfolgen durch die Übermittlung von einzelnen Bytes. Wie viele Bytes ein bestimmter Vorgang dabei benötigt, hängt von letzterem ab. Wie in der Assembler-Programmierung üblich, so gibt es auch bei MIDI ein den Vorgang klassifizierendes Byte, welchem ggf. spezifizierende Daten-Bytes folgen. Die klassifizierenden Bytes nennt man Status-Bytes. Zur eindeutigen Identifizierung ist bei einem Status-Byte immer das MSB auf 1 gesetzt, während Daten-Bytes ausnahmslos eine 0 beim MSB aufweisen und somit einen gültigen Wertebereich von maximal 0–127 nicht überschreiten dürfen. Abb. 4.4 verdeutlicht diesen Unterschied. Jegliche MIDI-Message lässt sich in eine Art Stammbaum einordnen, welcher die Untergliederung aller Messages darstellt (Abb. 4.5). Zu erkennen ist eine Aufspaltung in zwei große Gruppen, nämlich der Channel Messages und der System Messages.

4.2.3 Channel Message

Um gezielt einzelne Instrumente ansprechen zu können, sind unterschiedliche Kanäle adressierbar. Es ist aber zu beachten, dass ein MIDI-Instrument bezüglich des Datentransfers nicht immer auf einen Kanal beschränkt sein muss. Die Adressierung des Kanals erfolgt immer im unteren Nibble des Status-Byte, während die Daten im oberen Nibble die Anweisung selbst spezifizieren (Channel Voice erlaubt hierzu nicht alle Daten, siehe Abb. 4.6). Die Kanalnummer liegt im Bereich 1...16, während der Code für das untere Nibble natürlich nur Werte von 0–15 annehmen darf. Das höchstwertige Bit muss ja bei

Abb. 4.4 Zustände des MSB beim Status- und Daten-Byte sowie die gültigen Wertebereiche

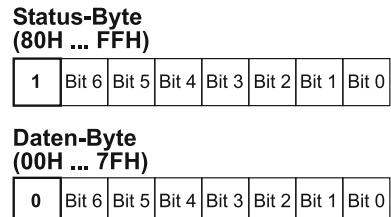


Abb. 4.5 Gliederung von MIDI-Messages

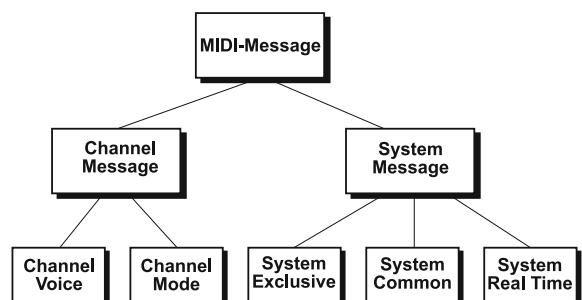


Abb. 4.6 Status-Byte für Channel Voice

oberes Nibble				unteres Nibble			
Bit 7	6	5	4	3	2	1	0
1	a	a	a	n	n	n	n
Anweisung				Kanal			

Tab. 4.1 Anweisungen zu Channel Voice (Status-Byte)

Hex	Binär	Anzahl der Daten-Bytes	Vorgang
8n	1000nnnn	2	Note aus
9n	1001nnnn	2	Note an
An	1010nnnn	2	Tastendruck nach dem Drücken der Taste (Aftertouch)
Bn	1011nnnn	2	Steuerungsänderung (Control Change)
Cn	1100nnnn	1	Programmänderung
Dn	1101nnnn	1	Tastendruck (Aftertouch) für gesamten Kanal
En	1110nnnn	2	Tonhöhenänderung (Pitch)

Status-Bytes immer 1 sein; die verbleibenden drei Bits dürfen die Konstellationen 000–110 (binär) annehmen, der Zustand 111 ist hier nicht erlaubt (dies würde einer System Message entsprechen).

Channel Voice Zur Steuerung der Stimmen eines Instruments dienen die Daten des Channel Voice. Das jeweilige Instrument empfängt die Daten über den Voice Channel. Tab. 4.1 gibt Aufschluss über die einzelnen Werte des Status-Byte.

Zu den oben genannten Anweisungen gehören unmittelbar folgende Daten-Bytes, die als solche durch eine führende Null im höchsten Bit erkennbar sind. Tab. 4.2 zeigt die Bedeutung der einzelnen Werte. Bei der Anweisung Note an wäre auch eine Anschlagsdynamik von 0 möglich; in diesem Fall entspricht dies der Anweisung Note aus.

Steuerungsänderungen. Diese beeinflussen gewisse Parameter einer Note, so z. B. Modulation, Sustain, Pan usw. Die Art der möglichen Änderung jedoch wird durch das Steuerungsgerät selbst definiert; so lässt sich z. B. bei einem Mischpult die Pan-Position eines Kanals einstellen, was ja die Stellung der Stimme im Stereoraum beeinflusst. Tab. 4.3 gibt eine Aufstellung der registrierten Kontrollfunktionen bzw. Steuerungen.

Dem Byte mit der Steuerungsnummer folgt mindestens ein Byte mit dem Steuerungswert. Einige Positionen in Tab. 4.3 stellen Schaltfunktionen dar, die also lediglich zwei Steuerungswerte einstellen können. Byte-Werte bis 3F werden dabei normalerweise als *aus* und Werte von 40–7F werden als *ein* interpretiert. Daneben gibt es Steuerungsarten,

Tab. 4.2 Werte zu Channel Voice (Daten-Bytes)

Vorgang	Daten-Byte	Hex	Beschreibung
Note aus	0kkkkkkk 0vvvvvv	0–7F 0–7F	k=Notennummer v=Ausklinggeschwindigkeit
Note an	0kkkkkkk 0vvvvvv	0–7F 0–7F	k=Notennummer v=Anschlagsdynamik
Tastendruck nach dem Drücken der Taste	0kkkkkkk 0vvvvvv	0–7F 0–7F	k=Notennummer v=Tastendruck
Steuerungsänderung	0uuuuuuu 0vvvvvv	0–78 0–7F	u=Steuerung (s. Tab. 4.3) v=Steuerungswert (s. Tab. 4.4)
Programmänderung	0vvvvvv	0–7F	v=Programm-Nummer – 1
Tastendruck (Aftertouch) für alle Kanäle	0vvvvvv	0–7F	v=Tastendruck
Tonhöhenänderung (Pitch)	0vvvvvv 0vvvvvv	0–7F 0–7F	v=Änderung (LSB, MSB)

Tab. 4.3 Steuerungsparameter zu Steuerungsänderung

Steuerung (Hex)	Zu ändernder Parameter	Steuerung (Hex)	Zu ändernder Parameter
00	Bank-Auswahl (-Select)	43	Soft-Pedal
01	Modulation	44	Legato-Fußschalter
02	Atem-Steuerung	45	Hold 2
03	undefiniert	46–4F	Sound-Steuerung 1–10
04	Fuß-Steuerung	50–53	allgemeine Steuerungszwecke
05	Portamento-Zeit	54	Portamento-Steuerung
06	Ankündigung für MSB	55–5A	undefiniert
07	Hauptvolumen	5B	Tiefe externer Effekte
08	Balance	5C	Tremolo-Tiefe
09	undefiniert	5D	Chorus-Tiefe
0A	Panorama (Pan)	5E	Celeste-Tiefe (Verstimmung)
0B	Expression (Ausdruck)	5F	Phaser-Tiefe
0C–0D	Effektsteuerung 1–2	60	Daten-Inkrement
0E–0F	undefiniert	61	Daten-Dekrement
10–13	allgemeine Steuerungszwecke	62	unregistrierter Parameter LSB
14–1F	undefiniert	63	unregistrierter Parameter MSB
20–3F	wie 00–1F, jedoch mit zwei folgenden Werte-Bytes	64 65	registrierter Parameter LSB registrierter Parameter MSB
40	Dämpfungspedal	66–77	undefiniert
41	Portamento	78–7F	reserviert für Channel Mode
42	Sostenuto		

welche kontinuierliche Steuerungswerte erfordern; manche davon ermöglichen sogar eine doppelte Auflösung und verlangen 2 folgende Daten-Bytes (bei den Steuerungen 20–3F). Das entspricht einer Auflösung von 14 Bit (je 7 Bit pro Byte), was einer Rasterung von 16384 entspricht. Während der niedrigste Steuerungswert (00) dem Fehlen des Effektes und der höchste Wert dem Effektmaximum entsprechen, so nehmen die drei Steuerungsparameter Balance, Panorama und Expression eine gewisse Sonderstellung ein, weshalb wir hier die Steuerungswerte anhand der Tab. 4.4 etwas näher erläutern wollen.

Die Steuerungsparameter 64H und 65H (siehe Tab. 4.3) kündigen das LSB bzw. MSB der registrierten Parameter an, die mit dem Pitch (Tonhöhenlage eines Instruments) zu tun haben. Tab. 4.5 zeigt einen Ausschnitt registrierter Parameter und ihren Bedeutungen.

Während der Pitch-Bereich lediglich die Einstellweite für eine individuelle Tonhöhenänderung definiert, basieren Feine Stimmung und Grobe Stimmung auf einer festen relative Frequenzänderung; außerdem erlauben sie ein unmittelbar darauffolgendes Stimmen, indem nur noch Transponierungs-Werte anzuhangen sind. Bei der Feinstimmung bestehen diese aus zwei Bytes, die eine Schrittzahl von 8191 nach oben und 8192 nach unten zulassen, um damit eine Verstimmung um $\pm 100\%$ erreichen zu können. Die Grobstimmung hingegen erfordert nur ein Byte und besteht somit nur aus 63 Stufen nach oben und 64 Stufen nach unten für eine Bewegung von wiederum $\pm 100\%$. Daraus lassen sich folgende Beispiele für MIDI-Daten ableiten:

Tab. 4.4 Steuerungswerte für spezielle Steuerfunktionen

Parameter	Steuerungswert (Hex)	Wirkung
Balance	00	volle Lautstärke links
	40	gleiche Lautstärke links
	7F	volle Lautstärke rechts
Panorama	00	Position ganz links
	40	Position Mitte
	7F	Position ganz rechts
Expression	00	kein Lautstärkenzuwachs
	40	halber Lautstärkenzuwachs
	7F	voller Lautstärkenzuwachs

Tab. 4.5 Registrierte Parameter bei der Steuerungsänderung

Parameter-Nr. LSB	Parameter-Nr. MSB	Registrierter Parameter	Auflösung der folgenden Daten-Bytes
00	00	Pitch-Bereich	1/128 Halbton
01	00	Feine Stimmung	1/8192 der Frequenz
02	00	Grobe Stimmung	1/64 der Frequenz (nur 1 Byte)
03	00	Programm-Stimmung	
04	00	Bank-Stimmung	

Tab. 4.6 Besondere Werte für die Tonhöhenänderung

LSB (Hex)	MSB (Hex)	Wirkung
00	00	maximale Pitch-Änderung nach unten
00	40	keine Änderung
7F	7F	maximale Pitch-Änderung nach oben
01	40	1/16384 des Pitch-Bereichs nach oben
7F	3F	1/16384 des Pitch-Bereichs nach unten

B3 64 00 65 00 40 01 → Pitch-Bereich auf 1,5 Halbtöne einstellen (Kanal 3) B2 64 01 65 00 51 40 → Feine Stimmung um 1 % nach oben (Kanal 2)

Programmänderungen. Ein Wechsel der Programm-Nummer bedeutet eine Umschaltung zum voreingestellten Klang eines Instruments. Damit kann also ein Instrumentenwechsel verbunden sein, oder aber es wird einfach eine andere Klangeinstellung gewählt.

General MIDI hat – wie wir später noch sehen werden – genau 128 verschiedene Instrumente definiert. Damit korrespondieren die Programm-Nummern 0–127. Darüber hinaus erlauben jedoch z. B. auch Audiokarten eine Klangänderung der 128 Instrumente. Ferner kann eine Aufteilung der Instrumente bestehen, so dass die Instrumente 0–31 von Audiokarte A, die Instrumente 32–63 von Karte B usw. bereitgestellt werden.

Als Channel-Voice-Information stellt die Programmänderung somit eine Zuordnung des MIDI-Kanals und des verwendeten Instruments dar. Sehr häufig fällt in diesem Kontext der englische Ausdruck *Patch*, was man jedoch nicht auf ein einzelnes Instrument beziehen sollte. Ein Patch steht vielmehr für eine bestimmte Anordnung der Instrumente und deren Bezeichnungen innerhalb des Programmbereichs. So könnte man z. B. die Anordnung *General MIDI* nennen, wenn das Patch entsprechend zusammengesetzt ist.

Tonhöhenänderung. Zur Stimmung eines Instruments kann nach dem entsprechenden Status-Byte (1110nnnn, siehe Tab. 4.1) die gewünschte Transponierung über zwei folgende Daten-Bytes angegeben werden. Es handelt sich dabei um einen Einstellbereich, der in der Steuerungsänderung in registrierten Parametern zuvor definiert wurde (siehe Tab. 4.5). Die Bewegung innerhalb dieses Bereichs erfolgt durch die Angabe der beiden Werte-Bytes für die Tonhöhenänderung. Tab. 4.6 gibt Aufschluss über die Wirkung bestimmter Werte.

Tastendruck (Aftertouch). Nach den Tab. 4.1 und 4.2 gibt es zwei spezifische Interpretationen für den Tastendruck. Einmal kann für jede einzelne Note ein separater Wert für den Tastendruck angegeben werden. Dies soll i. Allg. die Dynamik nach dem Anschlag (also beim Halten einer Taste) beim Spielen von Soli bestimmen. Was jedoch letzten Endes der Tastendruck beeinflusst, ist von den Einstellungen bzw. Gegebenheiten des MIDI-Gerätes abhängig (dies kann Lautstärke, Timbre, Vibrato usw. sein).

Erfolgt eine Definition des Tastendrucks für den gesamten Kanal, so wirkt sich dies auf alle gespielten Noten des entsprechenden Kanals aus. Somit ändert sich das Verhalten des gesamten Instruments, welches diesem Kanal zugeordnet wurde.

Channel Mode Das Status-Byte für eine Channel-Mode-Message zeigt einen Aufbau nach Abb. 4.7 und gleicht damit dem Status-Byte für Channel Voice bei einer Steuerungsänderung.

Um eine Anweisung für Channel Mode klar von Channel Voice (mit dem Status-Byte Bn) zu unterscheiden, liegt das folgende Daten-Byte im Bereich 79H–7FH, der ja für Channel Voice nicht erlaubt ist (siehe Tab. 4.2/Wert für n).

Tab. 4.7 zeigt eine Liste aller Channel Modes.

Da der Omni-Mode unabhängig vom Zustand Mono oder Poly (beide Modi schließen sich gegenseitig aus) geschaltet werden kann, gibt es für Daten empfangende MIDI-Geräte 4 verschiedene Modi, die aus Tab. 4.8 hervorgehen.

Mono Mode (Omni aus). Ein in dieser Betriebsart arbeitendes MIDI-Gerät interpretiert die ankommenden Informationen für seine Stimmen auf zugeordneten, separaten Kanälen. Es wird dabei jede Stimme einem Kanal zugewiesen. Besitzt das MIDI-Gerät M Stimmen, so gibt es eine Zuordnung für empfangene Daten laut Abb. 4.8.

Abb. 4.7 Status-Byte für Channel Mode (hinreichend definiert jedoch nur mit nachfolgendem Byte mit Werten 79H–7FH

oberes Nibble				unteres Nibble				
Bit	7	6	5	4	3	2	1	0
	1	0	1	1	n	n	n	n
Anweisung (fest)				Kanal				

Tab. 4.7 Channel Modes

1. Daten-Byte (Hex)	2. Daten-Byte (Hex)	Wirkung
78	00	alle Klänge aus
79	00	Reset-Initial für alle Steuerungen
7A	00	Lokale Steuerung aus
7A	7F	Lokale Steuerung ein
7B	00	alle Noten aus
7C	00	Omni Mode aus (alle Noten aus)
7D	00	Omni Mode ein (alle Noten aus)
7E	M	Mono Mode ein (alle Noten aus, M = Anzahl der Stimmen)
7E	00	Mono Mode ein (alle Noten aus, die Anzahl der Kanäle entspricht der Anzahl der Stimmen)
7F	00	Poly Mode ein (alle Noten aus)

Tab. 4.8 Betriebsarten bei Channel-Mode-Einstellungen

Poly	Mono	Omni	Modus
ein	aus	ein	1
aus	ein	ein	2
ein	aus	aus	3
aus	ein	aus	4

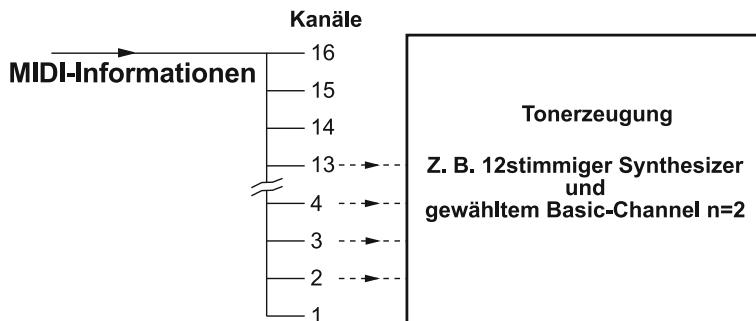


Abb. 4.8 MIDI-Gerät mit empfangenen Informationen im Mono-Modus mit Einstellung *Omni aus*

Jeder Kanal kann simultan nur eine Note am MIDI-Gerät steuern (ein-, ausschalten usw.). Der Bereich der benutzten Kanäle erstreckt sich von n bis $n + M - 1$. Die angegebene Kanalnummer bezeichnet in diesem Modus den sog. Basiskanal, über den exklusiv nur Modusänderungen vorgenommen werden können. Die angesprochenen Stimmen platzieren sich jedoch von einschließlich diesem bis maximal 16 oder bis der Stimmenvorrat des MIDI-Geräts erschöpft ist.

Der Mono-Modus ist vor allem dann empfehlenswert, wenn z. B. ein Gitarren-Synthesizer verwendet werden soll, bei dem ja die einzelnen Saiten jeweils monofone Klänge erzeugen soll, also bei einem Bundwechsel einer Saite der alte Ton unverzüglich zu verstummen. Andererseits dürfen alle Saiten für sich unabhängig klingen, was ja bei mehreren, auf die einzelnen Saiten verteilten Stimmen gelingt.

Mono Mode/Omni ein. Ein MIDI-Gerät, welches Daten in diesem Modus empfängt, reagiert auf alle Kanäle gleichermaßen. Die Steuersignale können jedoch nur eine einzige Stimme beeinflussen, alle weiteren Stimmen arbeiten nicht. Es wird hier also stillschweigend vorausgesetzt, dass $M = 1$ ist. Die Situation wird von Abb. 4.9 veranschaulicht. Welche Stimme des MIDI-Geräts für $M > 1$ tatsächlich angesprochen wird, hängt von ihm selbst ab.

Poly Mode. Soll ein MIDI-Gerät Informationen für mehrere Noten empfangen und korrekt verarbeiten, so muss es im Poly Mode arbeiten. Das Gerät kann verständlicherweise jedoch nur maximal M Noten zur gleichen Zeit spielen, weil eben jede Stimme für sich betrachtet nur monofon arbeitet. Befindet sich das Gerät im Omni-Modus (Modus 1),

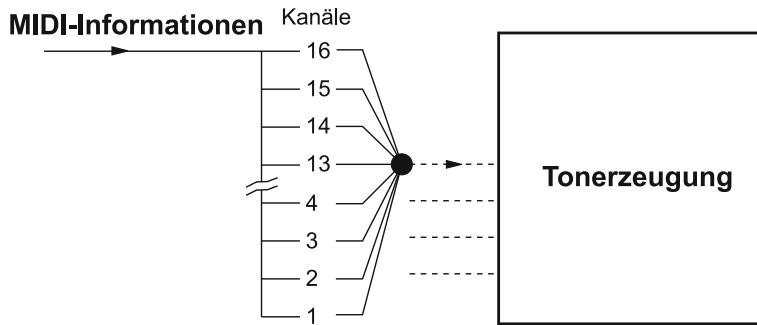


Abb. 4.9 MIDI-Gerät mit empfangenen Informationen im Mono-Modus mit Einstellung *Omni ein*

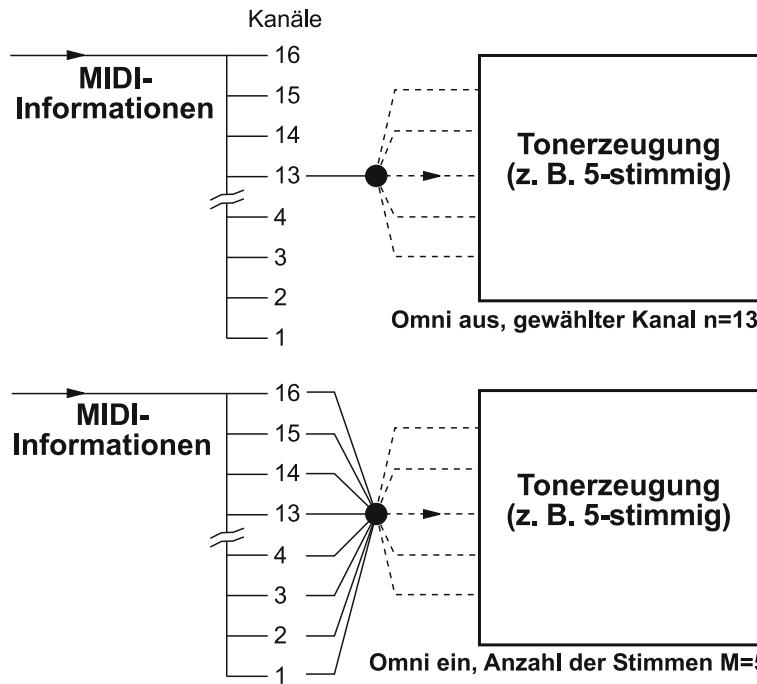


Abb. 4.10 MIDI-Gerät im Poly-Modus

so werden die Informationen aller Kanäle berücksichtigt, während bei ausgeschaltetem Omni-Modus (Modus 3) nur ein Kanal wirksam ist. Abb. 4.10 verdeutlicht das Schema der Schaltung im Poly-Modus.

Ein Aufsplitten nach dem Knoten ist nur erforderlich, wenn mehrere Noten zur gleichen Zeit initiiert werden (bzw. sich zeitlich überlappen), so dass das Gerät mehrere Stimmen aktivieren muss. Eine klare Unterscheidung erfolgt hier durch die Spezifikation des Voice Channel.

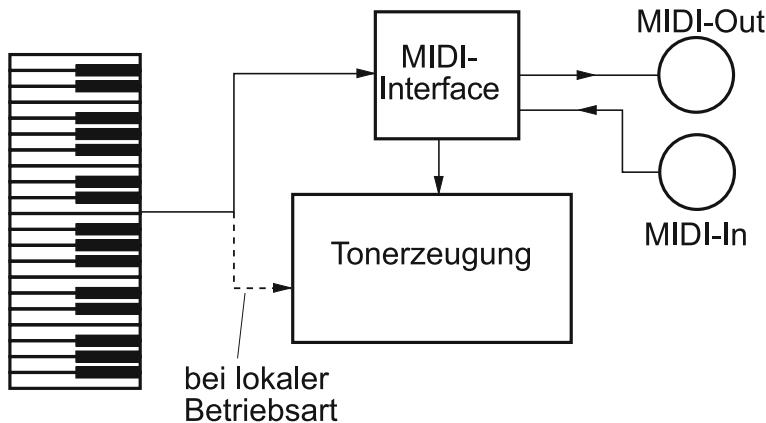


Abb. 4.11 Trennungsmöglichkeit des Keyboards von der internen Tonerzeugung

Lokale Steuerung. Soll ein MIDI-Gerät keine Daten von seinem eigenen Keyboard oder Griffbrett an die internen Generatoren senden, so muss die Lokalbetriebsart ausgeschaltet werden (siehe Tab. 4.7). Nach wie vor gelangen jedoch Daten an die MIDI-Out-Buchse und auch von der MIDI-In-Buchse in das Gerät, wo die Tonerzeuger eine entsprechende Umsetzung vornehmen (siehe Abb. 4.11).

Alle Noten aus. Diese Anweisung schaltet unverzüglich alle Noten des gewählten Kanals aus. Manche MIDI-Geräte schalten bei einer solchen Operation alle Stimmen aus; selbst eine lokale Steuerung über das eigene Keyboard hat dann keine Wirkung mehr.

Multimode. Viele MIDI-Geräte können Klänge unterschiedlicher virtueller Instrumente simultan ausgeben. Man spricht hier auch von multtimbraler Tonerzeugung, was auch von vielen modernen Audiokarten beherrscht wird. Die spezifische Ansteuerung muss dabei über einzelne Kanäle erfolgen.

4.2.4 System Exclusive

Nach den Channel-Vorgängen kommen wir nun zu den System-Operationen (siehe Abb. 4.5). Eine System-Exclusive-Anweisung wird von einem speziellen MIDI-Gerät des Netzes erkannt und verarbeitet, falls die zur Message gehörende ID-Nummer Übereinstimmung zeigt mit der Hersteller-ID des Geräts. Die entsprechenden Anweisungen sind also gerätebezogen und nicht kanalbezogen. Zweck solcher System-Exclusive-Messages ist es, Einstellungen am Gerät oder auch Sequenzer-Dumps zu übermitteln. Ferner können auch beliebige Instruktionen und Daten übertragen werden. Zum Beispiel ist es auch möglich, die Lautstärke oder das Panorama eines bestimmten Kanals vorzustellen. Tab. 4.9 zeigt den prinzipiellen Aufbau einer System Exclusive Message.

Tab. 4.9 Instruktionen-Folge von System Exclusive

Byte (Hex)	Byte (binär)	Bedeutung
F0	11110000	Beginn von System Exclusive
<80	0qqqqqqq	Identifikationsnummer des Herstellers
<80	0ddddddd	Daten (beliebige Anzahl von Bytes)
<80	Daten
<80	Daten
F7	11110111	Ende von System Exclusive

Alle Bytes zwischen dem Status-Byte F0 und dem Abschluss F7 müssen eine 0 als MSB aufweisen. Ist die ID-Nummer 0, so folgen zwei weitere Bytes als Erweiterungen. Einen Ausschnitt der derzeit festgelegten ID-Nummern zeigt Tab. 4.10.

Bei Mehr-Byte-IDs nimmt das zweite Byte nur diskrete Werte 00, 20 usw. an. Normalerweise ist es innerhalb einer Instruktion für System Exclusive nicht erlaubt, weitere Status-Bytes zu senden. Es gibt drei Byte-Werte, die eine Sonderrolle spielen: 7D (für nicht-kommerzielle Zwecke), 7E (für Non-Realtime) und 7F (für Realtime). Sie dienen Spezialzwecken und erweitern die Möglichkeiten für MIDI. Das entsprechende Status-Byte wird wie eine ID behandelt und folgt unmittelbar auf das den Beginn von System Exclusive anzeigenende Status-Byte F0. Die nachfolgenden Bytes spezifizieren weiter die folgende Meldung; jedoch sind dies i. Allg. noch keine Daten-Bytes. Man kann also die Byte-Folge bei Non-Realtime und Realtime wie folgt aufschlüsseln:

- F0 7E [Kanal] [Sub-ID 1] [Sub-ID 2] [Daten-Byte] ... F7 (Non-Realtime)
- F0 7F [Kanal] [Sub-ID 1] [Sub-ID 2] [Daten-Byte] ... F7 (Realtime)

System Exclusive/Non-Realtime. Solche Informationen werden außerhalb des eigentlichen Abspielens übermittelt. Es handelt sich in erster Linie um den Transfer von Sampler-Dumps oder speziellen Timecode-Daten. Tab. 4.11 zeigt die Werte der Sub-IDs für Non-Realtime.

Die letzten beiden Kennzeichen geben Aufschluss über die Übermittlungsqualität der letzten Datenpakete. Das NAK ist gleichzeitig auch eine Aufforderung an den Sender, das letzte Paket nochmals zu senden. Es werden zwar keine zweiten ID-Bytes benötigt, dafür jedoch die laufende Nummer des Pakets. Die anderen beiden Handshaking-Kennzeichen Warten und Abbruch dienen der vorläufigen bzw. endgültigen Aussetzung der Datenübertragung. Warten zeigt an, dass der Empfänger vorläufig keine weiteren Daten aufnehmen kann; ein ACK von diesem ausgehend setzt dann die Übertragung fort. Die setzbaren Timecode-Angaben ermöglichen eine spätere Synchronisation.

System Exclusive/Realtime Informationen über Timecode und andere Spezialfunktionen lassen sich auch als Echtzeitdaten übertragen. Tab. 4.12 zeigt einen kurzen Überblick hierzu. S. 187 [2, 3]

Tab. 4.10 Hersteller-ID-Nummern für System Exclusive

ID-Nummer (Hex)	Hersteller	ID-Nummer (Hex)	Hersteller
47	Akai	48	JVC
0A	AKG Acoustics	40	Kawai
00 00 0E	Alesis	42	Korg
11	Apple Computer	07	Kurzweil
20	Bontempi/Farfisa	06	Lexicon
41	Boss	04	Moog
44	Casio	10	Oberheim
1F	Clarity	41	Roland
13	Digidesign	01	Sequential
00 20 04	Dr. Böhm	26	Solton
30	Dynaccord	4C	Sony
18	Emu Systems	3A	Steinberg
14	Fairlight	4E	TEAC
08	Fender	03	Voyetra/Octave-Plateau
51	Fostex	0C	Waveframe Corp.
7E	General MIDI	3B	Wersi
24	Hohner	43	Yamaha

4.2.5 System Common

Sollen alle sich im System befindlichen MIDI-Geräte eine Anweisung erhalten, so darf diese weder geräte-, noch kanalbezogen sein. Es handelt sich um Informationen über Synchronisation, Liedauswahl und Stimmung. Tab. 4.13 listet alle System Common Messages auf.

Die ersten beiden Status-Bytes dienen der Synchronisation von Sequenzern oder auch Rhythmusgeräten. Eine Liedauswahl ist ebenfalls bei manchen MIDI-Instrumenten oder -Geräten möglich; diese ermöglichen meist per Tastenfeld eine Selektion, und über die MIDI-Schnittstelle kann auch ferngesteuert eine Liedauswahl erfolgen. Da dieselbe Nachricht zu allen Geräten des Netzes gelangt, muss jede Liednummer unter den Geräten auch tatsächlich einheitlichen Stücken zugeordnet sein, sonst ist eine simultane Liedauswahl auf allen Geräten wenig sinnvoll.

Das Stimmen ist nicht bei allen Geräten möglich und auch nicht bei allen nötig (digital arbeitende Geräte beziehen sich i. Allg. auf eine interne Quarzreferenz, die nur unmerklich von Nennfrequenzen abweicht). Auch das Endekennzeichen von System Exclusive zählt zu den System-Common-Messages, obwohl letztere kein solches Endekennzeichen benötigen.

Tab. 4.11 Identifikations-Bytes bei Non-Realtime-SysEx-Messages

ID 1 (Hex)	ID 2 (Hex)	Bedeutung
00	–	nicht verwendet
01	–	Sample-Dump-Header
02	–	Sample-Datenpaket
03	–	Sample-Dump-Anfrage
04	00	MIDI-Timecode, spezial
04	01	MIDI-Timecode, Punch-In-Punkte
04	02	MIDI-Timecode, Punch-Out-Punkte
04	03	MIDI-Timecode, Punch-In-Punkt löschen
04	04	MIDI-Timecode, Punch-Out-Punkt löschen
04	05	MIDI-Timecode, Startpunkt des Event
04	06	MIDI-Timecode, Stoppunkt des Event
04	07	MIDI-Timecode, Event-Startpunkte mit Zusatzinfo
04	08	MIDI-Timecode, Event-Stopppunkte mit Zusatzinfo
04	09	MIDI-Timecode, Event-Startpunkt löschen
04	0A	MIDI-Timecode, Event-Stopppunkt löschen
04	0B	MIDI-Timecode, Cue-Punkte
04	0C	MIDI-Timecode, Cue-Punkte mit Zusatzinfo
04	0D	MIDI-Timecode, Cue-Punkt löschen
04	0E	MIDI-Timecode, Event-Name im Zusatzinfo
05	01	Sample-Dump-Erweiterungen, Mehrfachschleifenpunkte
05	02	Sample-Dump-Erweiterungen, Anfrage für Schleifenpunkte
06	01	Allgemeininformation, Identitätsanfrage
06	02	Allgemeininformation, Identitätsantwort
07	01	File-Dump, Header
07	02	File-Dump, Datenpaket
07	03	File-Dump, Anforderung
08	00	MIDI-Tuning-Standard, Bulk-Dump-Anfrage
08	01	MIDI-Tuning-Standard, Bulk-Dump-Antwort
09	01	General MIDI, System ein
09	02	General MIDI, System aus
7B	–	Dateiende
7C	–	Warten
7D	–	Abbruch
7E	–	NAK (Kennzeichen für fehlerhaften Empfang eines Datenpakets)
7F	–	ACK (Kennzeichen für fehlerlosen Empfang eines Datenpakets)

Tab. 4.12 Identifikations-Bytes bei Realtime-SysEx-Messages

ID 1 (Hex)	ID 2 (Hex)	Bedeutung
00	–	nicht verwendet
01	nn	MIDI-Timecode
02	nn	MIDI Show Control
03	nn	Informationen über Notation
04	nn	Gerätesteuerung
05	nn	MIDI-Timecode-Cueing
06	nn	Befehle für MIDI-Maschinensteuerung
07	nn	Rückmeldungen der MIDI-Maschinensteuerung
08	nn	MIDI-Tuning-Standard

Tab. 4.13 System-Common-Messages

Status-Byte (Hex)	Daten-Byte (binär)	Bedeutung
F1	0qqqvvvv	MIDI-Timecode Viertelbild (q = Datenart, v = Wert)
F2	0bbbbbbb	Lied-Positionszeiger (LSB)
	0bbbbbbb	Lied-Positionszeiger (MSB)
F3	0zzzzzzz	Liedauswahl (z = Liednummer)
F4	–	undefiniert
F5	–	undefiniert
F6	–	Stimmen
F7	–	Ende von System Exclusive

4.2.6 System Realtime

Diese Messages sind ebenfalls nicht kanalbezogen und besorgen ein Aufrechterhalten der Synchronisation. Die Taktsignale werden hierbei 24 mal pro Viertelnote gesendet. System-Realtime-Informationen haben Vorrang vor allen anderen Daten, welche notfalls zurückgestellt werden müssen. Tab. 4.14 zeigt die möglichen Anweisungen.

Nicht alle MIDI-Geräte reagieren auf die oben genannten Realtime-Messages. In diesem Falle werden die entsprechenden Informationen ignoriert.

4.3 General MIDI

Neben dem üblichen MIDI-Standard bietet General MIDI einige ergänzende Festlegungen, die Kompatibilitätskriterien besser erfüllen sollen. Das Haupthandikap bestand bei den Synthesizern darin, dass Programm-Nummern nicht einheitlich mit einem Klang

Tab. 4.14 System-Realtime-Messages (Status-Bytes)

Byte (Hex)	Bezeichnung	Beschreibung
F8	Timing Clock	Regelmäßiges Taktsignal 24 mal pro Viertelnote
F9	–	undefiniert
FA	Start	Startinitial für einen Sequenzer oder Rhythmusgerät am Anfangspunkt eines Lieds
FB	Continue	dto., jedoch am letzten Haltepunkt
FC	Stopp	Stoppinitial (mit Verbleiben des Liedpositionszeigers)
FD	–	undefiniert
FE	Active Sensing	Erkennung, dass trotz längerer Pause (>300 ms) der Datentransfer ordnungsgemäß arbeitet
FF	System Reset	Rücksetzen aller Geräte in Ausgangsstatus (wie nach dem Einschalten)

eines virtuellen Instruments korrespondieren. Synthesizer A führt z. B. die Streicher auf Programm-Nummer n , während Synthesizer B die Streicher auf Nummer m platziert hat. S. 187 [4]

4.3.1 Festgelegte Programm-Nummern

Nach den Vereinbarungen von General MIDI werden allen Programm-Nummern von 1–128 definierte Instrumentenklänge zugeordnet, was aus Tab. 4.15 hervorgeht.

Eine solche Festlegung garantiert natürlich nicht, dass eine MIDI-Datei, die verschiedene Systeme steuern soll, auch tatsächlich immer identisch klingt. Die Klangerzeugung ist nach wie vor individuell in Bezug auf das Funktionsprinzip und somit auch in Bezug auf das Klangergebnis. Feste Verknüpfungen zwischen Programm-Nummern und virtuellen Instrumenten ermöglichen lediglich, dass die einzelnen Instrumente korrekt identifiziert werden können.

4.3.2 Festlegungen für Schlaginstrumente

Schlaginstrumente können als einziges virtuelles Instrument betrachtet werden; Schwierigkeiten bei der Notation begegnet man damit, dass mit den einzelnen Komponenten bestimmte Tonhöhen der Noten (= Key) korrespondieren. Beim Komponieren oder Spielen sind daher die einzelnen Trommeln usw. immer an den gleichen Stellen der Notation bzw. des Keyboards zu finden. Tab. 4.16 zeigt diese Notenzuordnungen.

Notennummern erstrecken sich allgemein von 0–127, wobei das mittlere C (C4) die Nummer 60 erhält und hier dem hohen Bongo entspricht (siehe auch Tab. 4.2). Alle Notennummern außerhalb der oben genannten Reihe sind reserviert für weitere Festlegungen.

Tab. 4.15 Programm-Nummern nach General MIDI (Originalbezeichnungen)

Nr.	Instrument	Nr.	Instrument	Nr.	Instrument
1	Acoustic Grand Piano	44	Contrabass	87	Lead 7 (fifths)
2	Bright Acoustic Piano	45	Tremolo Strings	88	Lead 8 (bass + lead)
3	Electric Grand Piano	46	Pizzicato Strings	89	Pad 1 (new age)
4	Honky-tonk Piano	47	Orchestral Harp	90	Pad 2 (warm)
5	Electric Piano 1	48	Timpani	91	Pad 3 (polysynth)
6	Electric Piano 2	49	String Ensemble 1	92	Pad 4 (choir)
7	Harpsichord	50	String Ensemble 2	93	Pad 5 (bowed)
8	Clavinet	51	Synth Strings 1	94	Pad 6 (metallic)
9	Celeste	52	Synth Strings 2	95	Pad 7 (halo)
10	Glockenspiel	53	Choir Aahs	96	Pad 8 (sweep)
11	Music Box	54	Voice Ooohs	97	FX 1 (rain)
12	Vibraphone	55	Synth Voice	98	FX 2 (soundtrack)
13	Marimba	56	Orchestra Hit	99	FX 3 (crystal)
14	Xylophone	57	Trumpet	100	FX 4 (atmosphere)
15	Tubular Bells	58	Trombone	101	FX 5 (brightness)
16	Dulcimer	59	Tuba	102	FX 6 (goblins)
17	Drawbar Organ	60	Muted Trumpet	103	FX 7 (echoes)
18	Percussive Organ	61	French Horn	104	FX 8 (sci-fi)
19	Rock Organ	62	Brass Section	105	Sitar
20	Church Organ	63	Synth Brass 1	106	Banjo
21	Reed Organ	64	Synth Brass 2	107	Shamisen
22	Accordion	65	Soprano Sax	108	Koto
23	Harmonica	66	Alto Sax	109	Kalimba
24	Tango Accordion	67	Tenor Sax	110	Bag Pipe
25	Acoustic Guitar (nylon)	68	Baritone Sax	111	Fiddle
26	Acoustic Guitar (steel)	69	Oboe	112	Shanai
27	Electric Guitar (jazz)	70	English Horn	113	Tinkle Bell
28	Electric Guitar (clean)	71	Bassoon	114	Agogo
29	Electric Guitar (muted)	72	Clarinet	115	Steel Drums
30	Overdriven Guitar	73	Piccolo	116	Wood Block
31	Distortion Guitar	74	Flute	117	Taiko Drum
32	Guitar Harmonics	75	Recorder	118	Melodic Tom
33	Acoustic Bass	76	Pan Flute	119	Synth Drum
34	Electric Bass (finger)	77	Blown Bottle	120	Reverse Cymbal
35	Electric Bass (pick)	78	Shakuhachi	121	Guitar Fret Noise
36	Fretless Bass	79	Whistle	122	Breath Noise
37	Slap Bass 1	80	Ocarina	123	Seashore
38	Slap Bass 2	81	Lead 1 (square)	124	Bird Tweet
39	Synth Bass 1	82	Lead 2 (sawtooth)	125	Telephone Ring
40	Synth Bass 2	83	Lead 3 (calliope)	126	Helicopter
41	Violin	84	Lead 4 (chiff)	127	Applause
42	Viola	85	Lead 5 (Charang)	128	Gunshot
43	Cello	86	Lead 6 (voice)		

Tab. 4.16 Schlaginstrumente und deren Notennummern

Noten-Nr.	Schlaginstrument	Noten-Nr.	Schlaginstrument
35	Basstrommel (akustisch)	59	Ride-Becken 2
36	Basstrommel 1	60	Bongo (hoch)
37	Side Stick	61	Bongo (tief)
38	Snare-Trommel (akustisch)	62	Konga (hoch, gedämpft)
39	Händeklatschen	63	Konga (hoch, offen)
40	Snare-Trommel (elektrisch)	64	Konga (tief)
41	Stand-Tom-Tom (tief)	65	Timbale (hoch)
42	Hi-Hat (geschlossen)	66	Timbale (tief)
43	Stand-Tom-Tom (hoch)	67	Agogo (hoch)
44	Pedal-Hi-Hat	68	Agogo (tief)
45	Hänge-Tom-Tom (tief)	69	Cabasa
46	Hi-Hat (offen)	70	Maracas
47	Tom-Tom (tief-mittel)	71	Pfiff (kurz)
48	Tom-Tom (mittel-hoch)	72	Pfiff (lang)
49	Crash-Becken 1	73	Guiro (kurz)
50	Hänge-Tom-Tom (hoch)	74	Guiro (lang)
51	Ride-Becken 1	75	Claves
52	China-Becken	76	Klangholz (hoch)
53	Ride-Glocke	77	Klangholz (tief)
54	Tambourin	78	Cuica (gedämpft)
55	Splash-Becken	79	Cuica (offen)
56	Kuhglocke	80	Triangel (gedämpft)
57	Crash-Becken 2	81	Triangel (offen)
58	Vibraslap		

Diese Key-Map kann jedoch z. B. innerhalb des Windows-Betriebssystems durch eine eigene Map modifiziert werden, wie wir weiter unten sehen werden.

4.3.3 Weitere Festlegungen

General-MIDI-Geräte sollen bestimmte Leistungsmerkmale erfüllen in Bezug auf Stimmen-, Instrumentenanzahl usw.

Stimmenanzahl Für melodische und rhythmische Klänge sollen mindestens 24 unabhängige Stimmen simultan verfügbar sein. Mindestens aber müssen 16 hiervon für Melodie-Instrumente und 8 für Schlaginstrumente zuteilbar sein.

Das bedeutet konkret, dass bei einer zeitlichen Überlappung von 16 aufgerufenen Noten alle wiedergegeben werden müssen. Ferner gilt dies auch für 8 weitere Noten im Perkussionskanal.

Kanäle Es werden 16 MIDI-Kanäle unterstützt, wobei jeder Kanal mehrere Stimmen polyphon (zeitlich unabhängig bei verschiedenen Notennummern) ansteuern kann. Jedem Kanal können unterschiedliche Instrumente zugeteilt werden, wobei jedoch Kanal 10 immer für Perkussionsinstrumente vorbehalten ist.

Somit beherrschen General-MIDI-Geräte einen erweiterten Poly-Modus, den sog. Multimode. Dieser setzt voraus, dass das Gerät mehrere Instrumentenklänge (Timbres) gleichzeitig erklingen lassen kann. Die Instrumente sind beliebig auf die 16 MIDI-Kanäle verteilt; lediglich die Perkussionsklänge sind auf Kanal 10 fixiert. Das Key-Mapping erfolgt nach Tab. 4.16.

Instrumente 128 Programm-Nummern spezifizieren die verfügbaren Instrumentenklänge nach Tab. 4.15. Dabei sind mindestens 16 unterschiedliche Instrumentenklänge (Timbres) gleichzeitig spielbar.

Sonstige Ausstattung Ein General-MIDI-Gerät verfügt über einen Regler für die Gesamtlautstärke (Master Volume). Anschlüsse: Audio- und Kopfhörerausgang sowie MIDI-In-Anschluss (andere Anschlüsse optional).

Symbole Zur Kennzeichnung erhalten Systeme ein Symbol, welches die Einhaltung der oben genannten Spezifikationen gewährleistet. Es ist als Aufkleber am MIDI-Gerät bzw. -Instrument zu finden.

Solche Symbole bedeuten aber keinesfalls immer zwangsläufig, dass die MIDI-Geräte oder -Instrumente auf den Standard nach General MIDI beschränkt sein müssen – sie lassen sich lediglich in einen Modus bringen, der den entsprechenden Auflagen genügt.

4.3.4 Richtlinien für General MIDI Level 2

Die bisher genannten Spezifikationen und Richtlinien genügen dem sog. Level 1; eine Weiterentwicklung dieses Standards wird für die Zukunft erwartet. Dabei kann jedoch davon ausgegangen werden, dass die bisherigen Richtlinien nicht ungültig werden, sondern lediglich eine Ergänzung erfahren.

4.3.5 Das GS-Format

Zur Erweiterung des GM-Standards hat die Fa. Roland ein Format geschaffen, das den Klangeigenschaften von MIDI-Geräten einen größeren Raum bietet. Dadurch soll auch

erreicht werden, dass Programm-Nummern nicht nur formal mit definierten virtuellen Instrumenten korrespondieren, sondern dass auch der Klang dieser Instrumente standardisiert bzw. vereinheitlicht werden. Die Abwärtskompatibilität zum GM-Standard soll dabei uneingeschränkt gewährleistet sein. Als Erweiterung lassen sich einige Festlegungen nennen, die im Folgenden erwähnt sind.

Polyphonie und multitimbrale Eigenschaften Beim GS-Format ist eine Mindestanzahl von 24 polyphonen Stimmen vorgeschrieben, d. h. auf einem MIDI-Kanal können mindestens 24 unterschiedliche Töne eines virtuellen Instruments simultan erklingen. Außerdem sind für den Multimode 16 virtuelle Instrumente vorgesehen, d. h. alle 16 MIDI-Kanäle sind beim Ausschöpfen dieser Betriebsart durch die virtuellen Instrumente belegt. Es erfolgt eine dynamische Zuordnung der Stimmen in den Kanälen, was bei einer Überlastung eines Kanals zu einem Ausweichen auf einen anderen Kanal bedeutet. Der Kanal für die Schlaginstrumente erfährt dabei eine bevorzugte Behandlung.

Klangbeeinflussung über Steuerungsänderungen Das GS-Datenformat sieht spezielle Möglichkeiten zur Veränderung des Klanges eines virtuellen Instruments vor. Nach dem einleitenden Präfix Bn für die Steuerungsänderung (siehe Tab. 4.1) folgt die Steuerungsart (siehe Tab. 4.2 und 4.3), nämlich die unregistrierten Parameter 62H (LSB) bzw. 63H (MSB).

Bank- und Programmwechsel Der Klang einer Stimme wird definiert über zwei Vorgaben, nämlich über die Bank oder des Programms. Während letzteres die Auswahl des virtuellen Instruments trifft, beinhalten unterschiedliche Banks Variationen der Klangdetails. Dabei ist für Bank 1 ($uu = 0$) die Ordnung der Programme nach den Festlegungen der Gremien vorgeschrieben.

Die Datenfolgen sind im einzelnen:

Variation der Bank-Nummer : $Bn\ 00\ uu$ bzw. $Bn\ 20\ uu$

Variation der Programm-Nummer: $Cn\ vv$

n MIDI-Kanal (0–F), uu Bank-Nummer (00–3F), vv Programm-Nummer (00–7F).

Während ja auch nach GM 128 Programme (00–7F) vorhanden sind, hat man lediglich 64 Banks vorgesehen. Somit stehen theoretisch immerhin 8192 unterschiedliche Klangfarben zur Verfügung, wenn man die Zusammenfassung in virtuelle Instrumente nicht vornehmen will. In Wirklichkeit sind jedoch meist bei weitem nicht alle der 64 möglichen Banks realisiert.

4.4 MIDI-Mapping

In älteren Betriebssystemen z. B. bis Windows 3.1 gab es ein separates Systemprogramm – den sog. MIDI-Mapper –, welches strukturelle Einstellungen gemäß Abb. 4.12 vornehmen konnte. Das MIDI-Mapping per Systemeinstellung ist in moderneren Betriebssystemen eigentlich verschwunden. Die verschiedenen Programme, die MIDI nutzen, nehmen individuelle Einstellungen vor.

Ziel solcher Setups ist die Anpassung verschiedener Klangerzeuger und Audiokarten aufeinander und auf besondere Gegebenheiten sowie die Möglichkeit, sowohl Kanal- und Programm- als auch Tonhöhenzuordnungen vorzunehmen. Dadurch sind also gezielte Abweichungen vom General-MIDI-Standard denkbar.

4.4.1 MIDI-Setup

Obwohl wie oben beschrieben keine Systemstruktur für die Gerätezuordnung mehr besteht, betrachten wir im Folgenden eine für die Applikation *Cubase LE* spezifische Einstellmöglichkeit (siehe auch Abschn. 4.9 am Ende des Kapitels). Allgemein besteht nach Abb. 4.13 die Option, für jede MIDI-Spur den Ausgabekanal (*chn*) für das Wiedergabegerät zu definieren. Steht hier z. B. eine 2, so muss der Klangerzeuger (auch *MIDI-Expander* genannt) auch auf Kanal 2 empfangen, sonst bleibt er stumm. Wählbar ist

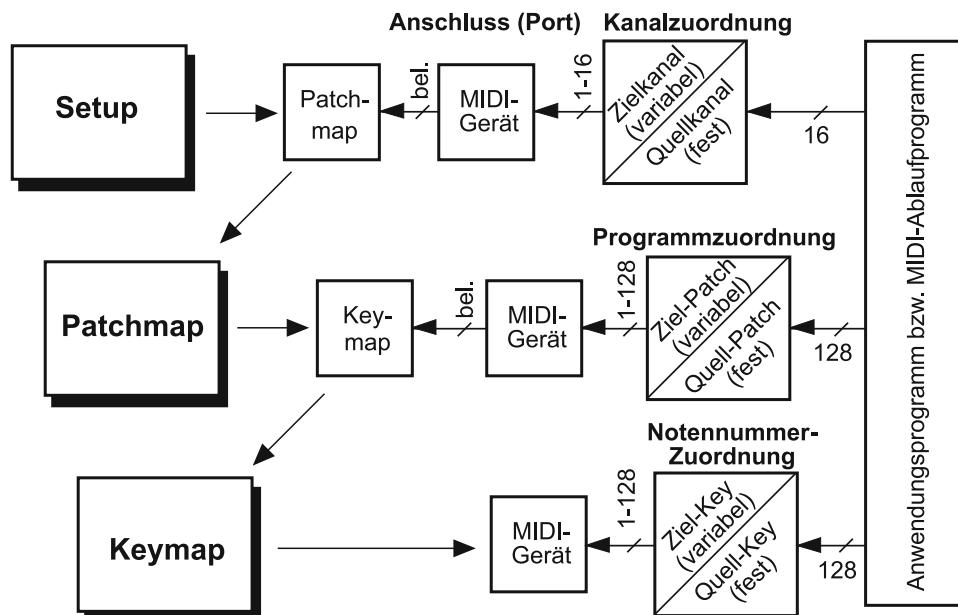
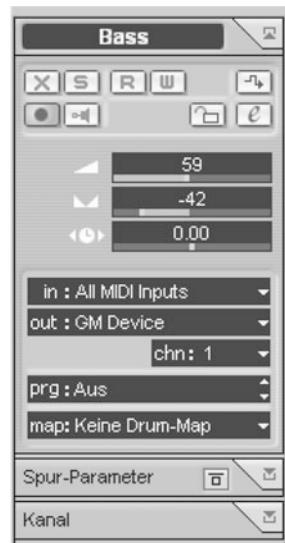


Abb. 4.12 MIDI-Mapping und Funktionen

Abb. 4.13 MIDI-Setup.
Dargestellt ist die Zuordnung von Spur, Kanal und Ausgabegerät



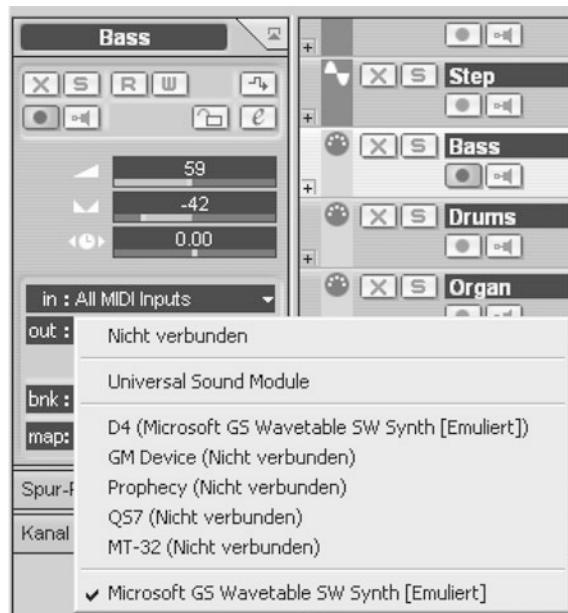
hier auch *Alle*, so dass dann alle Events dieser Spur an die spezifischen Kanäle gesendet werden (Ein- und Ausgabekanäle bleiben also unverändert). Der MIDI-Expander muss dann auch alle Kanäle empfangen können.

Da bei bestimmten Konstellationen verschiedene Ein- und Ausgabegeräte nicht immer einen diskreten Kanal bearbeiten, sondern z. B. Gerät₁ die Kanäle 1 und 2, gleichzeitig aber Gerät₂ die Kanäle 2 und 3, gibt es Verteilstrukturen, die ähnlich dem Setup in Abb. 4.12 sind. Die gesamte Wahl und Zuordnung dieser Geräte ist in den Menüs *in:* und *out:* zu treffen. Ein sehr einfaches System wäre, wenn man beispielsweise lediglich ein Gerät selektiert hätte, welches gleichzeitig MIDI-Daten zur Aufnahme und für die Wiedergabe aufbereitet. Benötigt der Anwender lediglich eine einfache Wiedergabe, ist im Feld *in:* keine Verbindung eingetragen, während im Feld *out:* das interne Soundsystem des PCs genügt (also quasi ein emulierter MIDI-Expander, siehe Abb. 4.14).

4.4.2 MIDI-Patchmaps

Bei der Patchmap ist es möglich, von der üblichen Aufteilung der Programme (siehe Tab. 4.15) abzuweichen. Sinnvoll kann dies z. B. werden, wenn man lediglich fünf verschiedene Programme benötigt, so dass diese einfach auf die Zahlen 0...4 gelegt werden. In Abb. 4.15 wird dies deutlich.

Abb. 4.14 MIDI-Setup, siehe oben. Hier kann das Gerät (oder die Software-Emulation) zur Ausgabe gewählt werden



4.4.3 MIDI-Keymaps

Eine individuelle Zuordnung von Notennummern ist nur bei Schlaginstrumenten vorgesehen, bei anderen Instrumenten wäre eine solche Manipulation weniger sinnvoll (Abb. 4.16 und 4.17).

4.5 Verschiedene Prinzipien der Klangerzeugung

MIDI-Geräte und -Instrumente setzen unterschiedliche Klangerzeugungen ein. Die Praxis richtete sich hier vollständig an Erfahrungen und natürlich auch am technologischen Standard aus. Während Untersuchungen und die damit verbundenen Erfahrungen zunächst Erfolge im Bereich der künstlichen Nachbildung von Klängen einbrachten, war der technologische Fortschritt dahingehend hilfreich, wenn es um die Verwirklichung speicherintensiver Methoden ging.

4.5.1 Analoge Klangsynthese

Eine Synthese von Klängen auf analoger Ebene war die Praxis ganz am Anfang der Ära elektronischer Instrumente. Die entsprechenden Erfahrungen auf diesem Gebiet sind heute mehr oder weniger nur noch historisch interessant.

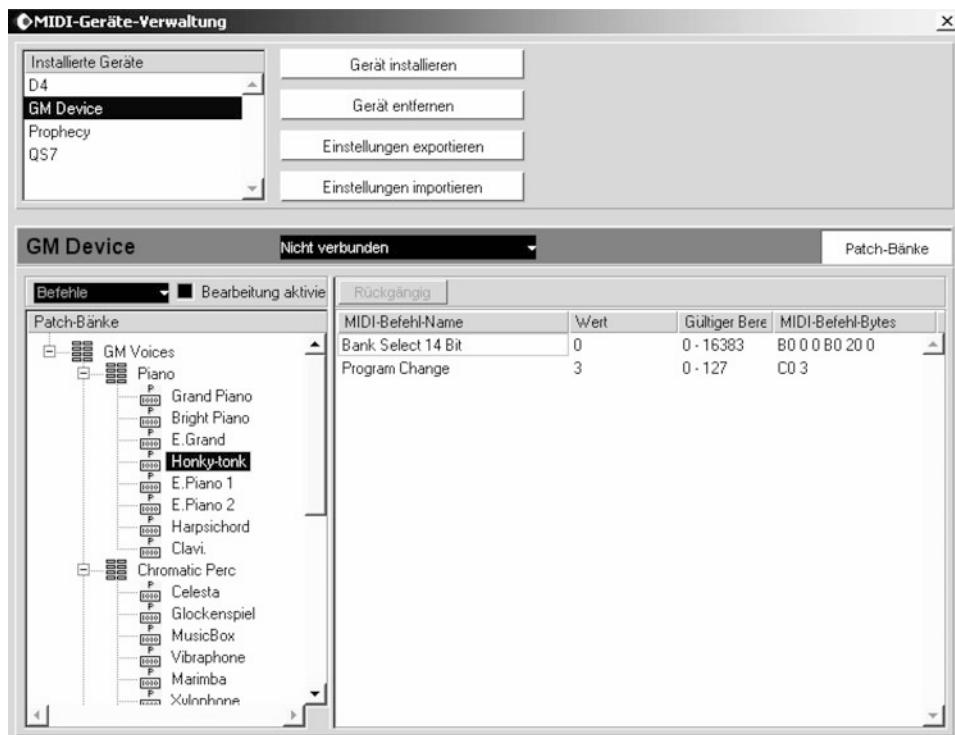


Abb. 4.15 Patchmap. Die Liste zeigt die Zuordnung der Programmnummern. Bei General MIDI ist beispielsweise das Electric Grand Piano auf Platz 2 (bei 0...127). Diese Nummer ist hier innerhalb des Bereiches beliebig veränderbar

Abb. 4.16 Keymap-Auswahl



Tonhöhe	Instrument	Quantisierung	St	E-Note	A-Note	Kanal	Ausgang
C1	Bass Drum	1/16 Note	C1	C1	10	Standard	
C#1	Side Stick	1/16 Note	C#1	C#1	10	Standard	
D1	Acoustic Snare	1/16 Note	D1	D1	10	Standard	
D#1	Hand Clap	1/16 Note	D#1	D#1	10	Standard	
E1	Electric Snare	1/16 Note	E1	E1	10	Standard	
F1	Low Floor Tom	1/16 Note	F1	F1	10	Standard	
F#1	Closed Hi-Hat	1/16 Note	F#1	F#1	10	Standard	
G1	High Floor Tom	1/16 Note	G1	G1	10	Standard	
G#1	Pedal Hi-Hat	1/16 Note	G#1	G#1	10	Standard	
A1	Low Tom	1/16 Note	A1	A1	10	Standard	
A#1	Open Hi-Hat	1/16 Note	A#1	A#1	10	Standard	
B1	Low Middle Tom	1/16 Note	B1	B1	10	Standard	
C2	High Middle Tom	1/16 Note	C2	C2	10	Standard	
C#2	Crash Cymbal 1	1/16 Note	C#2	C#2	10	Standard	
D2	High Tom	1/16 Note	D2	D2	10	Standard	
D#2	Ride Cymbal 1	1/16 Note	D#2	D#2	10	Standard	
E2	Chinese Cymbal	1/16 Note	E2	E2	10	Standard	
F2	Ride Bell	1/16 Note	F2	F2	10	Standard	

Abb. 4.17 Keymap-Inhalt. Hier lassen sich die Notennummern verschieben, d. h. bei gleicher Notennummer ist letztlich ein anderes Schlaginstrument hörbar

Elektronische Orgeln arbeiteten am Anfang auf analoger Basis, wobei die notwendigen Chöre meist durch analoge Treppenstufenteiler realisiert wurden.

Die ursprünglichen Synthesizer arbeiteten durchweg monofon, d. h. es wurde nur auf eine einzige gedrückte Taste des Keyboards reagiert, vornehmlich war dies entweder die niedrigste oder die höchste Tonhöhe, die durch die Tasten initiiert wurde. Bei manchen Synthesizern war diese Auswahl umschaltbar. Die Klaviatur lieferte hierzu eine Steuerspannung, die auf VCOs (Voltage Controlled Oscillator) gelangte. Obwohl von monofoner Konzeption, konnten einige dabei durchaus mit einer großen Klangfülle aufwarten, weil einfach mehrere dieser VCOs gleichzeitig auf die Steuerspannungen der Klaviatur reagierten.

Die Wellenform ist bei jedem VCO normalerweise zwischen einigen Grundarten umschaltbar, also Sinus, Sägezahn, Rechteck usw.

Der Nachteil bei einer solchen Klangerzeugung ist natürlich der mit der Klangfülle steigende Aufwand, weil ja für jede Stimme (die bei monofonen Instrumenten nicht unabhängig betrieben werden können) ein VCO zur Verfügung stehen m. Die Ausgänge der einzelnen VCOs lassen sich über Filter verändern, wobei diese mit einer Steuerspannung eine Veränderung erfahren, so dass nach dem Betätigen einer Taste ein fest einstellbarer Zeitablauf der Filtercharakteristik ausgelöst wird. Solche sog. VCFs (Voltage Controlled Filter) erhalten von der Klaviatur einen Trigger- bzw. Gate-Impuls. Zur Steuerung der Anschlagsdynamik stehen steuerbare Verstärker, sog. VCAs (Voltage Controlled Amplifier) zur Verfügung, die ihre Steuerspannung von einem Hüllkurvengenerator erhalten, dessen Hüllkurve einstellbar ist und der wiederum durch einen Gate-Impuls getriggert wird.

Moderne MIDI-Instrumente bzw. -Geräte setzen diese herkömmliche Methode der Klangerzeugung nicht mehr ein, weil eine Annäherung an die Klänge natürlicher Instrumente kaum möglich ist.

Die hier beschriebene Methode ist natürlich keineswegs spezifisch von analogen Kriterien bestimmt. Selbstverständlich ließe sich ein solches Verfahren auch mit digitalen Mitteln realisieren. Diese jedoch ermöglichen weitaus bessere Methoden mit überzeugenderen Klangresultaten. Die Analogtechnologie jedoch erlaubte in diesem frühen Stadium nur die oben beschriebenen Verfahren der Klangerzeugung.

4.5.2 FM-Klangsynthese

Das Streben nach synthetischer Nachbildung bestehender Instrumentenkänge führte zu der Erkenntnis, dass auch einfache Wellenformen zur Anwendung kommen können, wenn man die Generatoren einer Frequenz-Modulation unterzieht. Das Prinzip ist sehr einfach und stellt eine wesentliche Senkung des Aufwands dar in der Verwirklichung naturgetreuer Klänge.

Für eine Stimme benötigt man einen Trägergenerator und einen Modulator. Letzterer wirkt mit seinem Schwingungsverlauf auf die Frequenz des Trägers. Dabei sind sowohl die Frequenzen beider Generatoren als auch die primären Signalformen (vor der Frequenzmodulation) wählbar. Ferner lässt sich der Modulationsgrad einstellen, der die Frequenzvariation des Trägers definiert. Die prinzipielle Funktionsweise ist in Abb. 4.18 dargestellt.

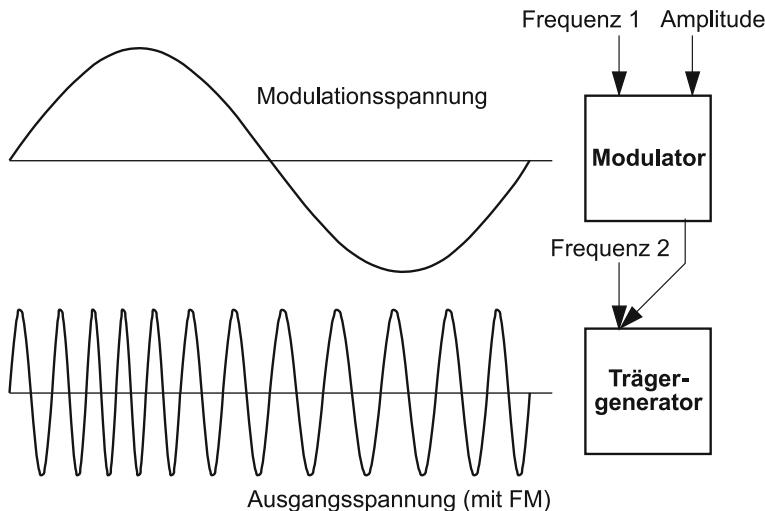


Abb. 4.18 Funktionsprinzip der FM-Klangsynthese

Nun wiederholt sich bei sehr niedrigen Modulationsfrequenzen das Bild des Ausgangssignals nach jeder Periode des Modulators; dies gilt natürlich genau genommen nur dann, wenn das Ausgangssignal an beiden Punkten der Modulationsperiode exakt dieselbe Phasenlage aufweist. Sind bei der Frequenzmodulation zwei Sinusschwingungen der Frequenzen ω_0 (Trägerfrequenz) und ω_m (Modulationsfrequenz) beteiligt, so ergibt sich eine endgültige Funktion der Gleichung:

$$f(t) = A \cdot \cos(\omega_0 t - \frac{\Delta\omega_0}{\omega_m} \cdot \sin(\omega_m t)) \quad (4.1)$$

Dadurch ergeben sich zahlreiche Seitenschwingungen, die sich um ganzzahlige Vielfache der Modulationsfrequenz von der Trägerfrequenz entfernt befinden. Die Amplituden dieser Seitenschwingungen übertreffen dabei die der Hauptschwingung. S. 187 [5]

Durch den enormen Gewinn an nichtharmonischen Schwingungen erhält der Klangcharakter eine sehr starke Fülle. Ist die Schwingungsform des Modulators nicht sinusförmig, so ist das Spektrum nochmals anders geartet. Interessant wird es ebenfalls, wenn die Modulationsfrequenz nicht mehr klein gegenüber der Trägerfrequenz ist. Außerdem wird teilweise ein geschlossener Modulationskreislauf praktiziert, bei dem sich zwei Generatoren gegenseitig frequenz-modulieren. Das ergibt ein chaotisches System, welches bereits durch kleinste Störungen starke Änderungen im Kurvenverlauf zeigt.

Nun belässt man die erhaltene Kurve nicht in ihrer ursprünglichen Form, sondern unterzieht sie einer Hüllkurvensteuerung, womit ein dynamisches Verhalten der Amplitude erhalten wird. Auch der Modulator kann in Bezug auf Amplitude bzw. Modulationstiefe gesteuert werden. Erst dadurch entsteht das endgültige, farbige Klangbild.

Das Verfahren der FM-Synthese beschränkt sich nicht auf eine analoge Technik; heutige Signalprozessoren ermöglichen Echtzeitberechnungen solcher Kurven, so dass weitere Generatoren nicht real verdrahtet werden müssen, sondern lediglich entsprechende Ergänzungen für den Prozessorcode bedeuten.

Das gezielte Herausfinden aller Parameter für die FM-Klangsynthese ist praktisch nicht möglich, weil bereits feinste Nuancen das Klangbild entscheidend verändern können. Wenn passende Parameter jedoch gefunden sind, ist nur die Speicherung letzterer notwendig, um jederzeit den gewünschten Klang reproduzieren zu können.

Die Klangvielfalt steht bei der FM-Synthese einem relativ geringen Aufwand gegenüber, so dass dadurch für elektronische Musikinstrumente ein entscheidender Fortschritt zu verbuchen war. Auch einfachere und billigere Audiokarten setzen dieses Verfahren ein; fertig integrierte Synthesizer-Chips ermöglichen die Nutzung eines festen Standards. Dennoch gibt es auch hier gewisse Grenzen, die für bestimmte Klangstrukturen zum Vorschein kommen und eine zufriedenstellende professionelle Nutzung im MIDI-Bereich einschränken. Im kommerziellen Bereich setzt man deshalb die FM-Synthese nur noch vereinzelt bzw. innerhalb eines Systems nur teilweise ein. Die bestehenden funktionellen Lücken müssen von Einheiten anderer Klangerzeugungsmethoden gefüllt werden.

4.5.3 Waveshaping

Beim Waveshaping durchläuft ein Sinussignal ein nichtlineares Glied, das entsprechend seiner Kennlinie harmonische Verzerrungen hervorruft. Durch geeignete Wahl dieser Kennlinie kann man den Gehalt und die Intensität der Obertöne sehr genau steuern, was mittels Tschebyscheff-Polynome erreichbar ist. S. 187 [6]

Die Erzeugung beliebiger Obertöne deckt jedoch nicht alle erforderlichen Klangstrukturen ab, weshalb eine auf diese Technik beschränkte Klangsynthese für MIDI-Geräte und -Instrumente zumindest nicht ausschließlich verwendet werden kann. Wie man sich leicht vorstellen kann, weisen die erzeugten Signale eine eindeutige Periodendauer auf, was für die Simulation vieler natürlicher Instrumenten natürlich nicht ausreicht. Vor allem Schlag- und Blasinstrumente erzeugen ein Tonsignal, das einen gewissen Rauschanteil zeigt.

Das Waveshaping eignen sich ebenfalls vorzüglich zum Einsatz in digitale Systeme, weil sich nichtlineare Kennlinien auf einfache Weise durch eine Tabelle simulierbar sind und somit die Werte der Grundfunktion schnell und bequem in gewünschter Form modifiziert werden können.

4.5.4 Fourier-Synthese

Das Zusammenstellen einer Grundschwingung und ihrer Obertöne kann auch über eine Fourier-Synthese geschehen. Bekanntlich lässt sich jegliche periodische Schwingung als trigonometrische Fourier-Reihe darstellen:

$$f(t) = \frac{a_0}{2} + a_1 \cos(\omega t) + a_2 \cos(2\omega t) + \dots + b_1 \sin(\omega t) + b_2 \sin(2\omega t) + \dots \quad (4.2)$$

Mit Δt als angenommene Periode ergeben sich hierbei die Fourier-Koeffizienten:

$$\begin{aligned} a_k &= \frac{2}{\Delta t} \int_0^{\Delta t} f(t) \cos(k\omega t) dt \\ b_k &= \frac{2}{\Delta t} \int_0^{\Delta t} f(t) \sin(k\omega t) dt \end{aligned} \quad (4.3)$$

mit $k = 0, 1, 2, 3, \dots$

Für spezielle Schwingungsformen fallen die Glieder a_k oder b_k weg, und unter gewissen Umständen können auch die Koeffizienten mit ungeradem oder die mit geradem Index null werden. Besteht ein Klang aus einer periodischen Schwingung, so lässt sich also diese auch durch Addition von Grundschwingung und Harmonischer erzeugen. S. 187 [7]

Für die Reproduktion einer gewünschten Schwingung ist nicht der Verlauf letzterer, sondern nur deren Fourier-Koeffizienten wichtig. Leider ist auch hiermit nur eine statische

Schwingung beschrieben; für die Nachbildung natürlicher Klänge jedoch sind natürlich auch gesteuerte Amplituden mittels Hüllkurven notwendig.

Die Fourier-Synthese kann wie das Waveshaping mit Signalprozessoren in Echtzeit durchgeführt werden; primär sind dem entsprechenden Prozess lediglich die Koeffizienten vorzugeben. Eine Berechnung kann selbstverständlich nach der Abarbeitung einer vollen Periode abbrechen, falls alle entsprechenden Schwingungswerte zwischengespeichert wurden.

Aufgrund der unzureichenden Klangresultate insbesondere in Bezug auf natürliche Instrumentenklänge wird die Fourier-Synthese kaum mehr als moderne Klangerzeugung herangezogen.

4.5.5 Lineare Arithmetik

Die Charakteristik von Instrumentenklängen ist vor allem während der Anklingzeit ausgeprägt. Würde man beispielsweise den Beginn von Saiteninstrumenten einfach entfernen, so ist oft eine sichere Zuordnung der Klänge nicht mehr möglich. Dieses die Klangcharakteristik prägende Anfangssignal zeigt zwangsläufig einen relativ komplexen Kurvenverlauf.

Bei der sog. Linear-Arithmetik verwendet man daher Sampling-Klänge von Originalinstrumenten für die anfängliche Phase des Tones. Beim Ausklingen hingegen bestimmen synthetisch erzeugte Signale mehr und mehr den Klang des Instruments. Daher stellt dieses Verfahren quasi eine Mischform von künstlich erzeugten und digitalisierten Klängen dar und bietet somit den Vorzug eines echten Klangbildes bei wenig Speicherbedarf.

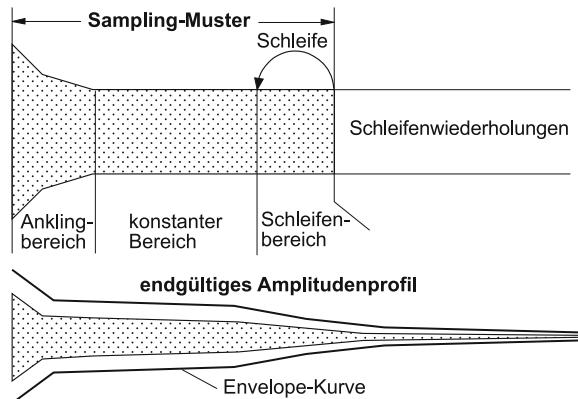
4.5.6 Sampling-Synthese

Alle Unzulänglichkeiten, die bei einer künstlichen Klangerzeugung nicht gänzlich auszumerzen sind, können durch Originalklänge von Instrumenten beseitigt werden. Hierzu gelangen letztere über Aufzeichnungen in einem digitalen PCM-Format auf ein digitales Speichermedium, von wo aus nach sorgfältiger Nachbearbeitung eine Kopie in ein ROM (Read Only Memory) gelangt. Um Speicherplatz zu sparen, kommen hierbei effiziente Kompressionstechniken zur Anwendung, die teilweise spezifisch auf den Klang zugeschnitten sind.

Bei vielen Instrumenten besteht der kritische zeitliche Teil des Klanges ganz am Anfang beim Anklingen und kurz danach. Selbstverständlich lässt sich nach dem Sampling dieser Abschnitt ebenfalls nachbearbeiten und effektvoll verstärken.

Grundsätzlich erhebt sich die Frage, welche zeitliche Ausdehnung der gespeicherte Klang einnehmen soll. Dabei sind Speicherkapazität und die Fähigkeit, eine Note von beliebiger Länge zu gestalten, wichtige Aspekte. Um beiden gerecht zu werden, erfolgt

Abb. 4.19 Sampling-Intervall mit Schleife und Envelope-Kurve



das Sampling nur über eine sehr kurze Zeit, an deren Schlussphase eine Schleife zu setzen ist. Abb. 4.19 zeigt das Prinzip einer solchen Wiedergabe mit Schleife.

Um die Natürlichkeit eines Instrumentenklanges (Timbre) zu wahren, wird meist eine einzige Sampling-Sequenz nur einem geringen Tonumfang gerecht; bei der Wiedergabe außerhalb dieses Intervalls erscheint der Klang mehr und mehr unnatürlich – eine spontane Identifizierung des Instruments ist bei zu großem Tonabstand sogar überhaupt nicht mehr möglich.

Die Envelope-Kurve formt zusätzlich den Verlauf der Amplitude, der ja ohne sie völlig gleichförmig wäre. Die Charakteristik sollte dem jeweiligen Instrument angepasst sein und kann prinzipiell natürlich ein beliebiges Profil aufweisen.

Die Wahl des Schleifenintervalls erfordert größte Sorgfalt. Die obligatorische Forderung, dass der Endpunkt der Schleife und der Anfangspunkt so zu liegen haben, dass keine Phasensprünge zu erwarten sind, kann meist relativ leicht erfüllt werden. Das ist gewährleistet, wenn an beiden Punkten die Schwingungen im Nulldurchgang bei gleicher Steigung sind. Dagegen sind Amplitudensprünge zwischen Schleifenausgang und -eingang nur sehr schwer zu vermeiden, weil der Klang eines natürlichen Instruments (vor allem eines ausklingenden) immer mit Amplitudenschwankungen verbunden ist. Ist die Länge des Intervalls zu kurz, so ist der Klang innerhalb des Schleifenablaufs zu starr und somit unnatürlich. Wählt man die Länge hingegen zu groß, ist die Gefahr von hörbaren Amplitudensprünge gegeben. Hinzu kommen hörbare, periodische Störungen, die ja bei der kleinsten Unregelmäßigkeit innerhalb des Schleifenintervalls mit deren Periodizität erscheinen. Amplitudensprünge können u. U. durch spezielle Maßnahmen kompensiert werden, hingegen sind periodische Schleifenstörungen kaum oder nur mit großem Aufwand zu beseitigen.

Zum Auffinden geeigneter Schleifenpunkte gibt es spezielle Software, die das Springen auf Nulldurchgänge garantiert und die Schleifenlänge variabel gestalten sowie das Schleifenintervall bewegen kann. Nach jedem Vorgang lässt sich dann die sich ergebende Schleifenfolge anhören, so dass auf diese Weise schnell eine Optimierung erfolgen kann.

Abb. 4.20 Verlauf der Schwingung einer angeschlagenen Gitarrensaite sowie das Setzen des Schleifenbeginns

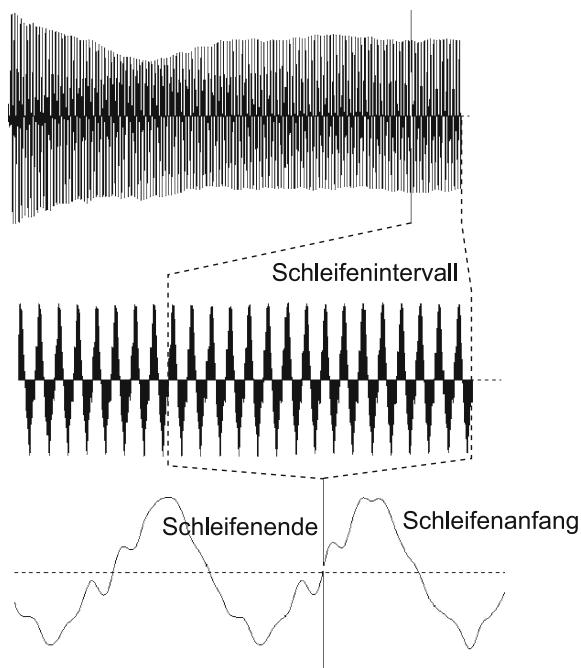


Abb. 4.20 repräsentiert ein Beispiel für das Setzen eines Schleifenpunktes. Es handelt sich hier um ein Sampling-Intervall einer angeschlagenen Gitarrensaite. Sehr deutlich ist der Anschlagbereich erkennbar, der ja zum Setzen des Schleifenbeginns ungeeignet ist. Innerhalb der Schleife befinden sich hier $n = 16$ Schwingungsperioden. Das Aneinandersetzen des Schleifenanfangs an des -ende wird mit dem Verlauf unten in der Abbildung erkennbar, was jedoch nur ein fiktiver Vorgang ist, der lediglich mittels Adressierung der Sampling-Werte realisiert wird. An den Scheitelpunkten der positiven Halbschwingungen sind durchaus Unterschiede feststellbar, die sich möglicherweise kontinuierlich durch das gesamte Schleifenintervall fortsetzen und somit mit einer Frequenz von $1/16$ der Saitenfrequenz zu Buche schlagen (ohne die Obertöne zu berücksichtigen). Vorteilig ist es jedenfalls meist, mit n eine 2er-Potenz zu wählen, damit Störfrequenzen nur als Subharmonische oder Harmonische in Erscheinung treten. Zu berücksichtigen ist allerdings, dass bei zu langen Schleifenintervallen möglicherweise Knackgeräusche trotzdem identifizierbar sind. Die gesamte Bearbeitung des Sampling-Intervalls erfordert jedenfalls auch mit komfortabler Software mitunter sehr viel Arbeitszeit, wenn die aufgerufenen Töne später ohne hörbare Störungen sein sollen.

Auf die oben beschriebene Weise können Samples jedoch nicht nur in ein unveränderliches ROM gelegt werden – spezielle Sampler-Geräte, die sogar auch als Computer-Steckkarten erhältlich sind, ermöglichen eigene Aufzeichnungen von Klängen oder Geräuschen, um sie dann nachher über eine Klaviatur oder entsprechende MIDI-Anweisungen als Noten spielen zu können. Sind beim jeweiligen MIDI-Gerät die gewünschten Klänge nicht verfügbar oder nicht zufriedenstellend, so erlaubt ein Sampler eine entsprechende Ergänzung. Die Einschränkungen unter General MIDI können mittels einer Wavetable-Karte

umgangen werden, indem man die gewünschten Klänge über Sampling aufgezeichnet und nachbearbeitet und sie dann als Substitution für die festgelegten Programme einsetzt.

4.6 Synchronisation von MIDI-Geräten

Für Sequencer und ähnliche Geräte gibt es einfache Möglichkeiten, über eine System-Realtime-Message auf ein MIDI-Steuergerät zu synchronisieren, indem letzteres Informationen nach Tab. 4.14 sendet. Eine derartige Synchronisation unter MIDI-Geräten und -Steuergeräten erfolgt über den reinen MIDI-Timecode oder nur den MIDI-Sync (MIDI-Taktsignal und Lied-Positionszeiger); sofern die Geräte eine solche Einrichtung besitzen bzw. solche Funktionen auslösen können, genügt eine einfache Verbindung der Interface-Anschlüsse.

Sollen alle MIDI-Geräte unbedingt ihres Typs und ihrer Bindung an einen spezifischen Kanal eine Synchronisation erfahren, so geschieht dies über eine System-Common-Message. Wie Tab. 4.13 zeigt, gibt es hierfür eine Anweisung, die den MIDI-Timecode betrifft.

4.6.1 Zweck einer Synchronisation

Die Zielsetzung bei der Verwendung mehrerer Geräte, die zu einem audiovisuellen Gesamtprogramm beitragen sollen, besteht letztlich aus der Forderung, zu definierten Spielzeitpunkten immer exakt die gleichen diskreten Einzelereignisse ihres Repertoires oder ihrer gespeicherten Ablaufprogramme zu erhalten. Diese Eigenschaft muss bei jedem erneuten Abspielen des Gesamtsystems erhalten bleiben.

Das zeitliche Verhältnis zwischen Bild- und Tonspur bei einer einfachen Videoaufnahme möge ein Beispiel hierfür sein; die zeitliche Bindung ist hier zwangsläufig durch die physische Fixierung beider Spuren auf ein Medium gewährleistet. Gleichgültig, wie oft das Band abgespielt wird, das gesprochene Wort stimmt immer mit den Lippenbewegungen überein. Eine solche physische Verknüpfung liegt jedoch nicht vor, wenn es sich um getrennte Geräte handelt. Stellen wir uns vor, ein Videorecorder müsste einen Computer zeitexakt steuern, so dass diskrete Stellen wieder absolut synchron zueinander sind.

Eine solche Anordnung ist u. a. notwendig, um Videosequenzen mit einem MIDI-Musikprogramm nachzuvertonen. Der Steuercomputer fungiert in diesem Fall als Master, der als Slaves mehrere MIDI-Instrumente oder -Geräte anzusteuern vermag. Er erhält von der Videobandmaschine Informationen, die eine einwandfreie Synchronisation ermöglichen.

Außer diesem zeitlichen Gleichlauf während des normalen Abspielens sind natürlich auch Suchoperationen von größter Wichtigkeit. Auch diese Option muss steuerungstechnisch möglich sein. Befinden sich alle Geräte in einer zeitlichen Konformität, so wird auch eine neu angefahrene Stelle korrekt wiedergegeben. Auf diese Weise ist es z. B. auch

möglich, in einer bestimmten Spur eines MIDI-Ablaufprogramms (Sequenzer-Programm) an einem vorgegebenen Zeitintervall frische Noten einzuspielen, ähnlich dem Vorgang, der bei herkömmlichen Audioaufnahmen im Mehrspurverfahren ja auch praktiziert wird.

4.6.2 Realisierung über Timecode

Um die oben genannten Forderungen erfüllen zu können, steht bei Verwendung von reinen MIDI-Geräten eine MIDI-Synchronisation zur Verfügung, welche über den MIDI-Takt und den Lied-Positionszeiger eine feste zeitliche Verbindung schafft. Zur Synchronisation von Nicht-MIDI-Geräten mit MIDI-Geräten benötigt man jedoch den sog. Timecode, der nach SMPTE ein definiertes Format aufweist und aus 80 Bits besteht. Die Nummerierung dieser Bits und ihre Bedeutung geht aus Tab. 4.17 hervor.

Die zeitliche Abfolge geschieht von Bit 0 bis Bit 79; das Synchronwort beginnt ebenfalls mit dem LSB und endet mit dem MSB. Die dazwischen liegenden Bits sind teilweise ungenutzt (0-Pegel) und teilweise als Userbits für Sondereintragungen des Anwenders verfügbar. Zur sicheren Lesbarkeit werden alle Nullen als einfachen Pegelwechsel übertragen, während Einsen mit doppeltem Pegelwechsel versehen sind.

Ein solcher Code gibt immer einen absoluten Zeitbezug innerhalb eines bespielten Bandes, welches natürlich entsprechend mit einem Timecode formatiert sein muss. Hierfür dient oftmals eine separate Bandspur, wenn ein Längscode (LTC) verwendet werden soll; alternativ besteht bei einigen Videorecordern die Möglichkeit, den Timecode in die Vertikalaustastlücke zu schreiben (VITC), was gewisse Vorteile gegenüber dem LTC bringt; beispielsweise kann bei beliebig langsamem Band – also auch bei Standbild – der Timecode immer sicher gelesen werden.

Tab. 4.17 Aufbau des SMPTE-Timecode

Bit-Nr.	Bedeutung	Stelle	Code	Bereich
0–3	Vollbild-Nummer (Frame)	Einer	BCD	
8–9	Vollbild-Nummer (Frame)	Zehner	BCD	0–24 (0–29)
16–19	Sekunden	Einer	BCD	
24–26	Sekunden	Zehner	BCD	0–59
32–35	Minuten	Einer	BCD	
40–42	Minuten	Zehner	BCD	0–59
48–51	Stunden	Einer	BCD	
56–57	Stunden	Zehner	BCD	0–23
64–79	Synchronwort	–	1011 1111 1111 1100	–

4.6.3 MIDI-Spezifikation für den Timecode

Um den SMPTE-Timecode in den MIDI-Timecode (MTC) zu übersetzen, bedarf es spezieller Schnittstellen-Geräte, die meist einfach zwischen die Geräte geschaltet werden. Dies ist nötig, da das Code-Format und auch die zeitliche Übertragungsrate zu beiden Seiten unterschiedlich sind.

Zu unterscheiden sind auf der MTC-Seite die Viertelbild-Messages (Quarter Frame) und die Voll-Messages.

Viertelbild-Messages Ein MTC, dessen Informationen nach jedem 4tel Videobild übertragen wird, benötigt die Dauer von zwei Videobildern für die komplette Übermittlung der aktuellen Zeitinformation, weil hierzu genau 8 Messages nötig sind, was auch aus Tab. 4.18 hervorgeht.

Da sich bei jeder Message nur ein Nibble der betreffenden Größe übertragen lässt, ist von dieser einfach der Hexadezimalwert und das entsprechende Nibble laut Tabelle zu nehmen. Das jeweils andere Nibble wird ignoriert und entspricht daher dem Zeichen x der Tabelle. Die Art des Timecode findet in der zweiten Stunden-Message Platz und ist codiert nach Tab. 4.19.

Abb. 4.21 zeigt ein Beispiel für eine Timecode-Konvertierung vom SMPTE- zum MIDI-Format.

Tab. 4.18 System-Common-Messages zur Übertragung des MTC

Status-Byte (Hex)	Daten-Byte (Hex/binär)	Bedeutung	Message-Nr.
F1	0b/0000 bbbb	Vollbild-Nummer xb (unteres Nibble)	0
F1	1b/0001 000b	Vollbild-Nummer bx (oberes Nibble)	1
F1	2s/0010 ssss	Sekunden xs (unteres Nibble)	2
F1	3s/0011 00ss	Sekunden sx (oberes Nibble)	3
F1	4m/0100 mmmm	Minuten xm (unteres Nibble)	4
F1	5m/0101 00mm	Minuten mx (oberes Nibble)	5
F1	6h/0100 hhhh	Stunden xh (unteres Nibble)	6
F1	7h/0101 0cch	Stunden hx (oberes Nibble), Timecode cc	7

Tab. 4.19 Timecode-Art und ihre Codierung

Code cc (binär)	Vollbilder pro Sekunde
00	24
01	25 (PAL)
10	30 (Drop-Frame, NTSC)
11	30 (Non-Drop-Frame, NTSC)

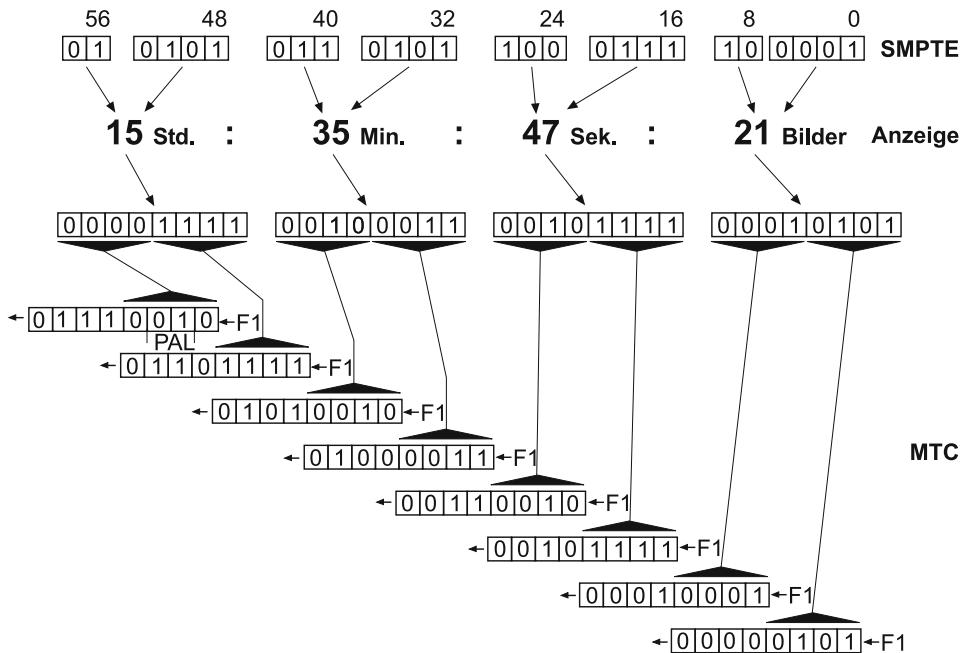


Abb. 4.21 Timecode-Wandlung von SMPTE nach MTC

Die endgltige Reihenfolge der Anweisungen hat von rechts unten nach links oben zu erfolgen, wobei zwischen jedem Daten-Byte natrlich das Status-Byte F1 zu setzen ist (aus Grnden der bersicht hexadezimal dargestellt). Außerdem wurde in dieser Abbildung auf das Synchronwort verzichtet, das im SMPTE vorhanden ist und bei MTC jedoch fehlt. Zu beachten ist ferner, dass in der letzten Anweisung das PAL-Format definiert ist.

Voll-Messages Viertelbild-Messages für den Timecode sind üblich während des Normalbetriebs eines Systems; sie eignen sich jedoch weniger gut beim schnellen Vor- oder Rücklauf, generell nicht für Betriebsarten, bei denen hohe Beschleunigungen auftreten können. Für diesen Zweck sind Voll-Messages gedacht, die mit einer einzigen Informationskette den kompletten Zeitpunkt vermitteln können.

Es handelt sich um eine Information der Kategorie System Exclusive (Realtime), welche laut Tab. 4.12 mit den beiden Bytes F0 7F beginnt. Die gesamte Message wird hier also (im Gegensatz zu den Viertelbild-Messages) nur durch zwei Status-Bytes begrenzt, nämlich F0 (Beginn von System Exclusive) und F7 (Ende von System Exclusive). Tab. 4.20 zeigt die Chronologie der gesamten Timecode-Message.

Es genügt, wenn ein MIDI-Gerät diesen Code einmal aussendet; die zu steuernden Geräte versuchen nun, den Zeitpunkt anzufahren. Im Normalbetrieb kann dann wieder auf Viertelbild-Messages umgeschaltet werden. S. 187 [3]

Tab. 4.20 Übermittelter Code bei einer Voll-Message für den Timecode

Byte (Hex)	Beschreibung
F0	Beginn von System Exclusive
7F	Spezifikation für Realtime
nn	nn = Kanal, 7FH für alle Kanäle
01	Spezifikation für MTC
01	Spezifikation für Voll-Message
hh	hh = Stunden (bei 24 Frames/Sekunde)
hh+20	hh = Stunden (bei 25 Frames/Sekunde)
hh+40	hh = Stunden (bei 30 Frames/Sekunde Drop-Frame, NTSC)
hh+60	hh = Stunden (bei 30 Frames/Sekunde Non-Drop-Frame, NTSC)
mm	mm = Minuten
ss	ss = Sekunden
bb	bb = Vollbild-Nummer
F7	Ende von System Exclusive

Bei der Viertelbild-Message sind insgesamt 16 Bytes zu übertragen, während eine vollständige Übermittlung der Zeitdaten mittels Voll-Messages 10 Bytes genügen.

4.7 MIDI-Geräte, Audiokarten und Zubehör

Für die Ansteuerung durch einen Computer gibt es unterschiedliche Möglichkeiten, die im wesentlichen darin unterscheiden, dass die MIDI-Geräte intern oder extern zu betreiben sind. Im erstenen Falle spricht man von der Audiokarte, im letzteren vom MIDI-Expander – eine etwas unglückliche Bezeichnung, weil der Begriff Expander bereits anderweitig in der Audiotechnik vergeben ist und weil im Grunde keine Expandierung stattfindet. Es handelt sich lediglich um eine elektronische Interpretation der MIDI-Anweisungen und um deren Umsetzung in Klänge. Vergleichbar ist dieser Vorgang mit der Arbeitsweise eines Druckers, der die elektronischen Daten des Computers in ein physisches Schriftbild verwandelt.

4.7.1 Ansteuerung externer MIDI-Geräte durch den Computer

Die Anwender-Software zur Ansteuerung von MIDI-Geräten kann so eingestellt werden, dass MIDI-Daten an den Interface-Ausgang des Computers gelangen. Dieses Interface kann dabei von vielerlei Gestalt sein; die Möglichkeiten reichen von MIDI-Interface bis zum Standard-Interface mit RS232- (V24) oder RS422-Format. Außerdem können natürlich auch Audiokarten mit MIDI-Out-Anschluss herangezogen werden.

Der hierfür notwendige Treiber ist entweder von der Anwender-Software selbst oder über ein übergeordnetes Betriebssystem (z. B. MS-Windows) anzusprechen. Letztere Situation ist i. Allg. günstiger, da ja die Anwender-Software im Normalfall nicht alle möglichen Interfaces kennt.

Soll die serielle Schnittstelle (RS232 oder RS422) zum Einsatz kommen, m auch hier ein MIDI-Treiber zur Verfügung stehen; ferner kommen in diesem Falle natürlich nur MIDI-Geräte in Frage, die eine entsprechende Datenrate verarbeiten können. Sie müssen diese Daten nämlich zunächst ins MIDI-Format wandeln oder zumindest anders behandeln als reine MIDI-Daten. Der Unterschied liegt in erster Linie in der Datenrate: Während MIDI eine feste Datenrate von 31250 Baud vorschreibt, ist bei RS232 ein maximaler Wert von nur 19200 Baud zulässig. RS422 ist in der Lage, noch höhere Baud-Raten zu bewältigen. Weitere Unterschiede betreffen die Übertragungsgröße, die bei V24 bipolare Spannungsimpulse und nur bei bestimmten Formen von RS232 auch Stromschleifenbetrieb zulässt, wobei dann allerdings 20 mA (bei MIDI 5 mA) verwendet werden. Da sich jedoch bereits die Datenraten unterscheiden, ist ohnehin ein Konverter notwendig, der dann zugleich die elektronische Anpassung vornehmen kann.

Liegt die Datenrate unter der des MIDI-Standards, kann es im Betrieb eher zu hörbaren Störungen oder Verzögerungen kommen als bei Verwendung der MIDI-Schnittstelle. Da die meisten Geräte jedoch die General-MIDI-Richtlinien nur wenig übertreffen, sind diesbezüglich wenig Qualitätseinbußen zu erwarten. Allerdings ist bei RS232 empfehlenswert, die höchste Datenrate von 19200 Baud einzustellen.

In Abb. 4.22 sind verschiedene Möglichkeiten zu sehen, ein externes MIDI-Gerät zu betreiben. Sollen mehrere Geräte zugleich angeschlossen werden, kann dies meist nur über verschiedene Schnittstellen erfolgen (eine Beschränkung der Kanäle ist i. Allg. nicht möglich und auch wenig sinnvoll, da die MIDI-Geräte nach General MIDI ja multitimbral sind und daher eine Kanalbeschränkung auch einer Reduzierung der gleichzeitig verfügbaren Instrumente gleichkommt).

4.7.2 Audiokarten und Betrieb unter MIDI

Der Einsatz von Audiokarten im Computer entspricht einer sehr komfortablen Lösung, wenn es darum geht, MIDI-Ablaufprogramme direkt in Klang umzusetzen und auch Audiosignale direkt aufzuzeichnen, falls eine Sampling-Einrichtung vorhanden ist. Doch auch Karten ohne Audio-Sampling-Ausstattung sind durchaus sinnvoll, wenn sie eine komplette MIDI-Schnittstelle aufweisen. Eine Verarbeitung von MIDI-Daten ist dann nicht nur intern möglich, sondern auch extern, d. h. die Karte kann MIDI-Daten an extern angeschlossene MIDI-Geräte oder -Instrumente abgeben oder auch von einem MIDI-Steuergerät (Klaviatur) MIDI-Daten entgegennehmen. Letztere Betriebsart stellt ein sog. MIDI-Sampling dar (was nicht zu verwechseln ist mit Analog-Audio-Sampling). Ein solches Aufzeichnen von mehreren MIDI-Quellen ist ohne weiteres möglich – man benötigt lediglich mehrere unabhängige MIDI-In-Anschlüsse, was über mehrere entsprechende Karten oder über eine Mehrfach-MIDI-Schnittstelle gelöst werden kann.

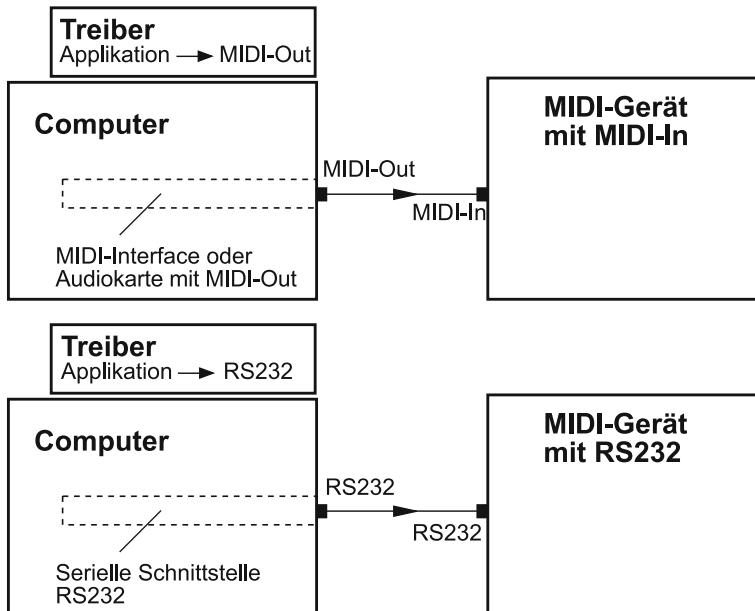


Abb. 4.22 Prinzipielle Ansteuerung externer MIDI-Geräte

Die Ausführung der MIDI-Steckkarte hängt von den jeweiligen Erfordernissen ab. Für hochqualitative Produktionen sollte sie die in General MIDI aufgestellten Konventionen übertreffen, zumindest was die Stimmenanzahl und die gleichzeitig verfügbaren Instrumente (*multitimbral*) anbelangt. Dariüber hinaus ist natürlich die Qualität der Instrumentenklangen ganz entscheidend; sie kann jedoch nicht immer mit Sampling-Klängen befriedigt werden. So gibt es durchaus Karten, bei denen das Resultat durch FM-Synthese besser scheint als bei anderen durch Sampling-Synthese. Hier entscheidet einzig und allein eine Hörprobe. Kartenhersteller bieten hierzu oftmals Referenzen in Form von CDs oder dergleichen an.

4.7.3 MIDI-Steuergeräte

Im oberen Abschnitt sind bereits MIDI-Steuergeräte angesprochen worden, die z. B. mit den MIDI-In-Anschlüssen von MIDI-Karten zu verbinden sind. Es handelt sich u. a. um Klaviaturen, die keine eigene Klangerzeugung aufweisen und auch mitunter Master-Keyboards genannt werden.

Durch den Verzicht auf die interne Klangerzeugung sind diese Klaviaturen meist recht preisgünstig. In der Ausstattung unterscheiden sie sich jedoch in puncto Steuerungsfunktionen, abgetasteter Anschlagsdynamik (Key-Velocity) und natürlich Tastenumfang. Die Mechanik der Tasten nimmt ebenfalls Einfluss auf die Bespielbarkeit; ein Keyboard mit dem akustischen Klavier nachgebildeter Mechanik lässt sich natürlich viel differenzierter spielen als eines mit vollkommen leichtgängiger Kunststofftastatur.

Neben Klaviaturen sind auch Blas- und Gitarrenkonverter erhältlich. Während Blaskonverter recht einfach eine MIDI-Umsetzung anhand der betätigten Klappen und der Luftströmung bewerkstelligen, ist die Realisierung aller Spielnuancen an einem Gitarrengriffbrett nicht ganz so einfach. S. 187 [8]

4.7.4 MIDI-Software

Software, die ihren Datenaustausch im MIDI-Format bewältigt, arbeitet intern keineswegs mit solchen Daten. Normalerweise bevorzugt jede Anwender-Software ihr eigenes Format; ihre Schnittstelle jedoch müssen dem MIDI-Standard entsprechen, weil sonst keine Kompatibilität mit anderen Systemen besteht.

Ein solches Programm steuert die angeschlossene Hardware, d. h. eingebaute Audiokarten oder extern befindliche MIDI-Geräte. Seine Ausrichtung ist dem Zweck ausgerichtet; so gibt es beispielsweise Sequencer-Programme, die hauptsächlich auf die Bedürfnisse der Spurenverwaltung und Schleifensteuerung berücksichtigen und Notationsprogramme, die Priorität auf möglichst komfortables Komponieren auf Notationsebene setzen. Die meisten dieser Programme bilden jedoch Mischungen aus allen Anwendungskriterien. Abb. 4.23 zeigt ein Beispiel für ein kleines, aber recht leistungsfähiges Programm zur Musikproduktion, welches auch die Möglichkeit der Notation bietet.

Wir wollen hier etwas näher auf den Aufbau eingehen. Ganz oben links ist darin die Bedienungseinheit zu finden, wie sie bei normalen Aufnahmegeräten auch üblich ist. Damit sind Wiedergabe- und Aufnahmefunktionen und natürlich auch schnelle Bewegungen innerhalb des bestehenden Ablaufprogramms möglich. Das rechte obere Feld sind Anzeigen für Zeit- und Taktnummern. Unterhalb der Klaviatur, die sowohl der Noteneingabe als auch der -anzeige dient, erscheint hier das Fenster der Spuren des Sequenzers. Diese sind mit willkürlichen Namen zu versehen und den gewünschten MIDI-Kanälen zuzuordnen. Außerdem lassen sich in den Spalten *Key* und *Vel* die allgemeinen Notenparameter festlegen. Ferner kann für jede Spur das Programm (Instrumentenklang) und der Port des anzusprechenden Gerätes definiert werden. Dabei findet eine Überschreibung der im MIDI-Mapper festgelegten Zuordnungen zwischen Kanal und Port (Anschluss) statt (vgl. Abschn. 4.4.1), weil die Einstellungen im Anwendungsprogramm i. Allg. Vorrang besitzen. Die Spalte *Loop* ermöglicht Schleifendurchgänge einer Spur.

Das überlappende Fenster zeigt die Notation einer Spur an. Im Stopp-Modus können hier mit der Maus einzelne Noten gesetzt oder verschoben werden. Die Dauer der Noten (Duration) wird durch nachgezogene Farblinien angedeutet. Das Hauptmenü lässt auch die Festlegungen der Tonart und der Taktart zu, was später für die Notation von entscheidender Bedeutung ist. Um die Notation übersichtlich halten zu können, sind oftmals zwischenzeitliche Tempi- oder auch Taktwechsel erwünscht, welche sich hier beliebig einbauen lassen. Als weiteres Fenster ist hier ein Mischer vorhanden, der alle Einstellungen der Spuren in Bezug auf Lautstärke, Panorama und ggf. Spezialeffekte vornimmt.



Abb. 4.23 Beispiel für ein Programm für Musik-Produktion (Steinberg Cubase LE)

Als Besonderheit bei MIDI-Anwenderprogrammen ist die Möglichkeit zu sehen, stehende Takes beliebig langsam ablaufen zu lassen, wobei sich natürlich nur das Tempo, nicht aber die Tonhöhe ändert. Das ist eine enorme Hilfe, wenn Feinheiten in der Notation editiert werden sollen.

Editierungsfunktionen wie Kopieren, Einfügen, Ausschneiden usw. gehören in den meisten Programmen heute genauso zum Standard wie bei der Textverarbeitung. Das erleichtert Routinearbeiten ganz entscheidend.

Außer dem direkten Setzen der Noten oder über die Klaviatur bieten manche Programme auch noch z. B. ein Gitarrengriffbrett an; eine monofone Spielweise für jede Saite sollte hier wählbar sein, um eine möglichst naturgetreue Umsetzung zu gewährleisten.

Obwohl das Programm mit seinem eigenen internen Format arbeitet, lassen sich die Musikprogramme auch als MIDI-Datei abspeichern. Hier sind jedoch teilweise Einschränkungen, die im Originalformat nicht bestehen.

In puncto Speicherbedarf für MIDI-Dateien (oder anderen Anwenderdateien für MIDI-Programme) braucht man sich keine großen Sorgen zu machen. Wenige kByte genügen hier oftmals bereits für mehrere Minuten Spielzeit, so dass der Rechner nur am Anfang einen Festplattenzugriff durchführen muss. Beim Abspielen von Sampling-Dateien hingegen ist der Datenstrom i. Allg. so stark, dass der Disk-Puffer ständig nachgefüllt werden muss.

4.7.5 MIDI-Softwareinterface

Verschiedene Software-Schnittstellen sorgen dafür, dass MIDI-Geräte, die am Rechner angeschlossen sind, auf das Betriebssystem bzw. die gerade laufende Applikation zugreifen können. Für Apple Macintosh bis OS 9 gibt es die beiden Programme *OMS* (Open Music System) und *FreeMIDI*. Ab dem Betriebssystem OS X sind dem Betriebssystem bereits Treibertools für MIDI bereitgestellt, der sog. Audio-MIDI-Setup.

Windows-Betriebssysteme greifen auf Gerät-spezifische MIDI-Treiber zu. Das bis Windows 3.1 bekannte Systemprogramm MIDI-Mapper ist jedoch bei 32- und 64-Bit-Systemen nicht mehr vorhanden. Ein Tool namens *Sounds und Audiogeräte* erlaubt die Wahl des wiedergebenden MIDI-Gerätes, an das einfache Player eine MIDI-Datei senden. Dieses Gerät kann im einfachsten Falle eine im Rechner integrierte MIDI-Hardware sein, aber auch irgendwelche extern angeschlossene wiedergebende MIDI-Geräte.

Applikationen verwenden ihre eigenen Einstellungen (siehe Abschn. 4.4. MIDI-spezifische Treiber von über USB (oder andere intelligente Schnittstellen) angeschlossenen Geräten sind jeweils im System zu installieren und erscheinen dann unter den Gerätemanagern.

4.8 MIDI-Applikationen und Problemlösungen

In diesem Abschnitt sollen einige Aufgabenstellungen und deren prinzipielle Lösungsmöglichkeit angedeutet werden. Dabei ist hier Wert darauf gelegt, die Forderungen an das System zu definieren, ohne auf die derzeit verfügbaren Einzelprodukte einzugehen.

4.8.1 Nachvertonung von Videosequenzen

Bei der Nachvertonung von Videosequenzen mit MIDI-Musikprogrammen ist zunächst die Frage zu stellen, ob eine asynchrone Untermalung genügt, oder ob eine sehr genaue Synchronisation von Noten ist, die über die gesamte Sequenz hinweg aufrechtzuerhalten ist. Im ersten Falle kann eine Zuspielung von der Audiokarte oder sonstigen MIDI-Geräten auf die Videonachvertonungsspur eines Recorders erfolgen. Ungefähr zeitliche Konformität ist dadurch zu erzielen, dass im Musikprogramm das Tempo herabgesetzt oder Pausen hinzufügt werden. Beides, Musikprogramm und Videoprogramm, sollte genügend leeren Vorlauf besitzen, so dass ein gleichzeitiger Start erfolgen kann. Bei längeren Takes kann jedoch auch möglicherweise selbst eine grobe zeitliche Übereinstimmung nicht mehr gewährleistet werden.

Fehlende Synchronisation Nehmen wir an, beide Systeme, nämlich der Videorecorder und der Audiozuspieler in Form des Computers folgen beider unabhängigen Synchrontakten, die jedoch quarzgenau sind (das kann für beide Systeme tatsächlich angenommen

werden; der Videorecorder synchronisiert seine Kopftrommel sogar dann mit einem Quarztakt, wenn keine externe Synchronisation angeschlossen ist). Quarzoszillatoren, die keine Möglichkeit eines Abgleichs aufweisen, unterliegen einer Frequenztoleranz, die vom Swinger selbst herrührt. Hier muss mit Streuungen von bis zu $\mu = \pm 10^{-4}$ gerechnet werden. Im Extremfall können beide Systeme – Audio und Video – entgegengesetzte Toleranzen besitzen, das entspräche einer gegenseitigen Toleranz von $2 \cdot 10^{-4}$. Bei einer Gesamtaufzeit von $t_{\text{ges}} = 1$ Stunde ergibt sich eine Abweichung von:

$$\Delta t = 2 \cdot \mu \cdot t_{\text{ges}} = 2 \cdot 10^{-4} \cdot 3600 \text{ s} \quad (4.4)$$

Um diese Zeit können beide Systeme nach einer Stunde maximal zeitlich auseinanderlaufen. Es kommt einzig und allein auf den Anwendungsfall an, ob ein solcher Zeitfehler akzeptabel ist oder nicht. Ein bereits beim Start auftretender Offset ist natürlich ebenfalls zu berücksichtigen.

Feste Synchronisation Ist eine permanente Synchronisation erforderlich, die eine Drift unabhängiger Systeme verhindert, m das Sequenzer-Programm eine Möglichkeit der externen Synchronisation über SMPTE-Timecode besitzen. Da Audiokarten i. Allg. diesen Timecode nicht direkt verarbeiten können, m ihn ein externer Konverter in das MTC-Format überführen. Eine solche Verschaltung zeigt Abb. 4.24.

Die meisten modernen, professionellen Videorecorder für Schnittbetrieb liefern den SMPTE-Timecode auf einer 9poligen Sub-D-Buchse, und die elektrischen Daten der Schnittstelle entsprechen i. Allg. RS422 (seltener RS232). Das Datenprotokoll muss bei Nichtfunktion überprüft werden.

Das Sequenzerprogramm lässt sich normalerweise auf ein bestimmtes Timecode-Datum zum Start initialisieren. Dieser Zeitpunkt sollte so gewählt sein, dass sowohl die Videosequenz als auch das MIDI-Musikprogramm ein paar Sekunden später beginnt. Auf diese Weise ist gewährleistet, dass beide Programme erst beginnen, wenn beide Systeme bereits sicher aufeinander synchronisiert haben. Zur Verdeutlichung diene Abb. 4.25.

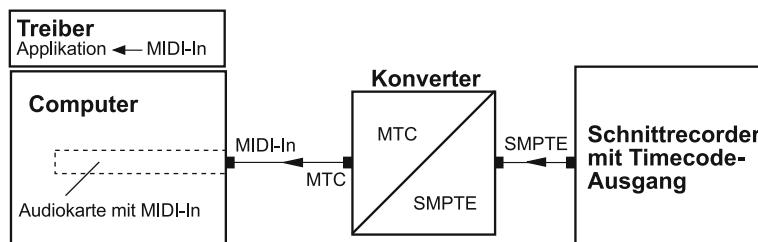


Abb. 4.24 Synchronisation des MIDI-Programms durch Videorecorder

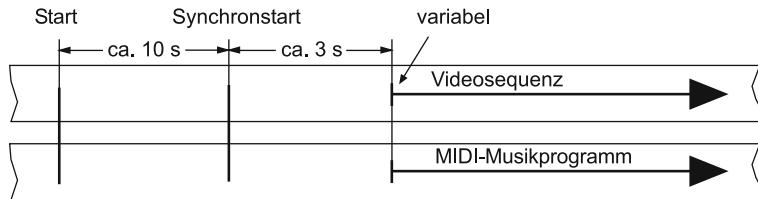


Abb. 4.25 Synchronzeitpunkt für Bild und Ton

Das Videoprogramm kann im Prinzip zu einem anderen Zeitpunkt einsetzen; aus Gründen der Übersichtlichkeit sollte man ihn jedoch übereinstimmend mit dem Beginn des Musikprogramms wählen. Das MIDI-Programm wird normalerweise bereits durch Starten in den Wartemodus gebracht werden; erreicht die Videobandmaschine den vorgegebenen Timecode, so erfolgt ein Einlocken in die Synchronisation von Band. Auch hier können natürlich die Zeitpunkte zwischen Videostart und Vorstart des MIDI-Programms differieren, solange gewährleistet ist, dass die Bandmaschine mindestens ca. 10 s und das MIDI-Musikprogramm noch sicher vor dem Synchronstart-Zeitpunkt gestartet werden.

Nachvertonungstechniken Üblicherweise lassen Videorecorder eine Nachvertonung nur auf ihren Längsspuren zu; eine eventuell vorhandene Möglichkeit einer Tonschrägspurenauflaufzeichnung entfällt, weil dort ja bereits Bildinformation steht, die nicht gelöscht werden darf. Ein mit Videokopf aufgezeichneter Ton kann nur genutzt werden, wenn man mit zwei Maschinen arbeitet: die erste liefert das Bildsignal und folgt oben genannten Synchronisationsmechanismen, und die zweite Maschine zeichnet dann Ton und Bild auf die Schrägspuren auf und nutzt somit die höhere Tonqualität. Als Zuspieler können natürlich auch mehrere Videobandmaschinen dienen, falls immer ein gültiger Timecode zur Verfügung steht und dieser das MIDI-System zu triggern imstande ist.

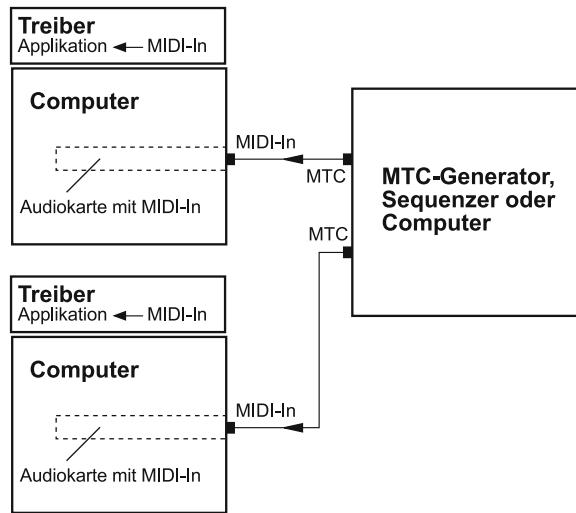
Natürlich können auch auf der Audioseite mehrere Zuspieler existieren. Eine entsprechende Synchronisation aller ist dadurch zu erreichen, dass das MTC-Signal nach Abb. 4.24 über die MIDI-Thru-Anschlüsse an die MIDI-Eingänge weiterer Systeme geführt werden, die ihrerseits wieder auf spezielle Timecode-Werte initialisiert wurden (die einzelnen Startpunkte können dabei durchaus voneinander abweichend sein).

4.8.2 Synchronisation mehrerer MIDI-Applikationen

Eine zentrale Synchronisierung mehrerer MIDI-Systeme ist relativ einfach möglich, indem das Signal eines MTC-Generators an die Slaves zu leiten ist. Lineare Verbindungen über die Thru-Anschlüsse sind bei wenigen Systemen zwar denkbar, sollten jedoch möglichst vermieden werden. Abb. 4.26 zeigt eine sternförmige Verbindung zur Synchronisation.

Als Quelle für den MIDI-Timecode können auch Sequenzer oder Computer mit entsprechender Ausstattung zur Anwendung kommen. Alternativ zum MTC kann allerdings

Abb. 4.26 Sternförmige Zuführung des MTC für Synchronisation mehrerer Systeme



auch der MIDI-Takt mit Lied-Positionszeiger (MIDI-Sync) genutzt werden, wenn sich alle Komponenten auf dieser Basis verständigen können (siehe auch Abschn. 4.2.6).

4.8.3 MIDI als Playback-Quelle für Musiker

MIDI-Musikprogramme lassen sich selbstverständlich auch als Übungs- oder Aufführungs-Playback verwenden. Eine Synchronisation ist nur beim Einsatz mehrerer Systeme notwendig (siehe Abschn. 4.8.2). Liegt das MIDI-Musikprogramm inmitten des zu spielenden Stücks, so muss das MIDI-System zum richtigen Zeitpunkt gestartet werden.

Da nun vor dem Einsatz des MIDI-Programms die gespielte Musik bereits ein vorgegebenes Tempo annimmt, ist eine Übereinstimmung beider Tempi keineswegs gegeben, wenn nicht vorher besondere Vorkehrungen getroffen werden. Die Folge davon ist entweder ein Auseinanderlaufen beider Metren oder ein abrupter Tempowechsel (siehe Abb. 4.27).

Das Tempo kann zwar über ein voreingestelltes Metronom dem MIDI-Musikstück angeglichen werden; sicherer sind jedoch die beiden unten gezeigten Methoden.

Verdrahtete Synchronisation Die aufwendigste Methode setzt wiederum MIDI-Sync oder MIDI-Timecode ein. Der Taktgeber gibt visuell Taktpunkte an und leitet gleichzeitig an das MIDI-System den MIDI-Sync oder den MTC.

Der Startzeitpunkt für das MIDI-System ist dann als MTC-Datum im Programm anzugeben. Der Start des Taktgebers jedoch erfolgt mit entsprechendem Offset vorher (siehe Abschn. 4.8.1). Selbstverständlich können auf diese Weise auch mehrere MIDI-Systeme

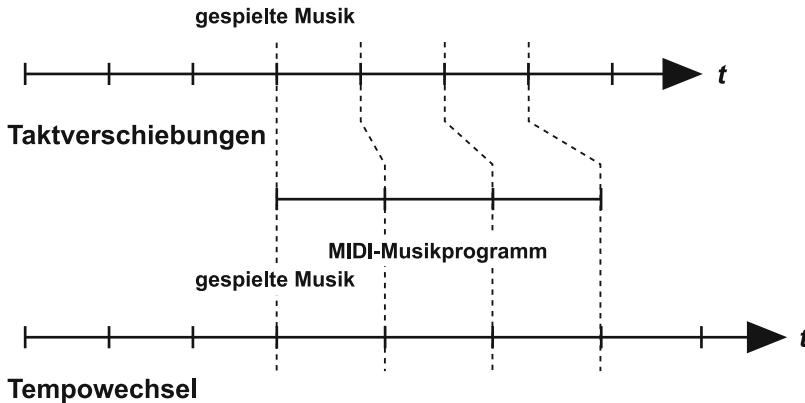
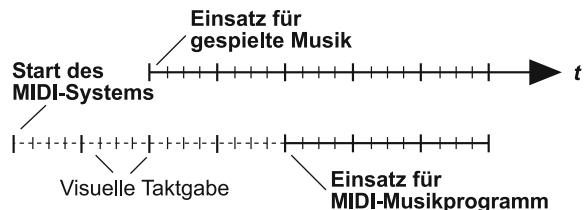


Abb. 4.27 Probleme bei asynchronem MIDI-Musikprogramm

Abb. 4.28 Einfache Synchronisationsmöglichkeit eines MIDI-Systems in der Live-Musik



gesteuert und -Musikprogramme eingebaut werden, deren Startpunkte sogar auseinander liegen können.

Im Falle einer Live-Aufführung sind Pannen natürlich besonders unangenehm, weshalb zwischenzeitliche Prüfungen einer Übereinstimmung markanter Zeitpunkte von Vorteil ist. Das kann anhand eines Countdown erfolgen, dessen Status noch vor dem Startinitial des MIDI-Musikprogramms zu eruieren ist. Sind hier Verschiebungen um ganze Takte oder mehr festzustellen, kann dies nur durch ein Nachstellen des Startzeitpunkts für das MIDI-System ausgeglichen werden. Für dieses Vorgehen, das sehr viel Geistesgegenwart, Reaktion und Taktempfinden erfordert, ist jedoch ein spezieller Operator angebracht, der z. B. auch für die Mischung zuständig ist.

Synchronisation mittels MIDI-System Eine Möglichkeit, die ohne separaten Taktgenerator auskommt, sei hier ebenfalls erwähnt. Hierzu übernimmt das MIDI-System mit Sequenzerprogramm die Aufgabe des Metronoms. Nach dem Start des letzteren erfolgt die Anzeige visueller Taktpunkte, die z. B. nur der Schlagzeuger wahrnimmt. Nach einer bestimmten Taktzahl kommt der Einsatz der Musiker und an der festgelegten Stelle der (hörbaren) Einsatz des MIDI-Musikprogramms (siehe Abb. 4.28).

Eine einfache Audiokarte kann die Aufgabe des Metronoms übernehmen, wobei die Taktimpulse auf eine separate Spur gelegt werden, die exklusiv mit der Audiokarte als Port korrespondiert (vgl. Abschn. 4.7.4). Die Überführung in optische Taktimpulse geschieht mittels kleiner Zusatzelektronik, wobei natürlich eine Differenzierung zwischen Takthauptschlägen und Zwischenschlägen möglich sein m.

Die Taktgabe kann u. U. auch über Kopfhörer realisiert werden, was im Einzelfall jedoch einem Test unterliegt. Zur Sicherheit sollte in jedem Falle eine visuelle Taktgabe vorhanden sein, wobei auch diese nicht immer auf Anhieb vom entsprechenden Musiker korrekt umgesetzt werden kann. Eine ganz einfache Schaltung für ein solches optisches Metronom ist in Abb. 4.29 zu sehen. Es ist an den Stereoausgang der dafür speziell reservierten Audiokarte anzuschließen. Die Dauer der Lichtimpulse beträgt nach dieser Dimension ca. 0,1 s, was jedoch einfach zu modifizieren ist. Auch die Empfindlichkeit von ca. 1 V lässt sich durch Verkleinern des 120k-Widerstands verbessern. Zur Unterscheidung zwischen Haupt- und Zwischenimpulsen dienen beide Stereoanschlüsse der Audiokarte, wobei hier nur ein Kanal dargestellt ist, der andere ist nur angedeutet. Die optischen Impulse werden von LEDs abgegeben, was im Normalfall genügt, da ja im Live-Betrieb auffällige optische Zeichen meist nur stören.

Bei einem 4/4-Takt beispielsweise ist ein Muster nach Abb. 4.30 auf die Metronomspur zu legen. Wenn es die Voreinstellung der Karte erlaubt, sollte trotz des Stereoklanges nur ein MIDI-Kanal belegt werden.

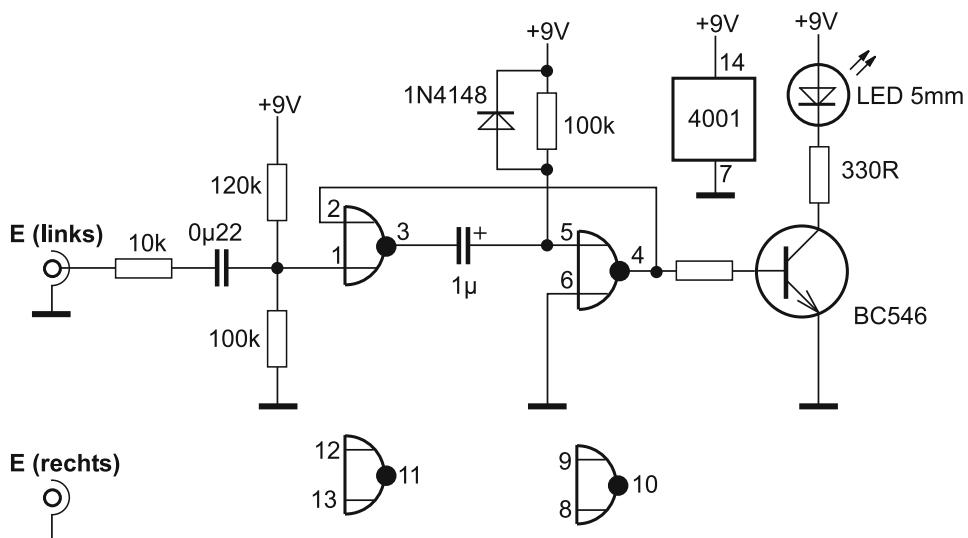
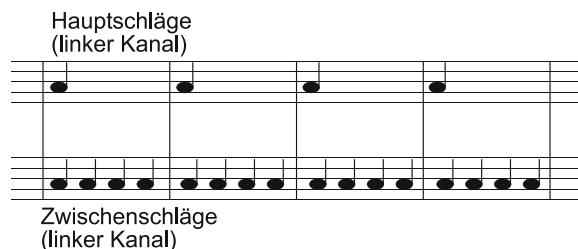


Abb. 4.29 Zusatzschaltung zur Realisierung eines visuellen Metronoms in Verbindung mit einer Audiokarte

Abb. 4.30 Notation der Metronomspur am Beispiel eines 4/4-Taktes



4.8.4 MIDI-Arrangement und Sampling-Dateien

In vielen Fällen wird wohl ein Arrangement, bestehend nur aus MIDI-Klängen, nicht ausreichend sein. Es gibt hierfür vielerlei Gründe; sei es bestehendes Sampling-Material, das unbedingt verwendet werden soll oder seien es überhaupt Klänge, die nicht annähernd über MIDI zu erzeugen sind – in solchen Fällen m die Möglichkeit einer Einbindung von Sampling-Dateien erwogen werden können.

Substitution eines MIDI-Instruments durch Sampling-Dateien Bereits in Abschn. 4.5.6 fand eine Schilderung darüber statt, wie die Klänge von Instrumenten über Sampling einer Notation für MIDI zur Verfügung stehen. Je nach Speichergröße der Wavetable-Karte ist jedoch auch das Sampling ganzer Passagen möglich, um sie dann als einzige Note innerhalb eines Sequenzer-Programms wieder ablaufen lassen zu können.

Obwohl das Setzen dieser einzelnen Note ja bereits die Genauigkeit des Einsatzes garantiert, trifft dies natürlich für den anschließenden Synchronismus nicht hinreichend auch zu. Um dem zu genügen, kann das Sampling simultan mit dem Abspielen der MIDI-Datei erfolgen. Meist genügt für diesen Vorgang sogar eine einzige Audiokarte, falls sie über einen MIDI-Synthesizer und eine Sampling-Einrichtung verfügt. Da Betriebssysteme wie Windows ein sog. Multitasking erlauben, können Sequenzer und Sampling-Programm sozusagen gleichzeitig (oder schnell abwechselnd) aktiviert werden. Nach dem Start des MIDI-Musikprogramms wird die entsprechende Sequenzer-Software jedoch eigentlich nicht mehr benötigt – den Ablauf besorgt normalerweise komplett der Prozessor der Audiokarte. Speicherintensive Operationen gibt es dann nur noch seitens Sampling-Programms, welches ebenfalls zur Aufzeichnung gestartet wird. Abb. 4.31 zeigt das einfache Schema bei diesem Vorgang.

Ein Beispiel für ein derartiges Verfahren bieten bereits Zubehör-Programme unter dem Betriebssystem MS-Windows; Abb. 4.32 zeigt beide Programme im geöffneten Zustand.

Natürlich ist eine solche Operation auch mit einem komfortablen Sequenzer- und einem Sampling-Programm machbar. Es sollte bei einem frühen Einsatz des Sampling darauf

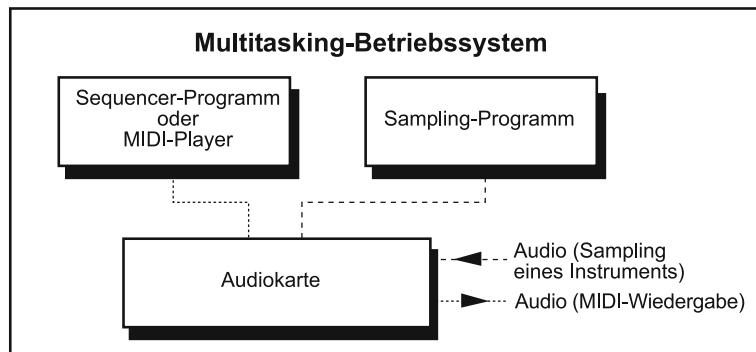


Abb. 4.31 Erzeugung einer zeitsynchronen Sampling-Datei



Abb. 4.32 Zubehör-Programme unter MS-Windows zum simultanen Handling der MIDI-Wiedergabe und des Audio-Sampling

geachtet werden, dass das MIDI-Musikprogramm einen genügend großen Vorspann hat, damit die Zeit zwischen dem Start von MIDI und der Aufzeichnung nicht zu kurz wird.

Bei der Substitution eines Instruments mittels Wavetable-Karte muss natürlich eine Notation unter dem Referenzton erfolgen, sonst entsteht eine veränderte Wiedergabegeschwindigkeit, was wiederum den Synchronismus und auch die Richtigkeit der Tonlage zwischen Sampling und MIDI zerstört.

Unabhängige Sampling-Datei Eine Substitution kommt unter zwei Bedingungen nicht in Frage: wenn die Datei für das RAM der Wavetable-Karte zu groß ist oder eine solche Karte überhaupt nicht verfügbar ist. Unter diesen Umständen müssen beide Dateien in zeitlichem Bezug zueinander gestartet werden, was diesmal jedoch nicht über die Notation geschehen kann. Eine solche zeitliche Verknüpfung kann auch nicht manuell realisiert werden, weil dies zu ungenau ist. Diesen Vorgang kann nur ein Spezialprogramm abwickeln, welches MIDI- und Sampling-Dateien synchron verarbeiten bzw. starten

kann. Ggf. können solche Probleme auch mit Makros gelöst werden, wobei allerdings zeitliche Verzögerungen dennoch nicht definiert bzw. konstant bleiben, wenn noch andere Anwendungen geöffnet sind oder der Prozessor viele Hintergrundaufgaben zu bewältigen hat.

4.9 MIDI-Praxisbeispiel

Ein Beispiel für gemessene MIDI-Daten zeigt uns Abb. 4.33.

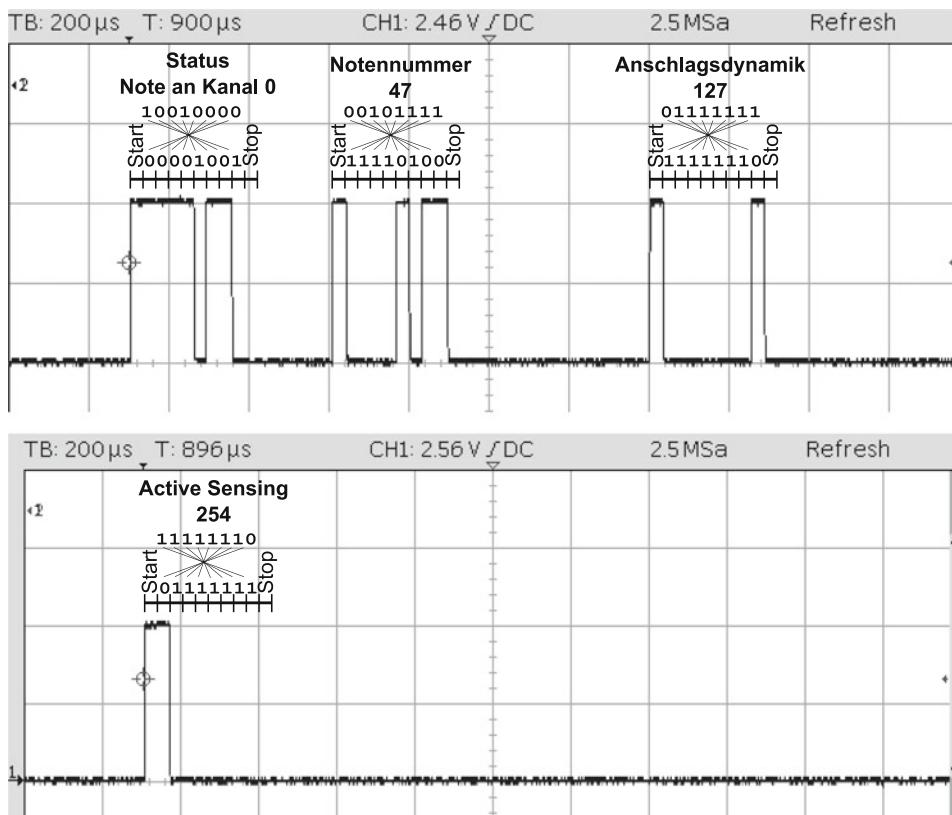


Abb. 4.33 MIDI-Daten mit einem Speicheroszilloskop aufgezeichnet. Gewonnen wurde dieses Digitalsignal am Ausgang eines MIDI-Keyboards beim Anschlagen einer bestimmten Taste. Im oberen Bereich sieht man, wie sich die Daten in drei Bytes aufschlüsseln: Statusbyte, Notennummer und Anschlagsdynamik. Wichtig ist zu wissen, dass laut Abb. 4.3 bei Stromfluss – also wenn H-Zustand auf dem Oszilloskop erscheint – der Zustand LOW zugewiesen wird. Außerdem kommen die Datenbits in umgekehrter Reihenfolge (zuerst Bit 0, dann Bit 7). Im unteren Bereich der Abbildung ist das Signal *Active Sensing* zu erkennen – die Abwesenheit dieses Signals wäre ein Zeichen für eine unterbrochene Verbindung. Formal erkennt man eigentlich nur zwei Bits mit L-Wert, der Byte-Wert ist FEh. Bei MIDI beträgt die Transferrate 31,250 kHz, ein einzelnes Bit dauert somit 32 µs

Es gibt viele Programme, die wie ein Recorder MIDI-Signale aufzeichnen und wiedergeben können. Eines davon ist Cubase LE. Das Einspielen von Daten kann über unterschiedliche Arten erfolgen. Reine MIDI-Keyboards sind in der Lage, MIDI-Daten auszugeben, ohne dass sie notwendigerweise auch Klangerzeuger sind. Letzteres wird i. Allg. von einer anderen Hardware bewerkstelligt. Die Daten sind auch per Maus oder Tastatur in einem Editorfenster setzbar und manipulierbar. Der Anwender kann jedoch nicht nur Noten bearbeiten, sondern auch alle Kontrollfunktionen, eben alles, was per MIDI möglich ist.

Das Ein- und Auslesen der Daten kann bei geeigneten Geräten auch über USB oder andere Schnittstellen erfolgen, damit unterliegt der Datenstrom nicht den begrenzten Transferraten wie bei MIDI.

Einige Fenster des Programmes *Cubase LE* seien in den Abb. 4.34, 4.35 und 4.36 dargestellt.



Abb. 4.34 Hauptfenster der MIDI-Software (Cubase LE). Der rechte Bereich zeigt die einzelnen Spuren eines MIDI-Stückes sowie das Transport-Fenster (Vor-/zurückspulen, Wiedergabe/Aufnahme). Links davon befinden sich die einige Schaltfunktionen für jede Spur, wie z. B. Stummschaltung, Soloschaltung, Aufnahmefreigabe und andere. Ganz links sind weitere Detaileinstellungen für die aktuell aktive Spur, z. B. Zuordnung zu welchem Ein- oder Ausgabegerät (diese kann der PC selbst sein, der über sein Audiosystem die MIDI-Daten in Klang übersetzt), welches Programm (also welches Instrument für diese Spur erklingen soll), eingeschleifte Effekte, Lautstärke und anderes mehr

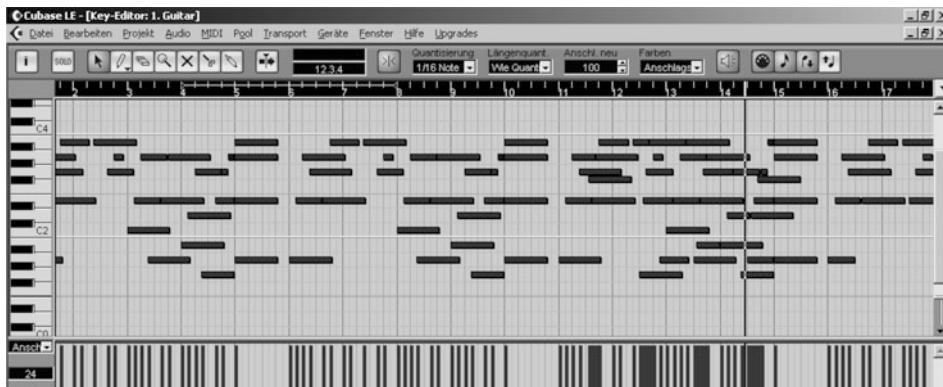
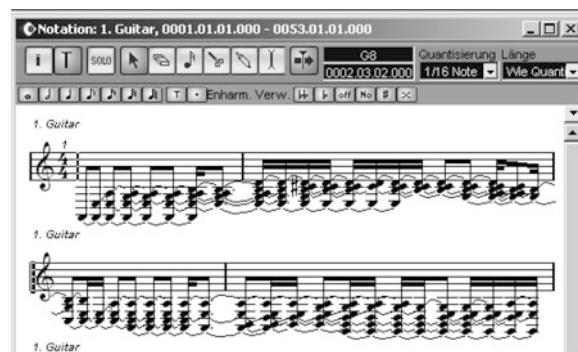


Abb. 4.35 Key-Editor. Durch Doppelklick auf eine der Spuren in Abb. 4.34 öffnet sich ein Editorfenster, welches alle gesetzten Noten als Rechtecke zeigt. Diese sind alle einzeln setzbar und in jeglicher Form veränderbar, die Werkzeuge hierzu befinden sich oben. Direkt unterhalb dieser Werkzeugpalette ist die Zeitlinie mit den Taktziffern zu erkennen und ganz links eine Klaviatur, deren Tasten mit der entsprechenden Tonhöhe im Raster kommunizieren. Ganz unten erscheinen die senkrechten Balken der Controller-Funktion, in diesem Falle spiegeln sich die Anschlagsstärke der einzelnen Noten wider

Abb. 4.36 Notationsfenster.

Für jede Spur lässt sich auch ein Notationsfenster öffnen. Die Darstellung des Notenbildes hängt von einigen Einstellungen ab



Literatur

1. International MIDI Association: Midi 1.0 Detailed Specification.
23634 Emelita Street, Woodland Hills, CA 91367.
2. International MIDI Association: Addendum to Midi 1.0 Detailed Specification.
3. International MIDI Association: MIDI Machine Control (MMC) 1.0.
4. International MIDI Association: General Midi Level 1 Specification.
5. Webers, J.: Tonstudientechnik. München: Franzis-Verlag 1989.
6. Jecklin, J.: Musikaufnahmen. München: Franzis-Verlag 1980.
7. Philippow, E.: Taschenbuch Elektrotechnik, Band 1.
Berlin: VEB Verlag Technik Berlin 1976.
8. Noll, J.: Multimedia, Midi und Musik. Frankfurt: Fischer Taschenbuch Verlag 1994.



5.1 Herkömmliche analoge Messtechnik

Spezielles Mess-Equipment für analoge Verfahren werden kaum noch hergestellt, und erhältlich sind sie fast nur noch über Internet-Auktionen. Manche Messungen sind mit universellen Messgeräten machbar, so z. B. dem Oszilloskop.

5.1.1 Frequenzgang

Zur Darstellung eines Frequenzgangs benötigt man einen Sweep-Generator, der ein Sinussignal mit stetig geänderter Frequenz ausgibt. Als Messgerät kann entweder ein Oszilloskop oder ein Schreiber dienen (inklusive einem Pegeldetektor). Eine entsprechende Anordnung ist Abb. 5.1 zu entnehmen. Ferner ist das entsprechende Oszillogramm zu sehen.

Ein von Hand durchgestimmter Generator und ein Pegelmessgerät wäre ebenfalls eine geeignete Kombination zur Bestimmung des Frequenzgangs. Eine weitere Messmethode besteht aus Rauschgenerator und Spektrum-Analyzer oder aber ein Spektrum-Analyzer mit Tracking-Generator. Letzteres Prinzip ist in Abb. 5.2 dargestellt.

5.1.2 Verzerrungen

Nichtlineare Verzerrungen im Sinne des Klirrfaktors konnte man in der Analogtechnik bereits sehr lange messen. Man benötigt hierzu eine Klirrfaktormessbrücke, von der ein Beispiel in Abb. 5.3 zu sehen ist.

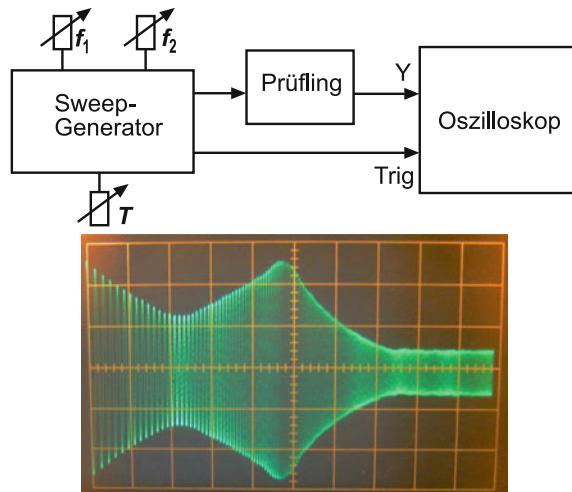


Abb. 5.1 Frequenzgangmessung auf herkömmliche Weise mit Sweep-Generator. Oben: Prinzipieller Aufbau. Unten: Resultat als Oszillosrogramm (Beispiel eines Frequenzgangs). Der Sweep-Generator gestattet die Einstellung der Start- und der Stopffrequenz, die Zeit für den Durchlauf sowie meist noch die Art des Sweeps (linear oder logarithmisch). Für jeden Durchlauf liefert er noch ein Triggersignal für das Oszilloskop. Alternativ könnte anstatt der Triggerung auch ein Signal für die X-Ablenkung verfügbar sein

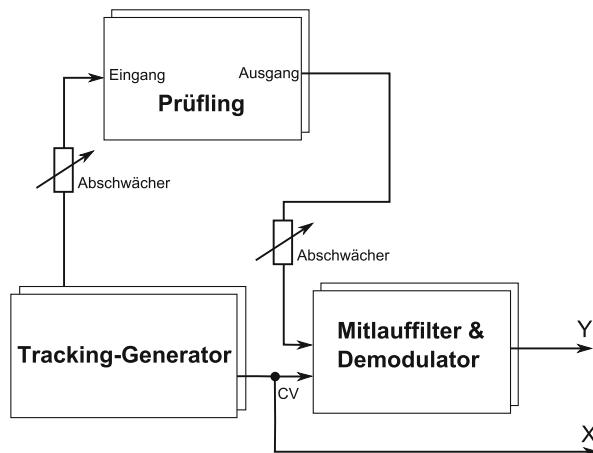


Abb. 5.2 Frequenzgangmessung mit Spektrum-Analyzator (und Tracking-Generator). Der Tracking-Generator gibt eine zur Frequenz proportionale Steuerspannung an den Eingang CV des Mitlauffilters. Die beiden Ausgänge führen auf die Horizontalablenkung und Vertikalablenkung eines Oszilloskops. Das Prinzip eines solchen Spektrum-Analyzers wird im HF-Bereich in modifizierter Form angewandt (statt des Mitlauffilters ein Mischer und ein bandbegrenzter Verstärker)

Abb. 5.3 Eine analog arbeitende Klirrfaktormessbrücke. (Foto: Hans M. Knoll)



Zur Bestimmung des Klirrfaktors erhält der Proband ein möglichst reines Sinussignal von definierter Amplitude, welches meist ebenfalls die Messbrücke bereitstellt. Der Ausgang des Probanden ist mit dem Eingang der Messbrücke zu verbinden. In der Stellung *kalibrieren* ist die Anzeige auf 100 % zu justieren. Nun schaltet man die Messbrücke auf *messen* und beide Steller sind solange zu justieren, bis die Anzeige optimal minimiert ist (der Messbereich ist dann ggf. zu verkleinern). Der verbleibende Anzeigewert ist der Klirrfaktor in Prozent.

Das Filter der Messbrücke sperrt bei optimalem Abgleich exakt die Frequenz des Messtones, alle anderen Signale können passieren. Normalerweise verbleiben nicht nur die restlichen Harmonischen, sondern auch andere Störgeräusche. Das Messergebnis ist demnach also THD & Noise (Summe der harmonischen Verzerrungen und Rauschen). Das Rauschen alleine könnte man nach Abschalten des Testsignals ermitteln, oder es wäre alternativ auch durch ein selektives Filter zu entfernen.

5.1.3 Pegelvergleichsmessungen

Unter der Rubrik *Pegelvergleich* finden sich u. a. Dynamikmessung, Geräuschabstandsmeßung und auch Vergleiche zwischen den Stereokanälen. Im Grunde ist hier lediglich ein Pegelmesser (NF-Millivoltmeter, Röhrenvoltmeter oder dergleichen) erforderlich. Der Proband wird hier mit einem Messton beaufschlagt, dessen Amplitude definiert ist und der den Referenzpegel (0 dB) darstellt. Schaltet man den Messton ab, so verbleibt der Pegel des Störgeräusches. Die Differenz beider Pegel entspricht nun dem Geräuschspannungsabstand, ebenfalls in dB. Während der Ermittlung des Referenzpegels herrscht ein Zustand, den man z. B. durch den Begriff *Vollaussteuerung* spezifiziert.

5.1.4 Gleichlauf

Bei der Messung des Gleichlaufs oder Wow & Flutter muss der Proband zunächst über ein Medium einen Referenzton wiedergeben, der üblicherweise bei 3000 oder 3150 Hz liegt. Bei einem Gerät, welches auch aufzeichnen kann, ist die Messung sozusagen über beide Wege – die Aufzeichnung und die Wiedergabe – messbar. Dabei ist zu beachten, dass die Maximalwerte der Abweichung bis auf das Zweifache gehen, als sie dies bei der reinen Wiedergabe eines Referenzmediums tun.

Gleichlaufschwankungen äußern sich nun formal durch eine Frequenzverschiebung und eine Frequenzmodulation. Der Blockschaltplan in Abb. 5.4 macht deutlich, wie daraus die Messwerte entstehen.

Der alleinige Wow-Wert spiegelt Schwankungen gegenüber der Referenzfrequenz wider, die langsamer sind als 10 Hz, während der Flutter-Wert die Schwankungen oberhalb 10 Hz repräsentiert. Der Drift-Wert stellt eine mehr oder weniger statische Abweichung dar.

Die Messung ist praktisch ausschließlich bei Geräten ein Qualitätskriterium, bei denen ein mechanischer Antrieb stattfindet und deswegen auch systembedingte Abweichungen durch Schlupf, Ungenauigkeiten und Schmutz vorkommen können. Heutige durch Quarz getaktete Wiedergabegeräte benötigen solche Messungen nicht mehr – die Werte sind so gut, dass sie praktisch unter die Messauflösung fallen.

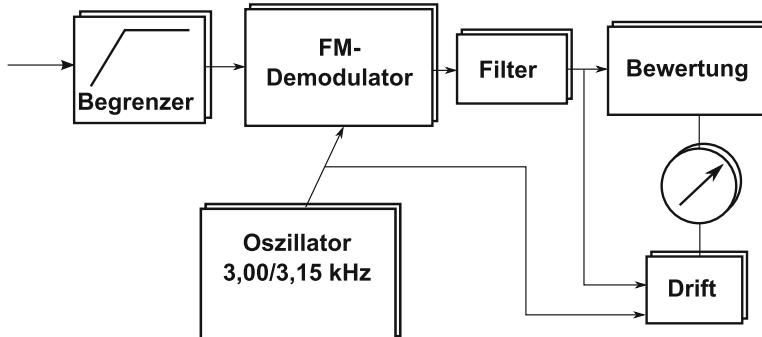


Abb. 5.4 Blockscheme eines Wow & Flutter-Messgerätes. Der vom Probanden kommende Signalton läuft über einen FM-Demodulator, der zur Demodulation auch den genauen Referenzton erhält. Der Momentanwert der gewonnenen Signalspannung ist ein Maß für die relative Abweichung gegenüber der Sollfrequenz, also damit auch ein Maß für die Abweichung der Wiedergabegeschwindigkeit (gegenüber der Aufnahmegeschwindigkeit). Dieses demodulierte Signal erfährt nun wiederum eine Bewertung über Filter, wodurch u. a. die Aufspaltung von Drift-, Wow- und Flutter-Wert stattfindet

5.1.5 Impuls-Messungen

Die Übertragung eines Impulses durch einen Verstärker oder eine andere Komponente lässt zunächst vermuten, es handele sich ebenfalls nur um die Betrachtung des Frequenzgangs im höheren Bereich. Jedoch bezieht sich die Messung des Frequenzgangs stets auf einen eingeschwungenen Zustand. Im Gegensatz dazu betrachtet man bei Impulsmessungen gerade die Zeit *während* des Einschwingens.

In Analogtechnik lassen sich beliebig geformte Impulse nur sehr schwer erzeugen. Meist beschränken sich Impulsgeneratoren auf wenige, mit Rechteckformen generierte Signale. Die Beurteilung erfolgt mit Oszilloskopen, die durch den Generator getriggert sind.

5.1.6 Korrelationsgrad

An dieser Stelle soll keine präzise Definition des Korrelationsgrades gegeben werden – eine grobe Beschreibung genügt (siehe auch Abschn. 5.2.5). Für die sog. *Mono-Kompatibilität* eines Stereosignals ist es von großer Wichtigkeit, dass die Ähnlichkeit beider Kanalsignale ausreichend ist, bevor man durch Addition das Monosignal erzeugt. Diese Ähnlichkeit besteht nur dann, wenn die sog. *Basisbreite* groß genug ist, wenn also viele Signalanteile auf beiden Kanälen identisch sind, und zwar bezüglich ihrer Phasenlage.

Man kann sich die Bildung des Korrelationsgrades in etwa so vorstellen wie die Bildung einer Leistungskurve aus der Multiplikation von Strom- und Spannungsverlauf. Auf das so gewonnene Signal wendet man eine zeitliche Mittelung an – eben ähnlich wie bei der Leistungsmessung. Statt auf Strom und Spannung wird auf die Signale aus beiden Signalen eine Multiplikation angewandt. Handelt es sich im einfachsten Falle um jeweils ein Sinussignal gleicher Frequenz, so ergibt sich bei Phasengleichheit ein positiver Maximalwert, bei Gegenphasigkeit zeigt der Korrelationsgrad negatives Maximum, und bei beliebigen Phasenwinkeln mit statistisch gleichverteiltem Charakter (also mit einem mittleren Winkel von 90°), so ergibt die Multiplikation und die zeitliche Mittelung den Wert null.

Bei einer reinen Laufzeit-Stereofonie besteht diese Ähnlichkeit ganz und gar nicht, weil die Signale auf beiden Kanälen mehr oder weniger phasenverschoben sind. Eine Verschmelzung zum Monosignal hätte fatale Folgen, denn je nach Frequenzlage käme es zu Auslöschen, und es läge ein kammartiger Frequenzgang vor (siehe Abschn. 1.9.2 und 3.3.3). Im Monosignal wären also möglicherweise viele essentielle Signalanteile verloren gegangen. Der Korrelationsgrad läge bei einem solchen Stereosignal dicht bei null.

Werden einzelne Instrumente dagegen als kohärente Signale mittels Panorama-Potentiometer auf beide Kanäle verteilt, und zwar ohne Extremstellung, so wäre der

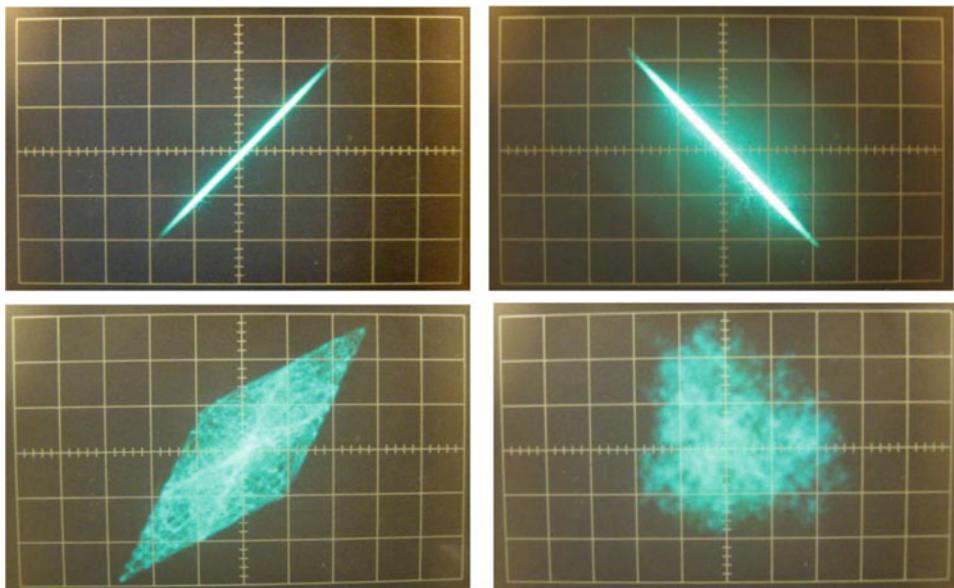


Abb. 5.5 Oszilloskop als Goniometer zur Darstellung der Korrelation. Oben links: Einzelnes Signal mit maximaler Korrelation, der Graph befindet sich im 1. und 3. Quadranten. Oben rechts: Dto, jedoch ist ein Kanal phasengedreht. Unten links: Mehrere Signale sind phasengleich auf beiden Kanälen (wie bei der reinen Intensitäts-Stereofonie), trotzdem erfolgt eine geringfügige Ausweitung in die Quadranten 2 und 4. Unten rechts: Keinerlei Phasenbezug zwischen den beiden Kanälen – entweder fehlende Kohärenz oder Laufzeitdifferenz zwischen den Kanälen

Korrelationgrad deutlich größer als null. Dies trifft bei der Intensitäts-Stereofonie, also XY- oder MS-Stereofonie zu.

Der gesamte Korrelationvorgang ist in analoger Technik nachzubilden. Statt der Anzeige eines Messwertes kann auch ein sog. *Goniometer* zur Beurteilung verwendet werden. Es handelt sich dabei eigentlich um ein Messgerät, das dem Oszilloskop ähnlich ist, allerdings ist die Bildröhre gegenüber letzterem um 45° im Gegenuhrzeigersinn verdreht. Ein zu beurteilendes Signal bezüglich seines Korrelationsgrades sei in Abb. 5.5 auf dem Schirm eines Oszilloskops gezeigt.

5.2 Digitale Messtechnik und PC-Systeme

Die folgenden Messverfahren beruhen auf der Weiterverarbeitung von digitalen Audio-Daten. Der Messaufwand ist normalerweise fast immer identisch, unabhängig von der zu bestimmenden Größe.

Bei digitalen Messmethoden findet die gesamte spezifische Signalverarbeitung meist auf Software-Ebene statt – nachdem das zu bearbeitende Signal in diskreter digitaler Form (also Samples) vorliegt.

5.2.1 Voraussetzungen für das Gelingen der Messungen

Für hochqualitative Messungen müssen Equipment und auch andere äußere Bedingungen gewisse Kriterien erfüllen, da andernfalls die Resultate fragwürdig sind.

Auflösung und Abtastrate Der interne oder externe Systemsound sollte eine Mindestauflösung von 16 Bit schaffen, besser sind 24 Bit. Beeinflusst werden durch diesen Wert möglicherweise die Dynamikgrenze, jedoch auch Messungen, bei denen Pegelvergleiche über einen weiten Bereich stattfinden, wie z. B. bei der Bestimmung von Klirrfaktor, Übersprechdämpfung, Störabstand und dergleichen.

Für die Abtastrate ist ein Mindestwert von 44,1 kHz erforderlich, sonst sind keine Messungen bis 20 kHz möglich. Höhere einstellbare Abtastraten von bis zu 96 kHz sind für ein Messsystem empfehlenswert.

Frequenzgang Der Eigenfrequenzgang des Systems sollte über den Bereich von 10 Hz bis 20 kHz möglichst linear und auch bei beiden Kanälen gleich sein. Vor allem zur Messung des Frequenzgangs eines Probanden ist dies wünschenswert, obwohl mal gewisse Schwankungen – wie wir weiter unten sehen werden – durch Ausgleichskurven kompensieren kann.

Grundgeräusche und Störspannungen Bei einer Eingangsspannung von 0 V – also bei kurzgeschlossenem Eingang des Audiosystems – sollte idealerweise bei der Aufzeichnung kein Störsignal digitalisiert werden. Eine Probe, wie gering diese Störung ist, sollte man grundsätzlich vor jeder Messung durchführen (Abb. 5.6).

5.2.2 Frequenzgang nach digitalem Verfahren

Die Wandlung diskreter Werte vom Zeitbereich in den Frequenzbereich erfolgt normalerweise durch Fast-Fourier-Transformation, obwohl bei der Generierung des eigenen Messtones auch eine Echtzeit-Pegelmessung genügen würde. Allerdings wäre dann ein Proband mit Laufzeit (z. B. Tonbandmaschine mit Hinterbandkontrolle) für die Methode nicht geeignet (Abb. 5.7).

Häufig erlauben solche Software-Systeme die Korrektur des Eigenfrequenzgangs durch Ausgleichskurven. Diese Methode zeigt uns Abb. 5.8.

5.2.3 Verzerrungen nach digitalem Verfahren

Verzerrungen nach digitalen Verfahren beruhen auf der Bewertung durch die Spektrum-Analyse. Unterschiedliche Parameter sind hier variierbar (Frequenz, Spannung bzw. Leistung). Der Proband befindet sich zwischen Ein- und Ausgang des Audiosystems des

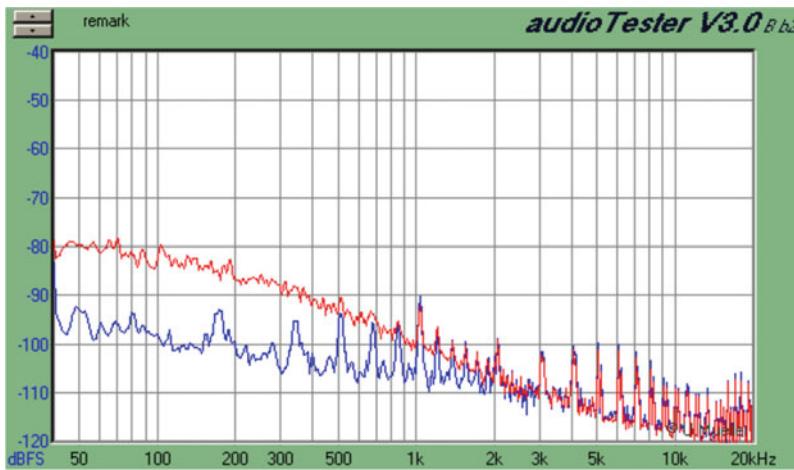


Abb. 5.6 Reststörspannung bei Null-Eingangsspannung. Definiert man den Eingangsspeigel für Vollaussteuerung mit 0 dB, so stellen die vorliegenden Spektren die Eingangsstörpegel beider Kanäle dar, und zwar für die rote Kurve offener Eingang rechter Kanal sowie für die blaue Kurve der mit GND kurzgeschlossene linke Kanal (S. 204 [1])

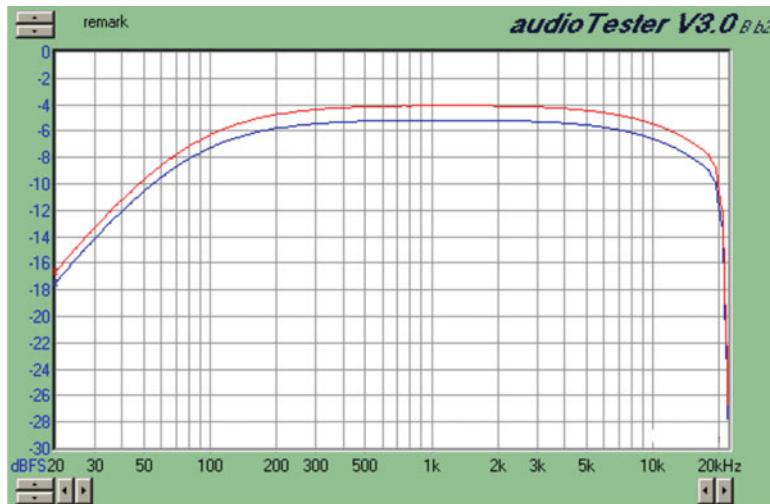


Abb. 5.7 Beispiel eines Frequenzgangs. Hier wurde die Messstrecke des Computers selbst bzw. der Sound-Karte zur Prüfung herangezogen. Man erkennt deutlich den kleinen Pegelunterschied der beiden Kanäle untereinander, der sich jedoch durch den gesamten Frequenzbereich zieht. Man kann die Einbrüche bei niedrigen und sehr hohen Frequenzen nicht zuordnen – sie können sowohl beim Ausgang als auch beim Eingang entstehen (S. 204 [1])

Rechners und wird durch dieses auch begrenzt. Man muss sich stets im Klaren darüber sein, wie groß die Reststörgrößen der Soundkarte selbst sind, um auf eine sichere Aussage zu kommen (Abb. 5.9, 5.10 und 5.11).

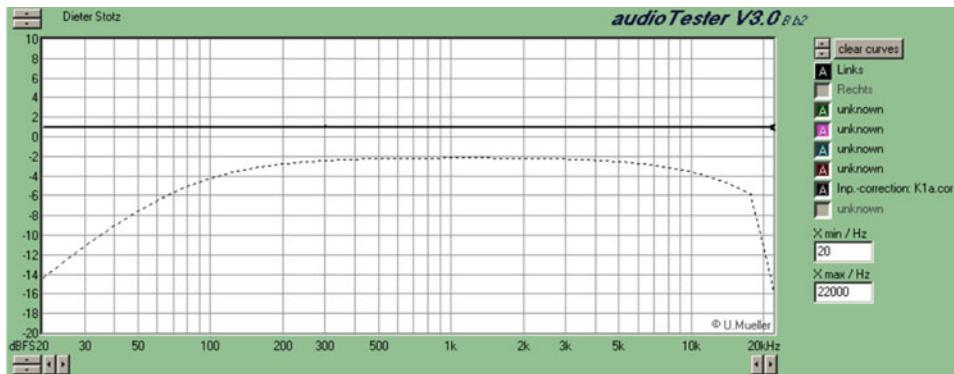


Abb. 5.8 Kompensation des Eigenfrequenzgangs des Soundsystems durch eine Ausgleichskurve. Die gestrichelte Kurve ist der ursprüngliche Frequenzgang des Systems, und diese dient intern gespiegelt als Korrekturkurve. Somit erhält man bei erneuter Aufzeichnung des Frequenzgangs einen nahezu horizontalen Verlauf. Man kann die Korrektur entweder für den Ausgang oder für den Eingang verwenden, so dass man optimalerweise nur auf den Bereich wirkt, der eine Frequenzgangsverzerrung aufweist (S. 204 [1])

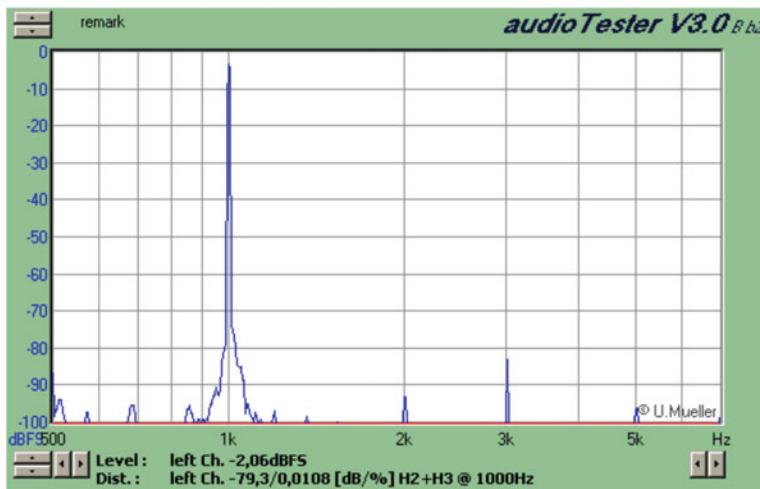


Abb. 5.9 Verzerrungsmessung an einem Präzisions-Sinusgenerator. Man kann hier deutlich zwischen K_2 und K_3 unterscheiden. In der Zeile unter dem Graphen ist der Gesamtwert von K_2 und K_3 angegeben (S. 204 [1])

5.2.4 Pegelvergleichsmessungen nach digitalem Verfahren

Die Messung der Pegelgleichheit soll prüfen, ob einzelne Kanäle hinreichend gleiche Signalstärke liefern. Die Frequenzgangsmessung deckt diese relativ triviale Prüfung eigentlich bereits ab (siehe Abb. 5.7). Der Pegelunterschied unter den Kanälen wird für eine bestimmte Frequenz angegeben. Weist das zu prüfende System einen Pilotgenerator auf,

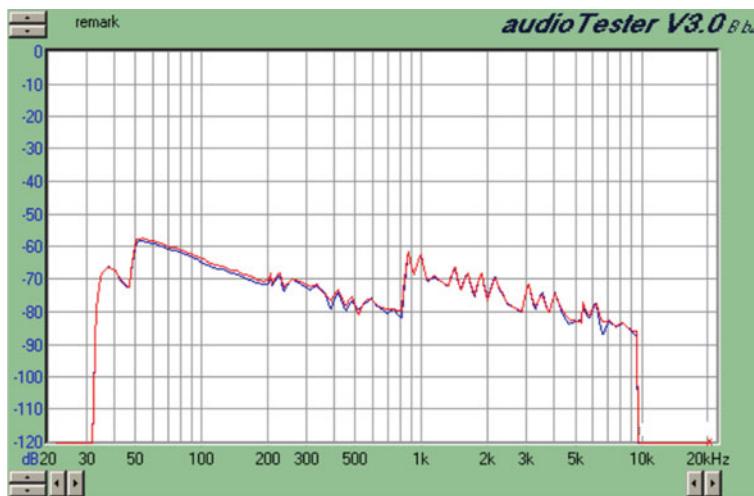


Abb. 5.10 Verzerrungs-Messung an einem Verstärker: Klirrfaktor über der Frequenz (S. 204 [1])

so genügt ebenfalls die Messung des Spektrums bei Betrachtung der Spitze bei der Pilotfrequenz.

5.2.5 Korrelations-Messungen nach digitalem Verfahren

Mit g_i = Werte Signal 1, h_i = Werte Signal 2 wird der Korrelationsgrad beider Signale zueinander:

$$\kappa = \frac{\sum_{i=0}^n [(g_i - \bar{g}) \cdot (h_i - \bar{h})]}{\sqrt{\sum_{i=0}^n (g_i - \bar{g})^2 \cdot \sum_{i=0}^n (h_i - \bar{h})^2}} \quad (5.1)$$

Da bei Signalen ohne Offset der Mittelwert bei hinreichend langer Probe gegen null geht, vereinfacht sich die Gleichung zu:

$$\kappa = \frac{\sum_{i=0}^n (g_i \cdot h_i)}{\sqrt{\sum_{i=0}^n (g_i^2) \cdot \sum_{i=0}^n (h_i^2)}} \quad (5.2)$$

Digital vorliegende Audiodaten sind prädestiniert für Analysen, die nach obiger Gleichung abzulaufen haben. Mehr noch, sie sind bestens geeignet für eine Echtzeitberechnung mittels DSP. Eine geeignete Software hierfür ist beispielsweise Sample Champion. Ein Screenshot zeigt Abb. 5.12 und Abb. 5.13. S. 204 [2]

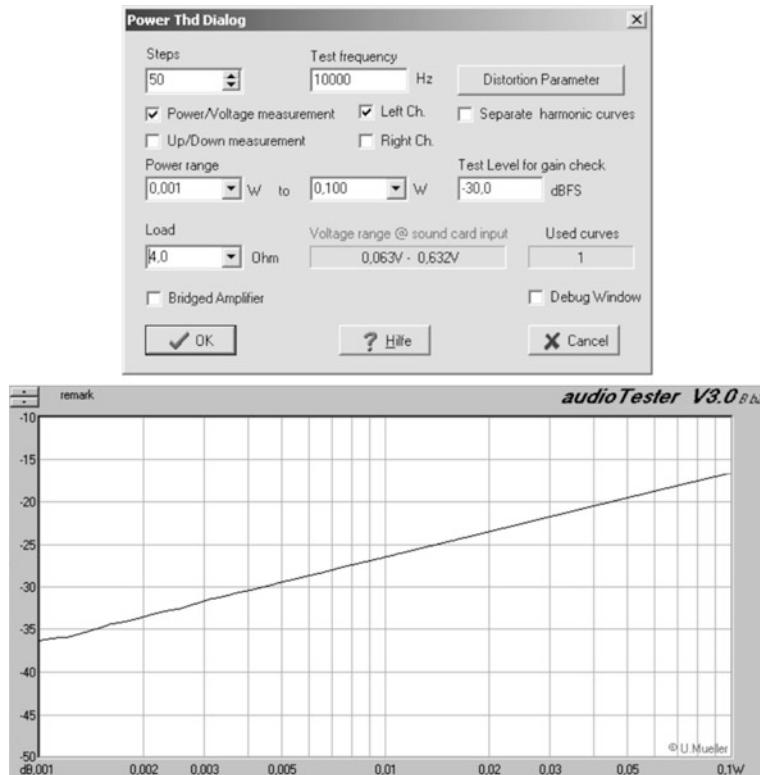


Abb. 5.11 Verzerrungsmessung an einem Verstärker: Klirrgrad über der Aussteuerung. Dieser sehr gleichmäßig ansteigende Verlauf ist eigentlich nicht der übliche Praxisfall, denn üblicherweise zeigt sich ein ab einem bestimmtem Pegel ein rapide einsetzendes Clipping und damit sprunghaft ansteigende Klirrgrade. Hier wurde einfach eine Siliziumdiode über einen Widerstand angesteuert (S. 204 [1])

5.2.6 Gleichlauf nach digitalem Verfahren

Handelt es sich um eine aufzeichnende Bandmaschine, so kann zunächst der Referenzton zur Aufnahme kommen. Zu beachten sind hier wiederum die einzustellenden Parameter, vor allem die korrekte Messfrequenz. Die Software digitalisiert das Ausgangssignal und wertet nach wie in Abschn. 5.1.4 besprochenem Verfahren aus (Abb. 5.14).

5.2.7 Impuls-Messungen nach digitalem Verfahren

Die Software generiert auf sehr flexible Art und Weise unterschiedlichste Impulsarten und -formen. Nach Verlassen des Prüflings gelangt der Impuls in den Audio-Eingang des Rechners. Als Auswertung zeigt das Programm das Bild in zeitlicher und frequenzmäßiger Auflösung, siehe Abb. 5.15.

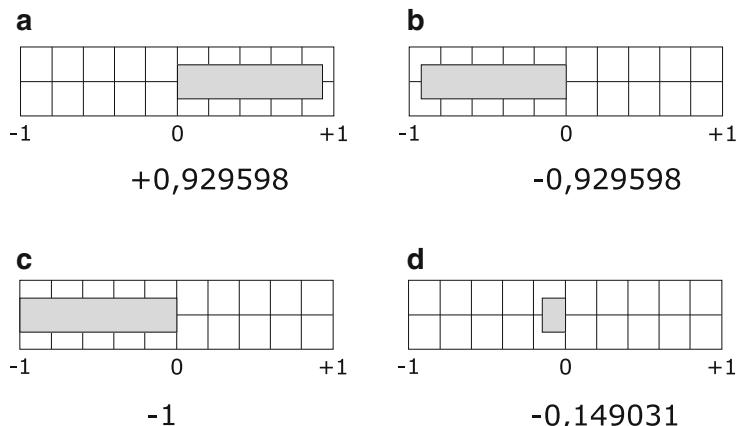


Abb. 5.12 Eine Stereo-Audiodatei kann mittels Software auf den Korrelationsgrad hin untersucht werden. (a) liegt bei einem normalem Stereosignal mit hohem Basisbreitenanteil vor, (b) ebenfalls, jedoch ist dort einer der Kanäle umgepolt (Phasendrehung um 180°), (c) zeigt ein reines Monosignal und (d) ein Signal, bei dem ein Kanal aus einer 90° -Phasendrehung (Hilbert-Transformation) des anderen Kanals entstanden ist (S. 204 [2])

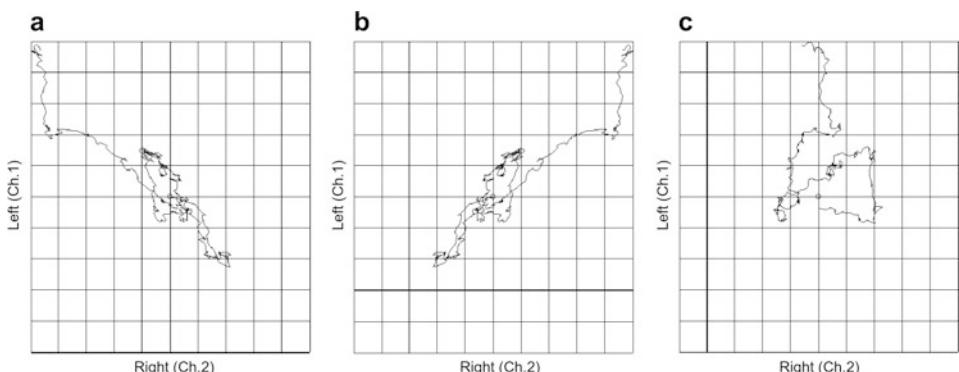


Abb. 5.13 Die Signale von Abb. 5.12 im Goniometer, welches die Auslenkungen der beiden Kanäle orthogonal darstellt (ähnlich XY-Betrieb beim Oszilloskop). Vgl. Abb. oben: (a) normales Stereosignal mit viel Basisbreite, (b) dto, jedoch Phasenumpolung eines Kanals und (c) Hilbert-Transformation eines Kanals (bei identischen Ursprungssignalen) (S. 204 [2])

5.2.8 Raumakustik-Messungen nach digitalem Verfahren

Zur Beurteilung eines Raumes in Bezug seiner akustischen Eigenschaften stehen zahlreiche Messmethoden zur Verfügung. Eine Methode zeigt die Nachhall-Eigenschaften, der *Schroeder-Plot*. Lautsprecher und Mikrofon sollten weit von reflektierenden Flächen entfernt und gleichzeitig möglichst nahe beieinander sein. Ein Schallimpuls des Lautsprechers breitet sich aus bis zum Mikrofon, wird an Wänden usw. reflektiert und als Hallsignal eine gewisse Zeit registriert (Abb. 5.16).

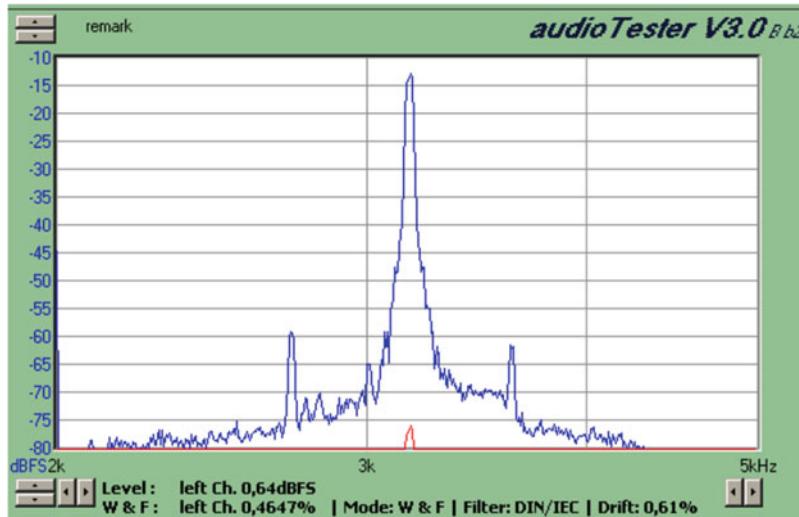


Abb. 5.14 Parameter und Ergebnis bei einer Wow & Flutter-Messung an einem Tonbandgerät. Der Drift-Wert gibt an, wie weit die gemessene (mittlere) Frequenz von der Referenzfrequenz abweicht (S. 204 [1])

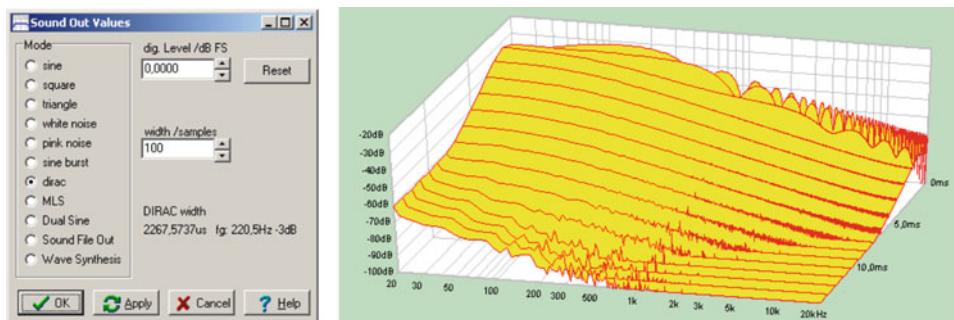


Abb. 5.15 Einstellung eines Dirac-Impulses durch Angabe der Samples und Ergebnis im 3D-Diagramm. Die Länge des Dirac-Impulses in Samples sowie seine Höhe sind einstellbar. Die Software bietet noch vielerlei anderer Impulse, deren Spektrum und zeitlicher Verlauf am Ausgang des Probanden anstehen (S. 204 [1])

5.2.9 Kompaktes Messlabor Red Pitaya

Viele verschiedene Einplatinencomputer sind auf dem Markt zu finden, wovon hier eines etwas näher betrachtet werden soll. Das Gerät heißt *Red Pitaya* und ist ein Produkt einer gleichnamigen slowenischen Firma (Abb. 5.17). Es besteht hauptsächlich aus zwei Messblöcken und zwei Generatorblöcken und kann folgende Daten aufweisen:

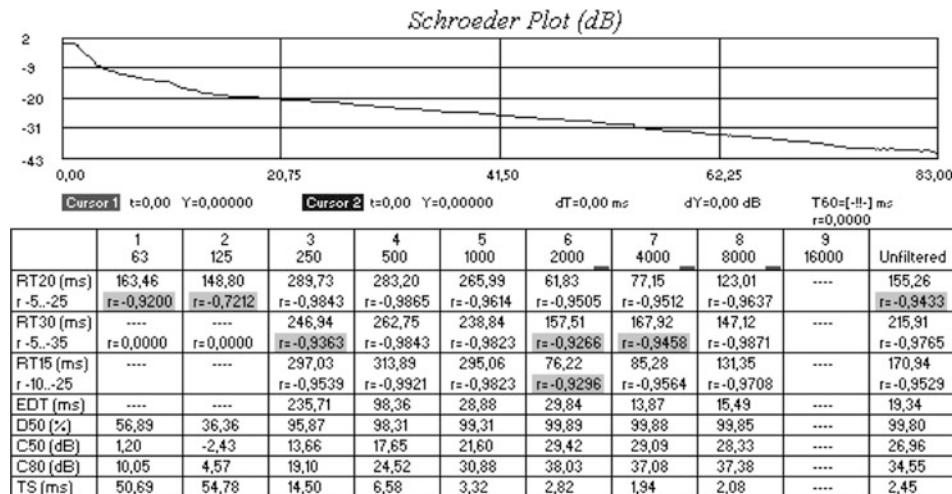


Abb. 5.16 Der Schroeder-Plot zeigt den Amplituderverlauf eines im Raum abklingenden Impulses. Die Zahlenwerte geben Aufschluss über die Abklinggeschwindigkeit. So bedeutet z. B. RT30 diejenige Zeit, in der die Hallamplitude um 30 dB abgeklungen ist (S. 204 [2])

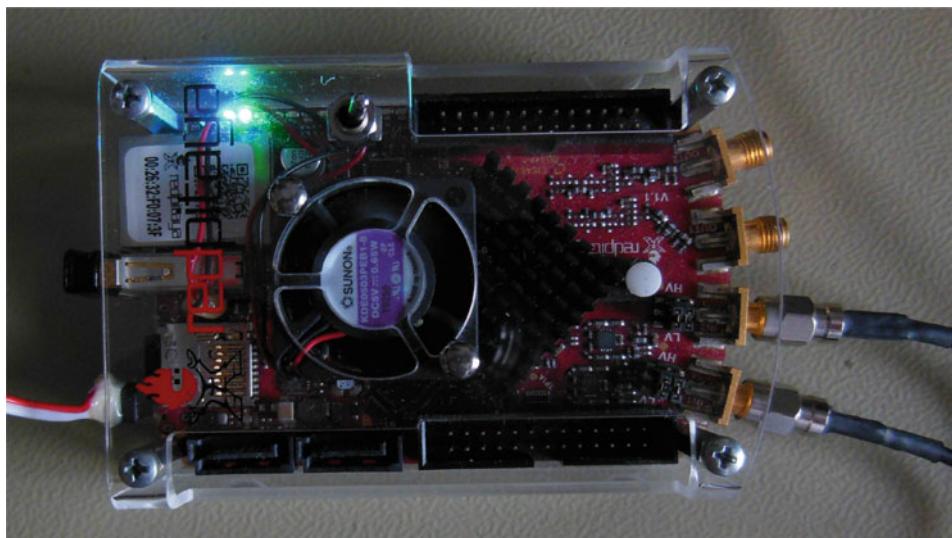


Abb. 5.17 Kompaktes Messmodul Red Pitaya. Ganz links ist das USB-WLAN-Steckmodul zu sehen. Ganz rechts sind vier SMA-Anschlüsse sichtbar, jeweils zwei für Messeingänge und Generatorausgänge. Oben und unten sind zwei weitere zweireihige Pfostenstecker für Zusatz Eingang und Ausgang vorhanden

Auflösung: 14 Bit

Takt-/Abtastrate: 125 MS/s

Bandbreite: DC – 50 MHz (–3 dB)

Zusätzliche Eingangskanäle: 4 mit 100 kS/s und Auflösung 12 Bit

Zusätzliche Ausgangskanäle: 4 mit PWM 4 ns

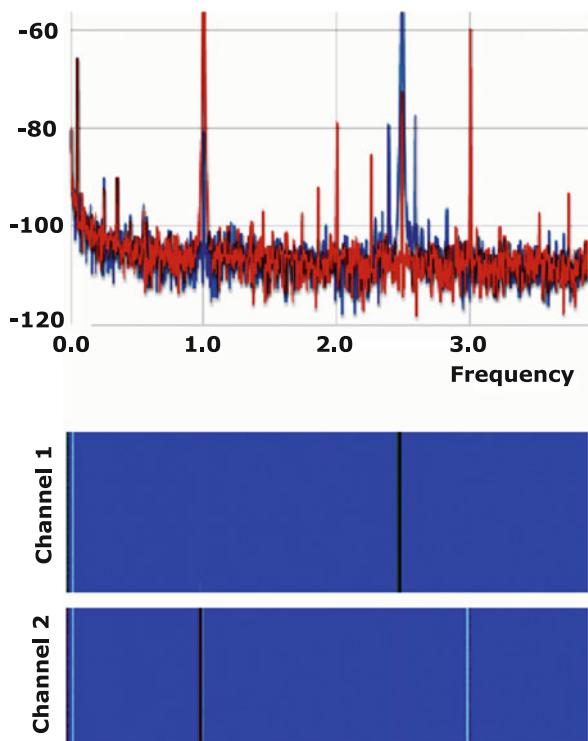
Zusätzliche Eingangs-/Ausgangspins: 16

Im Gegensatz zur üblichen Auflösung von 8 Bit bei Oszilloskopen liegt hier eine Auflösung von 14 Bit vor. Für Anwendungen der Spektralanalyse bedeutet dies, dass eine Darstellung von über 80 dB kein Problem ist. Allerdings setzt dies besondere Eigenschaften der Versorgung durch das Netzteil voraus: Häufig sind Störungen bereits weitaus höher als die kleinste Messgröße bei Verwendung von Schaltnetzteilen, sodass das Gerät optimalerweise durch ein kontinuierlich geregeltes Netzteil zu versorgen ist.

Das Gerät greift auf einen Webdienst zu, womit es seine Daten der jeweiligen Applikation auf jedem Computer oder Mobilgerät (Smartphone, Tablet oder Laptop) zur Verfügung stellen kann. Die Netzwerkverbindung kann entweder drahtlos über einen optionalen WLAN-Adapter erfolgen oder über eine LAN-Schnittstelle. Ferner besteht die Möglichkeit, auf die Daten des Gerätes über eine Micro-USB-Schnittstelle zuzugreifen, wobei dann ein Terminalprogramm und spezielle Befehle zu verwenden sind.

Mit dem Gerät lassen sich – durch verschiedene Programme – allerhand viele Messaufgaben bewältigen, sowohl im Zeitbereich als auch im Frequenzbereich. In Abb. 5.18 sei lediglich ein Beispiel für eine zweikanalige Spektrumanalyse dargestellt. S. 204 [3]

Abb. 5.18 Eine der vielen Applikationen, die mit dem RP möglich sind, in diesem Falle ein zweikanaliger Spektrumanalyzer



Literatur

1. Software *audioTester 3.0* von Ulrich Müller.
2. Software *Sample Champion PRO* von Paolo Guidorzi.
3. Stotz, D.: Ist RedPitaya das bessere Oszilloskop?
ELEKTRONIKPRAXIS Sonderheft Messtechnik 2015.



Vergleich mit herkömmlicher Audiotechnik

6

6.1 Schneidetechniken

Band- und Festplattenaufzeichnungen zeigen krasse Unterschiede in Bezug auf Schneidetechniken, was sich sowohl in der Arbeitsweise, der Endqualität und im Zeitaufwand niederschlägt.

6.1.1 Aufzeichnungen auf Analogband

Aufzeichnungen auf Band können grundsätzlich „geschnitten“ werden; der Ausdruck *Schnitt* ist beim Analogband durchaus wörtlich zu nehmen, sofern keine Überspieltechnik zur Anwendung kommen soll. Für letzteres sind ja mindestens zwei Systeme notwendig, die synchron laufen müssen (man kann sich leicht vorstellen, dass z. B. zusammenzustellendes Material vom Band des Zuspielers nur bei kontinuierlichem Lauf beider Maschinen überspielt werden kann, weil ein Start an der Schnittposition mit zeitlichen Verzögerungen bzw. Beschleunigungseffekten verbunden ist und deshalb nur unbefriedigende Resultate liefern kann). Einen kleineren technischen Aufwand bedeutet der physische Schnitt des Bandes; das Auffinden der relevanten Stellen hingegen erfordert dennoch etwas Übung. Eine derartige Prozedur, die spezielles Werkzeug und Klebebänder erfordert, lässt man natürlich nicht dem Originalband angedeihen, sondern einer Arbeitskopie davon.

In der Regel ist ein solcher Schnitt nur bei Live-Aufnahmen üblich; im Studio erfolgt meist eine Aufzeichnung, bei der die einzelnen Musiker zeitlich versetzt auf eine Mehrspurmaschine aufgenommen werden – der Schnitt entfällt in diesem Fall, weil er nicht nötig und im Prinzip auch nicht spurdiskret möglich ist.

Die Genauigkeit eines mechanischen Schnitts lässt zu wünschen übrig; sie hängt vom Gehör des Technikers und auch von der Bandgeschwindigkeit ab. Es gibt zwar Methoden,

akustisch markante Stellen auch sichtbar zu machen, davon wird jedoch i. Allg. kein Gebrauch gemacht, weil sie sehr umständlich sind.

6.1.2 Aufzeichnungen auf Computer-Festplatte

Diesen oben genannten Beschränkungen obliegt der Schnitt an einer Sampling-Datei nicht. Zwar wird man hier auch nicht am einzigen Original arbeiten, eine angefertigte Kopie jedoch entspricht genau der gleichen Qualität, weil die Vervielfältigung von digitalen Daten bei intaktem Datenträger keinerlei Verlust zeigt.

Die Vorgehensweise des Schnitts selbst ist denkbar einfach: wie bei einem Textverarbeitungssystem sind die Arbeitsschritte Kopieren, Ausschneiden, Einfügen, Löschen und Überschreiben möglich. Teile, die einmal im Kopierpuffer liegen, können beliebig oft und an beliebigen Stellen eingefügt werden. Die Überschreibe-Operation ist auch mit einer Option durchführbar, die ein Mischen mit dem bereits bestehenden Inhalt erlaubt. All diese Operationen sind Software-Bestandteile und können auf dieser Basis auch jederzeit ergänzt werden.

Ein Schnitt der hier beschriebenen Weise ist selbst bei Mehrspuraufnahmen spurdiskret möglich, d. h. bei einer Aufzeichnung, die beispielsweise aus 8 Parallelspuren besteht, ist die Bearbeitung einer einzelnen Spur ohne weiteres durchführbar.

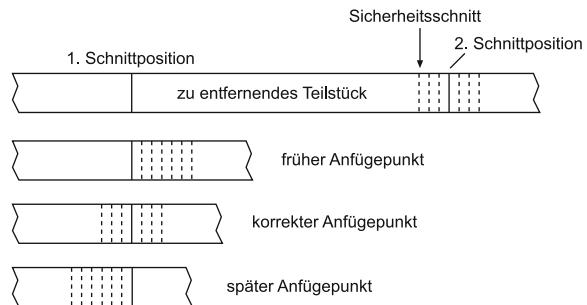
Die mögliche Genauigkeit des Schnitts ist sehr hoch, weil das Signal nicht nur hörbar, sondern auch dessen Verlauf sichtbar ist. Die Auflösung entspricht i. Allg. der Sampling-Abstände, das bedeutet bei 44 100 Hz etwa $23\ \mu\text{s}$. Selbstverständlich können auch hier beim Schnitt Fehler entstehen; nach dem Kriterium des Höreindrucks besteht jedoch die komfortable Möglichkeit, automatisch, halbautomatisch oder manuell die Schnittposition zu verändern. Nach dem Auffinden des optimalen Ergebnisses muss lediglich der Vorgang angehalten werden. Abb. 6.1 demonstriert diese Vorgänge. Es handelt sich um ein Vorhaben, ein Teilstück einer Aufnahme herauszutrennen und die offenen Enden wieder zusammenzufügen. Die 1. Schnittposition muss nicht unbedingt sehr genau getroffen werden (günstig ist jedoch eine leise oder stumme Stelle), wenn ab der 2. Schnittposition eine Wiederholung stattfindet. Die Aufzeichnung erfährt zunächst einen virtuellen Schnitt an einer sicheren frühen Stelle der 2. Schnittposition. In mehreren Schritten erfolgt nun ein Anfügen des rechten Teilstücks an das linke, bis ein optimaler Übergang gefunden ist.

Die Abstufung, mit der der Schnitt vorgenommen wird, ist variabel, so dass eine optimale Annäherung möglich ist. Außerdem sind Schnittkanten-Interpolationen einstellbar, so dass Knackgeräusche stark unterdrückt werden (siehe Abschn. 3.1.2).

6.1.3 Digitale Bandaufzeichnungen

Digitale Aufzeichnungen auf Band lassen sich normalerweise nicht mittels physischem Schnitt bearbeiten, weil erstens Schrägspräaufzeichnungen (DAT, Video 8 usw.) im

Abb. 6.1 Stufenweise Verschiebung beim Schnitt



Allgemeinfall eine Diskontinuität an der Schnittposition bedeutet und zweitens bei längsaufzeichnenden Digitalsystemen (DCC) ein Sprung in der Kontrollspur entsteht. Über Umwege ist dennoch ein Schnitt möglich, der jedoch eine spezielle elektronische Ausstattung erforderlich macht; ein solcher Schnitt kann demnach nicht am Material selbst ausgeführt werden, sondern erfordert wieder die Synchronisation mehrerer Systeme.

6.2 Filtertechniken

Das Filtern soll hier als nachträglicher Bearbeitungsschritt verstanden werden, der den Klang der Gesamtaufnahme oder einzelner Spuren verbessern oder verfremden soll. Oftmals ist damit auch eine Unterdrückung von Störgeräuschen erstrebenswert.

6.2.1 Filtern mit Spezialgeräten

Auf reiner Hardware-Basis geschieht das Filtern mit dem Grafik-Equalizer, dem parametrischen Equalizer und sonstigen Spezialfiltern. Je universeller die Einsatzgebiete für diese Filter sein sollen, desto aufwendiger sieht deren Realisierung aus.

Jede Filterung erfordert eine weitere Überspielung, die eine nicht vernachlässigende Verschlechterung bedeutet, falls der Vorgang sich auf Analogebene vollzieht. Ein spezieller Bereich der Aufzeichnung lässt sich nur mit Zusatzeinrichtungen separat bearbeiten; hier sich elektronische Schalter erforderlich, damit die Ein- und Abschaltung des Filters keine Knackgeräusche verursacht.

Der Echtzeiteinsatz eines physischen Filters macht i. Allg. keine Probleme, so dass beispielsweise auch bereits während der Aufnahme eine entsprechende Bearbeitung möglich ist.

Diese Hardware unterliegt einer gewissen Ausfallrate, und ein Ersatz ist immer mit Zeit- oder Geldaufwand verbunden.

6.2.2 Software-Filterung

Bei der Filterung einer Sampling-Datei genügt normalerweise eine übliche Bearbeitungs-Software; spezielle Filterungen lassen sich auch programmieren. Die Qualität der Filterung hängt nicht von der Hardware ab, sondern nur von der Filterimplementierung der Software. Natürlich ist das Argument, Software sei ebenfalls eine Ware und koste somit Geld, nicht ganz von der Hand zu weisen, jedoch liegt hier das Preis-Leistungsverhältnis zugunsten der Software. Bei Mehrplatzlizenzen oder gar kopierfreier Software erweitert sich der Sparfaktor noch mehr. Ein Ausbau des Filterprogramms lässt neue Möglichkeiten entstehen, ohne dass der Gerätelpark wächst. Nach diesen Ausführungen liegt der Vorteil klar auf der Seite der Software.

Diskrete Bereiche der Sampling-Datei können ohne weiteres separat bearbeitet werden, ohne dass davon die restliche Aufzeichnung betroffen wäre; ein korrekt arbeitendes Filterprogramm berücksichtigt ggf. auch die Nachbearbeitung der Nahtstellen, so dass auch keine Störungen an den Übergängen zu befürchten sind. Weiterhin ermöglicht die Kennzeichnung mehrerer Bereiche der Datei eine Filterung mit verschiedenen, individuellen Voreinstellungen (Presets) der Filtereigenschaften in einem einzigen Durchgang.

Für die Echtzeitbearbeitung ist eine reine Software-Lösung nicht geeignet; sie macht spezielle Hardware (Signalprozessoren) erforderlich.

Bei einem Ausfall des Computers ist meist schneller eine Ausweichmöglichkeit gefunden, als dies bei Verwendung von Hardware-Filtern der Fall ist.

6.2.3 Digitale Hardware-Filterung

Mit spezieller Hardware lässt sich auch in Echtzeit eine digitale Filterung durchführen. Hierzu eignen sich digitale Signalprozessoren (DSP), die über ein besonderes Programm gesteuert werden. S. 216 [1]

Sofern FIR- oder IIR-Filter nicht auf Hardware-Basis implementiert sind, muss der DSP über einige MAC-Operationen ein entsprechendes Filter emulieren (siehe Abschn. 3.2).

Bei weitem nicht alle Audiokarten verfügen über einen DSP, der eine Echtzeitfilterung erlaubt; außerdem ist natürlich ein entsprechendes Filterprogramm notwendig, falls keine feste Filterverdrahtung vorliegt. Geeignete Entwicklerwerkzeuge zum Ansprechen der Karte und zum Programmieren des Signalprozessors geben meist die notwendige Unterstützung.

Ein derartiger Aufwand lohnt nur, wenn in Echtzeit (also während der Aufzeichnung) gefiltert werden soll. Eine Nachfilterung kann jeder Prozessor – auch der des Computers – durchführen.

Eine Steuerung des Filters in Echtzeit ist ebenfalls denkbar, so dass an vorher zu definierenden Zeitpunkten das Filter seine Eigenschaften (Parameter bzw. Charakteristik)

ändert. Eine solche Kontrolle kommt jedoch selten vor, da bei einer Live-Darbietung die definierten Zeitpunkte noch nicht genau festliegen.

6.3 Einsatz von Effekten

Sollen Effekte wie Hall, Echo, Phasing usw. Verwendung finden, so ergeben sich grundsätzlich die gleichen markanten Unterschiede wie beim Filtern. Wir können uns hier also weitere Ausführungen sparen. Bei analoger Technik ist lediglich zu ergänzen, dass hier für jede Art von Effekt eine separate Schaltung notwendig ist. Eine generelle Mehrausnutzung von Hardware-Komponenten ist nur bei digitalen Geräten denkbar. Ebenfalls steigt der Software-Aufwand zur Realisierung verschiedener Effekte nur unerheblich an.

Im Zusammenhang mit dem Effekt Hall ist zu bemerken, dass zwar bei der herkömmlichen mehrere Möglichkeiten zur Realisierung bestehen; sie sind jedoch durchweg sehr umfangreich und teuer, wenn die Resultate zufriedenstellend sein sollen (Hallplatte und -folie, Hallraum, Torsionsfederhall usw.). Die Verwirklichung mit moderner Technik kann entweder auf hochintegrierten Schaltkreisen beruhen (z. T. sind nur Speicherbausteine notwendig), oder sie geschieht mittels Nachbearbeitung auf Software-Basis, wobei relativ einfache Programmalgorithmen notwendig sind; eine Qualitätssteigerung des Halleindrucks macht i. Allg. nur eine größere Rechenzeit erforderlich – die Anschaffungskosten bleiben unverändert.

6.4 Sonstige Signalmanipulationen

Eine Vielzahl an Veränderungen lassen sich nicht direkt irgendeiner Effektbearbeitung zuordnen. Da herkömmliche Technik aber praktisch alle Effektmöglichkeiten abdecken kann, bleibt die Frage, ob sie auch weitere Bearbeitungsarten unterstützt.

Es ist naheliegend, dass bei der Notwendigkeit komplizierter Algorithmen oder bei der Bearbeitung weniger, zeitbegrenzter Signalwerte die herkömmliche Technik versagt oder dass man zumindest mit ihr einen abnormalen Aufwand betreiben muss, um akzeptable Ergebnisse zu erzielen.

6.4.1 Rechenintensive Veränderungen

Spezielle Bearbeitungsmethoden sind Pitch- oder Zeitveränderungen des Signals, wobei natürlich jeweils exklusive Veränderungen stattfinden sollen, d. h. die Tonhöhe eines Signals wird verändert unter Beibehaltung des Zeitraums und umgekehrt.

Wir konnten bereits in Abschn. 3.3.6 bzw. 3.3.7 erfahren, wie solche Manipulationen bewerkstelligt werden können und in welcher Weise das Signal hier einer Bearbeitung unterliegt. Es handelt sich um komplizierte Algorithmen, die nur mit diskreten Signalwerten

vorzunehmen sind; diese Forderung kann lediglich mit digitalisierten Signalen erfüllt werden.

Komplizierte Geräte mit Signalprozessor können gar in Echtzeit Pitch-Veränderungen vornehmen (Harmonizer) – eine Zeitdehnung ist jedoch nur mit zusätzlichem Zwischen-speicher denkbar. Bei der Non-Realtime-Verarbeitung sind beide Vorgänge nur mit einem entsprechenden Programm zu bewerkstelligen.

Will man sich entsprechende Bearbeitungsvorgänge mit einem Analogsystem vorstellen, so besteht das Handicap vor allem darin, dass sich Analogsignale nur zeitdiskret mit geeigneten Analogprozessoren verarbeiten lassen. Dagegen sind Operationen, bei denen viele Einzelwerte heranzuziehen sind, in der Praxis kaum auf analoger Ebene machbar.

6.4.2 Diskrete Signalverformung

Die Bearbeitung eines Signals an einer diskreten Stelle nach visuellen Kriterien macht die Darstellung auf einem Monitor notwendig. Einerseits erleichtert dies in Verbindung mit einem Bearbeitungsprogramm die zeitliche Zuordnung, und andererseits ist die Veränderung des Signalverlaufs bei einem engen Zeitintervall durch Computerbearbeitung dadurch vereinfacht, dass sich die Darstellung durch Zoom-Funktionen auf den zu bearbeitenden Bereich konzentrieren lässt und jegliche Veränderungen – z. B. bewerkstelligt durch die Maus – sofort sichtbar werden und beliebige Korrekturen jederzeit anzubringen sind.

Im Gegensatz dazu lassen sich Signalverläufe bei Analogsystemen nur unter bestimmten Voraussetzungen derart diskret verändern. Zunächst stellt sich das Problem der einwandfreien Ortung einer Position der Aufzeichnung. Der Istzustand des Signals lässt sich noch relativ einfach mittels Oszilloskop darstellen. Eine Veränderung kann jedoch nur schwer vollzogen werden; hier ergibt sich die Problematik der unmittelbaren Darstellung, und – was u. U. noch wichtiger ist – die erfolgten Veränderungen wieder rückgängig (sog. Undo-Funktion) zu machen, ist praktisch unmöglich.

6.4.3 Zeitinversion

Das Spiegeln des zeitlichen Ablaufs eines Signals nennen wir hier Zeitinversion. Sie wird in der modernen Musik häufig als Ausdrucksmittel eingesetzt.

Bei der Signalverarbeitung auf dem Computer kann aus einer Sampling-Datei ohne weiteres eine Kopie angefertigt werden, bei der jedoch der Zeitablauf gespiegelt ist (siehe auch Abschn. 3.3.1). Auch einzelne Bereiche lassen sich auf diese Weise verändern, und bei mehrspurigen Sampling-Dateien sind sogar spur- und bereichsdiskrete Bearbeitungen dieser Art möglich.

In der Analogtechnik wird zur Realisierung einfach ein Band umgedreht bzw. die Abspielrichtung umgeschaltet. Verdrehungen, die nur innerhalb eines bestimmten Bereiches auftreten sollen, können dagegen nur unter Zuhilfenahme von Schneidetechniken

bewerkstelligt werden. Dasselbe gilt für spurdiskrete Zeitinversionen, wenn diese nachträglich verwirklicht werden sollen. Letzteres macht die Verwendung mehrerer Zuspieler erforderlich, die natürlich entsprechend zu synchronisieren sind.

Da zur Erzeugung von Inversechos ebenfalls der Vorgang der Zeitinversion notwendig ist, gelten die oben gemachten Aussagen entsprechend auch hierfür.

An dieser Stelle ist zusammenfassend zu sagen, dass bei herkömmlichen Analogverfahren und -aufzeichnungen zwar Zeitinversionen möglich sind, jedoch nur mit umständlichen Arbeitsmethoden. Alle Vorteile, die bei den Schneidetechniken für eine moderne, computergestützte Bearbeitung vorhanden sind, gelten auch in diesem Punkt.

6.5 Wartung der Systeme

Der Service und die regelmäßige Wartung für produktive Systeme ist ebenfalls ein entscheidender Faktor. Auch diesbezüglich gibt es wesentliche Unterschiede zwischen herkömmlicher und moderner Technik, wobei es sich hier hauptsächlich auch um Unterschiede zwischen Analog- und Digitalsystemen handelt. S. 216 [2]

6.5.1 Wartung bei Analogsystemen

Analog arbeitende Geräte zeigen grundsätzlich eine Signalqualität, die mit dem Abgleich steht und fällt. Es ist also ein kontinuierlicher Zusammenhang zwischen optimierbaren Parametern und endgültig erreichter Signalqualität vorhanden. Wir wollen hier das Beispiel einer Magnetbandmaschine nennen, deren Tonköpfe regelmäßig in puncto Azimut zu justieren sind, andernfalls sind Einbrüche beim Frequenzgang zu befürchten. Das gleiche gilt natürlich ebenfalls für die Reinigung. Nachlässigkeiten beeinflussen das endgültige Signal in negativer Weise.

Auch die Einstellung der Vormagnetisierung bei der Aufnahme ist eine kritische Sache, die für die Minimierung des Klirrfaktors wichtig ist. Diese Justage hat ebenfalls kontinuierlichen Charakter, und außerdem verlangt im Prinzip jedes Bandmaterial eine spezifische Einstellung, um optimale Qualität zu erreichen.

Oben genannte Justagen erfordern einen erheblichen Zeitaufwand, wenn sie gewissenhaft durchgeführt werden sollen. Daneben sind natürlich einige Geräte und Referenzaufzeichnungen erforderlich. Bei manchen modernen Maschinen wird sogar ein prozessorgesteuertes Einmessen ermöglicht; diese Fähigkeit hebt solche Geräte jedoch von der herkömmlichen Technik ab.

Der zeitliche Abstand von Azimut-Justagen beispielsweise hängt von äußeren Faktoren ab, zu denen auch der Transport der Maschinen gehört. Die entstehenden Vibrationen tragen sicher dazu bei, dass sich die Neigung der Köpfe verändern und somit die Tonqualität darunter leiden kann.

6.5.2 Wartung bei Digitalsystemen

Natürlich müssen auch digitale Aufnahmesysteme wie DAT justiert werden; das Ziel von Einstellungen dient i. Allg. hier jedoch nicht der Optimierung der Signalqualität, sondern der Erhöhung der Betriebssicherheit.

Jegliche Einstellungen, die sich auf die Qualität des gelesenen Digitalsignals auswirken, nehmen nicht auch zwangsläufig Einfluss auf die Qualität des endgültigen Ausgangssignals. Bei einer gewissen Dejustage setzt lediglich die Fehlerkompensation ein, und wenn diese die entstehenden Datenfehler nicht mehr ausmerzen kann, werden Fehler hörbar, oder die Maschine schaltet die Wiedergabe vollständig ab.

Findet eine Aufzeichnung auf die Festplatte eines Computers statt, so obliegt dem Anwender normalerweise keine Pflicht, das System durch regelmäßige Justagen zu warten. Festplatten sind im Vergleich zu Magnetbandmaschinen sehr zuverlässig; dennoch darf man nicht verschweigen, dass natürlich das Hard-Disk-Recording eine ständige mechanische Beanspruchung für die Festplatte bedeutet, wodurch es hie und da doch zu Ausfällen kommen kann.

6.6 Latenz-Zeit

In der digitalen Audiotechnik ist schon lange der Effekt der Latenz bekannt. Man versteht darunter die Durchlaufzeit eines Signals durch einen Rechner. In der Praxis ist dies z. B. eine Erschwerung für Musiker, gleichzeitig zu wiedergegebenen Spuren ihre eigene neue Spur einzuspielen, wenn diese im Kopfhörer verzögert erscheint. Zeitversätze von 100 ms machen das synchrone Zuspielen unmöglich. Erst wenn der Wert der Latenz-Zeit kleiner als 10 ms ist, wirkt sie physiologisch und technisch nicht mehr störend.

6.6.1 Monitoring-Latency-Zeit

Wie die Latenz zustande kommt, geht aus Abb. 6.2 hervor. Man kann dem Hauptgrund der Latenz durch Verringerung der Puffergrößen entgegentreten. Allerdings ist die CPU-Auslastung durch ständigen Zugriff auf Bus zwangsläufig größer. Das simultane Mithören des eingespielten Signals bei der Aufnahme – das Monitoring – ist nur dann praktisch unverzögert möglich, wenn direkt nach der Wandlung das Signal der Rückwandlung zugeführt wird. Diese Methode heißt *ZLM* (*Zero Latency Monitoring*) und wird nur bei moderner Audio-Hardware ermöglicht. Das Handling dieser Umlenkung bewerkstelligt spezielle Software.

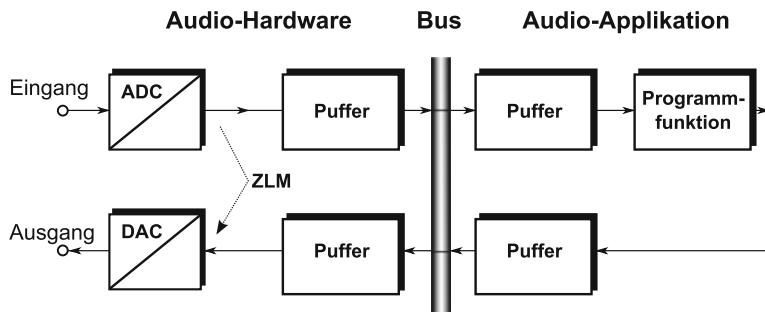


Abb. 6.2 Entstehung der Latenz. Bei der Eingabe und Ausgabe eines Audiosignals müssen Digitaldaten erst gewisse Pufferlängen durchqueren, damit die paketartige Abarbeitung durch Prozessor und Bus nicht zu Unterbrechungen führt. Erst durch einen Bypass ZLM lässt sich diese Verzögerung fast ganz unterdrücken

6.6.2 MIDI-Latenz

Eine Art der Monitoring-Latenz entsteht auch beim Datentransfer eines MIDI-Gerätes (z. B. Keyboard) und einem MIDI-Soundmodul (auch MIDI-Expander genannt). Wie bereits in Abschn. 4.1.3 gesehen, nimmt ein einziges Bit bereits eine Zeit von $32\ \mu\text{s}$ in Anspruch. Ein einziges Byte benötigt inklusive Start- und Stopppbit zehn Bits, also $320\ \mu\text{s}$. Ein einfacher Note-an-Befehl benötigt aber mindestens zwei Bytes, zwischen denen noch dazu eine kleine Pause bleibt. Damit kann man schon mit einer Übertragungszeit von $800\ \mu\text{s}$ rechnen. Eine gewisse Zeit ist fürs Decodieren hinzu zu rechnen, so dass etwa 1 ms zu veranschlagen ist. Laufen viele MIDI-Quellen zusammen, vervielfacht sich diese Zeit natürlich. Diese MIDI-Weichen können per Software puffergesteuert sein und müssen deswegen wieder erst den Puffer überwinden. Werden die MIDI-Eingänge parallel decodiert, entfällt diese Vervielfachung der Decodierzeiten und somit der Latenzzeiten. Zu vermeiden ist das Nadelöhr MIDI-Schnittstelle auch dadurch, dass man die Geräte mit USB- oder Firewire (IEEE-1394) zusammenschaltet.

6.6.3 Latenz-Jitter

Neben einer statisch auftretenden Latenzzeit gibt es auch die unangenehme Eigenschaft der Variation der Verzögerung. Man spricht hier von *Latenz-Jitter*. Diese Schwankungen stehen ursächlich mit regelmäßig oder sporadisch vorkommenden Aktivitäten der CPU und/oder Blockaden des Busses zusammen. Diese können wiederum verschiedene Gründe haben. Man vermeide deshalb bei kritischen Anwendungen generell irgendwelche Dinge, die „nebenher“ laufen wie z. B. Energiesparsysteme oder Screensaver, WLAN-Verbindung oder auch die automatische Abfrage des CD-Wechsels.

Ein weiteres Jitter entsteht durch inkohärente Quellen, also z. B. beim Drücken einer Taste eines MIDI-Keyboards und dem Abspielen bestehender Audiodaten von der Festplatte. Die Daten des Keyboards durchlaufen den Decoderprozess, und der Soundprozessor kann nach Start eines Puffers diesem Daten liefern. Steht der Pufferwechsel unmittelbar bevor, erfolgt mit nur einer Pufferlänge Latenz die Wiedergabe des MIDI-Tones. Ist der Pufferwechsel jedoch kurz vorher vollzogen, so ist der bereits laufende Puffer nicht mehr zugänglich für die Audiodaten des MIDI-Keyboards. Diese Pufferzeit ist abzuwarten, der folgende Puffer steht dann zur Verfügung. Doch erst nach Ablauf dieses zweiten Puffers erscheinen nun die MIDI-Klänge. Die Grundlatenz ist also eine Pufferlänge, kann jedoch bis zu zwei Pufferlängen betragen. Das Jitter beträgt demnach eine Pufferlänge.

6.6.4 Durchlauf-Latenz

Wir haben weiter oben die Latenz beim Monitoring und ihre Vermeidung betrachtet. Nach wie vor besteht aber ein Zeitintervall des wiedergegebenen und des aufgezeichneten Signals. Angenommen, eine erste Spur wird von einer Recording-Software (Cubase LE) abgespielt, gleichzeitig kommt eine zweite Spur zur Aufnahme. Laut Abb. 6.2 benötigt die Wiedergabe der ersten Spur mindestens die Latenzzeit, die durch den Wiedergabepuffer bedingt ist, während die Aufnahme entsprechend durch den Aufnahmepuffer verzögert ist. Messbar ist diese Gesamtlatenz einfach dadurch, dass man das Wiedergabesignal von Spur 1 auf Spur 2 aufzeichnet. Man betrachte hierzu Abb. 6.3. Der Test beinhaltet einen kurzen Signalimpuls. Die aufgezeichnete Spur 2 zeigt diesen Impuls um etwa 100 ms verzögert gegenüber dem Signal von Spur 1. Der Eingangspuffer betrug hier 4096 Samples, der Ausgangspuffer 512 Samples. Bei einer Samplingrate von 44,1 kHz beträgt die Latenz insgesamt somit:

$$T_{Latenz} = \frac{4096 + 512}{44100} \text{ s} = 104 \text{ ms} \quad (6.1)$$

Beim zweiten Versuch (Aufzeichnung auf Spur 3) wurde der Eingangspuffer auf 512 Samples zurückgesetzt, dadurch ergibt sich eine Latenz von:

$$T_{Latenz} = \frac{512 + 512}{44100} \text{ s} = 23 \text{ ms} \quad (6.2)$$

Die Beeinflussung der Puffer geschah hier übrigens über den ASIO-DirectX-Treiber.

Beim Zusammenmischen aller Spuren würde sich ein sehr verschwommenes Timing ergeben, denn bei der Wiedergabe der bereits vorhandenen Spuren gibt es einen Versatz des empfundenen Taktes. Natürlich ist eine Verbesserung durch die genannte Reduzierung der Pufferlängen möglich, der Systematik tut man damit jedoch keinen Abbruch. Besser ist die Generierung *einer* Pilotenspur, auf der alle relevanten Instrumente in korrektem Takt aufgespielt sind. Beim Overdubbing ist nur diese Spur und das neu einzuspielende

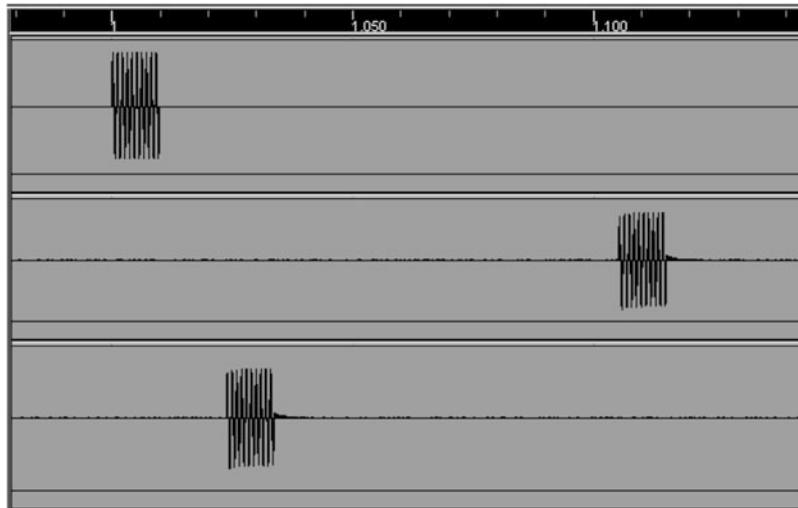


Abb. 6.3 Ein Test zur Demonstration der Latenz beim Overdubbing. Der Impuls in der ersten Spur (oben) gelangt auf Spur 2 und erscheint dort um mehr als 100 ms verzögert. In der dritten Spur (unten) wurde diese Überspielung wiederholt, jedoch bei verkleinertem Eingangspuffer. Der Versatz gegenüber der ersten Spur beträgt nur noch etwas mehr als 20 ms

Instrument (per ZLM) zu hören. Alle so eingespielten Spuren sind synchron in einem Zeitraster.

6.7 Gegenüberstellung Analogtechnik/Digitaltechnik

Wir wollen hier noch kurz einen tabellarischen Vergleich zwischen analoger und digitaler Verarbeitung von Audiosignalen anstellen (Tab. 6.1).

Es handelt sich hierbei jedoch lediglich um eine qualitative Gegenüberstellung; bei beiden Systemen wird von optimaler Qualität ausgegangen (im anderen Falle wäre es z. B. möglich, dass der Klirrfaktor beim Digitalsystem größer ist als beim Analogsystem). Quantitative Angaben würden sehr umfangreich ausfallen, weil hier immer eine Differenzierung zwischen den unterschiedlichen Geräten stattfinden müsste (Plattenspieler, Magnetbandgerät, DAT, Audiokarte usw.).

Als Fazit des gesamten Vergleichs ist zu sagen, dass die Vorteile der Digitaltechnik klar überwiegen; in naher Zukunft wird daher zu erwarten sein, dass die Analogtechnik auf den Gebieten der Aufzeichnung und der Weiterverarbeitung (Schnitt, Effekte usw.) verdrängt wird. Das geschieht jedoch nicht völlig zugunsten der computernahen Verarbeitung, sondern zunächst auch einzelner Digitalgeräte wie DAT usw.

Tab. 6.1 Analoge und digitale Verarbeitung im qualitativen Vergleich

Vorgang	Kriterium	Analoge Verarbeitung	Digitale Verarbeitung
Speicherung	Alterungseffekt	kontinuierlich	keiner
	Qualität bei Überspielungen	abfallend	gleichbleibend
	Mediumsbedarf	klein	groß
	Dropouts	häufig	selten (nur bei großen Mediendefekten)
Schnitt	Durchführung	umständlich	einfach
	Genauigkeit	niedrig	hoch
	Vervielfältigung	kompliziert	sehr einfach
	Schnittkanten	mit Sprüngen	interpolierbar
	Effekte	sehr kompliziert	einfach
Filterung	Aufwand	groß	klein (bei Software)
	Phasenverzerrung	vorhanden	bei FIR-Filtren keine
	Echtzeitbearbeitung	möglich	mit schnellen Prozessoren möglich
Reproduktion	Klirrfaktor	mittel	sehr klein
	Gleichlauf	befriedigend	sehr gut
	Störabstand	nicht hoch	sehr hoch
	Dynamikbereich	oft nicht ausreichend	groß
	Frequenzgang	große Toleranzen	kleine Toleranzen
Wartung	Häufigkeit	bei mechanischen Geräten regelmäßig	bei Computern selten
	Effekt	Optimierung der Tonqualität	Funktionssicherheit

Literatur

1. Prochaska, E.: Digitale Signalprozessoren. Ehningen: Expert-Verlag 1988.
2. Stotz, D.: Audio- und Videogeräte richtig einmessen und justieren. München: Franzis-Verlag 1994.

Teil II

Videotechnik

Anfang der 90er-Jahre, als der Boom mit den Video-Schnittsystemen begann, war der Aufwand für einen professionellen Schnittplatz enorm – benachbart zum eigentlichen Arbeitsplatz mit der Workstation stand ein mannshohes 19-Zoll-Rack, rand gefüllt mit Laufwerken. Die notwendigen Kapazitäten von mehreren Gigabytes konnten anders nicht realisiert werden. Heute, also etwa zwanzig Jahre später, ist der Platzbedarf immer noch recht groß, allerdings mit dem Unterschied, dass die notwendigen Datenmengen größer wurden.

Liegt Videomaterial in digitaler Form vor, so ist der Arbeitsgang *Schnitt* komplett an einem dafür eingerichteten PC-System möglich – es sind keine Spezialgeräte mehr nötig.

Das Buch führt den Leser im zweiten Teil von den Grundlagen bis hin zu praktischer Arbeit mit Schnittsystemen. Beleuchtet werden die Parameter von Videobildern, die neben Unterschieden der Standards auch viele Qualitätsstufen aufweisen. Selbstverständlich findet nicht nur der Film-, sondern auch der Grafik- und Animationsbereich Aufmerksamkeit.

Grundlagen und Verfahren der Videotechnik

7

7.1 Bildabtastung und -wiedergabe

Bei der Bildübertragung sind außer der Signalaufbereitung auch die Prinzipien der Abtastung und auch des Bildformats zu diskutieren. Dabei sind natürlich sowohl physikalische als auch physiologische Aspekte zu berücksichtigen, damit eine zufriedenstellende Bildqualität erzielt werden kann.

Zunächst einmal hat man festgelegt, dass das zu verarbeitende Bild in ein Raster zu zerlegen ist, welches dann in einer bestimmten Art und Weise abzutasten ist. Dieses Raster bestimmt auch, was für ein Bildseitenverhältnis vorliegt.

Damit unterschiedliche Geräte (z. B. Fernsehgeräte, Videorecorder usw.) betrieben werden können, ist die Einhaltung von Normen wichtig, die gleichzeitig die Kompatibilität einzelner Geräte gewährleisten soll. Jedoch auch andere Parameter werden durch die Normen festgelegt, wie z. B. die Modulationsart und Modulationsrichtung der Bildträgerschwingung, die Modulationsart der Tonträgerschwingung, Frequenzabstand zwischen Ton- und Bildträger, die Zwischenfrequenz der Tonübertragung usw. Die letzten Parameter sind für unsere Belange jedoch von untergeordneter Wichtigkeit, weil sie zur Fernsehübertragungstechnik und Sendertechnik gehören. Wir wollen uns dagegen hauptsächlich die Aufbereitung des Bildsignals selbst betrachten.

Die sog. Gerbernorm (benannt nach ihrem Erfinder Gerber) legt fest, dass ein komplettes Bild aus 625 Zeilen bestehen soll, welche in der Zeit 1/25 s zu durchlaufen sind. Nach dieser Norm (auch 625/50-System genannt) arbeiten die meisten westeuropäischen Systeme. Die Vollbildfrequenz von 25 Hz bzw. die Halbbildfrequenz von 50 Hz soll die Interferenzstörungen mit der Netzfrequenz möglichst klein halten. Außerdem hat man vereinbart, das Bildseitenverhältnis (Breite : Höhe) auf den Wert 4 : 3 zu setzen, was dem des normalen Kinofilms entspricht.

7.1.1 Zeilensprungverfahren (Interlaced)

Nach den oben genannten Vorgaben ergibt sich eine Bildwechselfrequenz von 25 Hz, was sich für das menschliche Auge als unangenehmes Flimmern auswirken würde. Anstatt nun die 625 Zeilen 25 mal in der Sekunde zu schreiben, setzt man besser 2 Halbbilder ineinander, die jeweils nur die Hälfte der Zeilenzahl, dafür aber die doppelte Ablenkfrequenz aufweisen. Ein solches ineinander Verzahnen der Bilder nennt man Interlaced (= verflochten) oder Zeilensprungverfahren. Hierzu betrachten wir Abb. 7.1.

Mit dieser Methode also erreicht man, dass dem Auge ein scheinbarer Bildwechsel von 50 Hz vorgetäuscht wird, obwohl es sich eigentlich um einen Halbbildwechsel handelt, der sich jedoch jeweils über den gesamten Bildschirm erstreckt. Somit ist relativ gute Flimmerfreiheit zu erzielen, ohne dass die Rate der zu übertragenden Information allzu hoch anzusetzen ist.

Während die Abtastung des Bildes bei der linken oberen Ecke beginnend bis zum unteren Rand des verläuft, werden nacheinander die ungeraden Zeilen 1,3,5,...625 abgearbeitet (ungerades Halbbild). Genau genommen reicht die 625. Zeile noch in das gerade Halbbild hinein, wo ansonsten die noch verbleibenden, geraden Zeilen zur Abtastung

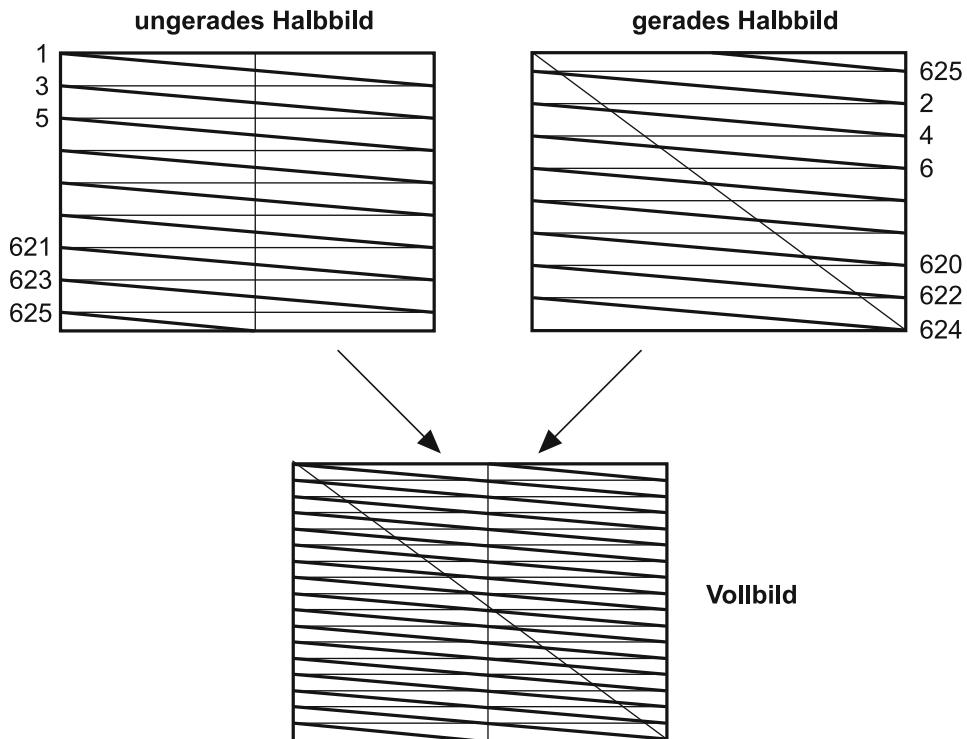


Abb. 7.1 Synthese zweier Halbbilder beim Zeilensprungverfahren

gelangen. Beides geschieht also zeitlich hintereinander; das Gesamtbild ergibt durch Verschmelzung beider Halbbilder, die bei der Darstellung durch den Videomonitor eigentlich nur durch die Trägheit des Auges als Gesamtbild empfunden werden (natürlich kann auch durch nachleuchtende Schichten der Bildröhre oder auch mittels Spezialelektronik dieser Effekt verstärkt werden).

Die dünnen Linien in Abb. 7.1 deuten den Rücklauf der Abtastvorgangs an, der ja wesentlich schneller vonstatten gehen muss als der normale Abtastlauf. Bei der ungeraden Zeilenzahl von 625 ist es notwendig, den Halbbildwechsel inmitten einer Zeile zu legen, nämlich der 625. Zeile. Würde man dies nicht tun, müsste die Vertikalablenkung alternierend mit zwei unterschiedlichen Zeitkonstanten arbeiten, je nachdem, welches Halbbild gerade aktuell ist.

Dieses Ineinandergreifen beider Halbbilder wäre andererseits bei einer geraden Zeilenzahl (z. B. 624) nicht ohne weiteres möglich, wenn wiederum mit einer gleichförmigen Vertikalablenkung gearbeitet werden soll. Die Ablenkphase des Vertikalsystems käme dann immer auf identische Zeilenphasen zu liegen, was zur Folge hätte, dass beide Halbbilder genau aufeinander zu liegen kommen – das versetzte Ineinandergreifen käme nicht zustande, und die effektive Zeilenzahl würde sich halbieren.

Nicht alle der 625 Zeilen gehören einem gültigen Bildinhalt an, weil eine gewisse Zeitspanne für die Vertikalsynchronisation verlorengeht; wir werden weiter unten noch genauer darauf zu sprechen kommen. Vielfach wird auch behauptet, die Zeit werde für die Reinitialisierung des Strahls an den oberen Bildrand benötigt. Das stimmt natürlich nur zu einem ganz geringen Teil, da ja theoretisch der Strahlrücklauf für die Vertikalablenkung genauso schnell sein kann wie für die Zeilenablenkung, also in maximal $12\ \mu\text{s}$, was somit einem Bruchteil einer Zeile entspricht. In Wirklichkeit aber werden zwischen den Halbbildern jeweils etwa 25 Zeilen, insgesamt pro Vollbild also 50 Zeilen ignoriert, die nicht für die Übertragung von Bildinformation verwendet werden. Der Bildrahmen für gültige Bildinformation beinhaltet mithin ca. 575 Zeilen.

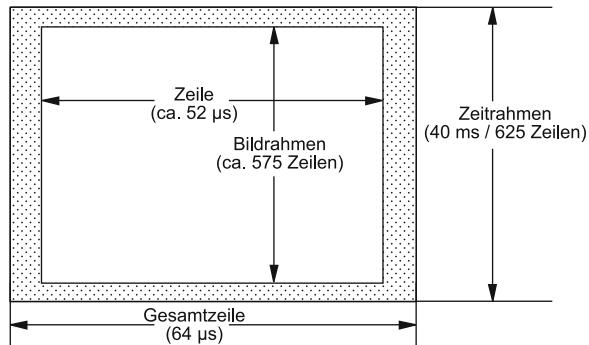
Aus den Werten für die Bildwechselfrequenz f_b und der (Brutto-)Zeilenzahl N_{Z0} ergibt sich eine Zeilenfrequenz von:

$$f_h = f_b \cdot N_{Z0} = 25 \cdot 625 \text{ Hz} = 15625 \text{ Hz} \quad (7.1)$$

Für die Abtastung einer vollen Zeile ergibt sich aus dem Kehrwert von f_h also $64\ \mu\text{s}$, wobei jedoch etwa $12\ \mu\text{s}$ dem Zeilenrücklauf zukommen, und die verbleibenden $52\ \mu\text{s}$ stellen also das zeitliche Maß des Bildrahmens für gültige Bildinformation (ungenaue Beschreibungen sprechen hier wiederum von der „sichtbaren“ Zeile, was genau genommen jedoch nicht zutrifft).

Während der Zeiten für den Zeilenrücklauf und der Halbbildlücke, in denen ja keine gültige Bildinformation übertragen wird, erfolgt eine Dunkeltastung des Strahls der wiedergebenden Monitore, weshalb man hier auch von Zeilen- bzw. Bildaustastung spricht.

Abb. 7.2 Rahmen mit gültigem Bildinhalt und Zeitrahmen



Wollte man den Bildrahmen innerhalb des Zeitrahmens darstellen, so ergibt sich für ein Vollbild das Diagramm nach Abb. 7.2.

In Wirklichkeit finden die Bildaustastungen natürlich zwischen den Halbbildern statt, nicht nur – wie hier dargestellt – nach jedem Vollbild. Letztere Betrachtungsweise ist dennoch naheliegend, wie wir vor allem im nächsten Kapitel bezüglich dem Overscan-Modus noch sehen werden.

Das hier vereinfacht beschriebene Verfahren entspricht sowohl bei der Bildaufnahme durch Kameras, als auch bei der Wiedergabe durch Monitore der angewandten Praxis. Selbst moderne hochauflösende Standards (HDTV) kommen ohne das Zeilensprungverfahren nicht aus.

Bei der Wiedergabe eines Computerbildes im Interlaced-Modus ist immer zu beobachten, dass starkes Flimmern einzelner Zeilen entsteht. Dies liegt vor allem an den oftmals feinen, auf einzelne Zeilen beschränkten Strukturen. Nehmen wir an, eine helle Zeile sei von zwei dunklen Zeilen umgeben. Es leuchtet ein, dass die einzelne helle und die dunklen Zeilen unterschiedlichen Halbbildern angehören. Somit leuchtet die helle Zeile nur mit einer Frequenz von 25 Hz auf, was bei den dunklen Nachbarzeilen natürlich besonders gut auffällt.

7.1.2 Non-Interlaced-Verfahren

Vor allem im Computerbereich kommt das sog. Non-Interlaced-Verfahren zur Anwendung, bei dem das ganze Bildraster in einem Zuge abgetastet wird. Hier muss man also keine Halbbilder (Fields) ineinander legen, sondern erhält sofort während eines Zyklus das gesamte Bild (Frame).s das gesamte Bild (Frame).

Damit ein flimmerfreies Bild entsteht, muss die Bildablenkung ebenfalls mit einer Frequenz von mindestens 50 Hz arbeiten, wobei diesmal jedoch auch eine ebenso hohe Vollbildwechselfrequenz erreicht wird. Gegenüber dem Interlaced-Verfahren bedeutet dies eine Verdopplung der Zeilenfrequenz, was nach Gl. 7.1 also 31 250 Hz ergibt. Der zu übertragende Informationsgehalt pro Zeiteinheit ist demzufolge ebenfalls doppelt so hoch; für Computersysteme besteht jedoch keine Beschränkung, die aus einer

Bandbreitenbegrenzung wie bei der Fernsehübertragung resultiert. Hier können vielmehr Monitor- und Grafikkartentechnologie die Grenzen bestimmen.

Natürlich sind Non-Interlaced-Zeilenfrequenzen nicht auf den oben genannten Wert begrenzt; in der Praxis sind sie oft noch wesentlich höher.

7.1.3 Synchronisation

Ein Bildsignal muss natürlich nicht nur die Information für die Bildpunkte beinhalten, sondern auch Impulse, die den Zeilen- und Bildanfang kennzeichnen. Diese Signale dienen also dem Bildstand oder der Synchronisation.

Zwischen jeder Zeile – also alle $64 \mu\text{s}$ – erfolgt beim Interlaced-Verfahren die Übertragung eines Zeilensynchronimpulses, der z. B. den Zeilenoszillator eines Wiedergabegerätes synchronisieren soll; ohne diese Impulse bestünde kein fester Phasenbezug zwischen dem zu übertragenden Signal mit seinen Zeileninformationen und dem wiedergebenden Gerät. Die Folge davon wäre ein völlig unkenntliches Bild.

Üblicherweise arbeitet man mit kurzen Impulsen von etwa $4,7 \mu\text{s}$ Dauer, die in negative Richtung zeigen. Das Potential der Dachkanten dieser Impulse liegt soweit unterhalb des üblichen Bildinhalts, dass eine klare Identifikation der Zeilensynchronimpulse stattfinden kann. Wir gehen hier einmal davon aus, dass nämlich beides, die Bildinformation und die Synchroninformation als gemeinsames Signal vorliegt. Dies ist durchaus nicht immer der Fall, vor allem dann nicht, wenn ein in Primärfarben getrenntes Signal (ein sog. RGB-Signal) verarbeitet wird.

Mit diesen Zeilensynchronimpulsen besteht eine eindeutige Kennzeichnung der Zeilenstartpunkte; um ein Bild jedoch richtig wiedergeben zu können, muss auch der Bild- bzw. der Halbbildanfang eindeutig markiert sein. Dieser als Vertikal- oder Bildsynchronisation bezeichneter Impulskomplex ist nicht ganz so einfach aufgebaut wie die Zeilensynchronisation.

Bei einem integrierten Signal wie dem BAS-Signal (siehe nächster Abschnitt) besteht der Bildwechselimpuls aus mehreren modifizierten Zeilensynchronimpulsen. In Abb. 7.3 sehen wir diese Impulskonstellation, die während des Übergangs zwischen zwei Halbbildern abgegeben wird.

Jeweils vor und nach den Bildwechseln ist noch ein Teil einer Zeile zu sehen, die gültige Bildinformation hat, obwohl sich noch etliche Dunkel-Zeilen zwischen Bildwechselimpulsen und gültigem Bild befinden.

Deutlich zu sehen ist die Frequenzverdopplung beim Eintreffen der Bildwechselimpulse; ohne die ja eine Bildumschaltung inmitten einer Zeile unmöglich wäre. Eine solche Situation ergibt sich ja immer zwischen ungeraden und geraden Halbbildern (Abb. 7.3 oben), während zwischen geraden und ungeraden Halbbildern der Zeilen- und Bildwechsel zusammenfallen (Abb. 7.3 unten).

Bei der Einbindung der Synchronimpulse in das Gesamtsignal (wie beim BAS-Signal, siehe nächster Abschnitt) muss die Bildsynchronisation aus der Zeilensynchronisation

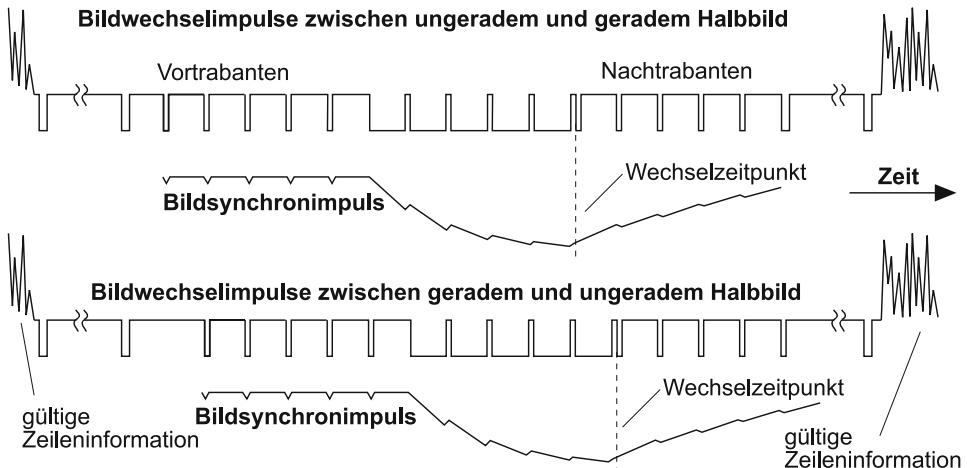


Abb. 7.3 Vertikalsynchronimpulse und daraus gewonnener Bildwechselimpuls

abgeleitet werden, weil ein einzelner, langer Impuls als Vertikaltriggerung die Zeilensynchronisation für kurze Zeit unterbrechen würde – und das darf natürlich unter keinen Umständen geschehen, weil sonst die Gefahr von Bildverzügen gegeben wäre. Ferner bestünde für die Farbsynchronisation ein Problem.

Die Gewinnung des Triggerimpulses für die Vertikalsynchronisation kann durch einfache Tiefpassfilterung bzw. Integration der Impulse erfolgen. Durch die Invertierung bzw. Änderung des Tastverhältnisses bei den mittleren 5 Impulsen erhält man ein allmähliches Abfallen der Integrationsspannung. Durch die Halbzeilenverschiebung zwischen den Wechseln stellt sich zwangsläufig auch eine Verschiebung des integrierten Bildwechselimpulses ein. Ein korrektes Ineinandergreifen beider Halbbilder ist somit gewährleistet.

Der qualitative Spannungsverlauf ist in Abb. 7.3 so dargestellt, wie es der Situation in der Videotechnik entspricht. Damit zeigen die Synchronimpulse immer in die negative Richtung; ihre absoluten Potentiale sind nebensächlich, da die Ankopplung meist eine DC-Trennung bewirkt (durch Kondensator).

7.1.4 BAS-Signal

Bereits im letzten Abschnitt wurde das BAS-Signal (Bild-Austast-Synchron-) erwähnt und in Abb. 7.3 angedeutet. Ein komplettes Signal einer Zeile ist in Abb. 7.4 zu sehen.

Die Festlegung der Spannungswerte beruht auf die Verteilung eines maximalen Spannungsintervalls von 1 V auf 140 IRE-Einheiten. Hiervon teilt man 100 IRE-Einheiten dem Bildluminanzwert U_Y zu, die restlichen 40 gehen an das Synchronintervall U_S . Damit ergibt sich:

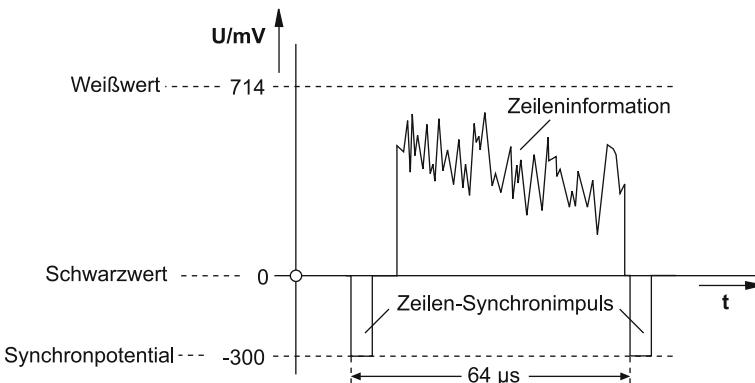


Abb. 7.4 BAS-Signal nach CCIR B/G

$$\begin{aligned}
 1 \text{ IRE} &= \frac{1 \text{ V}}{140} \approx 7,14 \text{ mV} \\
 U_Y &= 100 \cdot 7,14 \text{ mV} = 714 \text{ mV} \\
 U_S &= 40 \cdot 7,14 \text{ mV} = 286 \text{ mV}
 \end{aligned} \tag{7.2}$$

Der Schwarzwert wurde hier mit null angenommen, was heute in den meisten Fällen praktiziert wird, obwohl auch eine Anhebung um einige Prozent zulässig ist. Es sind sowohl hier als auch bei anderen Parametern Toleranzen zulässig. Für weitere Informationen über Zeitangaben sei hier auf Ausführungen verwiesen, die im Abschn. 7.2 erscheinen werden.

Es wurde bereits angedeutet, dass das Nullniveau in Abb. 7.4 nicht unbedingt auch einem Potential von 0 V gegenüber Masse entsprechen muss. Dies ist lediglich eine willkürliche Festlegung in der Abbildung.

Die nachfolgende Elektronik ist in der Lage, anhand der Synchronimpulse auf deren Schulterpotential zu schließen, welches dann individuell zum Nullpotential (Schwarzwert) deklariert wird. Auf diese Weise ist es beispielsweise auch möglich, homogene und konstante Bilder wiederzugeben, die ja keinerlei Wechselspannungskomponente mehr aufweisen. Allein die Synchronimpulse samt ihrer Schultern ermöglichen einer Klemmschaltung die Erkennung des Schwarzniveaus.

Dieses BAS-Signal ist lediglich imstande, ein Schwarzweißbild zu übertragen; damit auch eine Farbwiedergabe möglich wird, sind noch einige Spezialverfahren nötig, auf die wir ebenfalls im Abschn. 7.2 zu sprechen kommen.

7.1.5 Probleme und Fehlersymptome

Viele Probleme, die sich bei der Verarbeitung eines Bildes ergeben, stehen in kausalem Zusammenhang mit einer fehlerhaften Synchronisation. Das Erkennen solcher Symptome bzw. deren richtige Interpretation ist hilfreich für eine schnelle Beseitigung der Störung.

Zu unterscheiden ist bei der Entstehung von Synchronisationsproblemen zwischen fehlender Kompatibilität und mangelhaftem Signal.

Inkompatible Signale Dieser Sachverhalt trifft zu, wenn z. B. ein Non-Interlaced-Signal auf ein Gerät gelangt, das nur Interlaced-Signale verarbeiten kann (z. B. PAL-Encoder oder Videobandmaschine). Beispiele für solche Konstellationen und deren Auswirkungen zeigt Abb. 7.5.

Um die sich einstellenden Effekte einordnen zu können, muss man sich immer vor Augen halten, welche Diskrepanzen zwischen den Timings der Signale auftreten. So ist z. B. leicht einzusehen, dass geringe Unterschiede der Zeilenfrequenz aber relativ große Unterschiede der Bildfrequenz meist nur die Bildsynchroisation außer Tritt bringen, wie dies bei der Konstellation PAL/NTSC der Fall sein wird. Dagegen kann ein Monitor, der nur Interlaced-Signale darstellen kann, mitunter ein Doppelbild zeigen, das aber durchaus stabil steht.

Mangelhaftes Synchronsignal Fehlt die Synchronisation völlig oder ist ihr Signal zu schwach oder gestört, so können sich Effekte nach Abb. 7.6 einstellen.

Auch solche Fehler kommen in der Praxis der Bildverarbeitung relativ häufig vor – verstärkt bei der Einbindung von Computer-Applikationen, weil dort ebenfalls oft die Möglichkeit besteht, dass sich die Videokarten im Non-Interlaced-Modus befinden. Wir werden in den Kap. 9, 10 und 11 noch spezieller solche Möglichkeiten erörtern.

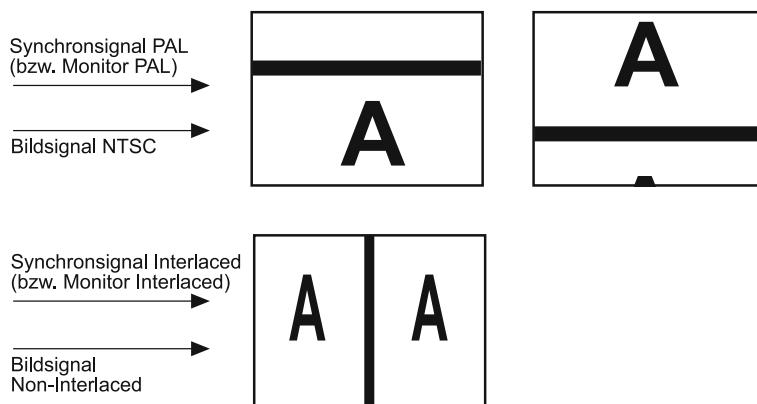


Abb. 7.5 Synchronisationsstörungen durch Inkompatibilität der Signale

Abb. 7.6 Signalmängel der Synchronisation

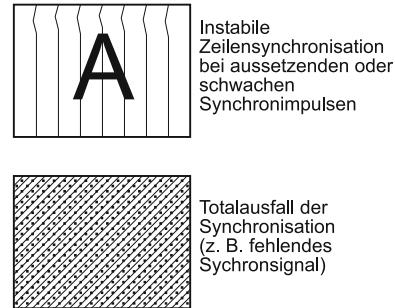


Abb. 7.7 Bildwiedergabe bei unabhängiger Quarzreferenz zwischen Bild-Timing und Synchronisation

Unabhängige Synchronreferenzen Dies ist eigentlich ein Spezialfall für fehlende Synchronisation, jedoch äußert sich diese anders als ein Synchronisationsausfall bei Geräten mit üblichen Freischwingen (z. B. die Ablenkgeneratoren bei Monitoren). Das Bildsignal hat auch hier ein unabhängiges Timing in Bezug auf das Synchronsignal, jedoch wird beides von getrennten Quarzoszillatoren gesteuert, so dass zwar das Bild langsam durchläuft, aber dennoch nicht unkenntlich ist, weil die relative Diskrepanz der Quarzreferenzen nur gering ist. Abb. 7.7 deutet die Situation an.

Je nachdem, wie stark beide Quarzreferenzen voneinander abweichen, stellt sich die Geschwindigkeit und die Richtung des horizontalen Wanderns des Bildes ein. Nach jedem Zyklus, dessen Frequenz im Bereich von etwa 0 Hz...2 Hz liegt, erscheint das Bild natürlich auch um eine Zeile in vertikaler Richtung verschoben. Ein sog. vorübergehendes „Einrasten“ ist bei völliger Unabhängigkeit nicht zu beobachten.

7.2 Farbübertragung

Das bisher erörterte Verfahren ist lediglich in der Lage, ein Bild mit Helligkeitsinformationen zu übertragen – Farbübertragungen benötigen hingegen spezielle technische Kunstgriffe, die in den folgenden Ausführungen etwas näher beleuchtet werden sollen.

Im Vordergrund steht – falls nicht ausdrücklich anders vermerkt – das bei uns eingesetzte PAL-Verfahren. Auf andere Verfahren kommen wir in Abschn. 7.4 zu sprechen.

7.2.1 Physikalische und physiologische Aspekte

Vor allen Gedanken der Farbcodierung bzw. der Verschmelzung von Farbinformation mit dem BAS-Signal, mussten Überlegungen und vor allem Untersuchungen des menschlichen Auges angestellt werden.

Nach den physiologischen Erkenntnissen richtet sich dann die prinzipielle Farbaufnahme und -wiedergabe; das Konzept für die Farbcodierung und -decodierung steht natürlich in diesem Stadium immer noch nicht fest.

Farbeindruck und -empfindlichkeit Das Farbensehen basiert vereinfacht gesprochen auf die Reizung verschiedener, für Farben empfindliche Sehzellen (Zapfen). Man nimmt an, dass dabei drei verschiedene Sorten von Zapfen auch auf bestimmte Bereiche von Lichtwellenlängen – also Farben – ansprechen. Farben, deren Wellenlänge dazwischen liegen, regen demnach zwei verschiedene Sorten von Zapfen an. Das den Farbeindruck verarbeitende Gehirn kann infolgedessen nicht unterscheiden, ob es sich um eine Mischfarbe handelt (also um eine Synthese aus Licht mit Komponenten unterschiedlicher Wellenlänge) oder nur monochromatisches Licht (mit nur einer diskreten Wellenlänge) vorliegt. Eine solche Trennung kann man optisch z. B. mit Filtern oder einem Prisma durchführen.

Teilt man ein aufzunehmendes Bild in drei Farbbereiche auf, so lässt sich bei entsprechender Wiedergabe des wieder zusammengesetzten Bildes der Originaleindruck annähernd rekonstruieren. Man nimmt eine solche Einteilung vor unter Verwendung der Primärfarben Rot, Grün und Blau. Das Prinzip dieser Farbzerlegung und -zusammensetzung geht aus Abb. 7.8 hervor.

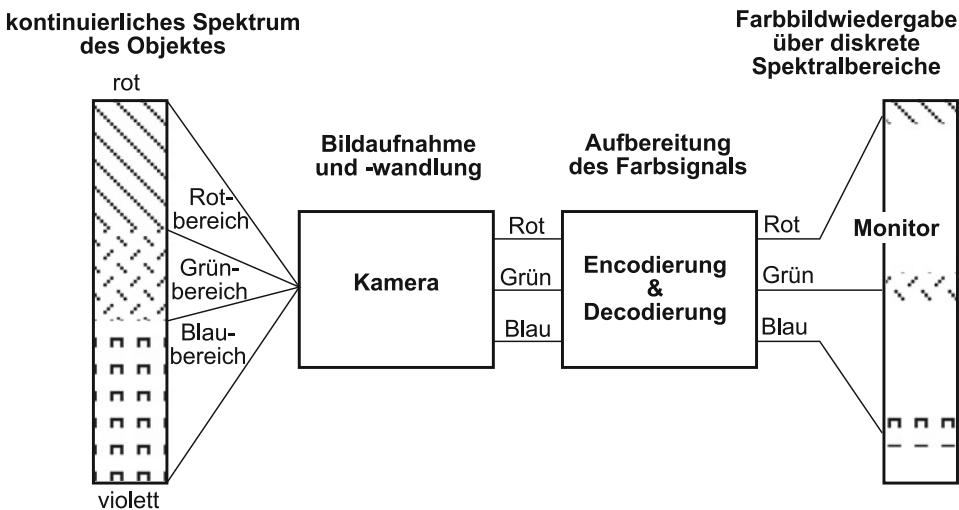
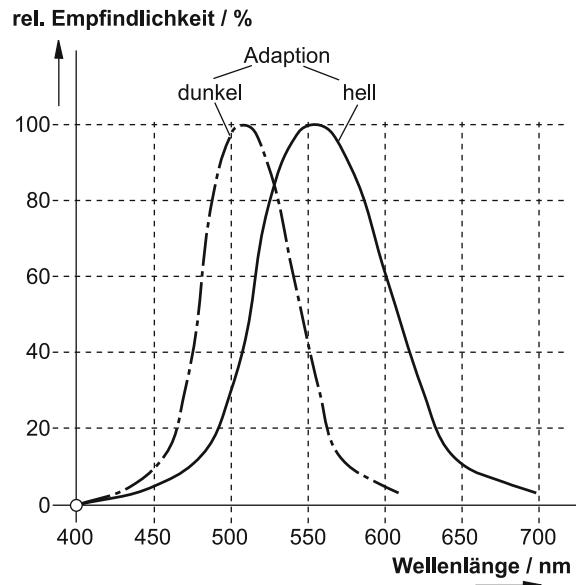


Abb. 7.8 Übertragung der Farbinformation über drei Primärfarben

Abb. 7.9 Spektrale Empfindlichkeit des menschlichen Auges



Bemerkenswert ist die Tatsache, dass das kontinuierliche Spektrum in Bereiche aufgeteilt wird, die dann bei der Wiedergabe fast nur noch als diskrete Wellenlängen erscheinen. Alle Zwischenfarben werden dabei wie erwähnt dadurch simuliert, dass die drei diskreten Wiedergabefarben eine Mischung generieren.

Die Empfindlichkeit des menschlichen Auges für verschiedene Spektralfarben bzw. Wellenlängen ist nicht konstant. Sie ist für die Farbe Grün am stärksten und verschiebt sich bei kleineren Helligkeiten in Richtung Blau. Abb. 7.9 zeigt den Verlauf der spektralen Augenempfindlichkeit.

Bei mittlerer und großer Helligkeit liegt das Maximum der Empfindlichkeit im grünen Bereich. Um aus einem Signal eines Farbbilds, das als separate Primärfarbsignale R , G und B vorliegt, ein Helligkeitssignal (Y) zu gewinnen, muss also das Grünsignal entsprechend höher bewertet werden als die anderen beiden Signale. Die Wandlung erfolgt also unter Berücksichtigung der Spektralempfindlichkeit des Auges nach der Gleichung:

$$Y = 0,299 R + 0,587 G + 0,114 B \quad (7.3)$$

Diese Gleichung ist außerordentlich wichtig für die Farbcodierung und -decodierung; sie erfüllt außerdem die Forderung, dass Farbsignale für eine Schwarzweiß-Wiedergabe geeignet sind, weil eben von vornherein der Helligkeitseindruck der Farben nach Gl. 7.3 in die Signalverarbeitung einfließt.

Farbart, Farbort und Farbsättigung Zur Definition einer Farbe hat sich u. a. für die Videotechnik die Festlegung nach Abb. 7.10 durchgesetzt. Daneben gibt es jedoch für

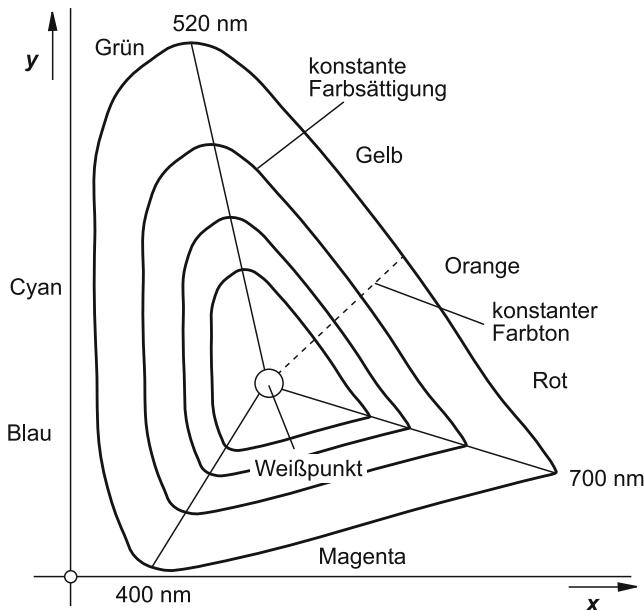


Abb. 7.10 Farbart-Darstellung nach Koordinaten (schematisches IBK-Farendiagramm)

andere Bereiche spezielle Definitionen, die hier jedoch unberücksichtigt bleiben sollen. S. 281 [6]

Es ergibt sich eine Aufteilung in Koordinaten x und y , wobei entlang der strahlenförmig verlaufenden Linien jeweils konstante Farbtöne und entlang der konzentrisch verlaufenden Linien jeweils konstante Farbsättigung herrschen. Demnach erhält man auf der äußersten Linie die maximale Sättigung und im Zentrum der Figur minimale Sättigung, nämlich der sog. *Weißpunkt*.

Die Darstellung des äußeren Kurvenzuges beruht darauf, dass zunächst ein gleichschenkliges Dreieck mit Einheitshöhe willkürlich festgelegt wurde, dessen Eckpunkte den zur Verfügung stehenden Grundfarben (Primärvalezenzen) entsprechen (beispielsweise die Grundfarben des Leuchtphosphors, die bei Monitoren der Farberzeugung dienen). Stellt man sich eine Farbmischung vor, indem man nun die Ecken entsprechend den Farbwerten beschwert, so stellt sich ein Schwerpunkt ein, dessen Ort die Farbart, also den Farbton und die Farbsättigung vollständig definiert. Diesen Ort nennt man auch den Farbort. Die Helligkeit wird hierbei jedoch nicht berücksichtigt, weil ein Schwerpunkt sich nicht ändert, wenn die Farbwerte immer dasselbe Verhältnis zueinander haben, wenn also die Farbwertanteile (normierte Farbwerte, also bezogen auf die Summe der Farbwerte) konstant bleiben. Die Abstände eines beliebig liegenden Farbortes zu den Seiten des Dreiecks (also die Lotlängen) definieren hinreichend die Farbart, weil sie nämlich genau den Farbwertanteilen entsprechen, deren Summe aufgrund der Normierung ja 1 ergeben muss. Diejenigen Lote, die von außen auf die Begrenzungsgeraden des

Dreiecks (verlängerte Seiten) treffen, sind negativ zu rechnen; dies ist so zu deuten, dass die Farbe nur virtuell über ebenfalls negative Farbwertanteile aus den zur Verfügung stehenden Primärvalenzen generiert werden kann.

Die Farborte sind hinreichend definiert, wenn nur zwei Lotlängen bzw. Farbwertanteile angegeben werden. Der dritte Wert ergibt sich schlicht als Ergänzung zu 1.

Über ein umfangreiches physikalisch-physiologisches Messverfahren erhält man beim Durchlaufen aller sichtbaren Wellenlängen die Punkte des Spektralfarbenzuges, also die äußere Linie des Farbdiagramms. Sie befindet sich komplett außerhalb des Dreiecks, wodurch feststeht, dass keine einzige Spektralfarbe mittels Monitor durch additive Farbmischung simuliert werden kann. Das Messverfahren, welches u. a. zwei Monitore und einen Monochromator (Gerät zur Erzeugung von Licht einer Wellenlänge) einsetzt, kann dennoch diese Punkte ermitteln.

Ein Normungsverfahren modifiziert die Werte der Kurve dergestalt, dass sie in ein orthogonales Koordinatensystem passt, wobei wiederum die Angabe von nur zwei Werten hinreichend zur Definition eines Farbortes und somit einer Farbart ist. Aus dem gleichschenkligen Dreieck wird durch dieses Verfahren zugleich ein rechtwinkliges Dreieck mit Seitenlänge 1, das von den Koordinatenachsen aufgespannt wird. Dieses Dreieck schließt gänzlich die Kurve der Spektralfarben ein, weil die Eckpunkte durch virtuelle Normfarbvalenzen gegeben sind. Eine solche Darstellung, wie wir sie dann mit Abb. 7.10 erhalten, besteht bei gegebenem Farbort und Helligkeit (Leuchtdichte) eine vollständige Aussage über den Farbwert.

Eine umfassendere Erklärung zur Farbmatrik würde den Rahmen dieses Buches sprengen; zur Vertiefung dieser Thematik sei auf entsprechende Literatur verwiesen. S. 281 [7]

7.2.2 Prinzip der Farbbildaufnahme

Um Objekte in Bildsignale umzuwandeln, bedarf es im Echtzeitbereich elektronischer Kameras. Man kann Videokameras in zwei große Gruppen aufteilen, deren Unterscheidungsmerkmal die Bildaufnahmeeinheit ist. Wir wollen hier nur das Funktionsprinzip umreißen.

Röhrenkameras Ursprünglich waren Bildaufnahmegeräte mit Vakuumröhren spezieller Bauart bestückt, deren Prinzip wir uns anhand von Abb. 7.11 klarmachen wollen. Trifft kein Licht auf den Schirmträger, so lädt der Elektronenstrahl den fotoelektrischen Film fast auf Katodenpotential auf. Zwischen Nesa-Film (Zinnoxid) und fotoelektrischem Film baut sich also fast die Target-Spannung auf. Am Lastwiderstand R_L fällt fast keine Spannung ab, da der fotoelektrische Film nach erfolgter Aufladung keine Elektronen mehr aufnimmt. Natürlich werden die Elektronen, die aus der Kathode nach wie vor austreten, weiterhin beschleunigt durch die Anodenspannung – von dieser Elektrode werden sie auch aufgenommen, wenn dies der fotoelektrische Film dies nicht mehr bewerkstelligen kann.

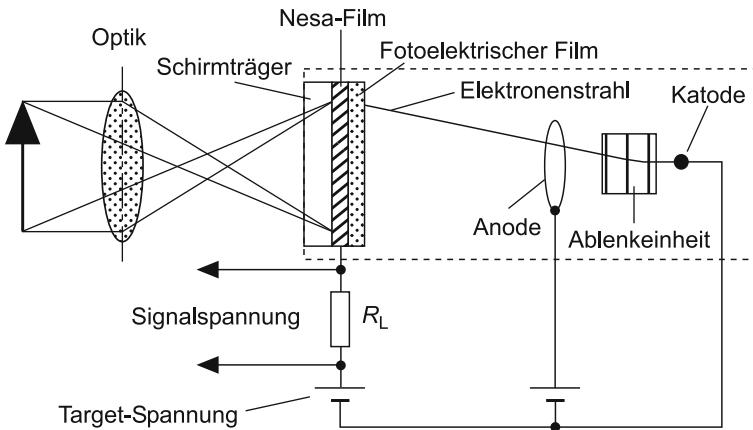


Abb. 7.11 Arbeitsweise eines Bildwandlers auf Röhrenbasis

Bei Lichteinfall entlädt sich an der betreffenden Stelle das Potential, und dadurch ist beim Auftreffen des Elektronenstrahls auch wieder ein Ladestrom möglich, der an R_L einen Spannungsabfall verursacht. Durch eine geeignete Rasterablenkung, welche mit einer magnetischen Ablenkeinheit erreicht wird, gelangt der Elektronenstrahl auf die gesamte Bildfläche des Sensors.

Bekannte Arten von Bildwandlern stellen z. B. das Vidikon oder auch das Plumbikon dar. Sie unterscheiden sich hauptsächlich durch die verwendeten Materialien des Bildelements und zeigen somit auch Unterschiede im Verhalten, was sich u. a. durch die Trägheit (Nachzieheffekte) äußert.

Halbleiterkameras Obwohl natürlich auch bei Röhrenkameras Halbleitermaterialien zum Einsatz kommen, benötigen Halbleiterkameras keinen Glaskolben mit Hochvakuum. Es handelt sich beim Halbleiter-Bildsensor um einen sog. *CCD-Wandler* (Charge Coupled Devices).

Die Technik der Hochintegration hat es ermöglicht, Fotosensoren in Form einer Matrix anzurufen und zugleich auf dem Chip Analogspeicherglieder aufzubringen. Jedem dieser Fotosensoren ist eine Speicherzelle zugeordnet. Abb. 7.12 zeigt schematisch diese Anordnung. Während der Zeit eines Halbbilds gelangen die Ladungen der Sensorzellen in die zugehörigen Speicherzellen. Innerhalb der Vertikalaustastzeit erfolgt eine Verschiebung aller gespeicherten Ladungen in den Zwischenspeicher, welcher über eine serielle Schieberegister zeilenweise ausgelesen wird. Eine spezielle Taktelektronik sorgt dafür, dass das Timing den Forderungen des Zeilensprungverfahrens gerecht wird. Es gibt unterschiedliche CCD-Verfahren, wobei hier ein modernes Prinzip erläutert wurde, nämlich das FIT-Verfahren (Frame Interline Transfer).

CCD-Kameras haben gegenüber Röhrenkameras den Vorteil, dass sie weniger stoßempfindlich sind. Außerdem ermöglichen sie natürlich eine weitaus kleinere Bauweise, wodurch spezielle Anwendungen (z. B. in der Medizin) erst möglich wurden.

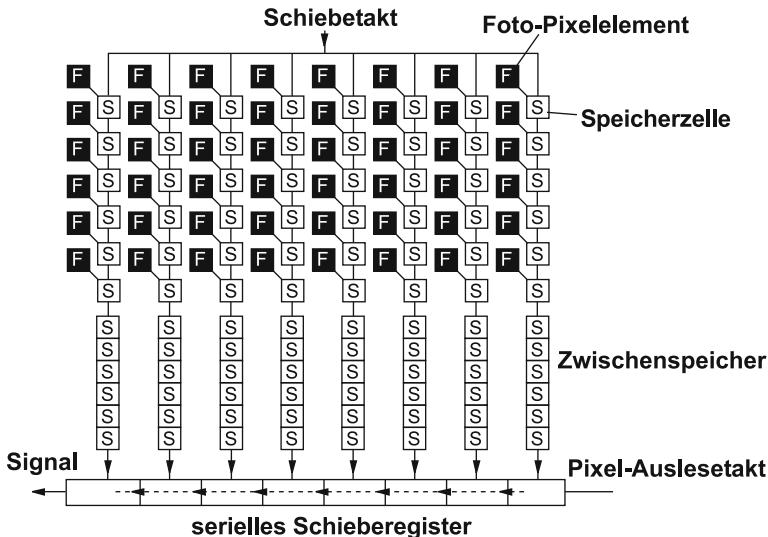


Abb. 7.12 Funktionsprinzip eines FIT-CCD-Wandlers

Kameras mit CMOS-Chip Im Gegensatz zur CCD-Technik sind beim CMOS-Kamerachip keine Ladungstransporte nötig. Jedes Pixel hat seine eigene Verstärkereinheit – man spricht deshalb auch von *Active Pixel Sensor* oder kurz APS. Während einer voreingestellten Belichtungszeit entlädt die Fotodiode ihre eigene Sperrsicht-Kapazität. Der erreichte Spannungseinbruch ist damit proportional zur Zeit und einfallender Leistungsdichte des Lichts.

Durch die Zeilen- und Spaltenansteuerung der Anordnung lässt sich jede Zelle separat ansprechen und auslesen. Der sog. *Blooming-Effekt* durch Übersprechen zu benachbarten Zellen bei Übersteuerung tritt im Gegensatz zur CCD-Technik kaum auf. Weitere Vorteile sind geringerer Stromverbrauch und schnelleres Auslesen der Bilddaten. Die CMOS-Technologie ist allerdings komplexer und somit teurer.

Kameras zur Farbbildverarbeitung Die oben beschriebenen Vorgänge erlauben zunächst nur die Umsetzung eines Schwarzweißbilds, da keine Farbselektion vorliegt. Hierfür sind spezielle Vorkehrungen zu treffen, wobei meist mit drei Systemen gearbeitet wird, die über Prismen oder dichroitische Spiegel das in Farbbereiche aufgeteilte Bild erhalten. Das so erhaltene RGB-Signal (siehe Abschn. 7.2.1) steht dann für eine weitere Verarbeitung zur Verfügung.

Eine andere Methode kommt mit nur einem Bildelement (Röhre oder CCD-Chip) aus; die Farbteilung geschieht hier mit einem Streifenfilter, welches eine feine senkrechte Unterteilung aufweist, die einer Sequenz von Farbfiltern darstellt. Das Bildsignal wird einem elektronischen Schalter zugeführt, der entsprechend der Abtastung und der Streifenteilung eine Separation der Farbsignale erreicht.

7.2.3 Prinzip der Farbbildwiedergabe

Liegt das Farbsignal nach der Übertragung (Sendung, Aufzeichnung/Wiedergabe usw.) wieder als Einzelsignale in RGB-Form vor, so kann eine Wiedergabe über eine geeigneten Einrichtung erfolgen.

Röhren-Monitor Die gängigste und auch älteste Methode, solche Signale in Form des reproduzierten Bildes wiederzugeben, besteht durch einen Farbmonitor auf der Basis einer *Braunschen Röhre*, bei der ein Elektronenstrahl mittels rasterförmiger Ablenkung und hoher Beschleunigungsspannung zum Auftreffen auf eine fluoreszierende Schicht gezwungen wird. Auf diese Weise sind zunächst jedoch nur monochromatische Bilder wiederzugeben (durch Wahl geeigneter Materialien ergibt sich dann ein Schwarzweißbild).

Um mit solch einem Monitor eine Farbbildwiedergabe zu ermöglichen, integriert man drei unabhängig steuerbare Katoden, deren austretende Elektronen mit einer besonderen Konvergenzeinrichtung dazu gebracht werden, bei allen Ablenkpositionen möglichst genau auf einen gemeinsamen Punkt der Monitorfläche aufzutreffen.

Um aus den Einzelstrahlen der Katoden einen entsprechenden Farbbeindruck zu erhalten, muss die Leuchtschicht in kleinste Farbtripel unterteilt und in dichtem Abstand davor eine Lochblende angebracht sein. Abb. 7.13 gibt den prinzipiellen Aufbau wieder.

Die Strahllinien einer Katode kann also immer nur auf ein Bildfarbelement der Tripel gelangen, dessen Leuchtfarbe mit der zu steuernden Farbe übereinstimmt. Entgegen der Darstellung handelt es sich eigentlich nicht um gerade Linien, weil auf dem gesamten Weg von den Katoden bis zum Leuchtpunkt statische und dynamische Ablenkungen erfolgen.

Konvergenz Damit beim Ablenkvorgang zu jedem Zeitpunkt alle drei Strahlen immer nur durch ein gemeinsames Loch gelangen, muss eine Konvergenzeinrichtung mit einem

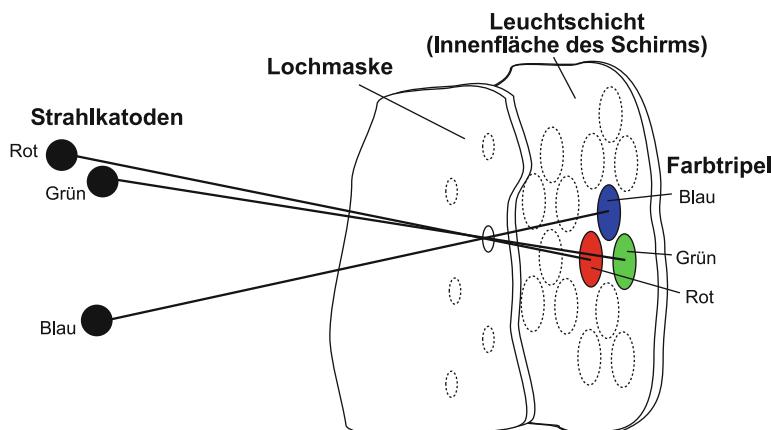


Abb. 7.13 Prinzip der Farbwiedergabe bei Röhren-Monitoren

Überlagerungsfeld die Strahlen separat beeinflussen. Man unterscheidet hier zwischen statischer und die dynamischer Konvergenz. Störungen zeigen sich hier durch die Entstehung von ungewollten Farben an Kontrastkanten eines Schwarzweißbildes. Das Symptom kann mit der Konvergenzeinheit beseitigt werden. S. 281 [8]

Farbreinheit Treffen Strahlen einer Katode auf Leuchtpunkte des Bildschirms einer falschen Farbe (also die Grünkatode z. B. auf Rotpunkte der Leuchtschicht), so liegt ein Fehler der Farbreinheit vor. Erzeugt sich vor allem bei der Wiedergabe einer homogenen Weißfläche. Mitunter kann dieser Effekt durch Justage beseitigt werden; es kann jedoch auch ein irreparabler Schaden der Bildröhre vorliegen. S. 281 [8]

Flimmern Vor allem im Interlaced-Modus und bei den niedrigen Vollbildraten von 25 Hz kann lästiges Flimmern auftreten. Es kann bedeutend reduziert werden, wenn die Bildröhre über eine Schirmschicht mit großer Nachleuchtdauer verfügt. Man ist jedoch immer gezwungen, einen Kompromiss zwischen Flimmerfreiheit und Nachzieheffekten einzugehen, die bei vergrößerter Nachleuchtdauer immer vorhanden sind, da die Bildröhre Bildpunkte zwar zum Aufleuchten, jedoch nicht zum Abdunkeln zwingen kann.

LCD-Monitor Als ein Resultat der Miniaturisierung und Gewichtsersparnis ist auch die Entwicklung von Flüssigkristall-Anzeigen (LCD = Liquid Crystal Display) zu sehen. Darüber hinaus ermöglicht diese relativ neue Technologie vor allem bei passiven Displays eine enorme Energieeinsparung.

TFT-Monitor Dabei handelt es sich um eine Weiterentwicklung von einfachen LCD-Bildschirmen. TFT bedeutet Thin Film Transistor, eine Technologie, bei der eine Matrix aus Transistoren direkt an der Leuchtschicht angesteuert werden. Moderne PC-Monitore sowie Fernseh-Wiedergabegeräte setzen häufig diese Technologie ein.

Zur Wiedergabe von Farbe wird weißes Licht mittels Farbfilter aufgeteilt und getrennt über DSTN-Zellen gesteuert. Bei der TFT-Technik erfolgt die Steuerung per Transistoren, die als Matrix an die aktiven LC-Zellen angeschlossen sind. Jede einzelne Zelle ist getrennt ansteuerbar, wodurch räumlich begrenzte Bildänderungen sehr viel schneller stattfinden können.

Nachfolgende Beschreibung gehört zu Abb. 7.14. Betrachten wir zunächst den linken Teil der Abbildung. Die Lichtquelle (meist Leuchtstoffröhre) sendet ihr Licht durch das Polarisationsfilter (a), so dass nur noch Licht mit einem linearen Feldvektor übrig bleibt. Dieses linear polarisierte Licht kann gleichwohl als Überlagerung zweier senkrecht aufeinander stehender Komponenten *ohne* Phasenverschiebung betrachtet werden (b). Die dicke Sinuslinie schwingt nur in einer Ebene. Das Licht gelangt weiter in die (spannungslose) aktive LC-Zelle (Liquid Crystal), wo es eine zirkulare (oder elliptische) Polarisierung erfährt. Dies geschieht dadurch, dass beide Schwingungskomponenten zueinander phasenverschoben sind. Der Feldvektor schreitet von nun an spiralenförmig im Raum weiter (c). Es folgt nun der Durchgang durch die passive LC-Zelle, die den

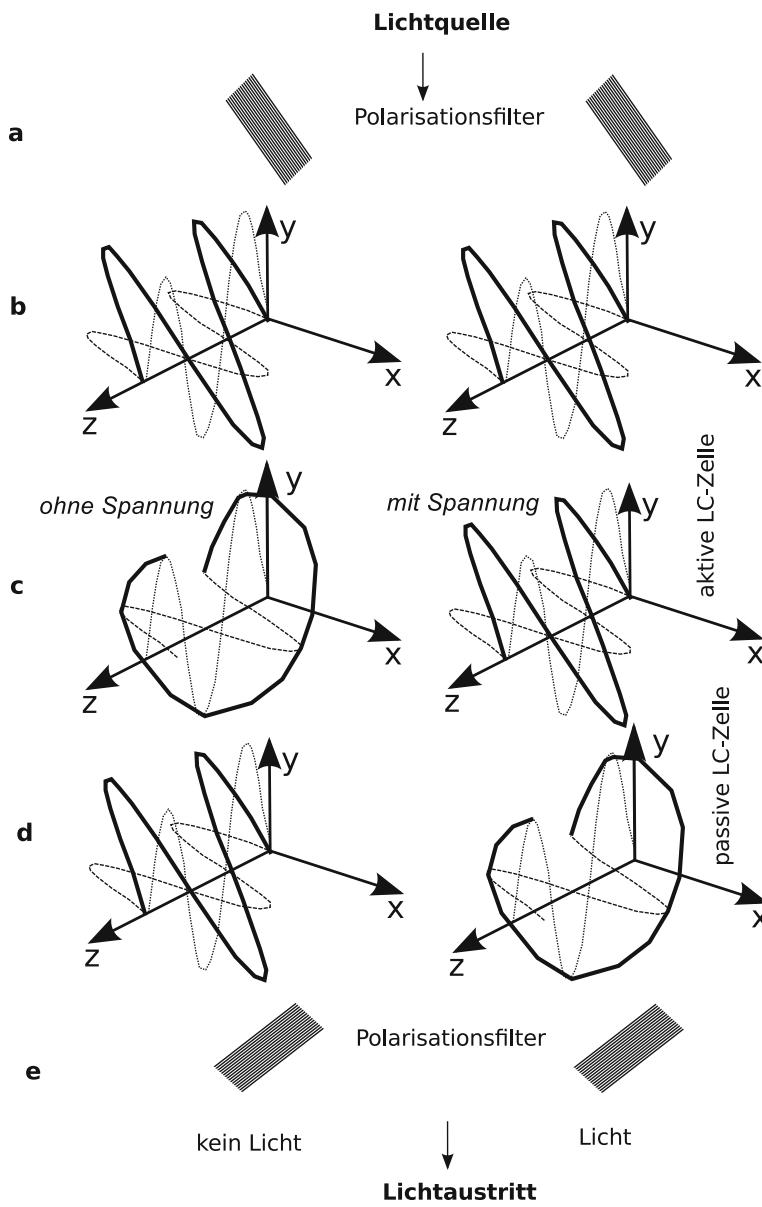


Abb. 7.14 Prinzip eines TFT-Monitors. Durch unterschiedliche Ansteuerung der Polarisations- schicht entsteht Undurchlässigkeit oder Durchlässigkeit für Licht. Siehe genauere Beschreibung im Text

Phasenunterschied in die entgegengesetzte Richtung betragsgleich vollzieht. Dadurch sind die Einzelkomponenten wieder in Phase und verhalten sich identisch zu dem vorherigen linear polarisierten Licht (**d**). Das untere Polarisationsfilter hat seine Ausrichtung exakt senkrecht zur Lichtpolarisation, somit tritt kein Licht aus (**e**). Durchlaufen wir den Weg im rechten Teil der Abbildung, so sehen wir als einzigen Unterschied den, dass die aktive Zelle jetzt unter Spannung steht (**c**). Das elektrische Feld darin bewirkt keine polarisierende Wirkung mehr. Die nachfolgende passive LC-Zelle macht wieder dieselbe Phasenverschiebung an den Komponenten wie vorher – da diese jedoch nicht phasenverschoben waren, erhalten sie jetzt einen Phasenwinkelversatz, sodass nunmehr zirkular polarisiertes Licht vorliegt (**d**). Für dieses Licht stellt aber das Polfilter keine Behinderung dar, da alle Schwingungsvektoren vorhanden sind. Dadurch wird am Polfilter jetzt Licht austreten (**e**)

Prinzip Da bei fehlender Spannung die Zelle lichtundurchlässig ist, spricht man von *Normally Black Mode*, während im umgekehrten Falle *Normally White Mode, NWM* vorliegt.

Es handelt sich hier um den Aufbau einer DSTN-Zelle (Double Super Twisted Nematic). Von den einfacheren TN- und STN-Zellen ausgehend war dies eine Weiterentwicklung, um vor allem den Kontrastumfang vergrößern zu können. Weiterhin gibt es eine Steigerung von DSTN in puncto Kontrast, Herstellungskosten und flacher Bauweise, nämlich die TSTN *Triple Super Twisted Nematic*.

Betrachtungswinkel Ein großes Thema bei LCD-Wiedergabeschirmen war die Abhängigkeit vom Blickwinkel. Inzwischen sind so vielfältige Verfahren entwickelt worden, um diese Effekte relativ gut unterdrücken zu können. Am meisten haben sich Multi-domain Vertical Alignment (MVA) und In-Plane Switching durchgesetzt. Es handelt sich dabei hauptsächlich um die Art der Anordnung der Flüssigkristallzellen, um eben die Abhängigkeit vom Blickwinkel möglichst optimal zu reduzieren. Wir wollen darauf nicht näher eingehen. S. 281 [9]

Hinterleuchtung Gleichmäßige Ausleuchtung der gesamten Bildschirmfläche geschieht mit einer lichtleitenden Kunststofffolie (Diffusor). Als Lichtquelle benutzt man meist Leuchtstoffröhren oder (meist bei Großanzeigen) Kaltkatodenröhren.

Daneben kommen auch LEDs bzw. LED-Gruppen als Element der Hinterleuchtung zur Anwendung. Auf dem Markt erscheinen diese Produkte häufig als *LED-Fernseher* oder *LED-Monitor*, obwohl die Bezeichnung freilich irreführend ist. Denn eigentlich handelt es sich weiterhin um LCD-Technologie und um keine *direktstrahlenden* LEDs. Es kommen zwei Arten des Aufbaus vor: Eine ordnet die LEDs außen an den Bildrand an (Edge-Slim-Technik) und verteilt das Licht über geeignete Diffusor-Folien, ähnlich wie bei den Lösungen mit Leuchtstoffröhren. Die zweite Art (Full-LED) verteilt mehrere hundert bis ca. 3000 Einzel-LEDs hinter der LCD-Maske, wobei dann Diffusoren entfallen. Vorteil beider LED-Hinterleuchtungen ist die Möglichkeit der lokalen Verdunkelung bis hin zu

Schwarz, so dass wirkliche schwarze Bereiche des Bildes auch nicht grau erscheinen. Der „Grobkontrast“ ist somit entscheidend besser als bei Leuchtstoff-Backlights, denn diese überstrahlen häufig die gesamte Schirmfläche – die LC-Zellen sind nicht in der Lage zu einer absoluten Sperrung.

Reaktionszeit Ohne TFT liegt eine passive Matrix vor, die Bildelemente werden ähnlich zum CCD-Chip einer Kamera im Zeitmultiplexverfahren angesteuert. Das verlangsamt die Reaktionszeit zufolge größerer Refresh-Zeiten. Die TFT-Technik lässt Änderungszeiten bis herunter nach ca. 5 ms zu, dies ermöglicht – ähnlich wie bei Röhrenmonitoren (CRT) – die Generierung von Zwischenbildern, um Bewegungsabläufe flüssiger und ohne Ruckeln zu gestalten (100-Hz- oder 200-Hz-Technik).

Auflösung Prinzipiell ist die Pixelstruktur bei einem TFT-LCD-Monitor fest, somit ist auch die native Auflösung Geräte-spezifisch. Weicht man davon ab, sind Interpolationsalgorithmen notwendig, damit das Bild keine Aliasing-Artefakte erhält. Der Monitor selbst muss diese Aktion leisten. Es wird zwar immer behauptet, Röhrenmonitore könnten verschiedene Bildauflösungen völlig problemlos verkraften, jedoch sind auch hier Schlieren zu beobachten, wenn das Pitch der Leuchtschicht abweicht vom Abstand der Pixel, die das Grafiksystem liefert. Die hauptsächliche Ursache liegt jedoch am nicht ganz kontinuierlichen (Glitch usw.) Signal des Grafiksystems.

Plasma-Schirm Für Großbildschirme wird auch häufig die Plasma-Technologie eingesetzt. Der Schirm besteht aus winzigen, geschlossenen Zellen, in denen ein Edelgas im Unterdruck eingefüllt ist. Durch Elektroden, an denen mehrere hundert Volt anliegen, zündet das Edelgas und sendet UV-Licht aus. Dieses trifft auf spezifische Leuchtschichten, die zur Wandlung in rotes, grünes oder blaues Licht geeignet sind. Die Steuerung der Leuchtstärke erfolgt durch Variation der Zündphase.

Durch zeilen- und spaltenweise Adressierung der entsprechend angeordneten Elektrodenstreifen ist es möglich, jede Zelle individuell zu anzusteuern.

LCD-Projektor Mit Flüssigkristallen können nicht nur feststehende Anzeigeflächen geschaffen werden; auch Projektionen sind möglich. Ganz analog der Bauweise eines Diaprojektors wird ein kleines, aber hell erleuchtetes Bild über eine Optik vergrößert auf eine Projektionsfläche geworfen. Das LCD-Element arbeitet normalerweise hier im Durchlichtbetrieb. Die Auflösung des Bildes hängt vor allem von der Pixelanzahl des LCD-Elements ab. Durch die Verwendung eines Strahlengangs kommt es zu keinerlei Konvergenzproblemen, d. h. die Bilder aller drei Farben kommen auf der gesamten Bildfläche zur Deckung. Ursachen für Geometrieverzerrungen sind hier rein optisch bedingt und können durch Wahl von hochwertigen Objektiven minimiert werden.

Der LCD-Projektor eignet sich gut für Großbildvorführungen und zeichnet sich i. Allg. durch gute Auflösung aus. Flimmerprobleme sind bei diesem System nicht zu befürchten, da die Pixelhelligkeit immer solange gehalten wird, bis sie aktualisiert wird (d. h. die

Nachleuchtdauer ist hier im Gegensatz zu Bildröhren mit Spezialschicht nicht länger als nötig).

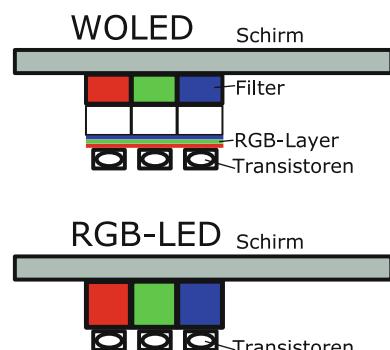
Licht-Beamer Zur Großbildprojektion gibt es bereits seit einigen Jahren sog. Beamer, die eine optische Trennung der Farbbilder aufweisen. Dadurch kann es bei diesen Systemen teilweise zu unangenehmen Konvergenzproblemen kommen, die meist nur mühsam und mangelhaft zu beseitigen sind.

Zur Erzeugung und Modulation des Lichts sowie der Bewältigung der Bildablenkung sind verschiedene Systeme bekannt, auf die hier nicht näher eingegangen werden soll. S. 281 [10]

LED-Bildschirm Eigentlich handelt es sich beim „richtigen“ LED-Wiedergabeschirm um einen OLED-Monitor (Organic Light Emitting Diode). Im Gegensatz zu den LED-hinterleuchteten Schirmen (siehe weiter oben) sind hier jeweils drei OLEDs für ein Pixel zuständig. Diese Technologie gestattet mikroskopisch kleine Anordnungen, die für eine Pixelmatrix nötig ist. Diese Direktstrahlung benötigt keine Hinterleuchtung und keine LC-Zellen. Der Preis liegt relativ hoch, und die Lebensdauer der OLEDs ist geringer als bei anderen Technologien. Trotzdem sind Lebensdauern von 15 Jahren Dauerbetrieb im Gespräch. Technisch gesehen ist die Bildqualität höher als bei den LCD-Bildschirmen, dies äußert sich in einer weiteren Kontraststeigerung und Erweiterung des Farbraumes. Der größte Vorteil gegenüber LCD liegt jedoch in der immens gesteigerten Reaktionsschnelligkeit. Ein einzelnes Pixel wechselt seine Helligkeit in weniger als einer Mikrosekunde.

Die Emission erfolgt in metallorganischen Verbindungen. Im Gegensatz zu den LCD-Monitoren ist hier keine LCD-Schicht notwendig. Die einzelnen Pixel verfügen über eine spontane Leuchtkraft, erzeugt durch mikroskopisch kleine, einzelne LEDs. Ohne näher auf die unterschiedlichen Technologien eingehen zu wollen, sind hauptsächlich zwei Arten am gebräuchlisten: Bei der RGB-SBS-AMOLED-Technologie (RGB,-Side-by-Side-Aktiv-Matrix-OLED) existieren je drei Farb-LEDs pro Pixel, während die WOLED-Technologie (White-OLED) aus weißen LEDs besteht, auf die Farbfilter folgen (siehe Abb. 7.15).

Abb. 7.15 Schematischer Aufbau von OLED-Displays



Das LED-Element selbst ist nicht einkristallin, sondern ist in Dünnschichttechnik gefertigt, was den Herstellungsprozess außerordentlich vereinfacht.

Vorteile Da bei dieser Technik keine globale oder pixel-übergreifende Hinterleuchtung nötig ist, ist eine scharfe Kontrastierung möglich, ja man kann sogar einzelne Pixel total in den schwarzen Zustand bringen. Außerdem ist die Reaktionszeit einzelner Pixel um Größenordnungen schneller (ca. $1 \mu\text{s}$) als bei LCD-Systemen.

Nachteile Leider ist die Technologie noch nicht so ausgereift wie bei LCD-Systemen, sodass die Ausschussquote noch merklich höher liegt. Außerdem ist die Lebensdauer der Displays geringer. Die Alterung der einzelnen Farben kann unterschiedlich schnell ablaufen, sodass sogar ein Farbstich auftreten kann. S. 281 [4]

7.2.4 RGB-Signal

Bei der einfachsten Übertragung eines Farbbildes über Leitungen bedient man sich des sog. RGB-Signals. Hierbei werden die Bilddaten für die drei Farben Rot, Grün und Blau getrennt auf jeweils einer Leitung übertragen. Dies geschieht meist in analoger Form, d. h. die Spannungen auf diesen Leitungen stehen in direktem und linearem Zusammenhang mit der Helligkeit, mit der die jeweilige Farbe wiedergegeben werden soll.

Beim PAL-Farbübertragungsverfahren hat man sich darauf geeinigt, dass die zulässigen Werte der Einzelspannungen im Bereich $0 \dots 714 \text{ mV}$ liegen sollen; wir kommen in Abschn. 7.1.4 noch genauer darauf zu sprechen.

Zur Synchronisation bei RGB verwendet man in der Videotechnik meist eine getrennte Leitung, die die Impulse für die Bild- und Zeilensynchronisation führt. In Abschn. 7.1.3 ist das Aussehen und die Wirkungsweise eines solchen Composite Sync beschrieben.

Manche Systeme und Geräte ermöglichen auch eine Kopplung des Synchronsignals auf der Grün-Leitung, wodurch man dann auf eine separate Synchronisationsleitung verzichten kann. Für manche Computeranwendungen werden dagegen Horizontal- und Vertikalsynchronisation getrennt geführt, so dass also insgesamt fünf Leitungen erforderlich sind.

EBU-Farbbalken Für Mess- und Testzwecke hat man verschiedene Testbilder eingeführt. Eine große Rolle spielt dabei das Signal der EBU-Farbbalken (European Broadcasting Union). Es dient vor allem der optimalen Pegeleinmessung bzw. -überprüfung. Die Signalverläufe für ein nach B/G-PAL festgelegtes Format sind in Abb. 7.16 gezeigt.

Solange es sich um ein reines RGB-Signal handelt, kann die Signalamplitude bei allen Farben bis 100 %, also bis 714 mV gehen. Sobald jedoch eine Wandlung in eine andere Signalart erfolgen soll (vor allem in eine solche mit Farbhilfsträger), darf bei reinen Farben jeweils nur ein Pegel von 75 % erreicht werden, sonst ist mit Übersteuerungen zu rechnen.

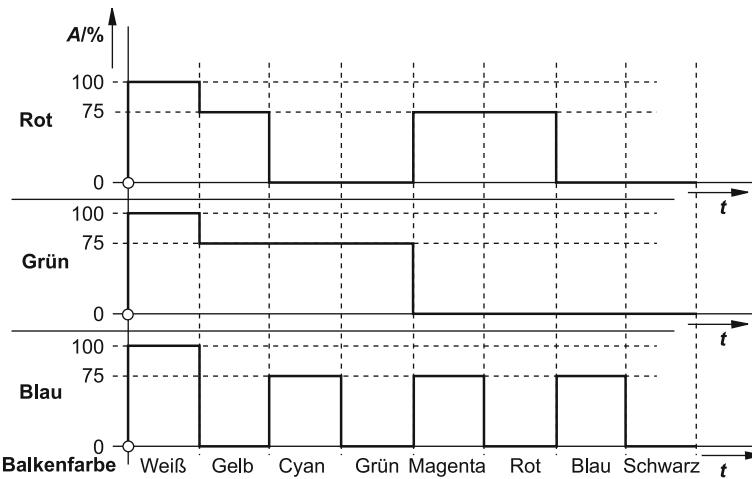


Abb. 7.16 Signalverlauf während einer Zeile bei EBU-Farbbalken (100/75)

Eine reine Farbe liegt auch dann vor, wenn nur zwei Primärfarben zur Mischung beitragen (z. B. für Gelb benötigt man Rot und Grün).

Merkmale der Signalqualität Das Signal liegt während der Bildaufnahme durch eine Kamera oder bei der Wiedergabe mit einer Grafikkarte normalerweise zunächst im RGB-Format vor. Da die Primärsignale in diesem Stadium in keiner Weise irgendeiner Codierung, Trägerung oder Filterung unterliegen, gelangen sie in voller Bandbreite von 5 MHz und höher zur weiteren Verarbeitung.

Monitore, die solch ein breitbandiges Signal verarbeiten können, zeigen selbst an vertikalen Kanten, die schnelle Helligkeits- oder Farbsprünge bedeuten, keinerlei Störungen wie Verwischung oder Schattenschwingungen. Dennoch gibt bei besonderen Konstellationen zwischen Bildmuster und Rasterung der Lochblende (bei Röhren-Monitoren) starke Moiré-Bildungen, die aufgrund von Aliasing-Störungen entstehen. Abb. 7.17 zeigt Beispiele solcher Effekte; die einzelnen Musterstreifen repräsentieren verschiedene Frequenzdifferenzen zwischen Raster- und Bildabtastung (der oberste Streifen zeigt das unverfälschte Schachbrettmuster).

Da sich hierbei bereits die kleinsten Exemplarstreuungen auswirken, kann dasselbe Signal bei mehreren Monitoren unterschiedliche Resultate liefern (je nach eingestellter Bildgröße). Zunächst handelt es sich dabei nur um ein Wiedergabeproblem – das Signal selbst braucht keinerlei Qualitätseinschränkungen aufweisen. Man wird bei derartigen Effekten jedoch gut daran tun, das Resultat nachfolgender Geräte für die Signalaufbereitung unter Kontrolle zu halten.

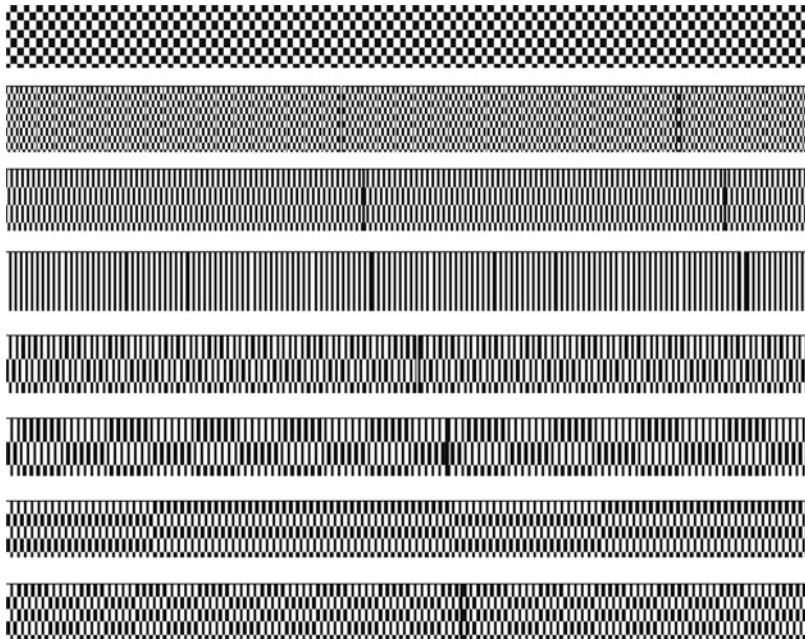


Abb. 7.17 Aliasing-Störung bei der Bildwiedergabe. Die Signalfrequenz kommt in die Nähe der durch die Überstreichung der Lochmaske wirksamen Abtastfrequenz

7.2.5 YUV-Signal

Auf dem Wege zur Signalaufbereitung zum FBAS-Signal durchläuft das RGB-Signal zunächst die Wandlung nach YUV, was auch Komponenten- oder Betacam-Signal genannt wird. Dieses in der Studiotechnik sehr gebräuchliche Signalformat kommt wie das RGB-Signal ohne Träger aus, da die Farbinformation durch arithmetische Operationen entsteht. Im wesentlichen kommen bei der Signalaufbereitung drei Gleichungen zum Tragen:

$$\begin{aligned}
 Y &= 0,299 \cdot R + 0,587 \cdot G + 0,114 \cdot B \\
 U &= 0,493 \cdot (B - Y) \\
 V &= 0,877 \cdot (R - Y)
 \end{aligned} \tag{7.4}$$

Die oberste Gleichung haben wir bereits weiter oben kennengelernt; legt man bei einem RGB-Monitor mit Weißabgleich an alle drei Farbeingänge den Maximalpegel mit 714 mV (100 %) gegenüber Schwarz, so entsteht Weiß, und Y wird genau 1. Die Klammerterme der unteren beiden Gleichungen in 7.5 stellen einfach Farbdifferenzen dar, welche durch die beiden Faktoren einer Bewertung unterzogen werden. In Abb. 7.18 ist die Entstehung der drei Signale aus RGB dargestellt. Die Zwischenstufe der reinen Farbdifferenzen $B - Y$ und $R - Y$ wurde hier bereits übersprungen.

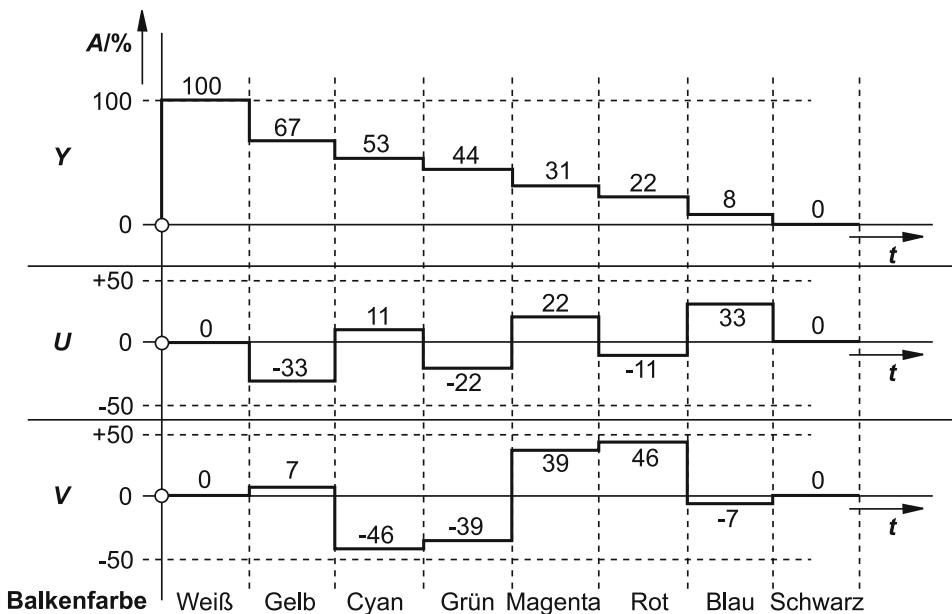


Abb. 7.18 Farbbalkensignal im Komponentenformat (YUV)

Das Synchronsignal ist in Y mit enthalten (hier nicht dargestellt) und folgt den Ausführungen in Abschn. 7.1.3. Auf diese Weise kann z. B. zu reinen Überwachungszwecken eine Schwarzweiß-Wiedergabe erfolgen, indem nur das Y -Signal genutzt wird.

Direkteinsatz des YUV-Signals Wie bereits erwähnt, kommt das Komponentensignal in der Studiotechnik sehr häufig direkt zum Einsatz. Es kann z. B. mit Betacam-Maschinen aufgezeichnet werden; doch auch viele andere Geräte wie Mischer und Kameras arbeiten auf diesem Standard.

Der 4-2-2-Standard (siehe Abschn. 7.3.10) sieht weiterhin die Möglichkeit vor, das YUV-Signal digital weiterzuverarbeiten. Hierzu müssen die modifizierten Farbdifferenzsignale U und V in ihrer Bandbreite begrenzt werden, da sie mit einer Rate von nur 6,75 MHz zur Abtastung gelangen. Die Bandbegrenzung ist festgelegt auf 2,75 MHz (sie muss nach dem Shannon-Theorem kleiner als die halbe Abtastfrequenz sein).

Signalqualität Ebenso wie das RGB-Signal kommt man bei YUV ohne Hilfsträger aus; die Farbinformation liegt lediglich in einer arithmetischen Codierung vor, nicht in einer modulierten. Die Helligkeits- und Farbauflösung ist deshalb sehr hoch, und es kommt im Regelfall zu keinen Verzeichnungen oder Schattenbildungen. Sogar nach einer Bandbegrenzung der Farbdifferenzsignale sind keine solchen Probleme zu befürchten.

Rück-Transformation nach RGB Wie aus den Gleichungen hervorgeht, kommt die Farbe Grün nicht bei den Farbdifferenzen vor. Dennoch wird sie natürlich indirekt übertragen, weil sie in der Luminanz-Gleichung enthalten ist. Die Rückbildung von RGB aus YUV geschieht durch die elektronische Nachbildung der Gleichungen:

$$\begin{aligned} R &= \frac{V}{0,877} + Y \\ G &= Y - 0,394 U - 0,581 V \\ B &= \frac{U}{0,493} + Y \end{aligned} \tag{7.5}$$

Selbstverständlich kann, falls bei U und V eine Bandbegrenzung vorgelegen hat, das rücktransformierte RGB-Signal ebenfalls keine höheren Bandbereiche der Farbwiedergabe liefern.

7.2.6 Y/C-Signal

Das Y/C-Signal, das landläufig wegen seiner entsprechenden Weiterverarbeitung auch S-VHS-Signal genannt wird, setzt erstmals einen Farbhilfsträger zur Codierung der Farbinformation ein. Zur Übertragung sind zwei Leitungen erforderlich.

Das Y-Signal entspricht dem des YUV-Signals oder auch dem BAS-Signal. Es enthält somit lediglich die Helligkeitsinformation sowie die Signale für die Synchronisation (Bild- und Zeilenstand). Das Y-Signal muss gegenüber dem unten beschriebenen C-Signal eine Verzögerung von ca. $1\mu\text{s}$ erfahren, da letzteres durch die Aufbereitung und die Bandbegrenzung zeitlich dem Y-Signal nachhinkt.

Die Aufbereitung des C-Signals geht aus Abb. 7.19 hervor. Die bandbegrenzten (Tiefpass 1,3 MHz) Farbdifferenzsignale U und V gelangen jeweils zu einem Farbmodulator, wo eine Amplitudenmodulation stattfindet. Der zu modulierende Farbhilfsträger hat eine Frequenz von ca. 4,43 MHz; der Oszillator liefert dem U -Modulator den unveränderten Hilfsträger, während der V -Modulator einen um 90° phasenverschobenen Hilfsträger erhält. Außerdem wird letzterer nach jeder Zeile komplett umgepolt, so dass letztendlich eine alternierende Trägerschwingung von -90° und $+90^\circ$ entsteht. Versteht man diese beiden Trägerschwingungen U und $\pm V$ als Modulationsachsen, so entsprechen die die Modulatoren verlassenden Signale den gesteuerten Anteilen U_1 und V_1 mit denselben Phasenlagen. Diese Signale können als rotierende Vektoren betrachtet werden; eine Signaladdition kommt daher einer sie in Abb. 7.20 zu sehen ist.

Aus den statischen Gleichtspannungen U_1 und V_1 werden also Farbhilfsträgerschwingungen, die gegeneinander um 90° phasenverschoben sind. Die Addition der Vektoren ist äquivalent zur Addition der Schwingungen.

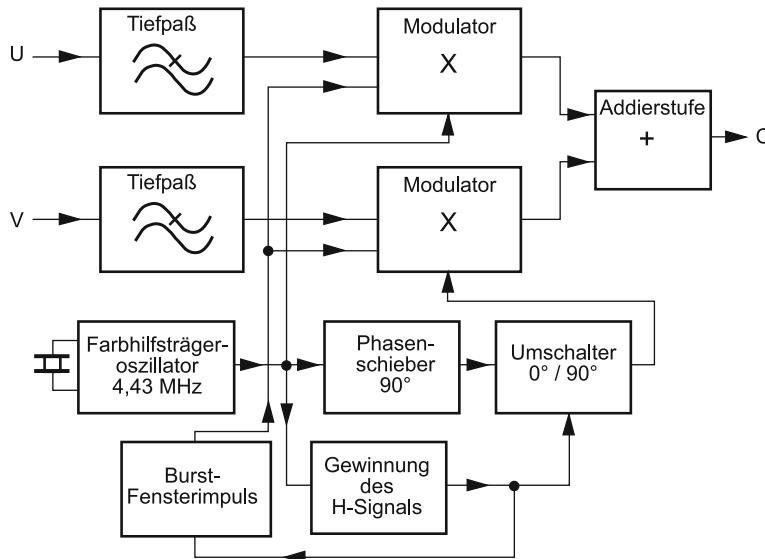
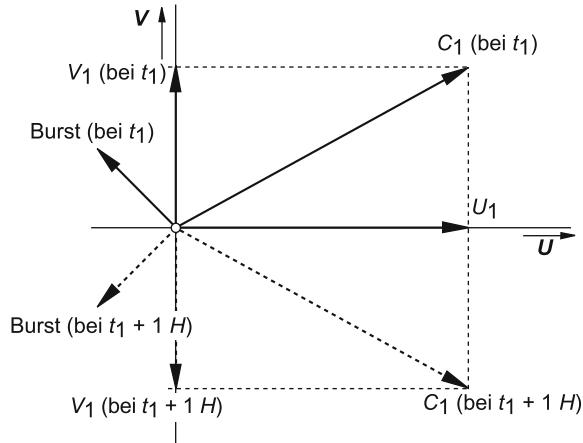


Abb. 7.19 Gewinnung des Farbartsignals

Abb. 7.20 Entstehung des Farbartsignals durch Vektoraddition



Burst und Farbsynchronisation Die Burst-Schwingungen, die nach Abb. 7.19 ebenfalls durch die Vorgabe von U- und V-Komponenten und durch Vektoraddition nach Abb. 7.20 entstehen, dienen als Referenz für die Farbvektoren C_1 . Zur späteren Rückführung in die nichtmodulierten Farbdifferenzsignale kann aus der Phasenlage des Farbvektors gegenüber dem Burst auf den Farbton geschlossen werden (siehe weiter unten). Der Burst besteht aus ca. 10 – 12 Schwingungen des Farbhilfsträgers, die während der Horizontalauastastung übertragen werden.

Farbton und Farbsättigung Zunächst ergibt sich Betrag und Winkel des Farbvektors zu:

$$|C_1| = \sqrt{U_1^2 + V_1^2}$$

$$\varphi = \arctan \left(\frac{V_1}{U_1} \right) \quad (+180^\circ \text{ für } U < 0) \quad (7.6)$$

Winkel φ bestimmt wie erwähnt den zu übertragenden Farbton, wobei er (entsprechend der Vektorspiegelung an der U-Achse) nach jeder Zeile alternierendes Vorzeichen erhält. Die Länge des Vektors ist ein Maß für die Farbsättigung – sie ist dennoch dadurch noch nicht hinreichend bestimmt. Betrachten wir die Farbe Magenta, die durch Mischung gleicher Teile von Rot und Blau entsteht. Verändert man beide im gleichen Verhältnis, so ändert sich die Farbsättigung nicht, sehr wohl aber Länge des Farbvektors. Als mitentscheidendes Kriterium für die Farbsättigung muss also die Leuchtdichte gewertet werden. Für eine maximale Sättigung S von 100 % ist jedoch noch ein Umrechnungsfaktor notwendig:

$$S = 0,7 \cdot \frac{C}{Y} \quad (7.7)$$

Für die Farbtöne der Farbbalken stellen sich die nach Tab. 7.1 angegebenen Phasenwinkel und (bei vollen Einzelpiegeln) Farbdifferenzwerte ein.

Chrominanz-Signalprofil Die Situation vor und nach der Addition der Trägersignale U und V ist in Abb. 7.21 zu sehen. Letztere erhalten ein Profil, das den ungeträgerten Signalen U und V nach Abb. 7.18 entsprechen. Erst nach dem Passieren der Modulatoren ergeben sich Farbhilfsträgerschwingungen mit entsprechender Amplitudenmodulation.

Rücktransformation ins Format YUV Die Rückwandlung geschieht mittels Decodern, wobei je nach Ausführung auch ins RGB-Format transformiert wird. Der im C-Signal enthaltene Burst (zwischen zwei Zeilen übertragen, in obiger Abbildung nicht dargestellt)

Tab. 7.1 Phasenwinkel und Farbdifferenzwerte der Farbbalken 100/75

Farbe	U	V	Phasenwinkel
Weiß	0	0	–
Gelb	−0,3276	0,08	+167,1°
Cyan	0,11	−0,4611	−76,5°
Grün	−0,2170	−0,3861	+240,7°
Magenta	0,22	0,39	+60,7°
Rot	−0,1106	0,46	+103,5°
Blau	0,33	−0,0750	−12,9°
Schwarz	0	0	–

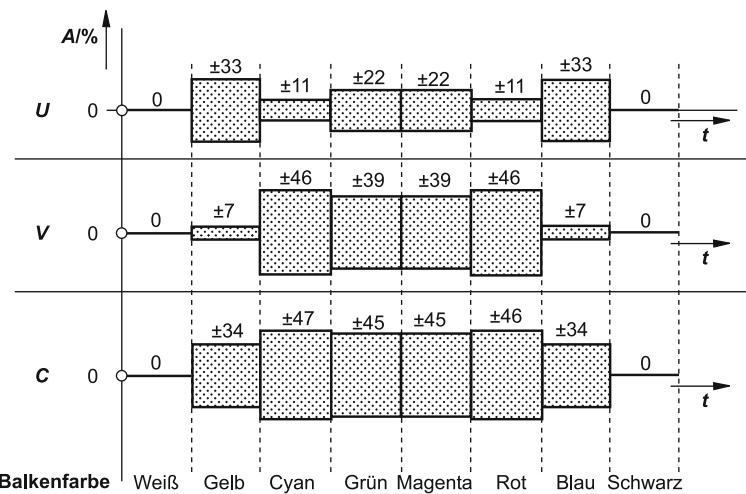


Abb. 7.21 Signale der Farbvektoren bei Verwendung von EBU-Farbbalken 100/75

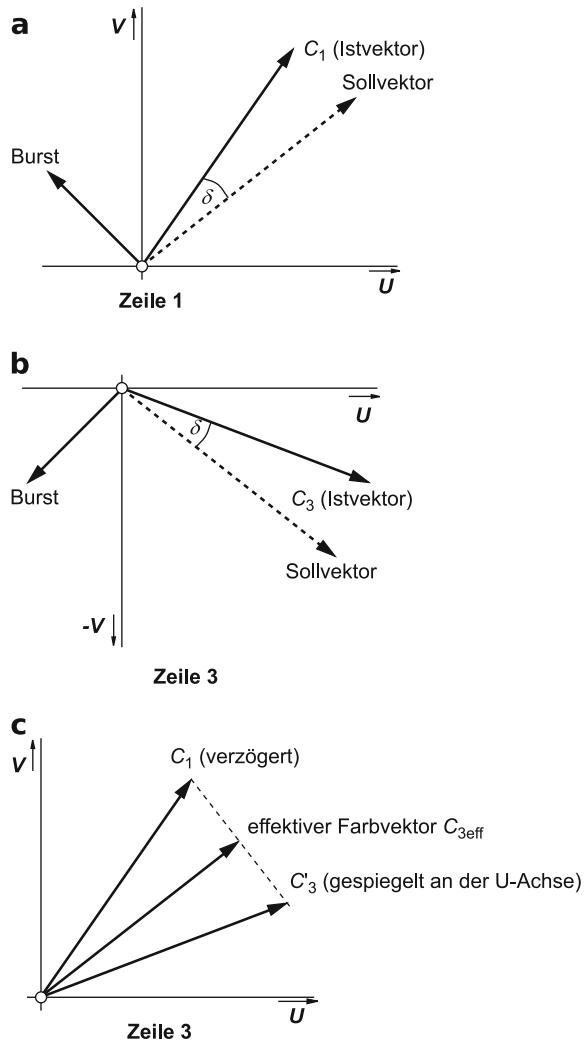
synchronisiert einen Quarzoszillator gleicher Frequenz. Das im Y-Signal vorhandene Synchronsignal triggert das sog. PAL-Flipflop; dabei erfolgt eine definierte Taktung, indem bei der Teilung noch die Phasenlage des Burst herangezogen wird.

Im PAL-Laufzeitdemodulator, bestehend aus Laufzeitleitung und zweifachen Synchrongeradialmodulatoren erfolgt die Wiedergewinnung der trägerlosen Farbdifferenzsignale U und V. Die Laufzeitleitung bewirkt, dass die laufende Zeile immer durch die zuletzt gespeicherte Farbinformation der vergangenen Zeile demoduliert wird. Damit der Synchrongeradialmodulator für das geträgerete V-Signal richtig arbeiten kann, muss dieser eine Umpolung seiner Demodulatorachse vornehmen, was durch das PAL-Flipflop geschieht. Zunächst findet die Trennung in die geträgereten Signale U und V statt, welche über Ratio-Detektoren in die niederfrequenten Signale U und V überführt werden. Eine ausführlichere Beschreibung der Demodulation ist in spezieller Literatur zu finden. S. 281 [8], S. 281 [11]

Farbfehlerkompensation Die Laufzeitleitung beim PAL-System und die damit verbundene Tatsache, dass die laufende Zeile immer durch die gespeicherte Farbinformation der vergangenen Zeile demoduliert wird, ermöglicht einen Ausgleich von Farbtonfehler. Betrachten wir hierzu eine Farbe, die in zwei Zeilen erscheinen soll – darauf basiert nämlich der Grundgedanke, dass sich die Farbinformation zweier zeitlich aufeinanderfolgender Zeilen kaum unterscheidet. Da die Phase der Modulationsachse V nach jeder Zeile alterniert, erscheint der entsprechende Farbvektor beider Zeilen also spiegelbildlich zur U-Achse (Abb. 7.22).

Signalqualität Die Bandbreite des Chromasignals ist mit 1,3 MHz geringer gehalten als bei YUV, was mit der Frequenz des Farbhilfsträgers (4,43 MHz) zusammenhängt. Als

Abb. 7.22 Ausgleich eines Farbtonfehlers beim PAL-System. (a) Vektor C_1 ist um einen Fehlwinkel δ gegenüber dem Soll verschoben. (b) In der nächsten Zeile besteht dieselbe Phasenverschiebung. (c) Bei der Demodulation wird C_3 zu C'_3 gespiegelt und der Durchschnitt mit dem Vektor C_1 (verzögerte Zeile) gebildet. Es entsteht der Farbvektor $C_{3\text{eff}}$, der die Phasenlage des Sollvektors hat und somit denselben Farbton ergibt; es ist lediglich eine leichte Kürzung und somit Entättigung eingetreten



Folge ergibt sich hier natürlich eine etwas schlechtere Farbauflösung, die jedoch noch akzeptabel ist. Das Leuchtdichtesignal Y wird nach wie vor mit der vollen Bandbreite von 5 MHz übertragen.

Verwendung von Y/C In vielen semiprofessionellen Bereichen wird heute das Y/C- bzw. das S-VHS-Signal eingesetzt. Viele Kameras, Recorder und auch Videografikkarten sind in der Lage, diese Signalart zu verarbeiten. Gegenüber Studiogeräten, die nur für das FBAS-Signal (siehe unten) geeignet sind, ist bei Verwendung von S-VHS-Geräten ein deutlicher Qualitätsvorsprung zu verzeichnen.

7.2.7 FBAS-Signal

Kommen wir zur letzten Signalart der Verarbeitungskette, dem FBAS-Signal. Gegenüber dem BAS-Signal kommt lediglich noch die Farbinformation und der Burst hinzu. Gegenüber dem Y/C-Signal sind lediglich die beiden Signale Y und C zu addieren, um das Format FBAS zu erhalten. Wie sich eine solche Addition im Aussehen des Signals niederschlägt, ist in Abb. 7.23 zu sehen.

Da es sich nur noch um ein Signal handelt, ist für dessen Übertragung auch nur eine Leitung notwendig. Die Mittelwerte der Schwingungen (gestrichelt gezeichnet im FBAS-Signal) entsprechen den Luminanzwerten. Die Synchronisation des Bildes und der Farbe ist aus dem Gesamtsignal in Abb. 7.24 ersichtlich – die Synchronimpulse stammen vom Y-Signal, während der Burst vom C-Signal kommt. Alle weiteren Zeitangaben sind ebenfalls enthalten.

Zunächst gelten für das FBAS-Signal die Eigenschaften in Bezug auf Farbfehlerkompensation, Farbwinkel, Sättigung usw. Dennoch gibt es natürlich diverse Unterschiede, die sich nicht nur darauf beschränken, dass zur Übertragung von FBAS nur noch eine Leitung notwendig ist.

Signalqualität Die Verschmelzung zwischen Y- und C-Signal zieht zwangsläufig Störungen nach sich, die irreparabel sind. Am unangenehmsten sind hier die Cross-Color-Störungen, die dadurch verursacht werden, dass bei der Trennung in die Signale Y und C Leckverhalten unvermeidlich sind.

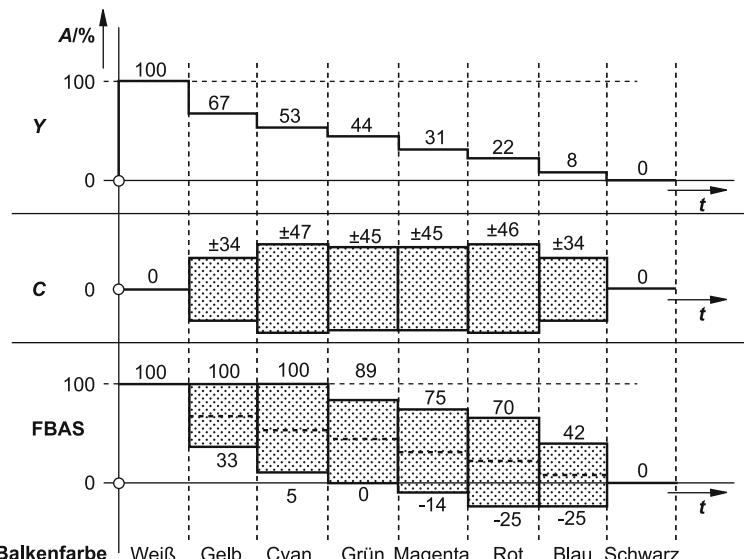


Abb. 7.23 Synthese des FBAS-Signals aus Y und C

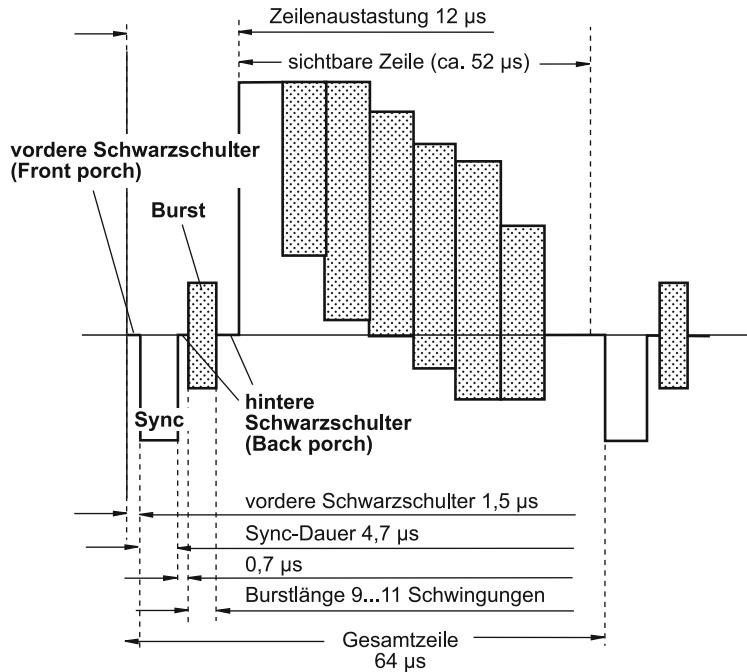


Abb. 7.24 Gesamtsignal FBAS mit Timings

Frequenzen, die als Helligkeitsschwankung in die Nähe der Farbhilfsträgerfrequenz kommen, werden auch nach der Rücktrennung in den C-Kanal gelangen, und dann erfolgt vom Decoder eine unerwünschte Farbinterpretation. Einmal als FBAS-Signal vorliegend, können also Frequenzen im Bereich 4...5 MHz nicht vollständig und richtig als Farb- oder als Helligkeitsinformation identifiziert werden. Ein Teil gelangt daher immer zur Farbdecodierung. Wenn im Originalbild Graumuster vorliegen, die einer hohen Signalfrequenz entsprechen, so sieht man bei der Reproduktion des Bildes daher Farbstörungen.

Bei der Cross-Luminanz-Störung tritt das Umgekehrte ein: Farbschwingungen werden als Helligkeitsschwankung interpretiert. Die Gewinnung der Helligkeitsinformationen aus dem FBAS-Signal geschieht durch Herausfiltern des Chromasignals mittels Frequenzsperre. Die Funktion dieser Sperre kann natürlich auch aufgrund von Seitenbändern nicht ein absolutes Fernhalten des Chromasignals vom Luminanzkanal bewirken. Je mehr die Chromaprodukte hier unterdrückt werden, desto stärker entsteht ein Einbruch im Frequenzgang des Luminanzkanals. Es ist also immer ein Kompromiss zwischen störendem Perlschnurmuster bei Cross-Luminanz und mangelhaftem Y-Frequenzgang zu schließen.

Die meisten Testbilder enthalten Bereiche, die variable Y-Frequenzen repräsentieren; das Ergebnis in Bezug auf Cross-Color lässt sich hiermit einfach prüfen. Die Cross-Luminanz-Eigenschaften lassen sich hingegen einfach nachweisen, indem durch Abschalten des Burst Schwarzweiß-Wiedergabe erzwungen wird. Das verbleibende Perl schnurmuster ist ein sicheres Indiz für schlechte Unterdrückung des Farbhilfsträgers im Luminanzkanal.

Die oben genannten Störeffekte sind unvermeidbar, wenn das Signal nur einmal in das Format FBAS gewandelt wurde; das Arbeiten mit Y/C vermag also keine Störunterdrückung zu bringen, wenn das Signal zuvor oder danach in FBAS gewandelt wird.

Das Erzeugen mehrerer Generationen von Videokopien verschlechtert die Qualität enorm; alle anderen Formate zeigen in dieser Beziehung wesentlich günstigere Eigenschaften.

Verwendung von FBAS Alle Analogsignale, die zur Zeit drahtlos übertragen werden, nehmen zunächst das Format FBAS an und werden dann jedoch über eine höhere Frequenz nochmals geträgert. Ganz allgemein erfolgt bei sequentiellen Analogsignalen immer eine Übertragung in FBAS.

VHS-Recorder können nur FBAS aufzeichnen, und auch bei vielen landläufigen Studiomaschinen war nur ein entsprechendes Handling möglich. Für Amateurzwecke ausgerichtete Videografikkarten verarbeiten ebenfalls nur FBAS-Signale. Auf der Produktionsseite reicht heute dieser Standard längst nicht mehr aus, wenn Bildqualität ein entscheidendes Kriterium ist.

Rücktransformation ins Format Y/C Wir hatten bereits weiter oben die Aufteilung in die Signale Y und C erwähnt. Prinzipiell sind hier lediglich geeignete Filter erforderlich, die als Frequenzweiche dienen, wobei im einfachsten Fall das Y-Signal durch einen Sperrkreis (abgestimmt auf Farbhilfsträgerfrequenz) und das C-Signal durch einen Bandpass gleicher Mittenfrequenz vom Gesamtsignal separiert wird.

Diese Filter sind Bestandteil des Decoders, der – je nach Bauart – weiter in RGB oder YUV wandelt. Eine optimale Abstimmung bzw. Dimensionierung entscheidet darüber, ob das systembedingte Optimum der Signalqualität für FBAS erzielt wird.

7.2.8 Viertelzeilenoffset-Verfahren

Da sich die Farbhilfsträgerschwingung bei der Rückgewinnung in das RGB-Signal nicht vollständig unterdrücken lässt, wird das sog. Viertelzeilenoffset-Verfahren (auch 8-Field-PAL-Sequence genannt) angewandt.

Beim PAL-System besteht folgender Zusammenhang zwischen Farbhilfsträgerfrequenz f_{sc} und Horizontalfrequenz f_h :

$$f_{sc} = \left(\frac{1135}{4} + \frac{1}{625} \right) \cdot f_h \Rightarrow \frac{f_{sc}}{f_h} = \frac{709379}{4 \cdot 625} \quad (7.8)$$

Nach dieser Gleichung vollzieht der Farbhilfsträger genau 709 379 Schwingungen, nachdem 4.625 Zeilen geschrieben sind, also genau nach 4 Vollbildern bzw. 8 Halbbildern. Betrachtet man also eine spezielle Stelle einer bestimmten Zeile des Bildes, so erreicht der Farbhilfsträger beim Schreiben dieses Bildpunktes immer genau nach 8 Halbbildern wieder Phasengleichheit.

Durch diese Tatsache vermindert sich die Störung der Chrominanzschwingung, so dass diese kaum mehr als Luminanzschwankungen wahrgenommen werden, das unangenehme Perlschnurmuster nicht mehr das Bild wandert. Damit dies zutrifft, muss allerdings der oben genannte Phasenbezug vorhanden sein.

Eine feste Phasenverkopplung zwischen Farbhilfsträger und Zeilenschwingungen hat außer der Störungsminderung noch einen weiteren wichtigen Zweck zu erfüllen, der mit dem Videoschnitt zu tun hat. Nehmen wir einmal an, zwei Videoszenen sollen durch elektronischen Schnitt aneinander gesetzt werden. Ergibt sich für die Nahtstelle dabei ein Phasensprung des Burst-Signals, welches ja für die Farbsynchronisation zuständig ist, so äußert sich dies bei der endgültigen Aufzeichnung sehr wahrscheinlich durch ein kurzes Aussetzen der Synchronisation beim Passieren der Schnittposition. Der Quarzoszillator des Decoders benötigt eine verhältnismäßig lange Zeit, um sich auf eine neue Phasensituation zu synchronisieren. Aus diesem Grunde sind die Phasenlagen des Farbhilfsträgers erstens in festem Bezug zur Horizontalschwingung und zweitens am Bildanfang jedes Halbbilds mit einer definierten Phasenlage versehen. Ein ordnungsgemäß arbeitendes Videoschnittsystem berücksichtigt diese Forderung auf Kosten einer freien Wahl der Schnittpositionen. Zu umgehen ist dieses Handikap nur mit entsprechendem elektronischen Aufwand.

Der Name des Verfahrens ergibt sich aus der Tatsache, dass der Farbhilfsträger nach einer Zeilendauer eine Phasenverschiebung von $1/4$ Schwingung (90°) aufweist.

7.2.9 Digitales Signal

Dem Arbeiten mit digitalen Signalen kommt immer mehr Bedeutung zu. Doch dem digitalen Format liegt zunächst ein analoges Signal zugrunde, welches meist dem YUV-Standard entspricht.

Eine Aufbereitung in die digitale Form hat meist die Zielsetzung, das Programm zu speichern und verlustfrei kopieren zu können. Wir werden daher in Abschn. 7.3.10 die Verfahren zur digitalen Speicherung etwas näher beleuchten.

Auch Videosignale, die aus Videografikkarten kommen, liegen primär in digitaler Form vor.

Das Format innerhalb des Speichers und das Aussehen der Aufbereitung werden wir in Kap. 9 und 10 kennenlernen.

Der wohl bekannteste Standard, nach dem Videosignale im Digitalformat verarbeitet werden, ist das D1-System. Die video- und audiotechnischen Daten gehen aus Tab. 7.2 hervor.

Tab. 7.2 Daten zum D1-System

Größe	Wert	Einheit
Video		
abzutastendes Analog-Videosignal	YUV (Komponenten)	–
Abtastfrequenz	13,5/6,75/6,75	MHz
Quantisierungstiefe	8	Bit
Störabstand (unbewertet)	56	dB
Datenstrom	216	MBaud
Abtastungen pro Zeile	720/360/360	–
Audio		
Abtastfrequenz	48	kHz
Frequenzgang	20...20000 Hz 0,5 dB	–
Quantisierungstiefe	16	Bit
Dynamikbereich	90	dB
Zahl der digitalen Audiokanäle	4	–

7.2.10 Probleme bei der Farbübertragung

Bei der Übertragung von Farbinformationen kann es zu Problemen kommen, die unterschiedliche Ursachen haben können und die sich auch verschiedenartig äußern.

Farbtonfehler Auch bei den Farbtonfehlern existieren verschiedene Ursachen und somit Wirkungen. Handelt es sich um einen eindeutigen Farbstich, der sich auf eine falsche Amplitude einer Primärfarbe zurückführen lässt, so ist i. Allg. auch recht einfach eine Kompensation möglich, indem an einer Stelle des Systems eine oder zwei Amplituden der Primärfarben variiert werden. Sofern es möglich ist, sollte natürlich diese Stelle möglichst nahe am Ursprung des Signals sein; im Extremfalle könnte somit statt einer Kompensation sogar eine Vermeidung von Farbtonfehlern stattfinden. Handelt es sich digital gespeicherte Bilder, so ist sowohl eine Nachbearbeitung der Farbbebalance über Histogramme als auch eine Echtzeitkorrektur mittels LUT-Einstellung möglich. Beschränkt sich der Fehler jedoch bei der Wiedergabe durch einen bestimmten Monitor, so ist dieser zu justieren, bevor man am Signal etwas verändert.

Eine andere Art von Farbtonfehlern kommt bei Systemen vor, die mit Farbdifferenzsignalen arbeiten (bei YUV, Y/C und FBAS). Unausgewogene Amplituden von U und V ergeben Wiedergabefehler, die man jedoch nicht als Farbstich erkennen wird. Man kann hier also nicht sagen, dass hier über das ganze Bild z. B. der Blauanteil zu groß ist. Der Fehler tritt hier nämlich nur in Farbbereichen auf, nicht aber an Graustellen des Bildes. Besteht ein solcher Verdacht, müssen die Amplituden der Farbdifferenz überprüft werden. S. 281 [8]

Fehler der Farbsättigung Erscheinen die Farben mit falscher Intensität, so ist dies meist die Folge von Fehlleistungen oder falscher Einstellung bei Systemen, die mit den Signalen FBAS, Y/C oder teilweise auch YUV arbeiten. Die Verstärker für Farbart bzw. Farbdifferenz weisen falsche Verstärkungsfaktoren auf. Auch hier ist es jedoch ratsam, vor jeglichen Veränderungen am System die Ursache aufzudecken.

Störung der Farbsynchronisation In Systemen mit PAL-Encodern oder -Decodern kann es mitunter zu Ausfällen der Farbsynchronisation kommen. Das FBAS- oder Y/C-Signal, die hier nur in Frage kommen, ist mit Messgeräten (Oszilloskop und Vektorskop) genau zu untersuchen; vor allem spielt die Amplitude und die Phasenlage des Burst eine entscheidende Rolle. S. 281 [8]

7.3 Aufzeichnungsverfahren

In diesem Abschnitt sollen verschiedene Aufzeichnungsverfahren für ein Videosignal auf ein Bandmedium beschrieben werden. Es handelt sich dabei keineswegs nur um die Aufzeichnung des FBAS-Signals; unabhängig von der Art des aufzuzeichnenden Signals (FBAS, Y/C usw.) ist eine Aufbereitung des Signals notwendig.

Die einzelnen Verfahren unterscheiden sich in den Bandformaten, der Aufteilung und Beschreibung des Bandes und in der Signalaufbereitung. In den folgenden Abschnitten werden die Verfahren für das PAL-System beschrieben. Die wichtigsten Verfahren sollen dabei etwas detaillierter zur Sprache kommen.

7.3.1 VHS-Verfahren

Beim VHS-Verfahren handelt es sich wohl um das populärste. Es wurde am Ende der 70er-Jahre von JVC entwickelt und auf den Markt gebracht. Eine rotierende Kopftrommel, die übrigens für alle Bandaufzeichnungssysteme obligatorisch ist, umschlingt das Band mit einem Winkel von ca. 190°; die beiden Videoköpfe liegen genau gegenüber, nehmen also einen Winkel von 180° gegeneinander auf der Kopftrommel ein. Auf diese Weise ist gewährleistet, dass immer mindestens ein Kopf das Band abtastet.

Geometriedaten Die Neigung der Achse der Kopftrommel gegenüber der Vertikalen beträgt 6°, somit werden also die Videoschreibspuren mit einer Schräge auf das Band gelegt. Die Drehzahl der Trommel beträgt 25 1/s, das kommt einem Kopfdurchlauf von 50 pro Sekunde gleich. Jeder Kopf schreibt bzw. liest beim Vorbeigleiten auf dem Band also ein Halbbild – eine ganze Rotation der Trommel entspricht der Zeit für ein Vollbild.

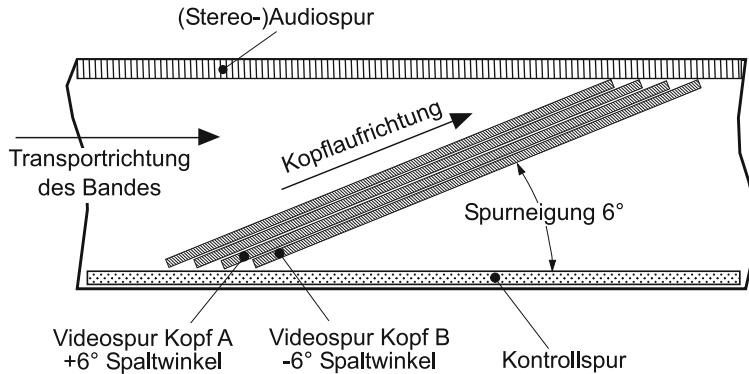


Abb. 7.25 Spurlagen beim VHS-System

Tab. 7.3 Daten des VHS-Aufzeichnungssystems

Größe	Wert	Einheit
Effektive Lesegeschwindigkeit	4,87	m/s
Bandgeschwindigkeit	2,34	cm/s
Kopftrommeldurchmesser	62	mm
Bandgröße	0,5/12,7	Zoll/mm
Bandumschlingung ca.	190	°
Kopfschräge	6	°
Azimut Kopf/Kopf	6	°

Da die Videospuren beider Halbbilder praktisch ohne Abstand nebeneinander liegen, könnte sich beim Lesen ein Übersprechen ergeben. Deshalb ist der Spalt des ersten Kopfes um $+6^\circ$, der des zweiten Kopfes um -6° gegenüber der Drehachse der Trommel versetzt. Zur Verdeutlichung der geometrischen Verhältnisse diene Abb. 7.25.

Die Audiosignale werden durch einen feststehenden Stereo-Tonkopf am Rande des Bandes geschrieben. Da die Bandgeschwindigkeit nur 2,34 cm/s beträgt, genügt die resultierende Tonqualität keinen gehobenen Ansprüchen. Tab. 7.3 gibt einen Überblick über die wichtigsten geometrischen Daten des VHS-Systems.

Verbesserte Tonqualität Eine Verbesserung brachte das VHS-HiFi-System, das den Ton über die Videokopftrommel auf Schrägspuren aufzeichnet. Außer den Videoköpfen befinden sich also auch noch zwei Audioköpfe, die ersteren vorausseilen. Jeder einzelne dieser Audioköpfe ist durch Verwendung einer FM-Signalabtastung des Audiosignals in der Lage, beide Stereosignale zu übertragen. Zur Wahrung der Kompatibilität zu Standard-VHS verzichtet man dennoch nicht auf die beiden Longitudinal-Tonspuren, die wahlweise verwendbar sind.

Bei einer Nachvertonung einer bestehenden Video-Aufnahme kann nicht die HiFi-Tonspur herangezogen werden, weil sie gemeinsam mit der Videospur aufgezeichnet wird – mehr noch, die Audiospuren liegen sogar ebenfalls auf den Videospuren. Damit es nicht zu gegenseitigen Störungen zwischen Ton und Bild kommt, arbeitet man mit einer größeren Eindringtiefe für die Audioschreibrspur und versieht die Audioköpfe wiederum mit einem anderen Spaltazimut als die Videoköpfe. Weiterhin liegt das Signalband für Audio von demjenigen für Video entfernt. Bei einem VHS-HiFi-System besteht meist außerdem die Einschränkung, dass Audio auf der Längsspur nur monaural ist.

Signalaufbereitung Der VHS-Recorder ist grundsätzlich nur für die Signalart FBAS ausgelegt. Bei der Aufzeichnung erfolgt zunächst eine Trennung in Luminanz- (Y) und Chromasignal (C). Das Amplitudenmodulation des Y-Signals wird in eine Frequenzmodulation (FM) überführt. Dieser Vorgang verhindert bei der späteren Wiedergabe, dass sich das Bandrauschen unangenehm bemerkbar machen kann; schließlich wäre bei AM die Gefahr, dass Signale mit geringem Pegel in die Nähe des Bandrauschen kommen. Für die FM dient eine Mittenfrequenz von 4,3 MHz mit einem Frequenzhub von $\pm 0,5$ MHz, wobei Weißpotential auf 4,8 MHz und Synchronpotential auf 3,8 MHz steuert. Das Chromasignal hingegen wird mit einer Hilfssignal gemischt und verlagert sich dadurch auf einen Frequenzbereich, der mit einer Mittenfrequenz von 627 kHz gerade unterhalb des unteren Seitenbandes des FM-Signals für Y liegt (Color-Under-Verfahren). Beide Signale werden addiert und gelangen zur Aufzeichnung.

Das Audiosignal erfährt eine der Tonbandaufzeichnung entsprechende Signalaufbereitung. Ein Tonsignal, welches im Schrägspurverfahren aufgezeichnet wird, erhält dagegen wie das Y-Signal eine Umsetzung als FM-Signal. Als Mittenfrequenz ist für den linken Kanal 1,4 MHz, für den rechten 1,8 MHz zugeteilt; damit liegt dieses Signal oberhalb des Chromaspektrums, jedoch noch innerhalb des unteren Seitenbandes des FM-Y-Signals. Eine sichere Trennung beider Signale ist aus oben genannten Gründen dennoch gewährleistet.

7.3.2 S-VHS-Verfahren

Das S-VHS-System arbeitet mit bereits getrennten Signalen für Luminanz und Chrominanz (Y/C-Signal). Dadurch werden Cross-Color- und Cross-Luminanz-Störungen wirksam vermieden, sofern Y und C bis zur endgültigen Wiedergabe immer getrennt bleiben.

In Verbindung mit Videobändern höherer Informationsdichte und Koerzitivkraft erhält man eine verbesserte Auflösung des Farbbildes. Ansonsten ist das Aufzeichnungsprinzip dem VHS-System gleich.

7.3.3 U-matic

Low-Band U-matic. Dieser Standard wurde von mehreren Firmen gemeinsam geschaffen. Viele Studios geben sich jedoch mit der erreichten Bildqualität nicht mehr zufrieden, weshalb solche Maschinen teilweise nur noch für Zwecke eingesetzt werden, in denen Bildqualität eine untergeordnete Rolle spielt. Wir wollen auf diese Technik daher nicht näher eingehen, sondern lediglich im nächsten Abschnitt in einer tabellarischen Gegenüberstellung die Unterschiede zum High-Band-Standard aufzeigen.

High-Band U-matic. Häufiger anzutreffen in Studios sind High-Band-Maschinen, das vom Prinzip her gleich arbeitet wie Low-Band U-matic, jedoch durch Wahl höherer Trägerfrequenzen einem höheren Qualitätsstandard gerecht wird.

Auch das U-Matic-System benötigen einen speziellen Kassettenstandard, der auch durch die Einfädelung bestimmt wird. Tab. 7.4 zeigt die Geometriedaten im Überblick.

Signalaufbereitung Das FBAS-Signal wird aufgeteilt in die Signale Y und C. Das Luminanzsignal erfährt eine Bandbegrenzung auf 3,8 MHz (bei High-Band SP, siehe Tab. 7.5). Die Mittenfrequenz dieser FM beträgt 6,4 MHz. Das Chromasignal wird auf eine Frequenz 924 kHz transponiert. Beide Signale, das FM-Y- sowie das transponierte C-Signal führt man einer Addition zu, und das Gesamtsignal gelangt sodann zur Aufzeichnung.

Bei der Bandgeschwindigkeit von 9,5 cm/s (doppelt so hoch wie beim Kassettenrekorder) kann man von den Audiospuren bereits HiFi-Qualität erwarten. Trotzdem ist hier natürlich ein Rauschunterdrückungssystem (Dolby oder DBX) vorgesehen.

Die High-Band-Maschine verfügt über eine getrennte Timecode-Spur, die wie die Audiospuren longitudinal verläuft. Bei Low-Band-Maschinen hingegen muss – falls eine Timecode-Einrichtung vorhanden ist – auf eine Audiospur verzichtet werden; bereits diese Tatsache bedeutet im Studiobetrieb eine enorme Einschränkung.

Tab. 7.5 zeigt die Unterschiede der Frequenzen bei den einzelnen U-Matic-Systemen, wobei anhand einiger Frequenzen bereits die Qualitätsabstufung deutlich wird. Bei allen Ausführungen handelt es sich um professionelle Maschinen, die jedoch wie erwähnt immer seltener in Studios zum Einsatz kommen.

Tab. 7.4 Geometrische Daten beim System U-matic

Größe	Wert	Einheit
Effektive Lesegeschwindigkeit	4,87	m/s
Bandgeschwindigkeit	2,34	cm/s
Kopftrommeldurchmesser	62	mm
Bandgröße	0,5/12,7	Zoll/mm
Bandumschlingung ca.	190	°
Kopfschräge	6	°
Azimut Kopf/Kopf	6	°

Tab. 7.5 Frequenzen bei U-Matic-Systemen (jeweils in MHz)

Größe	Low-Band LB	High-Band HB	Super-High-Band SP
Videobandbreite	3	3,5	3,8
Weißenwert	5,4	6,4	7,2
Synchronwert	3,8	4,8	5,6
Hubbereich	1,6	1,6	1,6
oberste Seitenbandfrequenz	7,6	9,1	10,2
unterste Seitenbandfrequenz	1,6	2,1	2,6
ZF-Träger	4,6	5,6	6,4
Farbzwischenträger	0,69	0,92	0,92

Tab. 7.6 Daten des Betacam-Systems

Größe	Wert	Einheit
Effektive Lesegeschwindigkeit	5,75	m/s
Bandgeschwindigkeit	10,15	cm/s
Kopftrommeldurchmesser	74,49	mm
Bandgröße	0,5/12,7	Zoll/mm
Bandumschlingung ca.	190	°
Kopfschräge	4,68	°

7.3.4 Betacam

Das Betacam-Verfahren ist eine Entwicklung der Fa. Sony auf dem Studiosektor. Die Konzeption zielt auf beste Bild- und Tonqualität ab. Zusammen mit dem universellen Schnittbetrieb hat sich Betacam zu einem der wichtigsten Standards für professionelle Anwendungen entwickelt.

Geometriedaten Die Forderungen an Betacam-Maschinen sind hoch, was sich zwangsläufig auch in den Daten für die Mechanik bzw. Geometrie niederschlägt. Einige dieser Daten sind in Tab. 7.6 zu finden.

Signalaufbereitung Bei diesem Verfahren werden die ungeträgerten Signale Y, U und V benötigt. Das Y-Signal erfährt eine Frequenzmodulation, wobei es jedoch in seiner vollen Bandbreite von 5 MHz verarbeitet wird. Die Farbinformationen U und V werden jeweils zeitlich im Verhältnis 2:1 komprimiert zeilenweise sequentiell wieder zusammengefügt nach dem CTDM-Verfahren (Compressed Time Division Multiplex). Sie haben beide eine Bandbreite von 1,5 MHz.

Für das frequenzmodulierte Y-Signal und für das Synthese-Signal der Multiplex-Signale U und V stehen jeweils zwei Videoköpfe zur Verfügung. Bevor jedoch das Synthese-Signal zur Aufzeichnung gelangen kann, durchläuft es ebenfalls einen

Frequenzmodulator. Die vollständige Trennung zwischen Farb- und Helligkeitsinformation verhindert Cross-Color- und Cross-Luminanz-Störungen.

Bei der Wiedergabe erfolgt eine Rückführung der Komponenten U und V in ihre ursprünglichen Zeitabläufe. Dieser Vorgang benötigt jeweils die Dauer einer Zeile, und die Chromasignale sind danach um die Dauer von zwei Zeilen (2 H) verzögert. Die zeitliche Angleichung übernimmt ein sog. Timebase-Corrector (TBC), der dem Wiedergabesignal ein quarzgenaues Timing verleiht.

Die relativ hohe Bandgeschwindigkeit ergibt für den Ton eine akzeptable Qualität. Außerdem ist wie bei manchen VHS-Maschinen auch eine Aufzeichnung des Tones auf der Videospur möglich (FM), was aber auch hier keine Nachvertonung erlaubt, weil die Videospur bereits beschrieben ist.

Neben den Videospuren (Y und C), den Tonspuren und der Kontrollspur verfügt das Betacam-System über eine separate Timecode-Spur, die ein bildgenaues Positionieren und Schneiden ermöglicht.

7.3.5 Betamax

Das Betamax-System arbeitet in ähnlicher Weise wie das VHS-System, es benutzt jedoch ein eigenes Kassettenformat. Die Schreibgeschwindigkeit der Videoköpfe beträgt 5,83 m/s, und die Bandgeschwindigkeit ist nur 1,87 cm/s. Ein Halbbild auf dem Band bedeutet eine Spurlänge von 122 mm. Betamax konnte sich neben VHS kaum durchsetzen.

7.3.6 Video 2000

Auch dieses System benötigt ein spezielles Kassettenformat. Um die Spieldauer zu verdoppeln, kommen wendbare Kassetten zur Anwendung. Bei einer Laufrichtung wird also nur die Hälfte des Bandes beschrieben. Die Schreibgeschwindigkeit beträgt 5,08 m/s, und die Bandgeschwindigkeit liegt bei 2,44 cm/s. Durch die Platz einsparung auf dem Band arbeitet dieses System mit extrem schmalen Videospuren. Deshalb verwendet man hier ein spezielles Spurnachführungssystem.

Video 2000 stellt wie VHS einen Standard für den privaten Gebrauch dar, wobei er natürlich bei weitem nicht so verbreitet ist wie VHS.

7.3.7 Video Hi-8

Dieses Aufzeichnungssystem wurde speziell für Camcorder entwickelt. Ein sehr kleines Kassettenformat ist hierfür Voraussetzung.

Als Weiterentwicklung des Camcorderformats Video-8 benutzt dieses System eigens dafür geschaffene ME-Bänder (Metal Evaporated), die die geringere Oberflächenrauigkeit

von nur 6 nm sowie eine größere Schreibdichte aufweisen. Ersteres verbessert den Band-Kopf-Kontakt. Eine Reduzierung der Spaltbreite auf nur 0,22 μm verbessert die obere Grenzfrequenz. Auf diese Weise erhält man eine Horizontalauflösung von mehr als 400 Pixel. Die Mittenfrequenz der FM beträgt 6,7 MHz, der Frequenzhub ist ± 1 MHz, und die Farbträgerfrequenz liegt bei 732,4 kHz.

Hi-8 genügt bereits semiprofessionellen Ansprüchen, weshalb es in diesem Bereich auch häufig eingesetzt wird.

7.3.8 1-Zoll-MAZ

Größere Bandformate, die i. Allg. für Studiozwecke gedacht sind, verwenden Spulenbänder – so auch das 1-Zoll-Format. Hier gibt es wieder unterschiedliche Standards, was die Belegung der Spuren und auch die Signalaufbereitung anbelangt. S. 281 [10]

Das modernste 1-Zoll-System stellt momentan das B-Format dar. Es verfügt über sog. *fliegende Löschköpfe*, mit denen auch diskrete Einzelbilder im Schnittbetrieb überschrieben werden können. Es stehen drei Tonlängsspuren zur Verfügung; eine Bandgeschwindigkeit von 24,3 cm/s ermöglicht eine sehr gute Tonqualität. Das aufzuzeichnende FBAS-Signal wird auf einen Träger der Frequenz 7,83 MHz transponiert. Bei der Wiedergabe wird das ZF-Signal entsprechend wieder in das FBAS-Signal überführt, wobei aufwändige Zusatzschaltungen (Dropout-Kompensator, Chroma-Regelung, Timebase-Corrector usw.) für eine stabile Bildqualität sorgen.

7.3.9 Laserdisc

Beim Zusammensetzen einzelner Szenen im sog. Assemble-Schnitt sind lange Suchzeiten bei Verwendung von Bandmaschinen kaum zu vermeiden. Einen Ausweg bietet die Speicherung auf optischen Platten – der Laserdisc.

Die Firma Sony hat mit dem Analogdraw-System (LVR-Serie) eine Möglichkeit geschaffen, Video- und Audiosignale in hervorragender Qualität aufzuzeichnen. Allerdings darf man nicht verschweigen, dass nur eine einmalige Aufnahme möglich ist (was allerdings bei der Aufzeichnung von Animationen nach dem Einzelbildverfahren kein Handikap bedeutet). Diese Geräte gibt es auch als reine Player-Ausführung, was für manche Zwecke durchaus ausreichend ist.

Je nach Ausführung können FBAS-, Y/C-, YUV- und sogar RGB-Signale aufgezeichnet werden. Die Signale werden Frequenzmodulatoren zugeführt, und die Ausgangssignale der Modulatoren durchlaufen eine Addition. Das endgültige FM-Signal passiert eine Amplitudenbegrenzung, und aus der Länge der Halbwellen ergibt sich die Länge der Pits. Dem Prinzip nach handelt es sich dann eigentlich um ein digitales System, das nach dem Pulsweiten-Verfahren (PWM) arbeitet. Da hierbei jedoch keine Quantisierung stattfindet, kann man zurecht noch von einem Analogverfahren sprechen.

Die Suchzeit einer bestimmten Stelle liegt unter 0,5 s. Die Maschinen verfügen über eine Timecode-Einrichtung, so dass Schnittbetrieb ohne weiteres möglich ist. Die Standbildqualität ist hervorragend und kann nicht vom normalen Lauf unterschieden werden.

7.3.10 Digitale Aufzeichnungsverfahren

Bei den digitalen Aufzeichnungsverfahren hat man sich weitgehend auf den 4-2-2-Standard geeinigt. Zum Einsatz kommt das Komponentensignal, welches aus den Signalen Y, U und V besteht (siehe Abschn. 7.2.5). Zur Wandlung in Digitalform wird das Y-Signal mit 13,5 MHz, die beiden Farbdifferenzsignale U und V jeweils mit 6,75 MHz abgetastet (siehe auch Anhang B (Abb. B.1, Tab. B.1)).

Das Zahlenverhältnis 4:2:2 spiegelt das Verhältnis der Abtastraten wider. Nach der Digitalisierung werden die digitalen Daten segmentiert und im Zeitmultiplexverfahren wieder zusammengefügt. Nach dieser Aufbereitung gelangen die Daten zur Aufzeichnung.

Der 4-2-2-Standard resultiert aus der CCIR-Empfehlung Nr. 601. Man kann den Unterschied zwischen den beiden Systemen 525/60 und 625/50 relativ einfach dadurch überwinden, indem man unter Beibehaltung der Mechanik die Anzahl der Segmente für die Aufzeichnung eines Halbbildes verändert. Die Segmentanzahl beträgt beim 625/50-System 12 und beim 525/60-System 10. Die Segmentvorwahl erfolgt durch Umschaltung am Digitalrecorder.

Der hohe Datenstrom von 216 MBit/s erfordert eine hohe Signalbandbreite. Dies wiederum setzt kleine Abmessungen der Magnetköpfe und der Spurbreite voraus. Um Schwankungen des Kopf-Band-Kontaktes und den daraus resultierenden, möglichen Bitverfälschungen zu begegnen, wird mittels einem sog. *Interleaving-Verfahren* eine Code-Spreizung erzeugt. Zeitlich benachbarte Digitalwerte werden auf Band weit auseinander gelegt. Bei der Wiedergabe erfolgt der umgekehrte Weg. Dabei lassen sich dann auch durch Störstellen fehlende einzelne Bits mittels Interpolation zwischen den Nachbarwerten sicher generieren.

Man hat die Möglichkeit, vier digitale Audiosignale aufzuzeichnen, die zwischen den Videospuren in vier Segmenten gespeichert werden. Die Sampling-Frequenz beträgt 48 kHz bei einer linearen Quantisierung von 16 Bit.

Moderne DV-Camcorder An dieser Stelle darf nicht unerwähnt bleiben, dass modernste Aufzeichnungsmaschinen in Echtzeit komprimieren und zur Speicherung unbewegliche Medien einsetzen.

7.4 Weitere Farbübertragungsverfahren

Neben PAL gibt es noch weitere Verfahren für die Farbübertragung. Wir wollen diese jedoch nur kurz tangieren und uns auf die markante Unterschiede beschränken.

Dass hier keine Weltnorm eingeführt werden konnte, liegt u. a. an trivialen Dingen wie der Netzfrequenz, teilweise jedoch auch an der Eigenständigkeit einzelner Länder und Ländergruppen.

7.4.1 NTSC

Das älteste Farbübertragungsverfahren stellt das NTSC-System dar (National Television System Committee). Es wurde im Jahre 1953 in den USA entwickelt und war auch Grundlage zur Entwicklung anderer Systeme, auch des PAL-Systems.

Bei NTSC findet keine Umschaltung der Phasenlage der V-Achse von Zeile zu Zeile statt – auf eine Farbfehlerkompensation wird also verzichtet. Es handelt sich um 525/60-Systeme, d. h. ein Vollbild weist 525 Zeilen auf, und die Bildablenkung arbeitet mit ca. 60 Hz (der genaue Wert beträgt 59,94 Hz). Die Zeilenfrequenz unterscheidet sich mit 15734,264 Hz nur um 0,7 % von 625/50-Systemen, was bei inkompatiblen Monitoren meist dazu führt, dass die Zeilenablenkung durchaus synchronisierbar ist.

Der Zusammenhang zwischen Farbhilfsträger und Horizontalfrequenz ergibt sich zu:

$$f_{sc} = \frac{455}{2} \cdot f_h \Rightarrow \frac{f_{sc}}{f_h} = \frac{455}{2} \quad (7.9)$$

Ein starrer Phasenbezug zwischen Farbhilfsträger und Horizontalfrequenz ist auch hier gegeben, um das Störmuster des Farbhilfsträgers möglichst gering zu halten und den Videoschnitt zu vereinfachen. Die Phasenlage des Farbhilfsträgers an einer definierten Stelle des Bildes wiederholt sich bei NTSC bereits nach 4 Halbbildern (bei PAL nach 8 Halbbildern), wie aus Gl. 7.9 hervorgeht.

Um Farbfehler, die bei diesem Verfahren immer wieder auftreten, ausgleichen zu können, haben NTSC-Decoder üblicherweise einen Einsteller, der mit Hue (= Färbung) bezeichnet ist.

Bei der Erzeugung des FBAS-Signals in NTSC wird das Luminanz-Signal in gleicher Weise wie bei PAL gewonnen (siehe Gl. 7.3). Die Generierung des Chrominanz-Signals jedoch zeigt einige Unterschiede. Das modulierte Farbsignal entsteht hier nicht durch vektorielle Addition der orthogonalen Komponenten U_1 und V_1 , sondern der durch Projektion entstehenden Vektoren Q_1 und I_1 auf die ebenfalls orthogonalen Achsen Q und I .

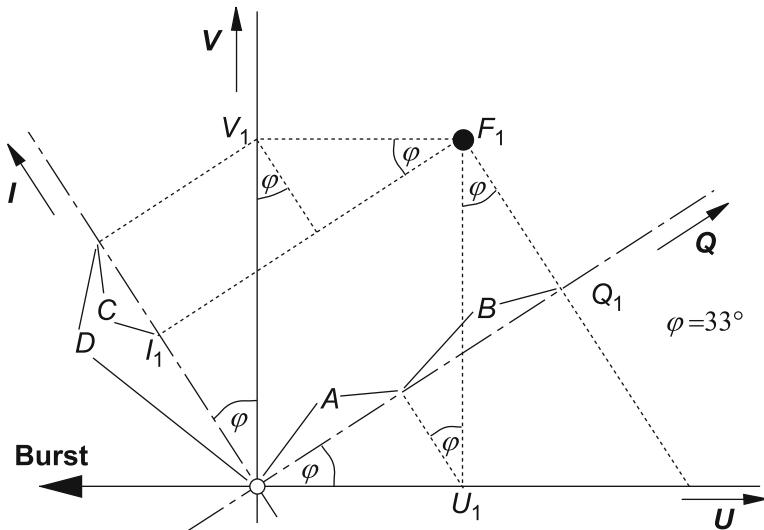


Abb. 7.26 Farbvektoren beim NTSC-System

Mathematisch ergibt sich also:

$$\begin{aligned}
 Q &= \underbrace{U \cdot \cos(33^\circ)}_A + \underbrace{V \cdot \sin(33^\circ)}_B \\
 I &= \underbrace{-U \cdot \sin(33^\circ)}_C + \underbrace{V \cdot \cos(33^\circ)}_D
 \end{aligned} \tag{7.10}$$

Zum Verständnis der Entstehung dieser Vektoren diene Abb. 7.26.

Unter Berücksichtigung der Reduktionsfaktoren 0,493 für $B - Y$ und 0,877 für $R - Y$, die im übrigen auch für PAL gelten, ergeben sich folgende Bestimmungsgleichungen:

$$\begin{aligned}
 Q &= 0,41 \cdot (B - Y) + 0,48 \cdot (R - Y) \\
 I &= -0,27 \cdot (B - Y) + 0,74 \cdot (R - Y)
 \end{aligned} \tag{7.11}$$

Der Burst wird bei NTSC ebenfalls nicht umgeschaltet und nimmt einen Winkel von 180° gegenüber der U -Achse ein. Nach der Modulation der Signale I und Q mit dem Farbhilfsträger (3 579 545 Hz) entstehen die Signale I' und Q' , deren Summe das Chromasignal F_C ergibt. Nach der weiteren Synthese mit dem Burst, Synchronsignal, Luminanz-Signal und Chromasignal ergibt sich das FBAS-Signal im NTSC-Format.

Die Hilfsachsen Q und I wurden eingeführt, weil man Farben im Bereich der I -Achse mit größerer Bandbreite übertragen wollte als die Farben im Bereich der Q -Achse.

Farbtonfehler, die sich infolge einer Phasenverschiebung zwischen Burst und Farbvektor einschleichen können, müssen manuell ausgeglichen werden, indem man die Decodier-Achsen gegenüber dem Burst verstellt. Das Kriterium für diese Korrektur ist jedoch ausschließlich subjektiver Art.

7.4.2 SECAM

Das SECAM-Verfahren (séquentielle à mémoire) wird in Frankreich und vielen osteuropäischen Ländern eingesetzt. Die Werte für Bild- und Zeilenfrequenz entsprechen denen von PAL, die Zeilenzahl ist ebenfalls 625.

Grundsätzliche Unterschiede bestehen jedoch in der Farbsignalaufbereitung bei FBAS und Y/C. Systeme ohne automatische oder manuelle Umschaltung sind daher nicht kompatibel.

Für jede Zeile erfolgt abwechselnd die Übertragung von U und V. Somit muss im Decoder das Farbsignal um eine Zeile verzögert bzw. zwischengespeichert werden. Dies geschieht wie bei PAL mit einer Ultraschall-Verzögerungsleitung ($1\text{ H} = 64\text{ }\mu\text{s}$); ebenfalls beruht das SECAM-System auf der Annahme, dass sich der Farbsignalinhalt zweier zeitlich aufeinanderfolgender Zeilen nur wenig ändert.

Auf der Übertragungsseite zeigen sich gegenüber PAL weitere Unterschiede, die die Modulationsart des Hochfrequenzträgers betreffen.

7.5 Hochauflösende Standards

Mängel der Bildqualität aller Systeme und wachsender technischer Standard haben die Komitees bereits in den 1980er-Jahren zu Verhandlungen über neue Übertragungsstandards bewogen.

7.5.1 PALplus und dessen Verbesserungen

Die Motivation von PALplus lag in der Erreichung eines verbesserten Filmformats und hochwertigeren Farbübertragung. Ein wichtiger Punkt war die senderseitige Unterdrückung von Fehlern bei der Farbwiedergabe. Durch ein spezielles Codierverfahren (*MACP = Motion Adaptive Colour Plus*) kommt es zu keinen Störungen zwischen Luminanz- und Chrominanzsignal. Ferner wird die volle Videobandbreite von 5 MHz genutzt, so dass auch bei der Vergrößerung der Bildbreite zufolge des 16:9-Formates keine Verschlechterung der Zeilenauflösung zu erwarten ist.

Die Kompatibilität zum normalen Fernsehgerät mit 4:3-Format ist etwas komplex. Es ist jedoch leicht einzusehen, dass bei einer einfachen Übertragung von Bilddaten eines Formats von 16:9 bei der Wiedergabe über einen 4:3-Gerät (bei identischem

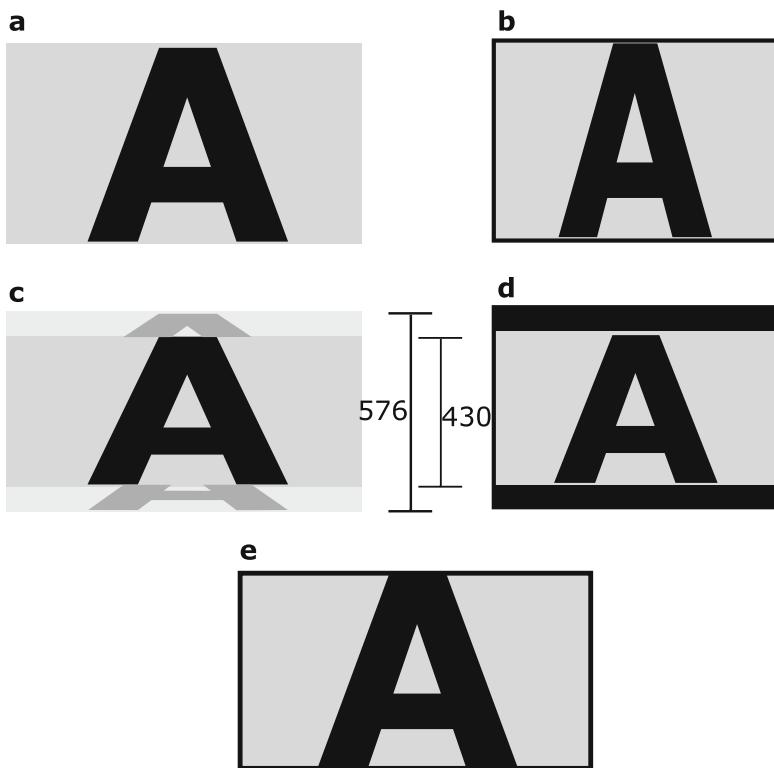


Abb. 7.27 Abwärtskompatible Bildübertragung eines 16:9-Bildes auf einem 4:3-Wiedergabegerät im Letterbox-Format. Das zu übertragende Bild im Format 16:9 (a) würde bei unveränderter Aufbereitung bei einem 4:3-Wiedergabesystem zu einem horizontal gestauchten Bild führen (b). Stattdessen erfahren die Bilddaten eine zeilenweise Modifikation (c), so dass in einem Streifen von ca. 430 Zeilen bereits die komplette Bildhöhe und in den Bereichen ober- und unterhalb die Information der dazwischen liegenden Zeilen enthalten ist. Diese Information gelangt als Modulation auf die Farbdifferenzachse $Q(B - Y)$, damit das vom 4:3-System in (d) empfangene Bild in diesem Bereich außerhalb der 430 gültigen Zeilen lediglich selten schwache blaue Schlieren auf einem sonst schwarz ausgetasteten Streifen sichtbar werden können. Durch den Bildausschnitt ist auch hier das Bildformat wieder 16:9 (Letterbox-Format), genauso wie auf dem 16:9-System in (e), welches mit dem PALplus-Decoder fast das volle Bild mit 574 Zeilen durch Interpolation rekonstruiert

Zeilen- und Bild-Timing) ein stark verzerrtes Format entstehen würde: Die horizontale Stauchung wäre unübersehbar. Als „normale“ Bilddaten, die auch ein 4:3-Empfänger darstellen kann, überträgt man lediglich die mittleren 430 Zeilen der insgesamt 576 eines üblichen PAL-Bildes. Man bezeichnet dieses Format auch als *Letterbox*. Wie es zu einer abwärtskompatiblen Bildwiedergabe kommt, geht aus Abb. 7.27 hervor.

Bei PALplus dienen die erste und letzte Zeile von den 576 Steuerungszwecken – darin ist u. a. enthalten, ob es sich gerade um ein 16:9-Bild oder einem 4:3-Bild handelt. Die Zwischenzeilen stellen übrigens keine Farb-, sondern nur Helligkeitsinformation bereit.

Die Unterdrückung von Farbartefakten geschieht senderseitig. In Abschn. 7.2.7 sind die Ursachen für Crosscolor-Störungen nachzulesen. Bildsignale für PALplus beinhalten eine Aufbereitung dergestalt, dass zu erwartende Störungen schon vor Codierung zum FBAS-Signal entfernt werden. Filter unterdrücken horizontale Muster, deren Signalfrequenz in der Nähe des Farbhilfsträger ist.

Da für die Decodierung eine Bildspeicherung erforderlich ist, drückte in der Entstehungszeit von PALplus diese Decoder-Elektronik den Preis der Geräte stark in die Höhe, weshalb sich das Verfahren nicht durchsetzen konnte. Analoges Fernsehen verschwindet immer mehr – die digitale Übertragung eröffnet sowohl bildtechnisch als auch tontechnisch ganz neue Wege.

7.5.2 HDTV-Historie und nach EUREKA 95

Ein Projekt namens EUREKA 95 war eine Initiative der EBU (European Broadcasting Union) und europäischer Gerätehersteller, einen zu PAL kompatiblen Standard zu finden.

In erster Linie erfolgt eine Änderung des Bildformats von 4:3 auf 16:9, was nach neuen Erkenntnissen dem Blickfeld des menschlichen Auges eher entsprechen soll. Ebenfalls wird die Zeilenzahl drastisch erhöht (bei Eureka 95 auf 1249 Zeilen statt vorher 625), und folgerichtig ist auch eine höhere Horizontalabtastrate bzw. -auflösung vorgesehen, zumal die Zeilenfrequenz verdoppelt werden muss. Auf das Zeilensprungverfahren kann natürlich nicht verzichtet werden, da die Vollbildfrequenz nach wie vor 25 Hz beträgt.

Da zunächst jedoch eine Kompatibilität zum bisherigen System gewährleistet werden soll, wirken sich derartige Änderungen nur auf den Produktionsstandard aus – der Übertragungs- und Wiedergabestandard bleibt vorläufig noch gleich. Einige Daten zum HDTV-System sind im Anhang B (Abb. B.2, Tab. B.2) zu finden.

7.5.3 Moderne HDTV-Technologie

Viele Entwürfe für Standards sind entstanden. Zwei davon sind 1080i/25 und 720p/50, welches sich mit hoher Wahrscheinlichkeit in Europa durchsetzen wird. S. 281 [12]

Bildformat und -auflösung Beim System 1080i ergeben sich als Bildformat 1080 Zeilen und 1920 Pixel/Zeile, Bildaufbau ist interlaced bei 25 Hz Vollbildfrequenz. Dagegen besteht das Format 720p aus 720 Zeilen und 960 Pixel/Zeile bei 50 Hz Vollbildfrequenz.

Datenmengen und Übertragung Die Daten für das Format liefern beim 1080i/25 rein rechnerisch etwa 155 MByte/s als zu übertragende Datenrate (bei Annahme von 24 Bit/Pixel). Da die Bilddaten digital vorliegen, sind sehr effiziente Kompressionsmethoden zur Anwendung kommen. Für MPEG-2 werden 27 Mbit/s empfohlen.

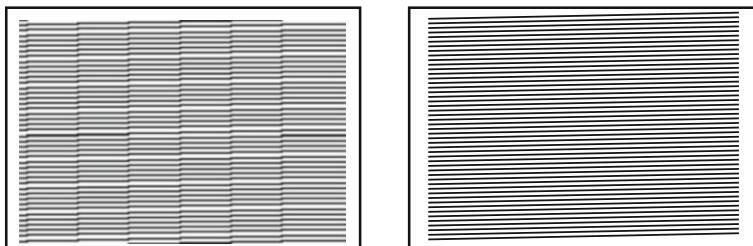


Abb. 7.28 Linienmuster bei konventioneller und gefilterter Abtastung

Technische Voraussetzungen Beim Fernsehen muss das Gerät entsprechende Hardware vorweisen, um die Daten von HDTV zu decodieren. Dabei ist zu beachten, dass nicht notwendigerweise jedes Gerät beide oben genannten Formate richtig verarbeiten kann.

Anforderungen an Videografikkarten Gerade der höhere Produktionsstandard in der RGB- bzw. YUV-Stufe verändert natürlich auch die Möglichkeiten bzw. die Anforderungen, die an Videografikkarten zu stellen sind. Der Bildspeicher muss nach diesen Richtlinien ca. 2 300 000 Bildpunkte erfassen können; bei einer Farbtiefe von 24 Bit/Pixel entspricht das einer Speichergröße von 6,9 MByte, wobei bei dieser Farbtiefe noch kein Overlay-Speicher berücksichtigt ist.

Der Pixeltakt beträgt bereits im Interlaced-Modus fast 60 MHz; soll eine Umschaltmöglichkeit für Non-Interlaced-Wiedergabe (nur während der Bearbeitungsphase) möglich sein, verdoppelt sich auch diese Frequenz. Aufgrund dieser erhöhten Anforderungen werden sicher teilweise auch noch in Zukunft einige Probleme zu erwarten sein.

Digitale Aufbereitung Neben Verbesserungen der Auflösung sind bei einer digitalen Aufbereitung natürlich auch Filterverfahren denkbar. Dadurch wird eine alias-freie Abtastung und eine zeilenfreie Bildwiedergabe möglich. Abb. 7.28 zeigt die Unterschiede zwischen herkömmlicher und progressiver Abtastung mit digitaler Vor- und Nachfilterung. Die hier vorgegebene Situation ist ein leicht schräg verlaufendes Linienmuster, wobei die Linienzahl etwas unterhalb der Zeilenzahl liegt. Die Folge davon ist eine störende Frequenz, die sich durch die Differenz zwischen Zeilenfrequenz und Signalfrequenz ergibt. Bei entsprechender Filterung jedoch sind quasi in beiden Ablenkungsrichtungen äquivalente Auflösungen erzielbar.

7.5.4 UHDTV und 4K-Standard

Der Werbeausdruck „4K“ bezieht sich auf eine hochauflösende Wiedergabetechnik, wobei damit latent die Pixelanzahl von 4000 in einer Zeile angedeutet wird. In Tab. 7.7 sind verschiedene neuere Auflösungsverhältnisse dargestellt. Für das menschliche Auge sind Auflösungen von durchschnittlich 2400×1800 anzunehmen, sodass die höchsten

Tab. 7.7 Hochauflösende Standards im Vergleich mit herkömmlicher niedrigauflösender Technik

Bezeichnung	horizontal	vertikal	Beschreibung
Ultra HDTV 1 (UHD-1)	3840	2160	4-fache Pixelanzahl gegenüber 1080i
Ultra HDTV 2 (UHD-2)	7680	4320	16-fache Pixelanzahl gegenüber 1080i
Full HD (1080i)	1920	1080	
HD (720p)	1280	720	
SD	720/768	576	720 nach CCIR bzw. 768 nach PAL
VGA u. NTSC	640	480	

Standards UHD-1 und UHD-2 dies mehr als abdecken und bei einer normalen visuellen Ausnutzung des Sichtwinkels nicht nutzbar ist. Sobald der Betrachter jedoch einen Bildausschnitt genauer betrachten möchte, wäre Teilbild immer noch in angemessener Auflösungsqualität, während bei herkömmlichen Standards (HD und darunter) schnell diskrete Pixel erkennen lassen. S. 281 [5]

Je mehr Pixel pro Bild und je höher die Bildwechselfrequenz, desto höher werden verständlicherweise die Datenraten, die zu übertragen sind. Dies ist insbesondere in der Fernsehtechnik von Belang, denn dort setzt die Trägerfrequenz und die Bandbreite Grenzen. Es ist daher stets üblich, den Datenstrom über spezielle Kompressionstechniken niedrig zu halten.

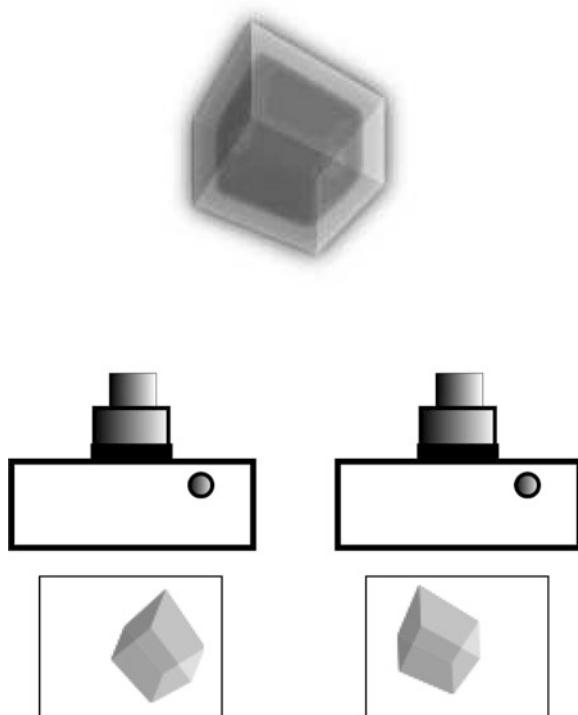
7.6 3D-Verfahren und Standards

Bilder und bildgebende Systeme haben systembedingt die Eigenschaft, zweidimensional zu sein, denn bei der Aufnahme sind keinerlei räumliche Zusatzdaten vorhanden, und bei der Wiedergabe oder Betrachtung befinden sich die Bilddaten ebenfalls lediglich auf einer Ebene. Da der räumliche Eindruck jedoch entscheidendes Gewicht bei der Wahrnehmung hat, ist auch im Bild- und Videobereich der Wunsch nach dreidimensionalen Bildern stark geworden.

7.6.1 Stereoskopie

Dreidimensionale Bilder oder räumliche Verfahren sind bereits kurz nach der Erfindung der Fotografie bekannt geworden. Es ging dabei zunächst um die *Stereoskopie*. Dabei werden einfach zwei Bilder gleichzeitig oder auch hintereinander (bei unbewegten Objekten) aus Positionen heraus gemacht, welche eine Distanz ähnlich des Augenabstandes aufweisen. Zum Betrachten können diese beiden Bilder als Dias in einen Diabetrachter gelegt werden. Auf diese Weise soll dem Auge derselbe Eindruck vermittelt werden wie zum Zeitpunkt der Aufnahme.

Abb. 7.29 Prinzip der Stereoskopie. Aufnahme eines räumlichen Objekts aus zwei verschiedenen Positionen, danach erfolgt Wiedergabe über einen Dia-Betrachter oder eine sonstige Optik mit für beide Augen getrennter Bildwiedergabe



Der dadurch entstehende räumliche Eindruck ist jedoch aus verschiedenen Gründen nicht besonders überzeugend. Die einzelnen Gegenstände der beiden Bilder, die sich in der Nähe des Betrachters befinden, haben Ähnlichkeit mit ausgeschnittenen Papierfiguren, welche zwar im Raum stehen, selbst aber flach und deshalb zweidimensional erscheinen (Abb. 7.29).

Auch im Plakatdruck gab und gibt es Beispiele für Stereoskopie, wobei dort keine echten zwei Bilder vorliegen, sondern nur eines. Der Betrachter muss sich darauf konzentrieren, seinen Fokuspunkt hinter die Betrachtungsebene zu legen. Es entsteht ein neuer Bildeindruck, bei dem Objekte räumlich erscheinen. Ein Beispiel hierfür sei in Abb. 7.30 gegeben. Der Effekt, der beim künstlichen Zusammenschieben der Bilder oder der Bildsegmente entsteht, wurde auch dazu genutzt, erst ganz neue Bildinhalte und -eindrücke entstehen zu lassen. Es handelt sich dabei um das *Single-Image-Stereogramm*.

7.6.2 3D-Technik durch Monochromfilter-Brille

Das Fernsehen und vor allem die Kinotechnik hatte bereits sehr früh begonnen, Raumeindruck zu übertragen. Die Technologie setzt eine Brille voraus, die der Betrachter tragen

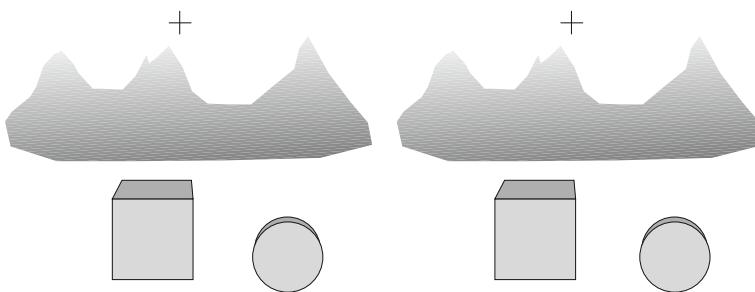
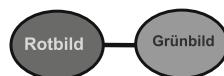


Abb. 7.30 Effekt der Stereoskopie bei einem Einzelbild. Das Auge muss versuchen, die beiden Kreuze durch Fokussierung in die Ferne zur Deckung zu bringen, dann sind die Berge im Hintergrund auch kongruent. In diesem Zustand der modifizierten Akkommodation erscheinen einige Objekte räumlich verteilt, nämlich der Würfel scheint mehr in der Nähe zu sein als der Zylinder. Verstärkt wird dieser Eindruck noch dadurch, dass man den Kopf etwas bewegt

Abb. 7.31 Prinzip der 3D-Technik mit Farbfilterbrille. Die beiden Monochrombilder, die aus verschiedenen Perspektiven gewonnen wurden, können gleichzeitig auf dem Schirm erscheinen. Für jedes Auge lässt die Brille nur den zugehörigen Teil durch



muss. Sie enthält auf der einen Seite ein rotes Farbfilter, auf der anderen Seite ein grünes (oder auch blaugrünes) Farbfilter.

Die Aufzeichnung des Films oder des Bildes erfolgt aus zwei nebeneinander liegenden Positionen, wie das bei der Stereoskopie auch der Fall ist. Von den Bildern werden jedoch nur die Rot- bzw. Grünanteile gespeichert. Bei der Wiedergabe erscheinen zwei farbige, überlagerte Bilder, deren Frontobjekte mehr oder weniger seitlich verschoben sind (weil diese ja auch seitlich verschoben aufgezeichnet wurden).

Es liegt auf der Hand, dass durch die optische Trennung der beiden Teilbilder kein Echtfarbenbild mehr übertragen werden kann, sondern dass der Bildeindruck eigentlich nur noch von monochromer Charakteristik ist (Abb. 7.31).

Kino und Fernsehen arbeiteten schon sehr früh mit derartigen Übertragungen. Das Handicap, dabei die Farbcharakteristik zu verlieren, ließ diese Technik jedoch nicht sehr erfolgreich werden.

Eine Abwandlung dieser Technik, die anstelle der Monochromfilter- eine Polarisationsbrille nutzt, sei hier ebenfalls erwähnt. Die Brille besteht aus zwei Polarisationsfiltern,

deren Polarisationsebenen gegenseitig um 90° gedreht sind. Die Bildprojektion erfordert zwei unabhängige Wiedergabegeräte, die jeweils eines der beiden Stereobilder auf einen Spezial-Projektionswand werfen. Beide Wiedergabegeräte schicken ihr Bild jedoch wiederum über um 90° verdrehte Polfilter auf diese Wand, welche nicht einfach nur mattweiß sein darf, sondern matt-metallisch spiegelnd. Die Reflexion an Metall verändert die Polarisationswinkel nicht. Somit ist gewährleistet, dass die Information beider Bilder auch getrennt für beide Augen vorliegt. Die Farübertragung ist somit auch möglich. Allerdings ist die Wiedergabe über Monitore – und somit fürs Fernsehen – technisch gesehen nicht praktikabel.

7.6.3 3D-Technik mit Shutter-Brille

Eine sehr viel moderne Methode, Stereobild-Information in Farbe zu übertragen, setzt eine *Shutter-Brille* ein. Im Gegensatz zum oben erwähnten Verfahren mit den Polarisationsfiltern erfolgt hier die Bildtrennung zeitlich für jedes Auge. Dies erledigt die Shutter-Brille, welche natürlich vom Wiedergabesystem synchronisiert werden muss. Das Prinzip geht aus Abb. 7.32 hervor.

Die Synchronisation erfolgt über Infrarot-Signale, die der Wiedergabemonitor an die Brille sendet. Die eigentliche Schließfunktion übernimmt ein Flüssigkristallfenster (LCD), welches elektronisch sehr einfach und auch energiesparend anzusteuern ist.

Normalerweise benötigt jeder Monitor, Projektor und jedes Fernsehgerät seine passende Shutter-Brille, da andernfalls die Synchronisation Timing-Probleme aufweisen könnte. Im Fehlerfalle kommt es zu zeitlichen Überlappungen, was den 3D-Effekt stört.

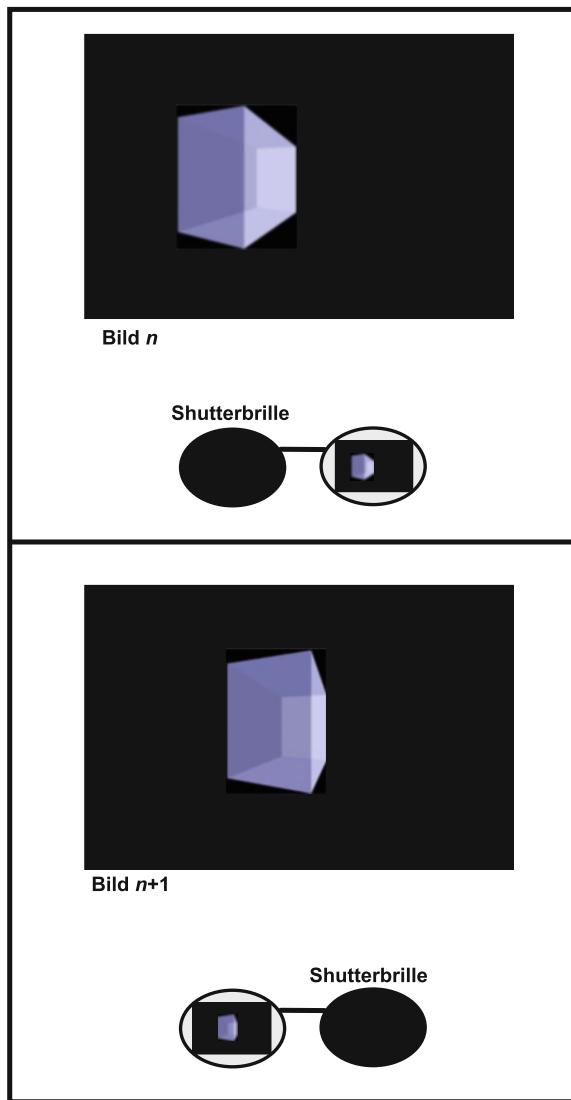
7.6.4 3D-Technik mit Polarisationsbrille

Gängige Methoden verwenden auf Betrachterseite Polarisationsbrillen. Diese Technik nutzt ebenfalls die Separation zweier Teilbilder für beide Augen des Betrachters. Es kommen hier zwei Verfahren in Frage – die lineare Polarisation und die zirkulare Polarisation. Die Trennung der Teilbilder geschieht ebenfalls mit Polfiltern, somit ist der Einsatz normalerweise auf Projektionswiedergabe beschränkt. Die Projektionsfläche muss metallisch reflektieren, sonst ginge die Polarisation verloren oder würde zumindest stark verfälscht werden.

Licht kann durch optische Filter polarisiert werden. Die Materialien hierzu sind optisch doppelbrechend.

Lineare Polarisation Gelangt diffus polarisiertes (also unpolarisiertes) Licht durch einen Polarisator linearer Charakteristik, so liegt danach Licht vor, welches nur noch eine einzige Schwingungsausrichtung aufweist. Wird ein zweiter Polarisator nachgeschaltet, dessen Schwingungsrichtung um 90° dem vorhergehenden verdreht ist, so ist an dessen

Abb. 7.32 Prinzip der 3D-Technik mit Shutter-Brille. Dargestellt sind zwei aufeinanderfolgende Bilder. Dies müssen nicht notwendigerweise zwei Vollbilder sein, der Bildwechsel für den Shutter kann je nach System schneller erfolgen. Phase des oberen Bildes: Die Optik für das rechte Auge ist offen. Phase des unteren Bildes: Die Optik für das linke Auge ist offen – vordergründige Objekte erscheinen aus perspektivischen Gründen dem linken Auge weiter rechts als dem rechten Auge



Ausgang im Idealfall kein Licht mehr vorhanden. Bei der Erzeugung des 3D-Bildes lässt man jeweils Licht der beiden Teilbilder durch jeweils ein Polarisationsfilter, wobei beide zueinander eine senkrecht ausgerichtete Polarisationsrichtung besitzen. Der Betrachter benötigt eine Brille mit ebenfalls orthogonal ausgerichteten Polfiltern, deren Schwingungsrichtungen möglichst exakt denen der projizierten Teilbilder entsprechen müssen, damit eine ideale Trennung der Bilder für das Auge gelingt. Diese Voraussetzung führt unmittelbar zu dem Nachteil dieses Verfahrens: Der Betrachter darf den Kopf

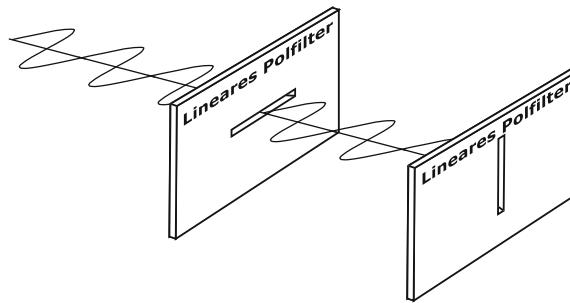


Abb. 7.33 Zur Anschauung der Durchlässigkeit oder Sperrung von Licht mit Polarisationsfiltern unterschiedlicher Polarisationsrichtung. Auf diese Weise ist es möglich, zwei Teilbilder, deren Schwingungsausrichtung orthogonal gewählt ist, zu separieren und beiden Augen des Betrachters zuzuführen

nur wenig neigen, um nicht die Trennung der Teilbilder durch falsch ausgerichtete Polarisationswinkel zu verschlechtern. Als Anschauung diene Abb. 7.33.

Zirkulare Polarisation Das Licht der Teilbilder ist links- bzw. rechtsdrehend polarisiert und wird über eine geeignete Polfilter-Brille den Augen zugeführt. Zirkulare Polarisation entsteht bei Licht, welches zuvor linear polarisiert war und zwar durch Überlagerung zweier linear polarisierter Wellen, deren Polarisationsrichtungen orthogonal zueinander stehen und gleichzeitig einen Phasenunterschied von 90° ($\pi/2$) aufweisen. Erreicht wird diese Konstellation durch die Verwendung eines $\lambda/4$ -Plättchens. Man kann sich in diesem Fall auch die Addition von Kosinus- und Sinuswert am Einheitskreis vorstellen: Als Summe entsteht der (rotierende) Zeiger. Sind beide Wellen von unterschiedlicher Amplitude, spricht man von elliptischer Polarisation. Was bei der Strahlumkehr durch einen Spiegel geschieht, macht Abb. 7.34 deutlich.

Die optische Achse des $\lambda/4$ -Plättchens besitzt gegenüber der Polarisationsrichtung des Linear-Polfilters einen Winkel von 45° . Im Plättchen erfolgt eine Aufspaltung in senkrecht zueinander polarisierte Strahlen, die unterschiedliche Ausbreitungsgeschwindigkeit haben und am Ende $\lambda/4$ Phasenversatz zueinander haben. Diese überlagern sich und generieren so den endgültigen, zirkular polarisierten Lichtstrahl. Die Drehrichtung hängt davon ab, ob die Linearpolarisation horizontal oder vertikal und somit $+45^\circ$ oder -45° zur optischen Achse des Lambda/4-Plättchens hat. Bei der Strahlumkehr (also bei der Filterung durch die Brille) durch das Plättchen entsteht wieder linear polarisiertes Licht. Demzufolge hängt es davon ab, welche Ausrichtung der Linearpolarisator hat (und somit welche Polarisationsdrehrichtung das Filter aufweist), ob der Strahl durchgelassen oder gesperrt wird. Wenn Wellenlänge nicht exakt passend ist für das Plättchen, entsteht elliptisch polarisiertes Licht. Bei Strahlumkehr ist folglich die Sperrung beim entsprechenden Polfilter nicht mehr ganz gegeben.

Der Vorteil bei diesem Verfahren gegenüber dem mit linearer Polarisation ist der, dass die Trennung der Bilder unabhängig ist von der Kopfneigung des Betrachters. Die Polarisierung ist hier bei jedem Winkel gleichbleibend. Zu nennen ist jedoch auch der Nachteil, dass die Zirkularpolarisation nur ideal gelingt für eine einzige Wellenlänge, nämlich derjenigen,

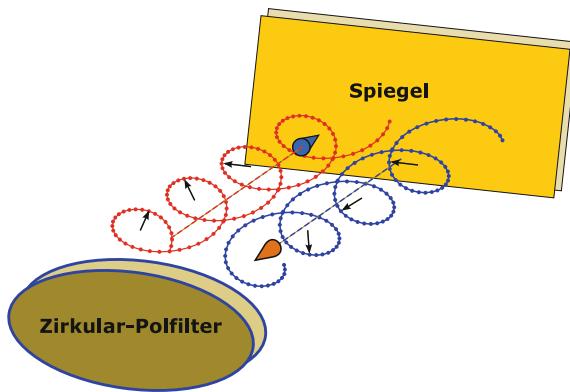


Abb. 7.34 Ein vom Spiegel zurückgeworfener Lichtstrahl kann nicht wieder dasselbe Zirkular-Polfilter passieren, deshalb erscheint es für das Auge dunkel (sofern kein seitliches Licht hinzukommt). Grund: Das Reflexionslicht kehrt seine Polarisationsdrehrichtung um, da die Rotationsrichtung (vom Betrachter aus gesehen) zwar gleich bleibt, die Ausbreitungsrichtung jedoch wechselt

für die die Phasenverschiebung genau 90° beträgt. Bei allen anderen Wellenlängen entsteht Licht mit elliptischer Polarisation – die Trennung beider Teilbilder gelingt hier nicht mehr vollständig. Das ist auch der Grund, warum bei Kaskadierung zweier Filter i. Allg. nicht Schwarz vorliegt, sondern ein farblich getöntes Grau (Abb. 7.35).

Die Verfahren mit Polfiltern ist mit speziellen Vorkehrungen auch mit Displays möglich. Hierzu wendet man i. Allg. eine spezielle Polarisationsfolie an, die auf dem Display-Glas aufgebracht ist. Diese Folie ist in horizontalen Streifen eingeteilt, die jeweils wechselnde Polarisationscharakteristik besitzen. Die Aufteilung erfolgt üblicherweise pro Zeile, sodass die Teilbilder vertikal jeweils nur noch halbe Auflösung haben und sozusagen *interlaced* sind, jedoch gleichzeitig erscheinen.

7.6.5 Weitere Verfahren für 3D-Wiedergabe

Wellenlängen-Multiplex-Verfahren Zur Aufteilung in zwei Bilder kommen hier Farbfilter zum Einsatz, deren Durchlassbereiche sich nicht oder wenig überlappen. Auf Projektorseite rotiert ein Filter, das je zur Hälfte die speziellen Filtereigenschaften aufweist. Die Bandbreiten für die drei Primärfarben sind mit ca. 50 nm recht eng gehalten und liegen bei den beiden Filtern auch nicht sehr weit auseinander: Filter 1 (rechts): $R=629\text{ nm}$, $G=532\text{ nm}$, $B=446\text{ nm}$ Filter 2 (links): $R=615\text{ nm}$, $G=518\text{ nm}$, $B=432\text{ nm}$.

Das System ist auf dem Markt unter der Bezeichnung *Dolby Digital 3D* zu finden.

Split-Image-Verfahren Natürlich besteht auch der Wunsch, auf Betrachterseite vollkommen ohne Spezialbrille auszukommen. Dies ist möglich, wenn auf dem Display feine

Abb. 7.35 Oben:

Demonstration mit zwei 3D-Brillen, die Zirkular-Polfilter enthalten. Die Außenseiten der Brillen stehen sich gegenüber, also derjenige Raum, wo im Betrieb die zirkular polarisierten Lichtsignale anstehen. Gleichartige Polfilter erscheinen hell, da sich der Strahlengang einfach umkehrt wie beim Gewindegang einer Schraube. Das andere Glas der fernen Brille hingegen schickt Licht zur nahen Brille mit zirkular polarisiertem Licht umgekehrter Drehrichtung – es erscheint dunkel. Unten: Eine 3D-Brille in Kombination mit einem Kamera-Polfilter aufgenommen. Die Farbstiche sind eigentlich normal, denn das obige Resultat erhält man nur durch Drehung des einen Brillenglases um 90°



senkrechte Prismen dafür sorgen, beide Teilbilder zu trennen. Die Blinkwinkel Augen bezüglich Display ist so unterschiedlich, dass jedes Auge nur jeweils eines der Teilbilder wahrnimmt. Nachteil ist dabei, dass die Voraussetzung für den richtigen Blickwinkel nur relativ dicht am Display gewährleistet und somit eine korrekte 3D-Wiedergabe meist nur für einen Zuschauer pro Display möglich ist.

7.6.6 Reales 3D-Sehen – Stand der Technik

Gegenüber der Stereoskopie – und von dieser Kategorie waren die bisher beschriebenen 3D-Verfahren – muss man ganz deutlich ein reales 3D-Sehen unterscheiden. Dieses nämlich muss die Möglichkeit bieten, um das betrachtete Objekt bzw. Bild herumzugehen, um es von allen Seiten ansehen zu können. Dies setzt voraus, dass tatsächlich Raumdaten eines Bildes oder eines Films vorhanden und verarbeitet werden müssen, während bei der Stereoskopie lediglich zwei unterschiedliche Bilder den Augen des Betrachters präsentiert werden, um dem Gehirn einen Raumeindruck *einer* Position zu vermitteln.

Der *Stand der Technik* bezieht sich auf das Jahr 2018. Das am Ausbau und möglicherweise auch Schaffung neuer Technologien weiterhin gearbeitet, steht außer Frage.

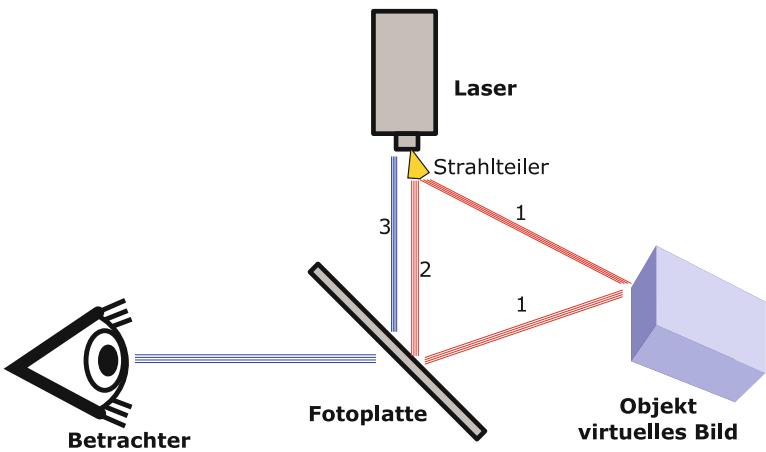


Abb. 7.36 Prinzip zur Erzeugung eines Hologramms. Bei der Herstellung (Strahlen in Rot) gelangen kohärente Wellen (1) zum Objekt, werden dort reflektiert und interferieren mit dem Referenzwellenzug (2), schwärzen dort eine hochauflösende Filmpalte, den *holografischen Film*. Bei der Rekonstruktion (Strahlen in Blau) des virtuellen Bildes führen Wellen (3) exakt gleicher Phasenlage wieder auf den entwickelten Film und nutzen diesen als Platte zur Erzeugung der Beugungsmuster. Es entsteht wiederum ein Bild durch Interferenz wie bei der Belichtung der Platte durch Überlagerung beider Wellenzüge. Somit entsteht ein virtuelles Bild, ein Hologramm

Holografie-Verfahren Holografie ist bereits seit den 1940er-Jahren bekannt. Allerdings hat sich seine Qualität enorm gesteigert, ferner sind spezielle Techniken entwickelt worden, wovon hier nicht alle zur Sprache kommen werden.

Bei der Herstellung des Hologramms (Abb. 7.36) wird durch sich überlagernde, kohärente Wellenzüge eine Bildplatte belichtet. Das Bildprodukt – welches übrigens keinerlei kenntliche Bildinhalte aufweist – kann später im entwickelten Zustand durch Interferenz das Bild selbst erzeugen, sobald wieder dieselbe kohärente Lichtquelle darauf scheint. Da zurzeit der Interferenz der Strahlwinkel entscheidend ist für Auslöschung bzw. Verstärkung, ist bei der späteren Reproduktion ebenfalls der Blickwinkel entscheidend dafür. Somit kann für einen weiten Betrachtungswinkel hinweg ein dreidimensionaler Bildeindruck entstehen. Die Laserlichtquelle benötigt einen Strahlspreizer, denn man benötigt keinen dünnen Strahl, sondern breit gefächertes Licht (was i. Allg. ein Laser ohne Kollimator-Optik zu liefern imstande ist).

Die beschriebene Technik hat keine Echtzeitcharakteristik – bereits das Belichten der holografischen Platte nimmt eine gewisse Zeit in Anspruch. Der Löwenanteil an Zeit jedoch benötigt das Entwickeln und Fixieren des Films. In verschiedenen Bereichen wie Medizin oder Maschinenbau ist man heute (Stand 2018) jedoch bereits so weit, Hologramme in Quasi-Echtzeit zu erzeugen und mit ihnen zu arbeiten. Von einer visuellen Videocharakteristik ist die Technik jedoch noch entfernt, denn der Bildlaufbau liegt noch im Zeitbereich von Sekunden, sodass man auch noch nicht von „flimmerfrei“ sprechen kann. Diese kommerziellen Anwendungen setzen jedoch *holografische Displays* ein,

die den holografischen Film ersetzen. Dieses spezielle Display erzeugt rechnergesteuert die feinen Beugungsmuster, die sonst eine Filmmplatte beinhalten. Diese Beugungsmuster zu berechnen stellt gewisse Ansprüchen an die Leistungsfähigkeit des Computers. S. 281 [1, 2]

Volumendisplay Die vielversprechendste Methode zur Wiedergabe von 3D-Bildern sind räumliche Projektionen in voluminöse Körper wie Würfel (*Visual Cube*) oder Kugeln (*Sphärendisplay*).

RGB-LED-Würfel Es gibt auch Bestrebungen, in einen Würfel sehr viele Dreifarb-LEDs zu integrieren. Ein solches System hat nicht nur technologische Grenzen, da die Größe der LEDs eine natürliche Mindestgröße verlangt und außerdem auch elektrische Zuleitungen benötigt, die das Gesamtbild stören. Es besteht auch eine systematische Grenze der Integration, denn die LED-Elemente müssten möglichst klein und dicht sein und dürfen dabei ihre Transparenz nicht verlieren.

Sphärendisplay Bei projizierenden Systemen wird die Oberfläche eines Volumenkörpers von außen durch Projektoren beleuchtet. Da es sich hierbei um eine Projektion auf die Oberfläche handelt, muss die Position des Betrachters möglichst genau erfasst werden, da das Projektionsbild davon abhängt. Ein leistungsfähiger Rechner muss die Bilder der Projektoren ständig korrigieren, wenn der Betrachter seinen Standpunkt variiert. Für verschiedene Anwendungen ist zusätzlich Interaktion möglich und gefordert, sodass der Rechner auch auf derlei Eingaben reagieren muss (Abb. 7.37).

Rotationsdisplay Ein interessantes System für 3D-Wiedergabe ist eine schnell rotierende Fläche, die von mehreren Laserlichtquellen angestrahlt werden. Die Projektionsfläche überstreicht dabei ein Raumvolumen, sodass innerhalb eines bestimmten Zeitintervalls alle Raumpunkte als Projektionspunkt aktiviert werden können und somit ein geschlossenes Raumbild entsteht. Voraussetzung für die genaue Wiedergabe ist exakte Synchronisation der Laserquellen mit der Drehposition der Projektionsfläche, was jedoch technisch keine Herausforderung mehr bedeutet. Ähnliche Systeme in der Werbetechnik setzen mitunter Rotationsflächen mit integrierten Mehrfarb-LEDs ein, womit dann Texte und einfache Figuren als 3D-Eindruck entstehen.

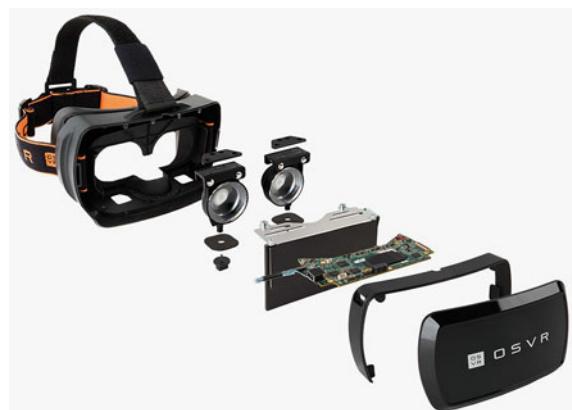
7.6.7 Virtual Reality – virtuelle Realität

Bei der virtuellen Realität handelt es sich um eine interaktive Wiedergabeform eines dreidimensionalen Bild- bzw. Videoeindrucks. Man benötigt einige spezielle Komponenten, die zusammen auch kostenintensiv sind.



Abb. 7.37 Beispiel für ein Sphärendisplay. (Photograph of the Perspecta volumetric display courtesy of Optics for Hire)

Abb. 7.38 Innerer Aufbau einer VR-Brille (mit freundlicher Genehmigung der Fa. Open Source Virtual Reality / OSVR)



VR-Brille Obwohl es virtuelle Realität auch ohne spezielle VR-Brille gibt, liegt der Fokus hier bei Systemen mit einer solchen Brille. Diese beinhaltet entweder zwei Displays (siehe Abb. 7.38) oder auch nur eines, welches beide Bilder entweder sequentiell oder simultan wiedergibt – über Shutter- oder Polfilter erfolgt dann wiederum die Separation der Bilder für beide Augen (siehe Abschn. 7.6.3 bzw. 7.6.4). Diese Brille wird möglichst passgenau am Kopf vor den Augen positioniert. Da der Abstand zu den Augen nur wenige Zentimeter beträgt, muss noch je eine Linse dafür sorgen, dass die Scharfstellung (Akkommodation) für das Auge überhaupt realisierbar ist.

Immersion Der Begriff stammt eigentlich aus den Bereichen Physik, Astrophysik und auch Mathematik und hat im weitesten Sinne etwas mit dem *Eintauchen* eines Körpers oder Mediums in ein anderes zu tun. Bei der virtuellen Realität spricht man dagegen von Immersion, wenn ein gewisser Grad von Realitätsempfinden erreicht ist, obwohl sich der Anwender *de facto* in einer künstlichen und nichtrealen Umgebung befindet.

Interaktivität Außer der VR-Brille steigert sich der Grad der Immersion dadurch, dass das Programm nicht nur das Visuelle steuert, sondern indem es auch die Position der Person, seine Augenstellung, seine Kopfneigung und andere Gesten verarbeitet. Dafür nötige Hilfsmittel sind Datenhandschuh, Flystick (eine Art Joystick), Laufband, 360°-Kamera, Force-Feedback, um auch Haptik zu integrieren.

VR-Programm Das Steuerprogramm, das auf einem Computer hoher Rechenleistung abläuft, muss nicht nur komplexe Aufgaben übernehmen, sondern auch über diverse Schnittstellen verfügen, um die oben genannten Interaktivitäten zu bedienen. Die Standardschnittstellen wie USB-3 und höher sind heutzutage in der Lage, diese Aufgaben zu übernehmen. Videoschnittstellen wie „HDMI“ 2 und höher sind ebenfalls geeignet für den Betrieb der VR-Brille. Nicht zu vergessen Audio: Hier ist wohl meist USB ausreichend. Bluetooth hat oftmals noch ein merkliches Latenzproblem.

Programme wie z. B. *Blender* haben viel Videoschnitt und Animationssoftware zu tun. Bearbeitungsprozesse reichen von Modeling, Tracking, Animation, Rigging übers Rendering zum fertigen Produkt (Abb. 7.39).

Das fertige Ablaufprogramm muss auf alle interaktiven Daten möglichst ohne Latenz reagieren, also auf die Position des Akteurs, auch auf die Augenstellung, auf die Kopfneigung und beim Einsatz von Datenhandschuh oder Flystick auf deren Daten. Ferner muss das Force-Feedback aktiviert werden, falls vorhanden.

Einsatzbereiche Ein VR-System hat vielerlei Einsatzmöglichkeiten, wovon sogar der Bereich der Unterhaltung als untergeordnet betrachtet werden kann:

- Raum- und Landschaftsplanung
- Flugsimulation
- Arbeitsschutz, üben an virtuellen Maschinen
- Produktentwicklung
- Spiele und Unterhaltung

Motion Sickness ist die Bezeichnung für eine Krankheitssymptom, bei dem der Benutzer in der virtuellen Realität ähnlich wie bei der Seekrankheit (oder auch Reisekrankheit) an Schwindelgefühl leidet. Die Ursache hierfür sind Unstimmigkeiten zwischen den Sinneseindrücken, beispielsweise zwischen visuell wahrgenommenen Bewegungen im Gegensatz zu den Reizen des Gleichgewichtsorgans. S. 281 [3]

7.7 Speichermedien

7.7.1 DVD – Digital Versatile Disc

Bei gleichen Abmessungen wie die CD weist die DVD einige Änderungen auf:

Es sind zwei Layer möglich, somit mehr Kapazität. Die Farbe des Lasers ist grün, somit ist eine bessere optische Auflösung möglich. Die Pits sind um 50 % kürzer, ebenfalls der Spur-zu-Spur-Abstand. Nach dem sog. Finalisieren sorgt eine spezielle Codierung für Video-Wiedergabe eine Menüführung, Sprachenauswahl usw.

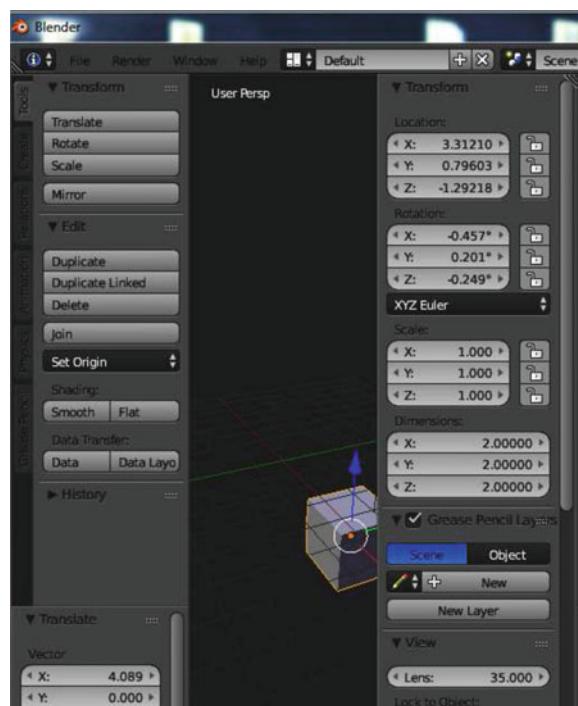
Die DVD wurde geschaffen, um im Bereich Unterhaltung ein mit Menü geführtes Medium zu erhalten. Außerdem sollte mindestens ein Film üblicher Länge auf eine Scheibe passen. Im Datenformat ist eine Timeline integriert, womit der Anwender sekundengenau auf eine Position des Films springen kann.

Kapazität: 4,7 GB bei Single-Layer, 8,5 GB bei Dual-Layer.

Typen von DVDs:

- DVD-R: Recordable, einmal beschreibbare DVD.
- DVD+R: Dto, jedoch mit Sonderfunktionen (DVD Alliance).
- DVD-RW: Wiederbeschreibbar (ca. bis zu 1000 mal).
- DVD+RW: Dto, jedoch mit Sonderfunktionen (DVD Alliance).

Abb. 7.39 Ein Ausschnitt aus dem Eröffnungsbildschirm: Das Open-Source-Programm *Blender*. Augenscheinlich hat es viel mit Animationssoftware zu tun, denn die Umgebung ist eine künstlich zu erstellende Landschaft



- DVD-RAM: Freier Lese- und Schreibzugriff möglich wie bei einem RAM.
- DVD-DL: Dual Layer, zwei Schichten.

7.7.2 Blu-ray

Gegenüber der DVD besitzt das Blu-ray-Medium eine wesentlich gesteigerte Speicherkapazität von ca. 25 GB im Singelayer und 50 GB im Duallayer (höhere Layer-Zahlen sind möglich). Die geometrischen Maße sind identisch zur DVD (12 bzw. 8 cm Durchmesser). Zur Realisierung dieser extrem hohen Schreib-/Lesedichte weisen die Geräte zwei entscheidende Punkte gegenüber denen für DVDs auf: Der Laser arbeitet mit einer kürzeren Wellenlänge (ca. 405 nm) – deshalb die Bezeichnung „Blu“. Ferner befindet sich die Optik wesentlich näher an der Datenträger-Schicht, dadurch gelingt noch eine weitere wesentliche Verbesserung der Fokussierung.

Das Medium selbst hat ist ebenfalls unterschiedlich aufgebaut – eine übliche Roh-DVD lässt sich nicht als Blu-ray beschreiben. Als wichtigster Unterschied ist die sehr viel dünner Schutzschicht zu nennen, welche auch aus anderem Material (*Durabis*) besteht.

Literatur

1. J. Eichler, G. Ackermann: Holographie, Springer-Verlag Berlin Heidelberg 1993.
2. Welt der Physik: Holographische Bilder fast in Echtzeit.
3. Manfred Brill: Virtuelle Realität, Springer-Verlag Berlin Heidelberg 2009.
4. Alistair J. Lees (Hrsg.): Photophysics of Organometallics.
Springer-Verlag Berlin Heidelberg 2010.
5. Robert Siebeck: Optik des Menschlichen Auges.
Springer-Verlag Berlin Heidelberg 1960.
6. Küppers, H.: Das Grundgesetz der Farbenlehre. Köln: DuMont-Verlag 1978.
7. Lang, H.: Farbmatrik und Farbfernsehen. München: Oldenbourg-Verlag 1978.
8. Stotz, D.: Audio- und Videogeräte richtig einmessen und justieren.
München: Franzis-Verlag 1994.
9. Wikipedia: Flüssigkristallbildschirm.
10. Webers, J.: Handbuch der Film- und Videotechnik. München: Franzis-Verlag 1993.
11. Häberle, H.: Elektronik, Teil 3: Nachrichtenelektronik.
Wuppertal: Europa-Lehrmittel-Verlag 1980.
12. Wikipedia: High Definition Television.



Parameter beim digitalen Videobild

8

8.1 Farbcodierung

8.1.1 Farbtiefe

Ein wichtiger Begriff bei der Verarbeitung und Wiedergabe einer Grafik ist die Farbtiefe. Sie ist in Bit/Pixel angegeben und definiert damit die Anzahl binärer Stellen, mit denen die Farbe eines Pixels beschrieben werden kann. Damit stellt die Farbtiefe das Äquivalent zur Auflösung eines digitalen Audiosignals dar.

Die Anzahl der möglichen Farbwerte (eigentlich ist es die Kombination aus Helligkeit und Farbton) ist somit:

$$c_{\max} = 2^D \quad (8.1)$$

Hierin ist D der Zahlenwert der Farbtiefe, wobei jedoch bestimmte, festgelegte Farbtiefe (32 und 16 Bit/Pixel) nicht immer alle Bits für die Farbtiefe heranziehen.

Die Feinheit des Farbrasters steigt somit also exponentiell mit der Farbtiefe.

In der Praxis häufig vorkommende Farbtiefe sind 8, 16, 24 und 32 Bit/Pixel. Die beiden letzteren unterscheiden sich dabei nicht in der sichtbaren Farbtiefe; bei 32 Bit existiert jeweils für jedes Pixel eine Zusatzinformation, die 8 Bit breit ist. Man kann hier jedem einzelnen Pixel ein sog. Attribut anhängen, das z. B. über Transparenz für ein zweites Bild entscheidet (Overlay-Technik) oder auch Schutzfunktionen bei der Bildbearbeitung übernimmt. Die professionelle Bildverarbeitung setzt meist 24 und 32 Bit/Pixel (16,8 Mill. Farbwerte) ein, während bei semiprofessionellen Anwendungen (teilweise auch für TV-Anwendungen) bereits 16 Bit/Pixel (32 768 Farbwerte) ausreichen.

8.1.2 Verschlüsselung der Farbdaten

Für die digitale Beschreibung von Farbdaten gibt es verschiedene Methoden. Die einfachste davon ist die Zuordnung einer paritätischen Anzahl von Bits der Farbtiefe den drei Primärfarben Rot, Grün und Blau. Man spricht hier auch von einer linearen RGB-Codierung. Abb. 8.1 zeigt diese Methode für die Farbtiefen 16 und 24 bzw. 32 Bit/Pixel.

Bei einer Tiefe von 24 bzw. 32 Bit/Pixel erhält jede Primärfarbe je 8 Bit zur Beschreibung der Helligkeit; das sind also immer die Werte 0...255, wobei 0 für Schwarz und 255 für größte Intensität steht.

Die Farbtiefe 16 Bit/Pixel nimmt ebenfalls eine gleichmäßige Aufteilung für die Primärfarben vor, jedoch mit der Einschränkung, dass hier jeweils nur 5 Bit zur Verfügung stehen. Damit können die Werte 0–248 in 8er-Schritten realisiert werden. Ob nun bei der Wiedergabe der Grafik die untersten drei Bits tatsächlich immer 0 bleiben, ist allein Sache der verarbeitenden Grafikkarte.

Damit der Weißwert nicht auf 248 beschränkt ist, koppeln manche Grafikkarten Bit 7 mit Bit 2, Bit 6 mit Bit 1 und Bit 5 mit Bit 0. Auf diese Weise wird der gesamte Bereich von 0...255 abgedeckt, wobei allerdings die Schrittweite verständlicherweise an einigen Stellen etwas variiert. Bei einer Umstellung der Grafikkarte von 24 auf 16 Bit/Pixel ist somit keine Verdunkelung der Weißbereiche festzustellen.

Applikationen, die nicht für den professionellen Grafikbereich gedacht sind, unterstützen möglicherweise nur geringere Farbtiefen, die normalerweise nicht linear codiert werden,

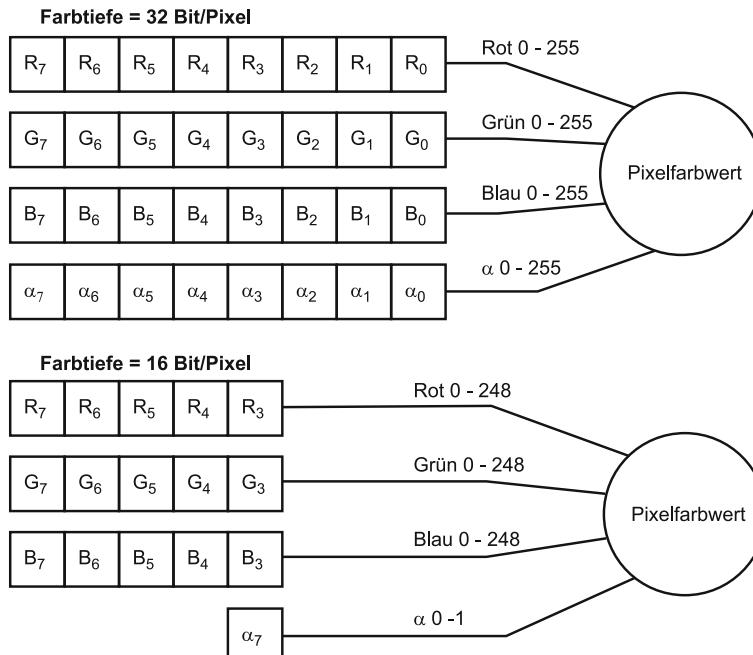


Abb. 8.1 Binärinformationen für ein Pixel bei verschiedenen Farbtiefen

Abb. 8.2 Indizierte Farbcodierung über eine Farbpalette

Index bzw. Adressierung der Farben

Hier z. B. Tabelleneintrag für Farbe C_{07} mit $R=140$, $G=80$ und $B=255$

Palette mit 16 möglichen Farbeinträgen

xx00	xx01	xx10	xx11	
00xx	C_{00}	C_{01}	C_{02}	C_{03}
01xx	C_{04}	C_{05}	C_{06}	C_{07}
10xx	C_{08}	C_{09}	C_{10}	C_{11}
11xx	C_{12}	C_{13}	C_{14}	C_{15}

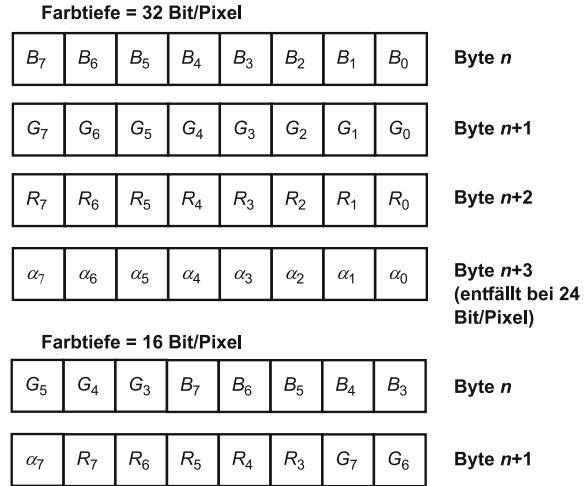
sondern eine sog. indizierte Codierung aufweisen. Basis für diese Indizierung ist die sog. Palette, die bei einer Farbtiefe von z. B. 8 Bit/Pixel 256 unterschiedliche Farben zur Verfügung hält. Welche Farben dies nun genau sind, hängt vom Eintrag der Farbpalette ab; letzterer kann normalerweise frei gewählt werden. Die Farbpalette ähnelt einer Tabelle, aus der über die Angabe der Zeile und Spalte eine definierte Farbe selektiert wird. Abb. 8.2 zeigt das Prinzip einer 4-Bit-Palette für 16 mögliche Farben.

Ob die Wiedergabe einer Grafik mit hoher Farbtiefe auf dem Monitor möglich ist, hängt von der verwendeten Grafikkarte ab. Sogenannte Truecolor-Karten unterstützen 24 Bit/Pixel, während HiColor-Karten lediglich bis 16 Bit/Pixel gehen. Herkömmliche VGA-Karten unterstützen lediglich 8 oder gar nur 4 Bit/Pixel (indizierte Codierung über Palette). Je nach eingestelltem Bildformat und Speicher bzw. physischer Speicheraufteilung sind unterschiedliche Farbtiefen möglich. Soll eine Grafik mit größerer Farbtiefe wiedergegeben werden, als dies mit der Grafikkarte möglich ist, so findet eine entsprechend reduzierte Farbwiedergabe statt, die die Anwendung zusammen mit dem Grafikkartentreiber teilweise in Echtzeit vollführt. Mitunter wird auch ein sog. Dithering durchgeführt, welches mit einem beschränkten Farvvorrat eine bestimmte Farbe anzunähern versucht, indem eine Mischung verschiedenfarbiger Pixel erzeugt wird. Letzteres ist natürlich nur ab einem bestimmten Grafik-Areal möglich, bei dem genügend benachbarte Pixel für diese Annäherung zur Verfügung stehen. Das Dither-Verfahren und auch andere, wie z. B. das von Floyd-Steinberg, die alle relativ rechenaufwändig sind, arbeiten auf reiner Software-Basis. S. 293 [1].

8.1.3 Datenanordnung der Farbinformation

Hier geht es um die Anordnung im Bildspeicher einer Grafikkarte, die ja verständlicherweise unabhängig vom verwendeten Grafikprogramm ist. Letzteres „liest“ Grafikdateien lediglich und lädt – unter Berücksichtigung der Regeln für die Datenanordnung im Bildspeicher – die Daten ein. Abb. 8.3 illustriert diese Regeln für zwei verschiedene Farbtiefen, wobei sich 24 von 32 Bit/Pixel ja nur in der Auslassung des a-Byte unterscheidet. Bei 16 Bit/Pixel wurde dieselbe Bit-Indizierung vorgenommen wie in Abb. 8.1, d. h. die Bits

Abb. 8.3 Datenanordnung im Bildspeicher bei RGB-Codierung



entsprechen jeweils ihrem interpretierten Wert und beginnen daher nicht mit dem Index 0. Eine Farbtiefe von 8 Bit/Pixel und darunter unterliegt ja wie erwähnt einer indizierten Palettencodierung und bietet somit in der Datenanordnung nichts Interessantes; hier werden lediglich chronologisch die Datenbytes der Datei in den Speicher gelegt – den Rest übernimmt eine spezielle Hardware der Grafikkarte.

Manche Grafikkarten verwenden keine RGB-Datenstruktur, sondern eine Aufteilung in Farbdifferenzwerte und Leuchtdichte (Helligkeit, Luminanz). Dies geschieht meist aus Gründen der Verringerung der Datenmenge bzw. der Speichergröße. Es handelt sich dabei meist um die sog. 4-1-1-Methode. Der Name sagt aus, dass die Abtastung der Luminanz Y prinzipiell viermal so häufig geschieht wie die der Farbdifferenzwerte U und V . Tab. 8.1 zeigt die chronologische Anordnung von vier aufeinanderfolgenden Pixeln innerhalb des Speichers. Pro Pixel sind im Speicher also (hier) zwei Bytes reserviert, bei denen nicht alle Bits mit definierten Werten belegt werden. Die jeweilige Zugehörigkeit zum entsprechenden Pixel wird vom Index des Luminanzsymbols Y angedeutet. Da ja bei vier Abtastungen der Luminanz innerhalb der insgesamt vier Pixel die Farbdifferenzwerte U und V nur einer Abtastung unterliegen, weisen diese keine Indizes auf.

Üblicherweise werden bei 4-1-1 einem Pixel zwei Bytes im Speicher zugeordnet. Der Luminanzwert Y ist dabei mit sieben Bits vertreten. Die Farbdifferenzwerte U und V erhalten ebenfalls je eine 7-Bit-Auflösung, jedoch verteilen sich diese auf vier Pixel (da ja die Abtastung mit nur 1/4 der Häufigkeit erfolgt). Die Darstellung der Doppelbytes darf nicht darüber hinweg täuschen, dass die Speicheranordnung prinzipiell byte-reversed ist, d. h. zunächst erscheint das Byte mit den Bits 0...7, dann das Byte mit den Bits 8...F.

Tab. 8.1 Aufteilung der Speicherdaten beim 4-1-1-System

	Bit-Nr.	F	E	D	C	B	A	9	8	7	6	5	4	3	2	1	0
Pixel 0	Komponente	<i>U</i>	<i>U</i>	<i>V</i>	<i>V</i>					Y_0							
	Wertigkeit	6	5	6	5					7	6	5	4	3	2	1	0
Pixel 1	Komponente	<i>U</i>	<i>U</i>	<i>V</i>	<i>V</i>					Y_1							
	Wertigkeit	4	3	4	3					7	6	5	4	3	2	1	0
Pixel 2	Komponente	<i>U</i>	<i>U</i>	<i>V</i>	<i>V</i>					Y_2							
	Wertigkeit	2	1	2	1					7	6	5	4	3	2	1	0
Pixel 3	Komponente	<i>U</i>		<i>V</i>						Y_3							
	Wertigkeit	0		0						7	6	5	4	3	2	1	0

8.2 Bildformat und Auflösung

Während sich die Farbtiefe ein Parameter jedes einzelnen Pixels ist, stellen das Bildformat und die Auflösung der Grafik Eigenschaften dar, die sich auf die Gesamtheit der Grafik beziehen.

8.2.1 Bildseitenformat

Bei der Entwicklung des Fernsehens ist man von einem Seitenverhältnis von 4:3 aus gegangen; heute weiß man, dass das Seitenverhältnis von 16:9 den physiologischen Gegebenheiten des Auges besser entspricht (HDTV). Wir wollen jedoch zunächst das herkömmliche Format als gegeben betrachten. Soll eine wiedergegebene Grafik bildfüllend wiedergegeben werden, so muss diese ebenfalls ein Seitenverhältnis von 4:3 aufweisen, sonst erscheint entweder ein Rand, oder die Grafik erscheint verzerrt.

8.2.2 Auflösung

Die Auflösung einer Grafik beschreibt durch zwei Werte die Anzahl der Pixel pro Zeile und die Anzahl der Zeilen des Gesamtbildes – oder kurz die Kolumnen – und die Zeilenzahl.

Solange eine Grafik nur mit einem Programm bearbeitet wird, kann sie theoretisch beliebige Auflösungen und somit Formate einnehmen. Bei einer Verwendung in Videoapplikationen jedoch sollten Grafiken grundsätzlich das entsprechende Videoformat aufweisen. Es ist hier vielfach ratsam, im Hinblick auf den späteren Einsatz im Videobereich das Grafikformat von vornherein entsprechend festzulegen.

Viele Applikationen für Videografik sind in der Lage, nachträgliche Änderungen in der Auflösung vorzunehmen. Die hierzu notwendige Technik der Interpolation bzw. Pixelauslassung verbessert jedoch die Grafik keinesfalls. Geht man z. B. von einer

ursprünglichen Auflösung 640×480 aus (das ist die übliche Auflösung einer Standard-VGA-Karte), so muss bei Verwendung unter PAL die Grafik „gestreckt“ werden, und das geschieht hier durch das Hinzufügen von interpolierten Pixeln, und zwar erfolgt eine Pixelergänzung nach jedem 5. Pixel in Zeilenrichtung und ebenfalls nach jeder 5. Zeile eine Zeilenergänzung. Eine derartige Umrechnung kann meist nur die Software unterstützen – eine entsprechende Hardware könnte dies zwar auch, jedoch ist dazu ein spezieller Grafikprozessor nötig.

8.2.3 Overscan

Bei herkömmlichen Fernsehen sind die Bildränder unsichtbar, d.h. das Bild wird über den Rand des Monitors hinaus abgelenkt. Der Grund hierfür liegt darin, dass eventuelle Störzonen (z. B. Kopfumschaltung bei Videorecordern) und Informationsbereiche (Videotextsignale in der Vertikalaustastlücke) im Verborgenen bleiben sollen. Aus Kap. 7 wissen wir, dass das B/G-PAL-System 625 Zeilen im Vollbild aufweist. Diese Zeilenzahl kann jedoch dargestellt werden, weil ja noch die Vertikalaustastlücke mit in die Ablenkzeit dieser 625 Zeilen fällt.

Als Brutto-Bildzeilenzahl verbleiben festlegungsgemäß noch 576 Zeilen, von denen jedoch aus oben genannten Gründen wiederum einige außerhalb der sichtbaren Monitorfläche beschrieben werden. Zum besseren Verständnis diene Abb. 8.4.

Der angegebene Zeilenzahlbereich erscheint also auf dem Schirm. Der tatsächliche Überlappungsgrad ist natürlich von der Einstellung des Monitors abhängig. Tatsächlich könnte man letzteren so einstellen, dass auch ein Bildformat mit 576 Zeilen mit Rändern dargestellt wird; dies tut jedoch keinen Abbruch daran, dass es sich um ein Overscan-Format handelt.

Wir wollen hier davon ausgehen, dass die zu bearbeitende Grafik als Vollbild (*Full Screen*) erscheint, was bei einer späteren Videoanwendung ja auch meist der Fall sein

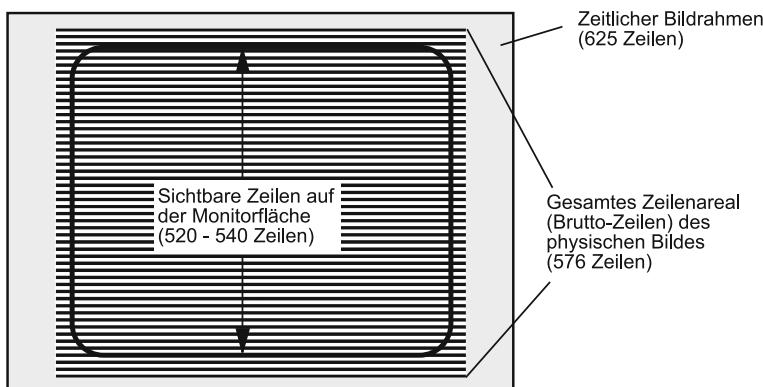


Abb. 8.4 Overscan-Betriebsart in der Videotechnik

wird. Die Grafik füllt also den gesamten Bildbereich aus und erscheint nicht etwa innerhalb eines Fensters. Die (eingestellte) Auflösung der Grafikkarte stimmt somit mit derjenigen der Grafik überein.

Die Anzahl der Kolumnen bzw. der Pixel pro Zeile ist nicht hinreichend entscheidend, ob eine Grafikauflösung die Overscan-Bedingung erfüllt oder nicht, denn die Ausdehnung einer Zeile hängt auch von der Geometrie der Pixel ab, wie wir im nächsten Abschnitt noch sehen werden. Dagegen legt die Zeilenzahl immer zugleich fest, wie lange das Schreiben des Bildes dauert und somit welcher prozentuale Anteil auf der sichtbaren Monitorfläche erscheint. S. 293 [2].

8.2.4 Underscan

Vornehmlich bei Vollbild-Computer-Grafiken für Non-Videoanwendungen hat die gesamte Grafikfläche Gültigkeit und ist somit auch sichtbar – alle Bildränder befinden sich innerhalb der Monitorgrenzen. Abb. 8.5 verdeutlicht diese Situation.

Da die Randzonen der Vollbild-Grafik sichtbar sind, werden hier oftmals auch wichtige Inhalte hineingelegt. Bei späterer Konvertierung auf ein Overscan-Format gehen diese Details dann zwangsläufig verloren. Günstiger ist es, bereits bei der Erstellung der Grafik eine Overscan-Auflösung zu wählen (bei PAL 768×576) und unter Overscan-Bedingungen zu arbeiten.

Wie bereits für den Overscan-Modus erwähnt wurde, ist allein entscheidend die Zeilenzahl einer Grafik – dasselbe trifft natürlich auch für die Underscan-Betriebsart zu. Als Grenzwert für den Übergang von Underscan nach Overscan kann etwa eine vertikale Auflösung von 540 (Zeilen) gelten.

Tab. 8.2 führt einige gängige Auflösungen auf.

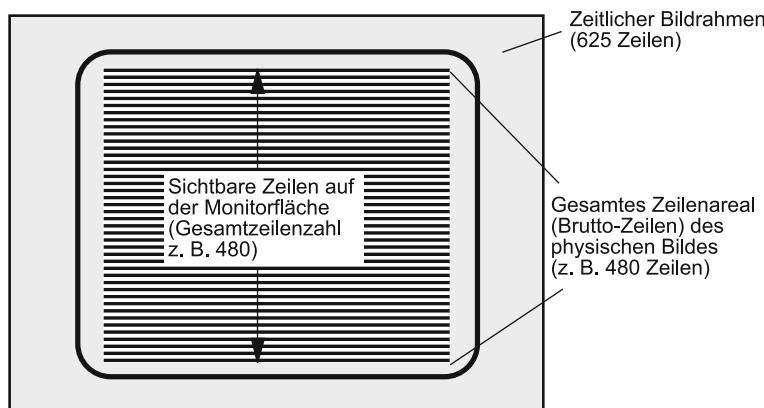


Abb. 8.5 Underscan-Modus einer Grafikauflösung

Tab. 8.2 Verschiedene Bildschirm-Auflösungen

Auflösung (Kolumnen × Zeilen)	Aspect ratio	Bemerkung
768 × 576	1	Overscan/PAL
720 × 576	1,07	Overscan/PAL; empfohlen nach CCIR 601
512 × 576	1,5	Overscan/PAL; oft günstig in Bezug auf Page-Begrenzungen
640 × 480	1	Underscan bei PAL (625/50) Overscan bei NTSC (525/60)
1024 × 768	1	HiRes-Modus; ungeeignet für PAL und NTSC
800 × 600	1	Unter bestimmten Bedingungen geeignet für PAL (Overscan)
756 × 486	0,86	Underscan bei PAL (625/50), Overscan bei NTSC (525/60)
1024 × 576	0,75	Overscan/PAL

Der Begriff *Aspect Ratio* wird im nächsten Abschnitt erklärt. Manche Auflösungen sind nur für PAL geeignet, nicht für NTSC. Das liegt an der Gesamtzeilenzahl, die ja bei NTSC mit 525 genau um 100 Zeilen kleiner ist als bei PAL.

8.3 Aspect Ratio

Der Begriff Aspect Ratio fällt im Bereich der Computergrafik immer wieder, vor allem wenn es darum geht, Grafiken verzerrungsfrei wiederzugeben.

Es handelt sich um das momentane Seitenverhältnis eines jeden Bildpixels, damit ist jedoch das physische Pixel gemeint, wie es theoretisch auf der Monitorfläche erscheint. Damit hängt die gültige Aspect Ratio primär mit dem Grafikkartenmodus zusammen. Die in Grafikdaten und -dateien gespeicherten Pixelwerte entbehren grundsätzlich einer Angabe der Aspect Ratio. Wir wollen uns durch Abb. 8.6 die Zusammenhänge etwas näher betrachten.

Die Aspect Ratio ist somit das Verhältnis aus Breite und Höhe eines Pixels, also:

$$\epsilon = \frac{\text{Breite}}{\text{Höhe}} = \frac{4}{3} \cdot \frac{N_{\text{lin}}}{N_{\text{col}}} \quad (8.2)$$

Zur Berechnung der Aspect Ratio mit den Daten der Bildauflösung ist immer die Kenntnis des Bildformats notwendig, welches unter den Standardgegebenheiten 4:3 ist. Neuere Festlegungen (nach HDTV) sehen ein Bildseitenverhältnis von 16:9 vor, auf welches wir später noch zu sprechen kommen.

Inkompatible Werte für die Aspect Ratio. Probleme treten bei der Wiedergabe von Grafiken auf, wenn für die Erstellung und die Wiedergabe unterschiedliche Werte der

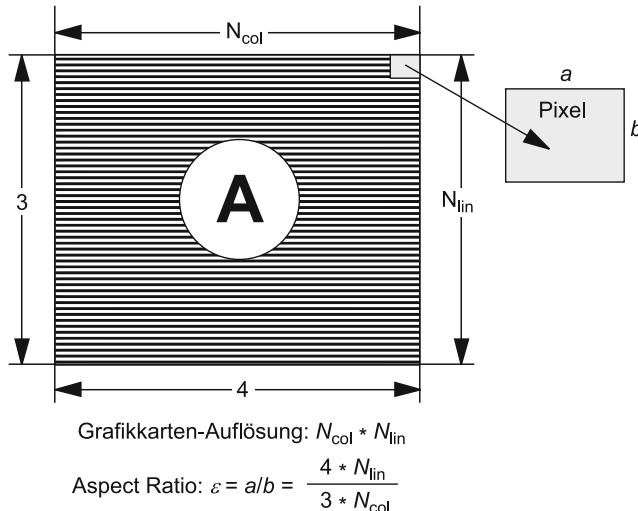


Abb. 8.6 Berechnung der Aspect Ratio bei einer Grafikwiedergabe

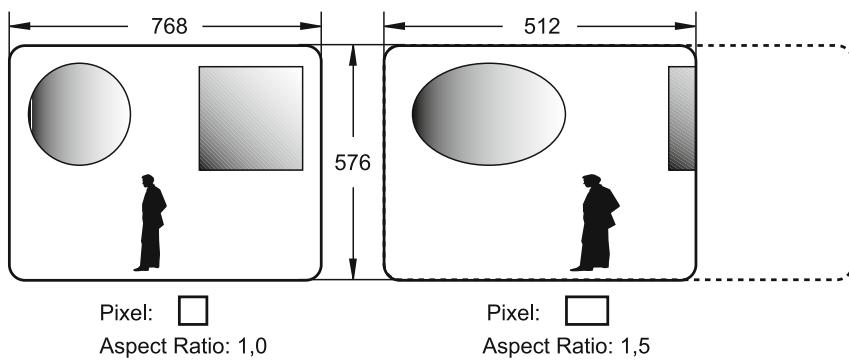


Abb. 8.7 Verzerrte Wiedergabe einer Grafik nach Veränderung der Aspect Ratio

Aspect Ratio angenommen werden. Das Resultat ist eine nicht zu übersehende Verzerrung, welche in Abb. 8.7 mit einem sehr deutlichen Beispiel zu sehen ist.

Zunächst wurde die Grafik in der Bildauflösung 768×576 als Vollbild generiert, wo Quadrat, Kugel und Figur die richtigen Proportionen aufweisen. Die Wiedergabe erfolgte auf einer Grafikkarte, die auf einen Treibermodus von 512×576 gesetzt war, wodurch eine deutliche Verzerrung auftritt. Während sich an der Zeilenzahl hier nichts ändert, werden durch den Treibermodus 512 Pixel einer Zeile auf die volle Bildbreite gebracht, denn jeder Treibermodus ist von der zeitlichen und geometrischen Aufteilung so ausgerichtet, dass beim Vollbild immer ein Seitenverhältnis von 4:3 eingehalten wird. Die virtuelle Bildbegrenzung ist gestrichelt eingezeichnet. Ob der Bildursprung (Origin) in der linken oberen Ecke liegt, hängt von den Daten des Grafikformats ab.

Der umgekehrte Fall ist natürlich ebenso möglich: bei Verkleinerung der Aspect Ratio durch einen alternativen Bildschirmtreiber erscheint die Grafik zu schmal, und die Grafikbegrenzung ist inmitten des Bildschirms sichtbar.

Sofern bei der Erstellung der Grafik der korrekte Bildschirmtreiber gewählt wurde, ist lediglich bei der Wiedergabe die entsprechende Grafikkarte mit dem gleichen Treiber auszustatten oder eine geeignete Videografikkarte zu wählen. Ist die Grafik jedoch unter einem Treiber generiert worden, der für die spätere Videoapplikation ungeeignet ist, so gibt es nur zwei Möglichkeiten aus dem Dilemma: Entweder muss die Grafik neu erstellt werden, oder die bestehende ist durch Pixel-Interpolation auf die neue Auflösung umzurechnen. Letzteres ist jedoch in den meisten Fällen mit einer Konturenverzerrung oder -unschärfe verbunden. Sofern das Grafikkartensignal analog weiterverarbeitet wird, ist meist die Verwendung einer unüblichen Spaltenzahl nicht von der Hand zu weisen. Schwierig wird es allerdings, wenn die Grafik mit einer Karte erstellt wurde, über deren Auflösung die Videografikkarte nicht verfügt.

8.4 Bildabtastung und Pixelraten

Das Auslesen des Bildspeichers erfolgt nach bestimmten Richtlinien, die u. a. auch von der Bildschirmauflösung bestimmt wird, wobei hier der theoretische Wert gemeint ist, den die Grafikkarte aufgrund ihres Treibers wiederzugeben vermag; es hat also nichts mit der Maskenauflösung des Monitors zu tun.

In der Videotechnik müssen die Bildsignale im Modus Interlaced zur Verfügung stehen; für die Zusammenstellung eines Halbbildsignals aus dem Speicher wird daher immer eine Zeile übersprungen. Bei der Adressierung des Videospeichers arbeitet man i. Allg. mit Teilern, wobei die Betriebsart Interlaced eine Zeilenumschaltung (Row-Address) mit doppelter Taktrate (halbem Teilungsverhältnis) im Vergleich zu Non-Interlaced erforderlich macht. Durch eine Manipulation eines Registerwerts ist bei den meisten Videografikkarten der Modus über die Software frei wählbar.

Die allgemein gültigen Beziehungen zwischen Vollbildfrequenz f_{vb} , Halbbildfrequenz f_{hb} , Gesamtzeilenzahl Z_{ges} , Zeilenfrequenz f_z , Pixelfrequenz f_p und Gesamtkolumnenzahl C_{ges} lauten:

$$f_z = Z_{ges} \cdot f_{vb} \quad (8.3)$$

$$f_{hb} = 2 \cdot f_{vb} \quad (\text{bei Interlaced}) \quad (8.4)$$

$$f_p = f_z \cdot C_{ges} \quad (8.5)$$

Beim B/G-PAL liegt ein 625/50-System vor, d. h. wir haben eine Gesamtzeilenzahl von $Z_{ges} = 625$ und eine Halbbildfrequenz (Vertikalfrequenz) von $f_{hb} = 50$ Hz. Außerdem sagt die CCIR-Empfehlung Nr. 601, dass pro Zeile 864 Abtastungen erfolgen sollen, d. h. also 864 Pixel/Zeile oder $C_{ges} = 864$. Nach den oben genannten Gleichungen ergeben sich daraus eine Zeilenfrequenz von $f_z = 15625$ Hz und ein Pixeltakt von $f_p = 13,5$ MHz.

Im Anhang sind hierzu unter Tabellen und Diagramme weitere Daten zu finden (auch für das NTSC-System). Hier sind dabei natürlich die Bruttowerte für Zeilenzahl und Pixelzahl relevant; die zum Bildspeicher und Grafikformat gehörenden Werte sind kleiner.

Da bei der Noninterlaced-Betriebsart ja die doppelte Zeilenfrequenz erforderlich ist, erhöht sich die Pixelrate ebenso dementsprechend. Ebenso kann es vorkommen, dass die Pixelzahl/Zeile erhöht ist (z. B. bei einer Bildschirmauflösung von 1024×576 ist $C_{\text{ges}} > 1100$). Damit erhöht sich natürlich nach Gl. 8.5 die Pixelrate in gleicher Weise.

Literatur

1. Holtorf, K.: Handbuch der Grafikformate. München: Franzis-Verlag 1994.
2. Wikipedia: Overscan



Videoausgabe einer Computergrafik

9

9.1 Anlegen von neuen Grafiken

Bei einer Einbindung einer Computergrafik ist es wenig sinnvoll, planlos mit der Arbeit zu beginnen. Vielmehr stehen die Aspekte Auflösung, Farbtiefe, Farbpegel und Grafik-Datenformat nach unterschiedlichen Kriterien der Weiterverarbeitung zur Disposition.

Ohne solch planvolles Vorgehen besteht immer die Gefahr, dass eine nachträgliche Bearbeitung der Grafik nötig wird, wobei keinesfalls gewährleistet ist, dass die Bildqualität aufrechtzuerhalten ist. Man sollte sich daher nie darauf verlassen, dass sich ergebende Probleme gänzlich durch eine Nachbearbeitung auszubügeln ist.

9.1.1 Auflösung

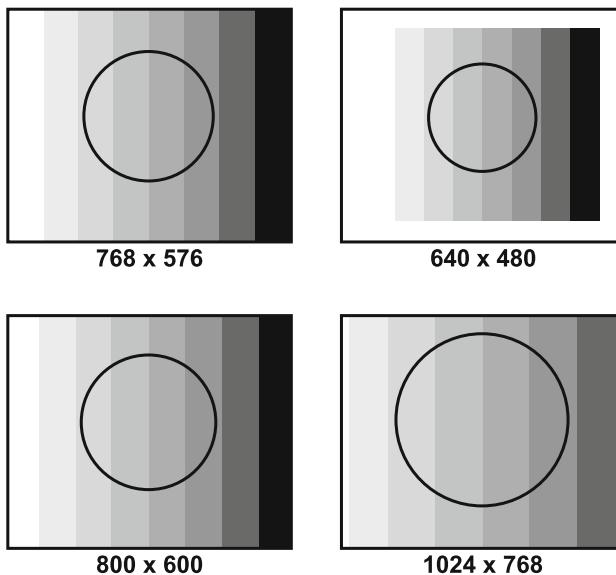
Die Auflösung einer Grafik wollen wir immer bei Vollbildwiedergabe betrachten, falls nichts anderes angegeben ist. Es handelt sich also um die Pixelzahl pro Zeile und um die Zeilenzahl.

Im letzten Kapitel haben wir den Zusammenhang zwischen den Auflösungsdaten und der Pixelform – der Aspect Ratio – kennengelernt. Aus Kompatibilitätsgründen ist immer ein Wert von 1 empfehlenswert, wodurch also eine quadratische Pixelform erreicht wird. In Tab. 9.1 sehen wir eine Auswahl verschiedener Auflösungen mit einer Aspect Ratio von 1. Es handelt sich dabei immer um Werte, die sich auf gültigen Bildinhalt beziehen.

Bei der Wahl der Grafikauflösung spielt die Auflösung der Videografikkarte eine entscheidende Rolle. Das Verhältnis des Grafikformats zur Auflösung der Grafikkarte also spiegelt die Situation bei der Wiedergabe wider. Für Videoanwendungen ist das Format 768×576 am günstigsten, weil es durch den Overscan-Modus Ränder vermeidet und dennoch alle gültigen Zeilen ausnutzt. Unter der Annahme einer solchen Auflösung

Tab. 9.1 Einige Auflösungen mit quadratischer Pixelform

Pixel pro Zeile	Zeilen	Verwendungszweck
640	480	NTSC-Format, bei PAL nur Underscan-Wiedergabe
768	576	PAL-Format Overscan
800	600	Höher auflösender VGA-Modus, unter Verwendung eines PAL-Formats erfolgt Beschniedung der Ränder
1024	768	Hochauflösender VGA-Modus, unter Verwendung eines PAL-Formats erfolgt starke Beschniedung der Ränder
1536	1152	Höchstauflösender VGA-Modus, Effekt wie oben

**Abb. 9.1** Wiedergabe einer Testgrafik der Auflösung 768×576 bei verschiedenen Bildschirmauflösungen

der Grafikkarte von 768×576 zeigt Abb. 9.1 die Bildwiedergabe bei verschiedenen Grafikauflösungen, falls der Bildaufbau konzentrisch verläuft.

Wohin der Ursprung des Bildes gelegt wird, ist jedoch von der Applikation abhängig und kann natürlich z. B. auch mit einem der bildentsprechenden Begrenzungspunkte zusammenfallen.

Die Auflösung, unter der die Grafik am Ende gesichert wird, sollte tunlichst mit der späteren Grafikkarten-Auflösung übereinstimmen, und hier ist für PAL-Systeme die Auflösung 768×576 zu bevorzugen. Unter bestimmten Bedingungen (z. B. digitale Weiterverarbeitung mit Geräten nach CCIR 610) ist auch die Auflösung von 720×576 in Betracht zu ziehen, wobei für 768×576 eine Konvertierung notwendig wird, da es sonst zu Bildverzerrungen in horizontaler Richtung kommt (siehe Abschn. 8.4).

Sollen Grafiken für Video und Druck verwendet werden, wäre die Wahl von z. B. der in beiden Richtungen doppelten Auflösung, nämlich 1536×1152 , angebracht. Eine spätere Konvertierung auf 768×576 bringt hier keine zusätzlichen Verluste, weil immer 4 Pixel zu einem Pixel zusammengefasst werden können.

9.1.2 Farbtiefe

Die Farbtiefe ist eine Größe, die ebenfalls nach verschiedenen Kriterien gewählt werden sollte. Einerseits geht sie linear in die Größe des benötigten Speichers ein, und andererseits bringt sie bei bestimmten Bildinhalten eine gewisse Kontinuität, ohne die vor allem Farbübergänge stufig erscheinen bzw. flächendiskrete Unterschiede zeigen.

Eine größere Farbtiefe einer Grafik geht meist mit einer deutlichen Geschwindigkeitseinbuße im Bildaufbau einher, auch wenn das einige neue Technologien der Grafikkarten durch eine allgemein hohe Verarbeitungsgeschwindigkeit zu kaschieren vermögen.

Bei einer Diskrepanz zwischen eingestellter Farbtiefe der Grafikkarte und der Farbtiefe der wiederzugebenden Grafik kann es u. U. ebenfalls zu einem verlangsamten Bildaufbau kommen. Solche Verzögerungen fallen z. B. dann ins Gewicht, wenn bei Präsentationen Bildwechsel vorkommen. Immer dann, wenn eine Echtzeitkonvertierung der Farbtiefe Rechenarbeit erforderlich macht, ist ein deutlicher Geschwindigkeitsabfall zu verzeichnen. Das hängt aber eigentlich nur indirekt mit der Farbtiefe zusammen, weil nach Abb. 9.2 bei direkter Farbcodierung eine hardware-nahe Konversion stattfinden kann. In diesem Beispiel gelangt ein Pixel mit 16-Bit-Farbtiefe vom Host in Speicherzellen der Grafikkarte, die für eine Farbtiefe von 24 Bit organisiert sind und umgekehrt. Im Prinzip handelt es sich lediglich um Bit-Verknüpfungen, die die Karte selbst nach Kenntnisnahme der zu übertragenden Farbtiefe herstellen kann.

Wesentlich ungünstiger verläuft die Konversion von Linearcodierung auf Palettencodierung. Die Farbwerte müssen zunächst mit den Einträgen der Palette (die ja wie eine Tabelle organisiert ist) verglichen werden; die Eintrag mit der kleinsten Diskrepanz zum tatsächlichen Farbton gelangt zur Auswahl. Obwohl es viele schnelle Algorithmen gibt, die solche Vergleichsentscheidungen bewerkstelligen, ist leicht einzusehen, dass dennoch mehr Zeit für derartige Operationen verstreicht, als dies bei direkten Bit-Verknüpfungen der Fall ist.

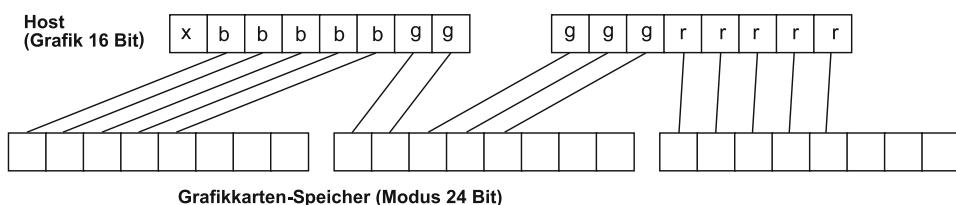


Abb. 9.2 Konvertierung von unterschiedlichen Farbtiefen bei Linearcodierung

Der Weg von Paletten- zur Linearcodierung ist wiederum einfacher, weil die Palette selbst die Übersetzung in Farbwerte vollzieht. Bei Konversionen von einer Palette zu einer anderen sind die Operationen allerdings noch komplizierter. Palettencodierung liegt normalerweise nur bei niedrigen Farbtiefen vor (z. B. 8 Bit oder 4 Bit). Während solche Farbtiefen oftmals für Computergrafiken mit einzelnen Farben völlig ausreichend sind, sollten höhere Farbtiefen mit 16 oder 24 Bit bei natürlichen Bildern oder Grafiken mit Helligkeits- Farbübergängen Verwendung finden. Eine Konvertierung auf höhere Farbtiefen ist zwar jederzeit möglich, jedoch bringt dies niemals einen Gewinn der Bildqualität – lediglich aus Kompatibilitätsgründen wird dieser Vorgang manchmal praktiziert.

Sind mehrere Grafiken beteiligt, so ist auch eine Mischung unterschiedlicher Farbtiefen denkbar; dies sollte jedoch nur bei absoluten Speicherplatzmangel in Betracht kommen, vor allem nur dann, wenn dabei keine Reduzierung der Farbtiefe notwendig ist. Ebenfalls ist zu erwägen, die Originalgrafiken mit den ursprünglichen Farbtiefen zu archivieren und nur für die Produktion eine angemessene Farbtiefenreduktion durchzuführen.

9.1.3 Farbpegel

Bei der Weiterverarbeitung und Signalwandlung von Computergrafiken in die Formate FBAS oder Y/C ist auf eine unbedingte Einhaltung von Farbpegelgrenzen zu achten. In den Abschn. 7.2.4 und 7.2.7 wurde festgestellt, dass ein Reinfarbpegel nur 75 % des Maximalwertes annehmen darf.

Solange nur mit RGB-Signalen gearbeitet wird, ist eine volle Ausschöpfung des Pegelumfangs bis 100 % möglich; Probleme entstehen erst bei der Umsetzung in Y/C oder FBAS. Das in Abb. 7.24 dargestellte FBAS-Signal beinhaltet Farbhilfsträgerschwingungen, die u. a. den Weißpegel von 100 % überschreiten würden, falls die Pegel für die Reinfarben (Farben, die maximal aus zwei Primärfarben zusammengesetzt sind) 75 % überschreiten. Betrachten wir in Abb. 9.3 z. B. ein FBAS-Signal, bei dem das Farbbalkenmuster durchweg Pegel von 100 % aufweist.

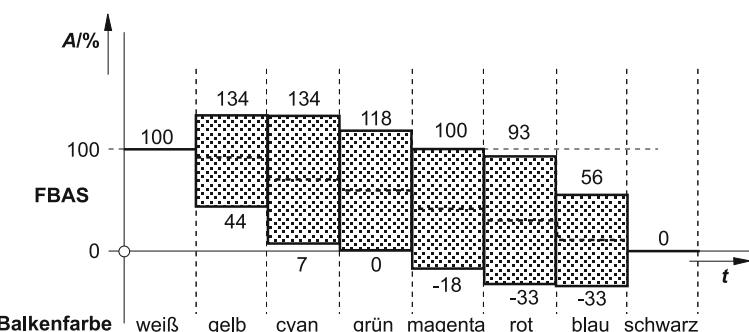


Abb. 9.3 FBAS-Signal mit zu großen Farbpegeln

Die ersten drei Farbbalken liegen mit ihren oberen Scheitelwerten deutlich über 100 %; doch auch die restlichen Farbbalken bereiten bei der weiteren Signalaufbereitung Probleme.

Um solche Übersteuerungen und die daraus resultierenden Probleme zu vermeiden, sollten die benutzten Farben entweder im Pegel oder aber im Sättigungsgrad reduziert werden.

Einige Grafikprogramme sind in der Lage, die Farbauswahl nach den oben genannten Aspekten auszurichten bzw. zu beschränken. Man sollte diese Möglichkeit nutzen, wenn die Weiterverarbeitung des Signals in Y/C oder FBAS in Betracht gezogen wird.

Sollten Grafiken mit zu hohen Farbpegeln angepasst werden, so ist auch das möglich, wie wir in Abschn. 9.2 sehen werden.

Für die gewählten Farben sollte folgende Faustformel erfüllt sein:

$$\left(\frac{-0,5R - G + 1,5B}{290} \right)^2 + \left(\frac{1,2R - G - 0,2B}{240} \right)^2 \leq 1 \quad (9.1)$$

Hierin liegen die Farbwerte R , G und B innerhalb des Bereichs 0 . . . 255, was bei vielen Grafikprogrammen üblich ist. Wird Gl. 9.1 nicht erfüllt, so können entweder die beteiligten Farbvalenzen reduziert werden, oder man mischt die Komplementärfarbe zu, also diejenige Farbe, die das Gemisch in Richtung Weiß führt.

Den Gesamtpiegel des FBAS-Signals abzusenken, ist wenig sinnvoll, weil dadurch auch die Luminanzwerte grauer Bereiche verändert werden. Bei einem Y/C-Signal dagegen wäre es denkbar, die C-Komponente abzuschwächen, wodurch jedoch auch Farbsättigungen eine Reduktion erfahren, die noch im erlaubten Bereich wären.

9.1.4 Grafik-Datenformat

Bei den heute gängigen Grafik-Datenformaten gibt es einige Ähnlichkeiten, wobei allerdings trotzdem keine Kompatibilität vorhanden ist. Grundsätzlich ist jedoch zwischen Vektor- bzw. Objekt-Grafiken und Pixel- bzw. Bitmap-Grafiken zu unterscheiden. S. 324 [1]

Zunächst spielt es signaltechnisch keine Rolle, um welchen Typ es sich bei der jeweiligen Grafik handelt. Jedoch gibt es vor allem Unterschiede in der Geschwindigkeit des Bildaufbaus und natürlich im Speicherbedarf. Doch kann in Bezug dieser beiden Kriterien auch keine kategorische Aussage getroffen werden, weil immer die Komplexität der Grafik eine große Rolle spielt.

Bis zu einem bestimmten Grade der Komplexität jedoch wird wohl eine Objektgrafik meist weniger Speicherbedarf entwickeln als eine (unkomprimierte) Pixelgrafik. Dies liegt darin begründet, dass die Pixelgrafik, selbst im Extremfalle einer leeren, weißen Fläche immer etwa soviel Speicher benötigt wie die Anzahl der Pixel multipliziert mit der Anzahl

der Bytes pro Pixel. Die Objektgrafik hingegen definiert einzelne Objekte wie Flächen, Linien usw. durch Angabe einzelner Zahlenwerte; dabei ist es völlig belanglos, wie lang z. B. eine Linie ist oder wo sie liegt – es ist immer dieselbe Anzahl von Parametern anzugeben. Natürlich wird das Grafik-Datenformat immer auch von der Applikation abhängen; so erzeugen reine Konstruktionsprogramme meist Objektgrafiken, während reine Malprogramme meist pixelorientiert sind.

Eine Konvertierung vom objekt- zum pixelorientierten Status ist durchaus möglich; der umgekehrte Weg ist jedoch nicht möglich.

Wenn es um Geschwindigkeit des Bildaufbaus geht, stellt eine pixelorientierte Grafik (ohne Kompression) meist die günstigste Voraussetzung dar, da keine oder nur geringfügige Veränderungen beim Transfer vom Rechner zum Videospeicher notwendig sind.

9.2 Veränderungen an bestehenden Grafiken

Liegen bereits Grafiken vor, die in irgendeinem Punkt nicht den gewünschten Wiedergabeparametern entsprechen, lassen sie sich mit spezieller Software verändern, wobei in manchen Fällen natürlich gewisse Verluste unvermeidbar sind. Aus diesem Grunde ist die Archivierung im Originalzustand immer empfehlenswert.

Bei den nachfolgend beschriebenen Vorgängen sollen lediglich videotechnische und verarbeitungstechnische Aspekte untersucht werden.

9.2.1 Konvertierung der Auflösung

Die wirksamen Algorithmen einer Software, die die Auflösung einer Grafik manipulieren soll, richten sich u. a. danach, ob sich auch die Zeilenzahl ändern soll oder nicht. Zunächst muss natürlich eruiert werden, welches Format die Grafik besitzt und unter welchem Bildschirmformat eine Wiedergabe erfolgen soll.

Interpolation bei konstanter Zeilenzahl Bleibt die Zeilenzahl der Grafik erhalten, so muss lediglich eine Interpolation innerhalb jeweils einer Zeile stattfinden. Es findet bei einer exklusiven Änderung der Pixelzahl pro Zeile zwangsläufig ein Wandel der Aspect Ratio statt. In Abb. 9.4 sehen wir, wie die Konvertierung einer Farbe über eine Interpolation zu realisieren ist.

Die mathematische Formulierung einer solchen Interpolation lautet:

$$Y_m = X_n + \frac{b}{a}(X_{n+1} - X_n) \quad (9.2)$$

Der Abstand b ist über eine Modulo-Berechnung zu ermitteln, wenn der Abstand a bekannt ist und der Abstand der zu berechnenden Pixel vom linken Rand der Zeile über

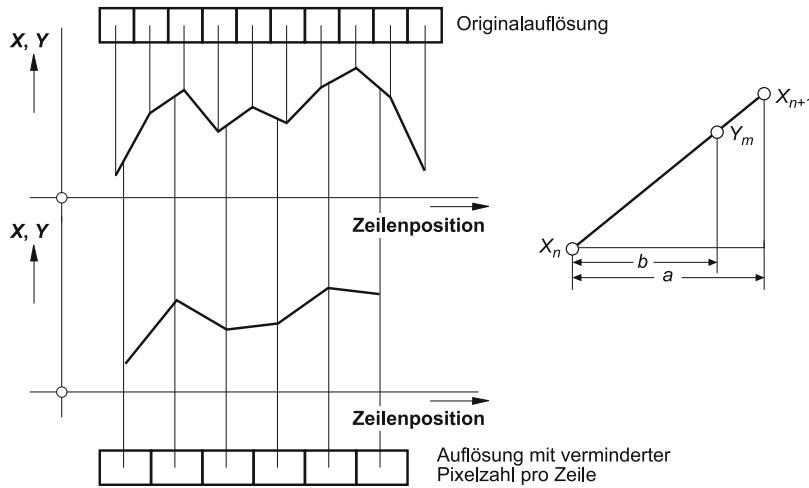


Abb. 9.4 Veränderung der Pixelzahl pro Zeile bei einer Grafik und Werteberechnung einer Farbe über Interpolation

ein Summenregister S_Y aufgefangen wird (a und b können als reine Maßzahlen betrachtet und behandelt werden):

$$b = S_Y \bmod a \quad (9.3)$$

Bei breiter werdenden Pixeln ist eine Verbesserung des Interpolationsalgorithmus dadurch zu erreichen, dass nicht nur die unmittelbar benachbarten Stützstellen zur Berechnung herangezogen werden, sondern auch die weiter entfernten Berücksichtigung finden. Die Wertung der weiter entfernten Stützstellen kann dann nach unterschiedlichen Methoden nachlassen.

Interpolation bei variierender Zeilenzahl Soll eine Grafik für Non-Video-Anwendungen an ein Videoformat angepasst werden, so ist oftmals auch die Zeilenzahl zu verändern. Das Prinzip einer Interpolation ist entsprechend zu Abb. 9.4 vorstellbar; das Werteprofil einer Primärfarbe ist diesmal jedoch dreidimensional. Das Raster der neuen Bildauflösung ist auf dieses Gebirge zu projizieren, und die gedachten Mittelpunkte der Pixel treffen auf die Niveaus, aus denen die neuen Werte gewonnen werden. Die Gl. 9.2 erweitert sich daher wie folgt zu:

$$Y_m = Y_1 + \frac{d}{c}(Y_2 - Y_1) \quad \text{mit} \quad (9.4)$$

$$Y_1 = X_n + \frac{b}{a}(X_{n+1} - X_n) \quad \text{und} \quad Y_2 = X_{n+h} + \frac{b}{a}(X_{n+h+1} - X_{n+h})$$

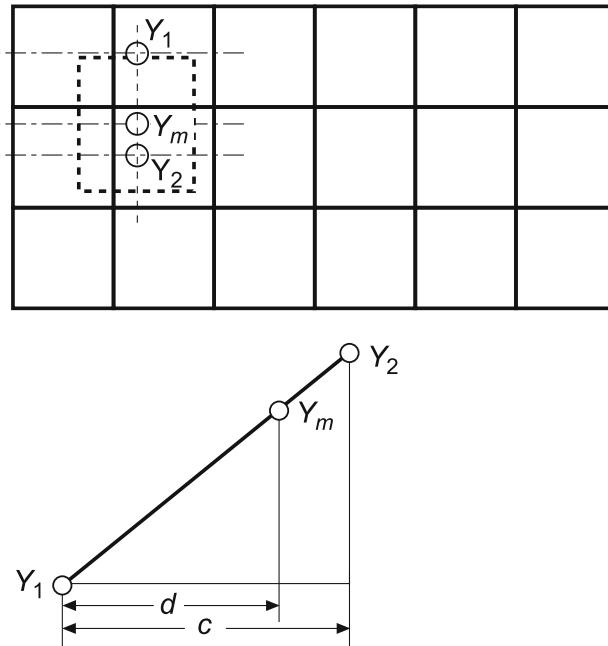


Abb. 9.5 Zweidimensionale Interpolation bei variierenden Zeilenpixeln und Zeilen

Das Verhältnis d/c entspricht hierbei dem Bruchteil der Zeilen, mit dem in vertikaler Richtung zu schreiten ist, und h ist die Kolumnenzahl. Auch bei der zweidimensionalen Interpolation besteht natürlich die Möglichkeit, weiter entfernte Pixel gewichtet mit einzubeziehen. Abb. 9.5 veranschaulicht die normale Berechnung.

Die Interpolationen vereinfachen sich, wenn die Verhältnisse von Kolumnen- und Zeilenzahl eine ganze Zahl darstellen. In diesem Falle sind lediglich die jeweiligen Pixel zusammenzufassen und das arithmetische Mittel zu bilden. Die Grenzen der neuen Pixel fallen dabei mit den Grenzen der alten Pixel zusammen.

Aliasing-Probleme bei der Interpolation Bei der Herabsetzung der Auflösung können unter gewissen Umständen Störungen bei der Wiedergabe des konvertierten Bildes entstehen, die man mit Aliasing bezeichnet. Aus feinen Pixelstrukturen, die im Verlaufe einer Zeile ein Signalprofil erzeugen, treten bei der Reduzierung der Auflösung und somit der Abtastfrequenz niedrige Störfrequenzen auf, die der Differenz von ursprünglichem und neuem Pixeltakt (bei angenommener konstanter Zeilenabtastzeit) entsprechen. Abb. 9.6 zeigt ein extremes Beispiel für eine hohe Signalfrequenz, die über die Interpolation mit niedrigerer Auflösung zur Abtastung gelangt.

Aus der sehr hohen Rasterfrequenz der ursprünglichen Grafik entsteht eine sehr niedrige Schwebung zwischen dem Originalsignal und der neuen Raster- bzw. Abtastfrequenz. Um diesem unerwünschten Effekt zu begegnen, muss vor der Interpolation – was in diesem

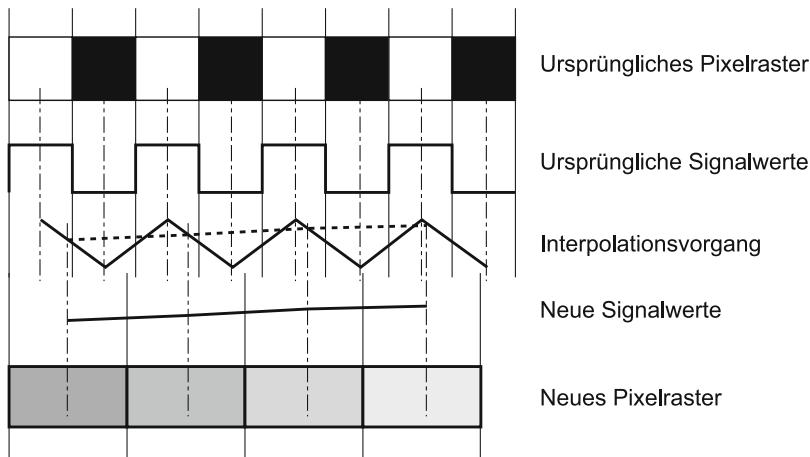


Abb. 9.6 Entstehung von Aliasing bei Herabsetzung der Horizontalauflösung

Falle einer Abtastung entspricht – eine Filterung stattfinden, die als Halbbandtiefpass für die neue Rasterfrequenz zu betrachten ist, also Frequenzen oberhalb der neuen halben Rasterfrequenz abschneidet.

Praktikabel sind hier z. B. FIR-Filter, welche jedoch von hoher Ordnung sein müssen (vgl. auch Abschn. 3.2.1), damit eine hohe Flankensteilheit erreicht wird. Es ist jedoch einleuchtend, dass in dem abgebildeten Beispiel dann nach der Filterung theoretisch keine Struktur mehr wiedergegeben wird, da die höchste Signalfrequenz der höheren, ursprünglichen Rasterung nach der Filterung ausbleibt. Das lässt sich systembedingt nicht vermeiden. Allerdings ist auch nur so eine wirksame Unterdrückung von Aliasing-Störungen zu erreichen.

9.2.2 Änderung der Farbtiefe

Eine Farbtiefenänderung geschieht normalerweise nur aus Kompatibilitäts- oder aus Gründen der Speicherplatzersparung. Auch hier ist keinesfalls eine Verbesserung der Bildqualität zu erwarten – meist tritt das Gegenteil ein.

Reduzierung der Farbtiefe Das Herabsetzen der Farbtiefe einer Grafik hat prinzipiell zur Folge, dass die Quantisierung und damit die Abstufung vergröbert wird. Helligkeits- oder Farbübergänge erscheinen nach der Reduzierung diskontinuierlich, was teilweise störend sein kann. Existieren solche Übergänge jedoch nicht, so ist z. B. eine Reduzierung von 24 Bit auf 16 Bit durchaus zu überlegen.

Bei einer linearen Farbcodierung für RGB bewirkt die Reduzierung der Farbtiefe, dass niederwertige Bits ignoriert werden. Besser ist jedoch eine Auswertung für die Rundung, d. h. das niederwertigste Bit der niedrigen Farbtiefe richtet sich nach den

unterdrückten Bits, die einem Rundungsalgorithmus zuzuführen sind. Auf diese Weise wird der maximale Fehler gegenüber einer vollständigen Ignoranz halbiert.

Steigerung der Farbtiefe Vergrößerte Farbtiefen bringen gegenüber dem Original keinen echten Informationsgewinn und somit auch keine Qualitätsverbesserung. Ein solcher Vorgang dient meist lediglich der Konformität mit anderen Grafiken oder der Kompatibilität. Begleitet wird das Ganze natürlich auch von einem größeren Speicherbedarf der Grafik.

Bei linearer Farbcodierung sind jeweils nur digitale Nullstellen an die Werte anzuhängen. Schon daran ist erkennbar, dass keine echte Verbesserung stattfindet.

9.2.3 Farbpegelanpassung

Grafiken, die ohne Rücksicht auf spätere Videoapplikationen erstellt wurden, zeigen oft zu große Farbpegel, die ein durch Farbhilfsträger codiertes Signal (FBAS oder Y/C) in puncto Amplitude zu groß werden lassen, und dadurch zu Übersteuerungen führen können.

Zur Modulation des Farbhilfsträgers werden Farbdifferenzspannungen herangezogen, die als Abszissen- bzw. Ordinatenwerte zu betrachten sind. Bei der Erzeugung von Farben mit maximal zulässigem Pegel ergibt sich in diesem Koordinatensystem näherungsweise eine Ellipse (siehe Abschn. 7.2.6). Farbvalenzen, die einen Punkt innerhalb dieser Ellipse generieren, haben das zulässige Limit noch nicht erreicht und bewirken somit auch keine Übersteuerung. Nachfolgend seien beispielhaft ein paar Algorithmen dargestellt, mit deren Hilfe sich eine Farbpegelanpassung durchführen lässt.

Wir haben eine Näherung für die entsprechende Ellipse mit Gl. 9.1 bereits kennengelernt. Eine etwas genauere Gleichung ergibt sich mit:

$$C_e = \frac{(B - Y)^2}{29404} + \frac{(R - Y)^2}{20223} - \frac{(R - Y)(B - Y) \cdot 0,00041}{(B - R)^2 + 100} \quad (9.5)$$

Für die Luminanz Y ist hierbei Gl. 7.3 zu verwenden, und alle Einzelwerte für R , G und B sind in der Skala 0...255 zu wählen. Die Ellipsenkonstante C_e ist ein Kriterium dafür, wie die Einzelwerte für die Farben verändert werden sollen.

Aus C_e wird ein Faktor k gewonnen, der die Primärfarben in Bezug auf deren Pegel reduziert. Hierfür sind verschiedene Modi denkbar, die wir hier etwas näher betrachten wollen. Der Standard-Modus sei so gestaltet, dass erst beim Überschreiten der Ellipsenlinie eine Pegelreduzierung stattfindet; auf diese Weise werden voll gesättigte Farben auf 75 % des Maximalpegels begrenzt. Im Advanced-Modus dagegen setzt bereits für $C_e \geq 0,8$ eine Pegelbegrenzung ein. Der Linear-Modus verkleinert den Multiplikator k für den gesamten Bereich von C_e linear. Der vierte Modus (S) belässt die Farbpegel und mischt lediglich die Komplementärfarbe zu, so dass eine Farbentsättigung erfolgt und auf diese Weise der Vektor innerhalb der Grenzlinie der Ellipse gehalten wird (hier erfolgt also eine Reduzierung der Farbdifferenzwerte durch k , wobei der Luminanzwert Y

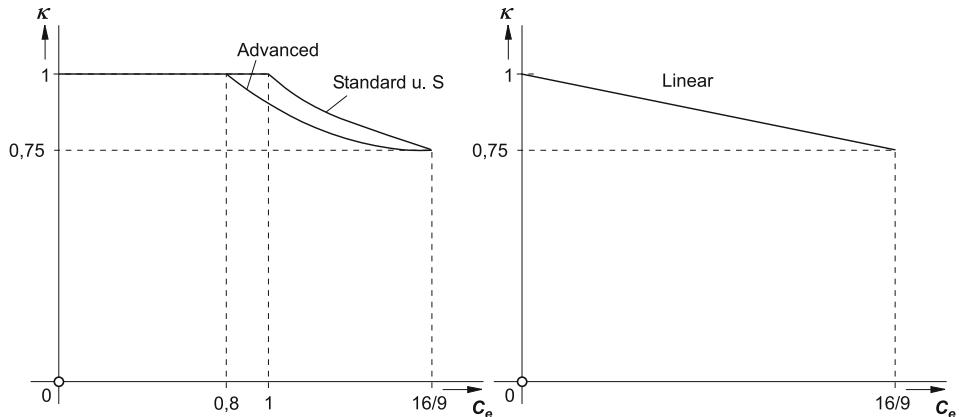


Abb. 9.7 Reduktionsfaktor in Abhängigkeit der Ellipsenkonstante

erhalten bleibt). Für die einzelnen Verfahren ergeben sich folgende Reduktionsfaktoren:

$$\begin{aligned}
 \text{Modus Standard und S: } \kappa &= \begin{cases} 1, & \text{für } C_e < 1 \\ \sqrt{\frac{1}{C_e}}, & \text{für } C_e \geq 1 \end{cases} \\
 \text{Modus Advanced: } \kappa &= \begin{cases} 1, & \text{für } C_e < 0,8 \\ 0,261 \cdot \left(\frac{16}{9} - C_e\right)^2 + 0,75 & \text{für } C_e \geq 0,8 \end{cases} \\
 \text{Modus Linear: } \kappa &= 1 - 0,1406 \cdot C_e
 \end{aligned} \tag{9.6}$$

Abb. 9.7 zeigt die Gewinnung des Reduzierungsfaktors, während Abb. 9.8 die erreichten Kennlinien einer voll gesättigten Farbe wiedergibt.

Modus **S** bewahrt die Luminanzabstufung und somit das Kontrastprofil, wobei die Entättigung der Farbe subjektiv kaum ins Gewicht fällt. Das Resultat nach der Konvertierung kommt dem Original am nächsten, ohne dass die Farbpegel ihre Grenzwerte überschreiten. Dagegen sind hier ein paar Rechenoperationen mehr notwendig, was die Konvertierungszeit der Grafik erhöht. Interpolationsalgorithmen im Zusammenhang mit angelegten Tabellen können jedoch auch hier Abhilfe schaffen. Im Standard-Modus hingegen wird ein Profil in helleren Bereichen vollständig unterdrückt, wodurch einige Nuancen verlorengehen. Der Linear-Modus verwendet einen einfachen und somit schnellen Algorithmus, wobei jedoch nachteilig ist, dass auch dunkle Bereiche einer Veränderung unterzogen werden.

Der Modus Standard lässt als einziger keine korrekte und eindeutige Umkehr der Operation zu, was natürlich ebenfalls ein gewichtiger Gesichtspunkt sein kann.

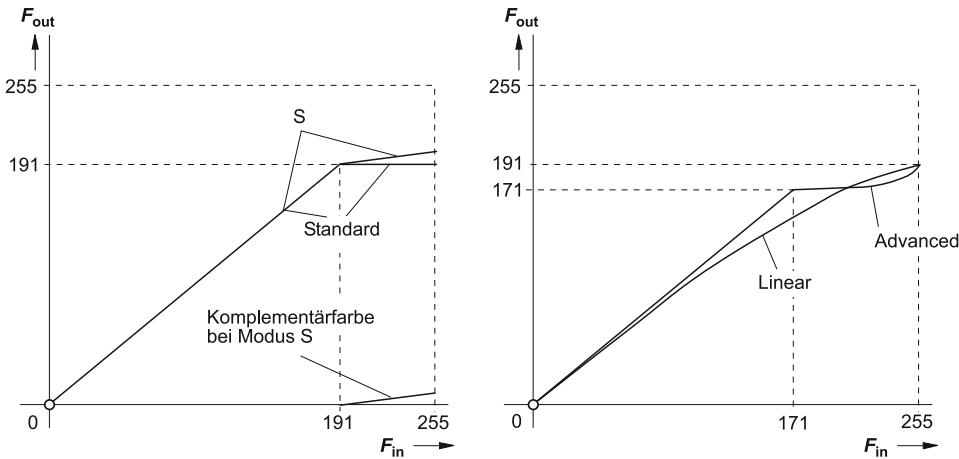


Abb. 9.8 Veränderung der Einzelfarbwerte einer voll gesättigten Farbe

9.2.4 Export in andere Grafik-Datenformate

Die Vielfalt an grafikverarbeitender Applikations-Software macht es nötig, aus Kompatibilitätsgründen die Datenformate zu übersetzen. Nur auf diese Weise ist es z. B. möglich, eine auf Software A erstellte Grafik mit der Software B auf einer bestimmten Grafikkarte wiederzugeben. Es gibt zwar inzwischen einige Formate, die recht häufig anzutreffen sind, aber dennoch fehlen hier ab und zu immer noch die Brücken.

Bei der Konvertierung in ein anderes Datenformat ändert sich normalerweise die Bildqualität nicht, sofern die Farbtiefe und die Auflösung keine Einbußen erleiden. Teilweise unterscheiden sich die Formate gar nur im Aufbau des Informationsblocks, der als Header bezeichnet wird, da er sich sozusagen am Kopf bzw. am Anfang der Datei befindet.

Generell sind gängige Formate exotischen vorzuziehen, weil hier einfach die Wahrscheinlichkeit größer ist, dass das Handling von den meisten Programmen möglich ist. Für Konvertierungen eignen sich viele Programme; auch manche Grafikprogramme ermöglichen das Sichern unter verschiedenen Formaten. Sind viele einzelne Dateien zu konvertieren, so ist die Möglichkeit einer automatisierten Konvertierung zu eruieren (z. B. mit DOS-Konvertierprogramm und Stapeldatei oder bestimmte Kommandoparameter oder auch mit Windows-Programm und Skriptprogramm).

Bei Kenntnis des Daten- und Header-Aufbaus ist auch eine Erstellung eines Programms nicht allzu schwierig. Im Anhang D werden wir etwas auf den Aufbau verschiedener Datenformate eingehen; dennoch sind Implementierungen einer Konvertierung für geläufige Formate wenig sinnvoll bzw. lohnend, da unzählige Zusatzprogramme bereits auf dem Markt erhältlich sind.

Betriebssysteme, die die Übergabe von Daten via Zwischenablage ermöglichen (z. B. Windows, OS/2, Apple usw.), verfügen sozusagen von Haus aus über ein Zwischenformat,

welches als universelles Bindeglied fungiert. Die Schnittstelle wird hier also von der Zwischenablage gebildet.

Viele Programme können zwar gängige Formate verarbeiten bzw. einführen und ausführen, konvertieren jedoch das aktuelle Material grundsätzlich in ein internes Format, welches vor allem in puncto Verarbeitungsgeschwindigkeit gegenüber den anderen Formaten Vorteile hat. Dies hat jedoch in den meisten Fällen primär etwas mit Programminternen zu tun – weniger mit generellen Nachteilen von bekannten Formaten.

9.3 Wiedergabeeigenschaften von Grafikkarten

Die Wiedergabeeigenschaften spielen eine maßgebliche Rolle bei der Beurteilung von Grafikkarten. Sie beeinflussen hauptsächlich die Grenzwerte für die Auflösung und die Farbtiefe. Die Kriterien sind außerordentlich wichtig bei der Kaufentscheidung.

Manche Grafikkarten sind in der Lage, durch den Ausbau ihres Videospeichers die Limits allgemein nach oben zu setzen.

9.3.1 Speicherorganisation

Je nach Speicherausbau sind manche Grafikkarten in der Lage, ihre Grenzen bezüglich Auflösung, Farbtiefe und auch der Seiten (Pages) zu ändern. Wir wollen uns die Speicherorganisation einer Grafikkarte anhand Abb. 9.9 etwas näher betrachten.

In diesen beiden Beispielen werden in den vorhandenen Videospeicher die gewünschten Auflösungen eingepasst; sofern das Auflösungsrechteck mehrmals hineinpasst, können entweder mehrere Seiten (Pages) oder aber eine höhere Farbtiefe angelegt werden.

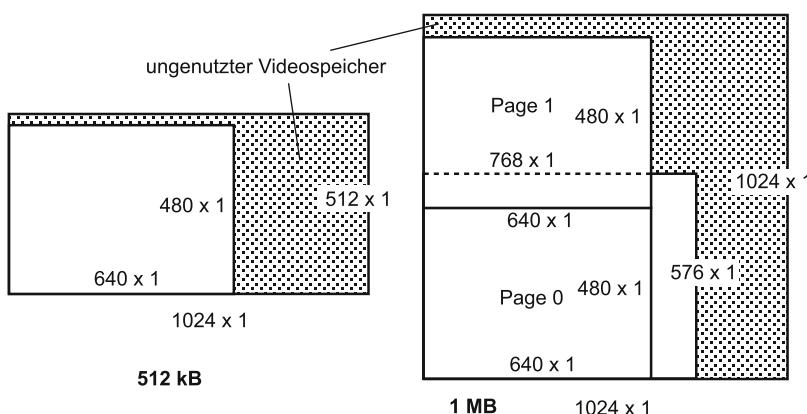


Abb. 9.9 Unterschiedliche Größen des Videospeichers und physische Anordnung im Raster

Bei einer Verdopplung der Speichergröße kann man statt $x1$ jeweils $x2$ schreiben, und die sich ergebende Farbtiefe (16 Bit) erscheint dann auch als physische Tiefe. Mehrere Bildschirmseiten können bei Videografikkarten für Spezialeffekte interessant werden, wenn z. B. auf einer Page im Verborgenen eingefroren wird, während der Inhalt der anderen Page zur Wiedergabe gelangt.

Die Darstellungen in der obigen Abbildung sind natürlich rein fiktiv; die tatsächlichen Verhältnisse einer Grafikkarte ist dem technischen Handbuch zu entnehmen. Hochwertige Ausführungen sind meist mit größerem Speicher ausgestattet, so dass grundsätzlich höhere Auflösungen und Farbtiefen erreichbar sind. Die Größe des Speichers zusammen mit der Vielfalt der sich ergebenden Aufteilungen machen den universellen Einsatz einer Videografikkarte aus. Im Hinblick auf die Steigerung des Videostandards mit HDTV und PALPlus sind immer höhere Erwartungen auch an die Grafikkarten zu stellen.

9.3.2 Signalarten der Ausgänge

Für die Aufzeichnung oder die Wiedergabe eines Bildsignals spielen natürlich auch die verfügbaren Ausgänge und deren Signalqualität eine bedeutende Rolle. Beschränkungen in Bezug auf die verfügbaren Ausgänge gehen oft mit einem Qualitätsverlust bei der endgültigen Wiedergabe oder zumindest mit einem größeren Aufwand an externen Zusatzgeräten einher. Wir werden in Abschn. 9.6 Möglichkeiten zum Test der Wiedergabequalität besprechen. Etwas mehr zu den Signalarten ist in Abschn. 7.2 zu finden.

RGB-Signal Das in Primärfarben zerlegte Signal kommt für herkömmlich Computergrafikkarten (VGA usw.) zum Einsatz, weil es eine ausgezeichnete Auflösung und Farbtreue aufweist. Für die direkte Wiedergabe an Computer-Monitoren ist das RGB-Signal unentbehrlich. Es ist vorläufig außerdem das einzige Signal, das von der Videonorm abweichende Zeilen- und Bildwechselfrequenzen sowie Auflösungen erlaubt. Wegen der somit möglichen Flimmerfreiheit ist RGB prädestiniert zum Erstellen einer Grafik.

Zur Weiterverarbeitung ist das RGB-Signal nur in manchen Fällen zu gebrauchen. Es gibt z. B. Videomischer, die auch RGB-Eingänge besitzen. Aufzeichnungsgeräte verfügen jedoch nur in Ausnahmefällen über solche Eingänge.

Alpha-(Key-)Signal Dieser Sonderausgang besitzen meist nur hochwertige Videografikkarten, die für Studiozwecke geeignet sind. Der Alpha-Speicher ist äquivalent zu den anderen Farbspeichern und kann als eigenständiges Signal zum Ausgang geführt werden. Es dient externen Kontrollfunktionen wie z. B. Stanzen für Laufschriften usw. Je nach der eingestellten Farbtiefe kann es sich um ein quasi-analoges Signal mit einer Quantisierung von 8 Bit oder ein digitales Signal mit nur zwei Zuständen handeln. Sind Key-Vorgänge beabsichtigt, müssen eine solche Fähigkeit und deren Bedingungen beim Kauf einer

Videografikkarte genau eruiert werden. In Kap. 11 werden wir Anwendungen des Alpha-Signals genauer kennenlernen.

FBAS-Signal Das für die meisten Hobbyanwendungen übliche Signal ist das FBAS-Signal. Es benötigt zur Übertragung nur eine Koaxleitung und kann bei den meisten Videorecordern als Aufzeichnungssignal dienen. Die erreichte Bildqualität unterliegt systembedingt starken Beschränkungen, wobei der von RGB auf FBAS wandelnde Encoder noch weitere Beeinträchtigungen verursachen kann. Da manche Studiobandmaschinen (z. B. Highband) nur FBAS aufzeichnen können, sollte die maximal erreichbare Qualität erzielt werden.

Y/C-Signal Zum gehobenen Hobbybereich sowie auch im Studiobereich gehört das Y/C-Signal, das landläufig auch als S-VHS-Signal bekannt ist. Die Signalqualität ist wesentlich besser als bei FBAS, weshalb das Y/C-Signal auf jeden Fall zu bevorzugen ist. Zur Übertragung sind zwei Koaxleitungen erforderlich, die jedoch meist als zweiadriges Kabel zusammengefasst sind. Auch hier können in der Signalaufbereitung der Karte Qualitätseinbußen auftreten, weil vor allem das Chromasignal nicht ganz unkritisch ist.

YUV-Signal Für professionelle Anwendungen und Aufzeichnungen (Betacam) ist das YUV- oder Komponentensignal geeignet. Zur Übertragung sind drei Koaxleitungen notwendig. Videografikkarten für den Studiobereich sollten unbedingt über YUV verfügen, obwohl auch eine externe Wandlung von RGB kaum Qualitätseinbußen bringt (Manche Grafikkarten verfügen auch über eine Umschaltmöglichkeit zwischen RGB und YUV, was dann meist über Software selektiert wird). YUV gewährleistet bei optimaler Hardware beste Bildqualität und stellt deswegen eine gute Möglichkeit dar, den Ansprüchen einer Studioproduktion zu genügen.

D1-Signal Neben den Analogsignalen sind im Studiobereich natürlich immer mehr digitale Formate gefragt. Hochwertige Videografikkarten bieten heute deshalb auch solche Ausgänge an, wobei es sich vorwiegend um das D1-Format handelt, welches ja bekanntlich ein digitalisiertes Komponentensignal (YUV) nach CCIR 601 (4:2:2-Abtastung) darstellt. Seltener sind D2-Ausführungen, welche das FBAS-Signal in die digitale Form bringen.

Es handelt sich um ein serielles Signal, das auf einer koaxialen Verbindung geführt wird. Aufzeichnungen mit entsprechenden Maschinen können verlustlos erfolgen, somit erzeugt auch das Kopieren über mehrere Generationen keine Qualitätseinbußen. Wenn man das ordnungsgemäße Arbeiten der Digitalaufbereitung voraussetzt, hängt letztlich die Bildqualität nur von den analogen Komponenten der Grafikkarte ab. Ist also das analoge YUV-Signal mit Mängeln, pflanzen sich diese naturgemäß auf das D1-Signal fort.

9.3.3 LUT-Manipulation

Zur Beeinflussung von Farben in Echtzeit dienen sog. Look-Up-Tables, über die die meisten Videografikkarten ausgangsseitig verfügen. Manchmal befinden sich die LUTs auch zusammen mit den D/A-Wandlern in einem Chip, dem sog. RAMDAC.

Es handelt sich bei einer LUT um einen Speicher, der als I/O-Tabelle fungiert, d. h. ein digitaler Eingangswert (Adresse) bewirkt einen definierten digitalen Ausgangswert (Daten). Auf diese Weise kann für jede Farbe eine Kennlinie erzeugt werden. Mit dieser Einrichtung sind nicht nur Farbkorrekturen, sondern auch Effekte wie Falschfarbenwiedergabe möglich.

Die Einträge der LUTs können üblicherweise in der Zeit der vertikalen Bildaustastlücke aktualisiert werden; somit sind auch schnelle Änderungen der Farbwiedergabe möglich. Abb. 9.10 zeigt die schematische Funktionsweise der LUTs einer Grafikkarte.

Dargestellt ist hier die Aufbereitung für nur eine Farbe; alle drei Farben durchlaufen i. Allg. solche eine LUT-Anordnung. Eine Kontrastkorrektur, die die Empfindlichkeitskennlinie von Kameras und Monitoren kompensiert, nennt man Gamma-Korrektur. Sie krümmt die drei Kennlinien gleichermaßen konvex bzw. konkav. Eine einfache Kontrastveränderung ergibt sich, wenn die Steigung aller Kennlinien einheitlich variiert. Helligkeitseinstellungen sind möglich, indem die Kennlinien einfach mit einem Offset verschoben werden. Neben all diesen Korrekturvorgängen sind aber wie bereits erwähnt auch Effekte erzielbar, wenn die Kennlinien beliebig willkürlichen Verläufen folgen. Beispiele für solche Manipulationen zeigt Abb. 9.11.

Für viele Grafikkarten gibt es Zusatzprogramme, mit deren Hilfe die LUT-Inhalte und somit die Kennlinien zu ändern sind. Interessant ist ebenfalls ein Effekt, der den LUT-Inhalt dynamisch verändert und somit eine Art Farbanimation bewerkstellt. Wir werden darauf in Abschn. 11.3 noch näher darauf zu sprechen kommen.

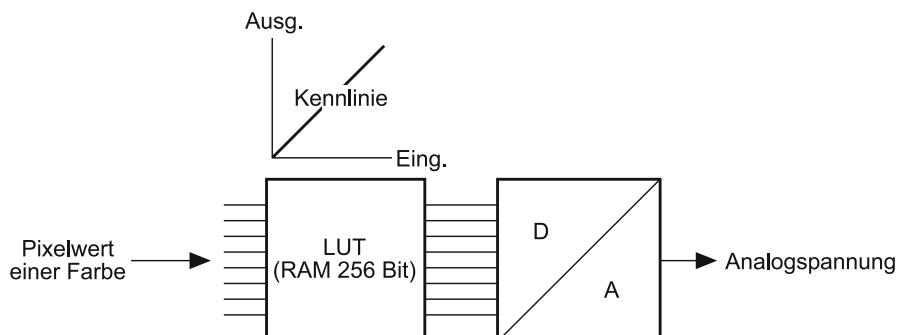


Abb. 9.10 Wirkungsweise von LUTs während der Wiedergabe einer Grafik

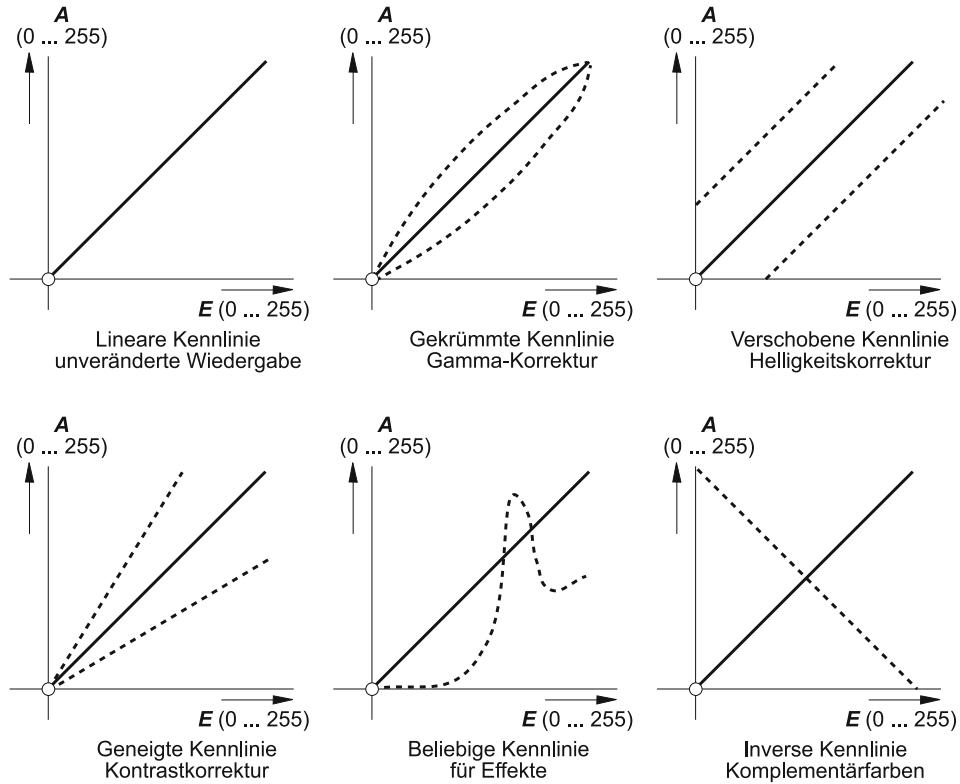


Abb. 9.11 LUT-Kennlinien und deren Auswirkungen bzw. Anwendungen

9.3.4 Durchschleifen von Signalen anderer Grafikkarten

Manche Videografikkarten verfügen über einen internen Anschluss – im PC-Bereich auch VGA-Feature-Connector genannt – zum Durchschleifen des digitalen Signals einer VGA-Karte oder anderer Grafikkarten.

Dieser Anschluss ist hierzu mit dem Pendant der VGA zu verbinden (leider verfügt nicht jede VGA-Karte über solch einen Anschluss); die sich so ergebende Doppelfunktion der Videografikkarte erübrigert die Notwendigkeit von zwei Monitoren. Die Verbindung zwischen den Karten überträgt außer dem 8 Bit breiten Digitalsignal der VGA auch noch deren Synchron- und andere Kontrollsignale. Voraussetzung für die Störungsfreie Einkopplung ist der korrekte Wiedergabemodus der Videografikkarte, der dem einer VGA-Karte entsprechen muss (was die Ablenkfrequenzen betrifft). Aus diesem Grunde ist i. Allg. kein Videomodus wählbar (landläufige Meinungen gehen oft davon aus, dass über den Feature-Connector eingekoppelte VGA-Signale plötzlich auf Videorecorder aufgezeichnet werden könnten, was natürlich nicht der Fall ist).

Über Hotkeys kann meist zwischen der Wiedergabe beider Karten gewählt werden. Auch Überblendeffekte sowie aus VGA und Videokarte physisch zusammengesetzte Bilder sind mitunter möglich.

Sofern die Videografikkarte keinen integrierten Normenwandler (von VGA nach FBAS) aufweist, ist kein ordnungsgemäßes FBAS-Signal verfügbar. Videoeigenschaften der Karte können in diesem Modus also i. Allg. nicht genutzt werden.

Leider verhalten sich VGA-Karten verschiedener Hersteller nicht einheitlich; bei Durchschleifen des Signals ist es daher mitunter schon zu erheblichen Qualitätseinbußen des VGA-Bildes gekommen, was jedoch meist nicht von der Videografikkarte verschuldet wird. Unterschiedliche Impedanzen der Daten- und Kontrollleitungen der VGA-Karten kommen hier meist als Ursache in Frage. Ist die Option des Durchschleifens zu erwägen, so kann mitunter die Dokumentation der Videografikkarte eine VGA-Referenzliste beinhalten, nach der man auf eine zumeist gründlich erprobte Konstellation zurückgreifen kann.

Eine auf diese Weise realisierte „Ein-Monitor-Lösung“ erfordert einen Monitor, der die verschiedenen Wiedergabemodi von VGA und Video ermöglicht – ein sog. Multisync-Monitor. Vorteilhaft sind hierbei Exemplare, die erstens eine automatische Umschaltung unterstützen und deren Anpassung der Bildgeometrie an die neuen Verhältnisse ebenfalls ohne besonderes Dazutun abläuft.

Mitunter kann das Durchschleifen sehr vorteilhaft sein, wenn die Arbeit am System einen ständigen Wechsel zwischen der Arbeit mit VGA und der Videografikkarte erfordert. Neben einer Hotkey-Umschaltung wären jedoch auch logische Umschaltoptionen wichtig, wo der Aufruf entsprechender Applikationen mit der Umschaltung verknüpft ist, sonst erwächst aus dem gewonnenen Komfort der „Ein-Monitor-Lösung“ eine ziemlich umständliche Handhabung.

9.3.5 Wiedergabe der Benutzeroberfläche

Neben dem Einschleifen des VGA-Signals sind natürlich auch Modussteuerungen praktikabel, welche über eine Treiber-Software die Videografikkarte nach dem Boot-Vorgang des Computers zunächst in einen Modus versetzt, in dem sie die normale Benutzeroberfläche wiedergeben kann. Hierbei ist im Prinzip zweitrangig, welcher Modus bezüglich der Wiedergabefrequenzen dies sein soll. Für grafische Arbeiten sollte man jedoch auf relativ flimmernde Videomodi verzichten.

Entscheidend für die Wiedergabe der Benutzeroberfläche ist der spezielle Treiber für die Videografikkarte, der zudem noch für den gewünschten Modus zu präparieren ist. Auch auf diese Weise ist eine „Ein-Monitor-Lösung“ zu verwirklichen, die sogar eine Standard-Grafikkarte überflüssig macht. Theoretisch kann sogar ein einfacher Videomonitor genügen, um alle Funktionen abzudecken (obwohl wie gesagt flimmerfreie Modi vorzuziehen sind, solange keine Videoanwendung gestartet ist).

9.4 Genlock-Wiedergabebetrieb

Der Genlock-Betrieb ist vor allem in Studioumgebung außerordentlich wichtig, da hier keine Unabhängigkeit zwischen den Systemen besteht, sondern eine zentrale Synchronisation aller Geräte wünschenswert ist. Die Computer-Einbindung macht es daher erforderlich, dass die Videosignale der Karten ebenfalls diesem Studiotakt gehorchen und daher Bilder im gleichen zeitlichen Raster zur Wiedergabe gelangen wie bei den restlichen Studiogeräten.

Eine Genlock-Einrichtung bei einer Videografikkarte setzt also auch voraus, dass ein entsprechender Eingang vorhanden ist, der zusammen mit der nachfolgenden Elektronik die internen Taktgeneratoren synchronisiert.

9.4.1 Synchronisation des Pixeltaktes

Der Eingang eines genlock-fähigen Gerätes erhält im allgemeinen Fall ein FBAS-Signal, aus dem Impulse gewonnen werden, die letztlich zur korrekten Abtastung des Videospeichers dienen. Ein sehr vereinfachtes Blockdiagramm für eine solche Anordnung sehen wir in Abb. 9.12.

Zunächst gelangt im ungünstigsten Falle ein FBAS-Signal auf einen Eingang der Videografikkarte. Von diesem Signal sind Horizontal- und Vertikalimpulse zu trennen. Die H-Impulse gelangen auf ein veränderbares Verzögerungsglied und dann auf einen Eingang eines Phasenkomparators. Der andere Eingang dieses Komparators erhält ebenfalls H-Impulse ähnlicher Frequenz. Die unterschiedliche und „durchlaufende“ Phasenlage beider Signale veranlasst den Phasenkomparator dazu, eine Regelspannung zu erzeugen und solange zu verändern, bis Phasenstarrheit erreicht ist. Diese gilt dann auch für den Pixelclock-Generator in Bezug auf das FBAS-Signal, jedoch erzeugt dieser das n -fache

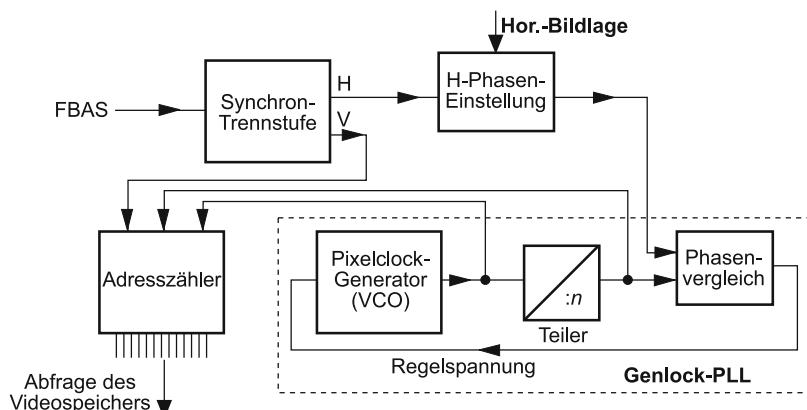


Abb. 9.12 Arbeitsweise einer Genlock-Einrichtung im Blockdiagramm

der Zeilenfrequenz, wo n die Zahl der Pixel während der Zeilenablenkung ist. Damit alleine aber können die richtigen Pixel im Speicher noch nicht angesprochen werden, weil auch hier eine zeilen- und bildentsprechende Initialisierung erforderlich ist. Auf den Adresszähler gelangen also auch die Horizontal- und die Vertikalimpulse.

Natürlich arbeitet die Bildwiedergabe auch ohne dieses angeschlossene FBAS-Signal, nur schaltet in diesem Falle der Takteingang des Adresszählers auf eine unabhängige Quarzreferenz um. Das Vorhandensein von gültigen FBAS-Impulsen am Eingang muss also über eine Kennungselektronik detektiert werden, die in obiger Abbildung nicht angedeutet ist.

9.4.2 H-Phasen-Einstellung

In Abb. 9.12 ist ein H-Phasen-Glied eingezeichnet, welches eine zeitliche Verschiebung des H-Impulses in beiden Richtungen möglich macht. Da hiervon ja auch die Pixelclock abgeleitet ist, lässt sich auch das ausgegebene Bildsignal zeitlich gegenüber dem FBAS-Signal verschieben. Betrachtet man letzteres als Zentraltakt, so ist bezogen auf diese Referenz eine Horizontalverschiebung des Bildes möglich. Zum besseren Verständnis betrachten die Signale nach Abb. 9.13.

Dargestellt ist ein ausgegebenes FBAS-Signal, wobei natürlich eine entsprechende Situation auch für RGB oder andere Signale denkbar ist.

Eine Horizontalverschiebung ist natürlich erst dann wirksam, wenn sich der Bildinhalt des Signals auf das Eingangssignal bezieht. Schließt man das Ausgangssignal inklusive Ausgangssynchronsignal hingegen komplett an einen Monitor an, so ist an diesem keine Verschiebung zu registrieren, weil die Bezugsreferenz außer Acht gelassen wird. Im Studiobetrieb besteht i. Allg. jedoch immer der Bezug zum Zentraltakt, so dass hier die H-Phasen-Einstellung Effekt zeigt.

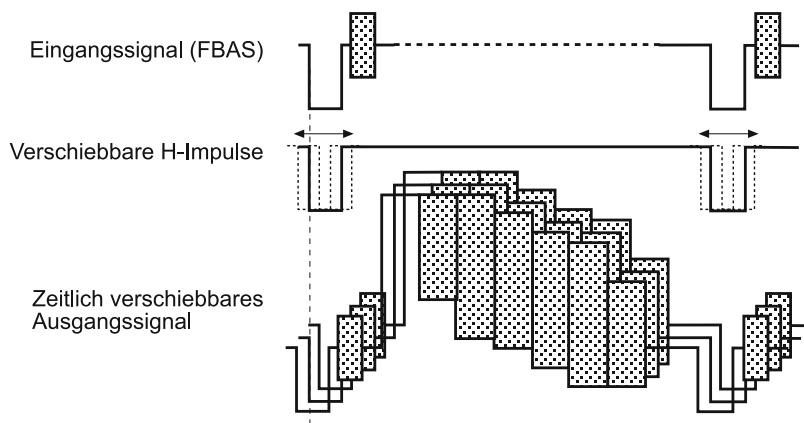


Abb. 9.13 Zeitlicher Bezug der Signale bei der H-Phasen-Einstellung (Display-Modus)

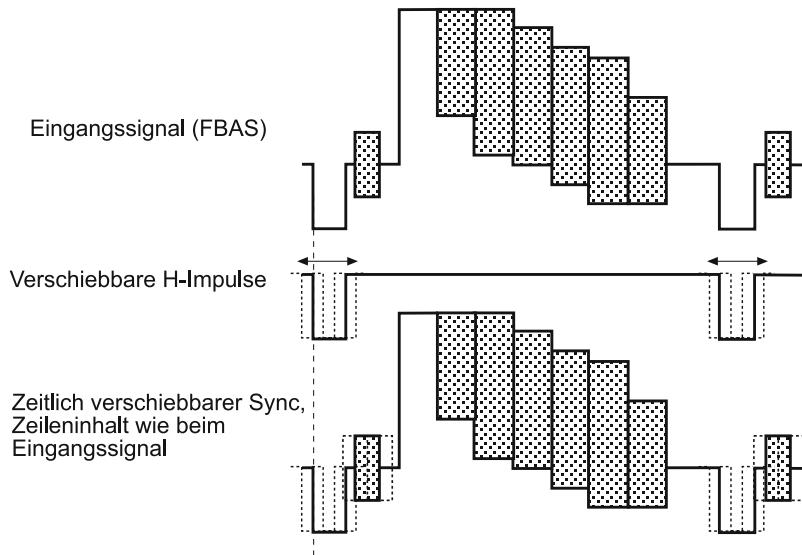


Abb. 9.14 Zeitlicher Bezug der Signale bei der H-Phasen-Einstellung (Live-Modus)

Solange die Videografikkarte das Ausgangssignal aus ihren Speicherdaten zusammensetzt (sog. Display-Modus), sind die Bildinhalte des Taktsignals von keiner Relevanz. Erst im sog. Live-Modus erfolgt ein Durchschleifen des am Eingang anliegenden Bildsignals. Da unter normalen Umständen das Live-Signal nicht beliebig zeitlich verschiebbar ist, kehrt sich der oben beschriebene Effekt um: Eine H-Phasen-Verstellung wirkt sich nur noch am ausgangsseitigen Sync-Impuls aus. Ein Monitor, der seine Synchronreferenz vom Zentraltakt erhält (also den Takt am Eingang der Karte), zeigt keinerlei Horizontalverschiebungen, während er solche sehr wohl aufweist, wenn er sich nur nach dem ausgangsseitigen Syncrontakt ausrichtet. Abb. 9.14 zeigt diese veränderte Situation.

Am Ausgangssignal ändert sich nur mehr die Sync-Lage – das Signal des Zeileninhalts gleicht (was vor allem das Timing anbelangt) dem des Eingangssignals.

Das Live-Signal kann nur über Spezialeinrichtungen (wie z. B. einem Timebase-Corrector) zeitlich verschoben werden. Zeitliche Vorverlegungen sind jedoch verständlicherweise nicht real durchzuführen, weil dies ja einer Zeitumkehr gleichkäme. Auch er kann lediglich um einen Bruchteil einer Zeile weniger als ein Vollbild verzögern, was sich in der Bildlage dennoch in einer Linksverschiebung äußert (siehe Kap. 14).

9.4.3 SC-Phasen-Einstellung

Verfügt eine Videografikkarte über einen FBAS- oder Y/C-Ausgang, so kann es wünschenswert sein, dass sich der Farbhilfsträger (SC = Subcarrier) an einer Eingangsreferenz

ausrichtet. Eine solche Eingangsreferenz ist üblicherweise ein FBAS-Signal mit schwarzem Bildinhalt – auch mit Black Burst bezeichnet.

Dieser Studiotakt synchronisiert alle Geräte auch in Bezug auf die Phasenlage des Farbhilfsträgers; mit dieser Referenz ergibt sich eine feste Relation, um das Signal zu decodieren bzw. die richtigen Farbtöne zu erhalten (Wie wir wissen, synchronisiert der immer wiederkehrende Burst einen Farbhilfsträgeroszillator, zu dessen Schwingungen die Farbartschwingungen des Zeilensignals einen Phasenunterschied aufweisen, der wiederum den Farbton festlegt.).

Da nun im Studio alle Geräte (vor allem der Mischer) mit dem Black Burst synchronisiert sind, muss auch die Videografikkarte ihr Farbhilfsträgersignal am Ausgang an dieser Referenz ausrichten. In dieser Konstellation ist es jedoch ratsam, den Phasenwinkel zwischen Eingangsreferenz (Black Burst) und Ausgangsreferenz (Burst) einstellbar zu machen. Mit der Ausgangsreferenz jedoch wandern natürlich auch die Phasenlagen der Farbinformation mit. Somit ändern sich bei einer Justage der SC-Phase die Winkel zwischen Farbvektoren und Eingangsreferenz, was letztlich einer Variation der Farbtöne gleichkommt. Abb. 9.15 zeigt diese Situation anhand der Vektorbilder für Farbbalken.

Die Phasenlage des Farbsignals bezüglich Ausgangs-Burst ändert sich analog zur H-Phaseneinstellung nicht; eine direkte Auswertung des Ausgangssignals brächte daher keinen Wandel der Farbtonwiedergabe.

Da im Studio mitunter mit sehr großen Leitungslängen zu rechnen ist und ferner einige Geräte insbesondere für den Farbhilfsträger nicht zu vernachlässigende Laufzeiten erzeugen, ist eine Einstellung der SC-Phase sehr zu empfehlen. Die Justage kann aber im Gegensatz zur H-Phasen-Justage nicht visuell vorgenommen werden, sondern bedarf einer korrekten Prozedur mit dem Vektorskop als Messmittel. S. 324 [2]

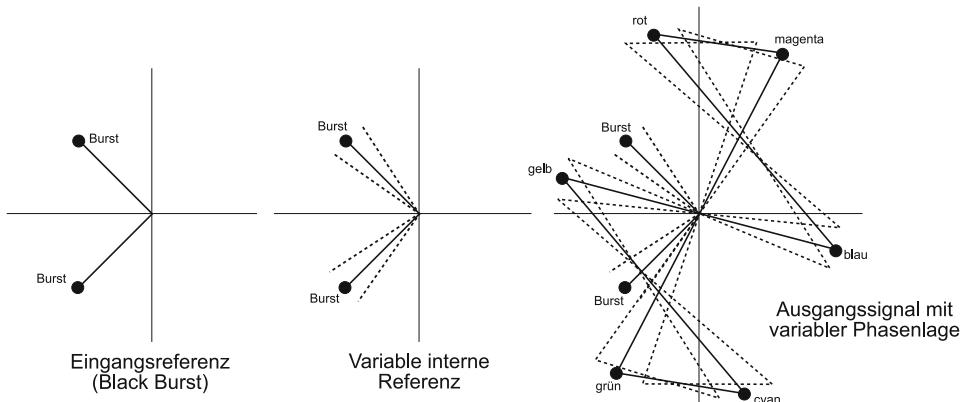


Abb. 9.15 Phasenlagen von Eingangsreferenz und Ausgangssignal bei einer SC-Phasen-Einstellung. Die Darstellung erfolgt hier als Vektorgramm, wie dies mit entsprechendem Messgerät erhalten wird. Allerdings wird eine Drehung erst dann sichtbar, wenn auch das Vektorskop als Referenz den eingangsseitigen Black Burst erhält

Besteht die Möglichkeit einer SC-Phasen-Einstellung, so ist diese entweder physisch über ein Potentiometer oder häufiger über eine Zusatz-Software vorzunehmen, wobei diese dann zweckmäßigerweise die letzte Einstellung abspeichern sollte.

9.5 Sichtbare Störungen

Treten bei der Wiedergabe einer Grafik über ein Videosystem Störungen auf, sind diese zunächst einzuordnen, bevor man sich an deren Beseitigung macht. Oftmals handelt es sich nämlich um vermeintliche technische Fehler, obgleich keine Verbesserungen zu bewerkstelligen sind.

Andererseits können natürlich auch vermeidbare Störungen vorkommen, die jedoch bei komplexen Verschaltungen häufig langwierige Fehlersuchen erfordern.

9.5.1 Systembedingte Störungseffekte

Systembedingte Störungseffekte können vom Prinzip her nicht beseitigt werden. Es bestehen jedoch mitunter Kompromisslösungen, die die Auswirkungen zumindest dämpfen können.

Interlaced-Flimmern Bereits in Abschn. 7.1 war von einem unangenehmen Effekt die Rede, der häufig dann auffällt, wenn ein Computerbild mit feinen Zeilenstrukturen über eine Videowiedergabegerät betrachtet wird. Einzelne helle Zeilen der Grafik, die von schwarzen Zeilen umgeben sind, gehören einem bestimmten Halbbild an, das sich beim Zeilensprungverfahren ja bekanntlich nur 25 mal in der Sekunde wiederholt. Die anderen Halbbilder, die zeitlich dazwischen liegen, schreiben diese helle Zeile nicht, und es erfolgt daher an dieser Stelle eine Abdunkelung. Was bei einzelnen Zeilen sehr stark auftritt, ist bei mehreren Zeilen ebenfalls – wenn auch in abgeschwächter Form – präsent. So ist besonders bei jeder ungeraden Zahl an hellen Zeilen und an den Übergangsstellen zu dunklen Zeilen immer mit derartigen Flimmereffekten zu rechnen. Abb. 9.16 verdeutlicht das Entstehen des Flimmereffekts.

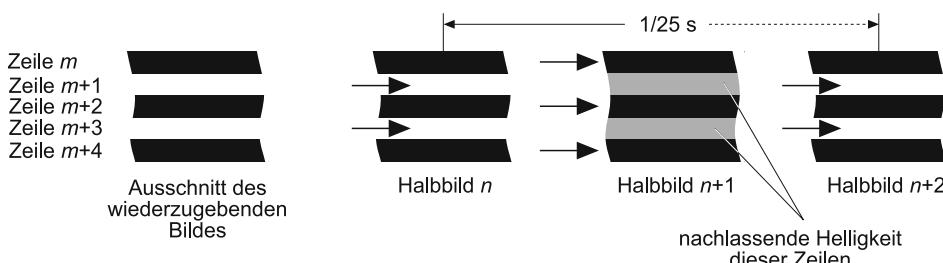


Abb. 9.16 Interlaced-Flimmern bei feinen Zeilenstrukturen

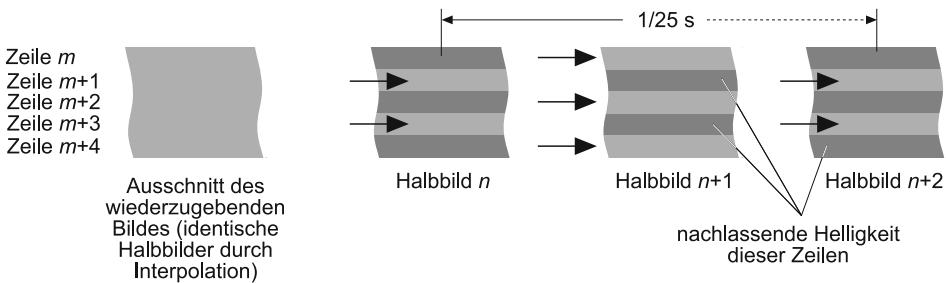


Abb. 9.17 Reduzierung des Interlaced-Flimmers durch Klonen von Halbbildern

Daraus geht hervor, dass erst nach $1/25$ s wieder der gleiche Bildeindruck entsteht. Zur Verhinderung oder Abschwächung dieses Störeffekts stehen einige Methoden zur Verfügung.

Das einfachste Verfahren, dem Interlaced- oder Halbbild-Flimmern wirkungsvoll zu begegnen, besteht in der Verdopplung eines Halbbilds oder der interpolierten Verschmelzung beider Halbbilder zu wiederum zwei Halbbildern, die jedoch identisch sind. Das Flimmern ist auf diese Weise wirkungsvoll reduziert, was jedoch auf Kosten der vertikalen Auflösung geht. Die in Abb. 9.17 dargelegte Situation hat das gleiche Ursprungsbild wie in Abb. 9.16, jedoch erfolgt hier die Wiedergabe zweier identischer Halbbilder, die durch Interpolation entstanden sind. Die Folge davon ist ein bedeutend kleinerer Wechsel des Kontrasts im Raster zwischen den einzelnen Halbbildern. Die Folge ist das Erscheinen einer homogeneren Fläche.

Eine andere Möglichkeit der Störungsminderung liegt in der Verwendung von trügerigen Leuchtschichten bei den Monitor-Bildröhren. Helle Stellen eines Halbbildes weisen so ein längeres Nachleuchten auf. Auch hier findet ein schwächerer Strukturwechsel zwischen den einzelnen Halbbildern statt. Allerdings bestehen hier keine Einbußen in der statischen Auflösung. Nachteilig wirken sich jedoch Trägheitseffekte der Leuchtschicht auf, die sich in Nachzieheffekten zeigen, die bei schnellen Bewegungsabläufen sehr störend sein können, zudem die Nachleuchtdauer für die drei Farben i. Allg. unterschiedlich ausfällt, wodurch Schmiereffekte farbig erscheinen.

Die Firma Philips hat auf der Monitorseite eine andere Technik beschritten, die ebenfalls die Trägheit ausnutzt, allerdings die des Auges, was ja den natürlichen visuellen Eindruck nicht verschlechtert. Das Bild wird dabei mit einer Halbbildfrequenz von 100 Hz (anstatt 50 Hz) abgetastet. Dies setzt natürlich eine verdoppelte Zeilenfrequenz voraus. Jedes Halbbild wird also zweimal und mit doppelter Geschwindigkeit geschrieben. Da die Bildinformation nicht in dieser zeitlichen Abfolge vorliegt, muss hier mit einem Bildzwischenspeicher gearbeitet werden. Wohlgemerkt, das Signal wird in geschilderter Weise nur für die Anzeige durch den Monitor aufbereitet.

Viele Videografikkarten arbeiten mit sog. Anti-Flicker-Filters. Die meisten dieser (teilweise optional ausgeführten) Baugruppen arbeiten über eine Interpolation zwischen den Zeileninformationen, was auch wieder zuungunsten der vertikalen Auflösung geht.

Cross-Color und Cross-Luminanz Bei Vorhandensein eines FBAS-Ausgangsignals sind Cross-Color- und Cross-Luminanz-Störungen systembedingt nicht zu vermeiden.

Cross-Color Man versteht unter Cross-Color den Effekt, welcher hochfrequente Luminanzschwingungen als Farbe erscheinen lässt. Man kann diese Störung nur dadurch abschwächen, indem bereits bei der Signalsynthese der Luminanz-Frequenzgang in der Nähe der Farbhilfsträgerfrequenz stark abfällt. Dies kann im Ausgangsschaltungskomplex einer Videografikkarte enthalten sein, was jedoch dann natürlich besagte Frequenzgangschwächen in der Luminanz-Übertragung verursacht. Deshalb ist eine solche Korrektur auch relativ selten anzutreffen.

Cross-Luminanz Sie wirkt in umgekehrter Richtung: Reguläre Farbhilfsträgerschwingungen gelangen in den Luminanzkanal und äußern sich nach der Decodierung in Form eines Perlschnurmusters. Gehorcht das FBAS-Signal einer festen Verkopplung zwischen SC- und H-Schwingungen, wie dies bei der 8er-PAL-Sequenz vorgeschrieben ist, so fällt dieses Muster bei der Wiedergabe weniger auf, weil es nicht wandert. Generell kann das Übersprechen der Farbhilfsträgerschwingung in den Luminanzkanal durch Sperrkreise bzw. Filter unterdrückt werden. Das hat jedoch zwangsläufig einen entsprechenden Einbruch im Luminanz-Frequenzgang zur Folge. Das gesamte Phänomen der Cross-Luminanz fällt jedoch in den Wirkungsbereich des Decoders, also in die eingangsseitige Signalaufbereitung und -konvertierung für FBAS in RGB. Mit einem durch eine Videografikkarte erzeugten FBAS-Ausgangssignal hat die Problematik also generell nichts zu tun – der Farbhilfsträger ist nun mal mit im Signal integriert.

Fazit Zur Überprüfung der oben genannten Störungen ist die Wiedergabe bestimmter Testbilder empfehlenswert. Man muss sich jedoch darüber im klaren sein, dass die Cross-Color-Störung auf der Kartenausgangsseite nur durch Kompromisse im Luminanz-Frequenzgang vermindert werden kann. Die Cross-Luminanz-Störung dagegen ist im Ausgangssignal (FBAS) nicht zu behandeln – dies kann erst bei der Separation in Y und C (also am Anfang der Decodierung in RGB) geschehen. Beide Störungen lassen sich nur durch Umgehung des FBAS-Signals vermeiden (also z. B. Verwendung von Y/C oder ganz ungeträgerten Signalen wie RGB, YUV usw.). S. 324 [2]

9.5.2 Bildlagefehler

Die Ausgabe einer Grafik als Videosignal kann vor allem im Genlock-Betrieb einen horizontalen Offset aufweisen, wenn das Wiedergabegerät den zentralen Synchronisationstakt erhält (siehe Abschn. 9.4).

Je nach Leitungslänge und im Signalweg zwischengeschalteten Geräten kann es zu verzögerten Signalen kommen. Einen Ausgleich kann die zeitliche Vorverlegung der Grafikausgabe gegenüber dem Referenztakt bringen.

RGB-Ausgangssignale von Videografikkarten unterliegen (direkt an der Ausgangsbuchse gemessen) aufbereitungsbedingt nur sehr geringen Verzögerungen: sie liegen i. Allg. unter der Periodendauer eines Pixeltakts (bei 13,5 MHz sind dies 74 ns) und sind daher auch kaum messbar. Dagegen entsteht bei einer Aufbereitung zum FBAS-Signal eine Verzögerung um ca. 1 μ s, die sich bei der Wiedergabe durch einen Monitor durchaus bemerkbar macht. Solange es sich nur um die Übertragung einer Grafik handelt, die vom Videospeicher der Karte stammt, kann ein Ausgleich durch Einstellung der H-Phase stattfinden.

Die H-Phasen-Einstellung geschieht entweder über Potentiometer, die meist von außen zugänglich sind, oder über eine Kontrollsoftware. Die Prüfung der korrekten Justage kann nach unterschiedlichen Kriterien erfolgen; das einfachste ist der visuelle Eindruck bei der Monitorwiedergabe.

9.5.3 Synchronisationsfehler

Bei komplexen Systemen treten immer wieder Synchronisationsstörungen auf, die sich meist auf eine falsche Signalverschaltung zurückführen lassen. Das Ausgeben einer Grafik als Videosignal setzt natürlich voraus, dass zunächst der Videomodus gewahrt ist, der verlangt wird (also z. B. 625/50 Interlaced). Zunächst ist bei solchen Störungen darauf zu achten, welcher Teil des Systems gestört wird. Vorausgesetzt ist bei den Fehlererörterungen hier immer das einwandfreie Funktionieren der Einzelkomponenten. Die Behandlung von Gerätedefekten würde den Rahmen dieses Buches bei weitem sprengen. Siehe hierzu auch Abschn. 7.1.5.

Synchronstörung bei externen Geräten Ist die Grafikwiedergabe direkt am Signalerzeuger, also der Videografikkarte, in Ordnung, so deutet dies auf eine falsche Wiedergabefreqenz hin. Sehr häufige Ursachen für Bildstandstörungen bei extern angeschlossenen Geräten sind die Modi Non-Interlaced oder auch NTSC der Grafikkarte. Messbar sind Unterschiede zum PAL-Signal sehr einfach und komfortabel, indem einem Zweikanal-Oszilloskop ein PAL-Testsignal eines Generators und das zu untersuchende Signal zugeführt wird. Nur wenn beide Signale dem 625/50-Standard entsprechen, sind die zeitlichen Ausmaße auf dem Oszilloskopschirm identisch, und das gegenläufige Wandern beider geschieht außerdem sehr langsam.

Synchronstörung bei der Videografikkarte Wird die Synchronisation des Ausgangssignals der Karte gestört, so geschieht dies meist durch das Referenzsignal, über das die Grafikkarte im Genlock-Modus (Slave-Mode) gehalten werden soll. Das Abhängen der Referenz bringt meist Klarheit darüber, ob Kartenmodus und Referenzfrequenz inkompatibel sind und die unabhängige Kartenfunktion gewährleistet ist (Master-Mode).

Entspricht der Referenztakt eindeutig den Kriterien von 625/50-Systemen, so ist mit großer Wahrscheinlichkeit der Wiedergabemodus der Karte falsch eingestellt.

9.6 Grafiken als Testbilder

Über einfache Grafiken sind Wiedergabeeigenschaften und -parameter von Grafikkarten prüfbar. Wir wollen uns kurz ein paar solcher Testbilder betrachten und erörtern, worauf es bei welcher Prüfung ankommt.

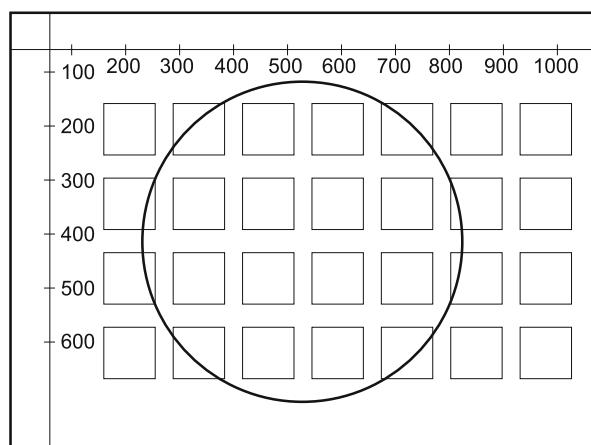
9.6.1 Prüfung der Bildgeometrie und Aspect Ratio

Bei Prüfungen der Auflösungen und der Aspect Ratio können Testgrafiken hilfreich sein, die einerseits geometrische Muster wie Quadrate oder Kreise und andererseits Kolumnen- und Zeilenskalen aufweisen. Ein solches Beispiel sehen wir in Abb. 9.18.

Die Erstellung einer solchen Grafik erfordert nicht einmal eine hochauflösenden Modus, wenn eine entsprechende Emulation mit der Grafik-Software möglich ist und die Cursor-Position angezeigt werden kann. Entsprechend dieser Position erfolgt die Beschriftung der Skalen. Die Quadrate umfassen bei einer Aspect Ratio von genau 1 (quadratische Pixel) horizontal wie vertikal gleich viele Pixel. Bei Karten mit Hardware-Zoom besteht sogar die Möglichkeit des Auszählens der Pixel in einem Quadrat.

Beim Testeinsatz einer solchen Grafik muss eine Vollbild-Darstellung erfolgen, sonst sind die Skalierungen nicht relevant; außerdem ist nicht gewährleistet, ob durch eine Fensterdarstellung keine Verzerrung stattfindet. Je nach eingestellter Auflösung der Karte werden die Ränder mehr oder weniger abgeschnitten – die Skalen zeigen somit die maximale Pixelzahl in den beiden Richtungen an. Die Rechtecke haben Quadratform, wenn die Aspect Ratio genau 1 und wenn der Monitor korrekt eingestellt ist (siehe auch Abschn. 8.2 und 8.3). Die verteilten Quadrate sowie der Kreis sind darüber hinaus geeignet, Unlinearitäten des Monitors aufzudecken.

Abb. 9.18 Testgrafik zur Prüfung der Aspect Ratio und der Bildauflösung



Nicht immer beginnt der Bildaufbau der Grafik an der linken oberen Ecke; befindet sich der Ursprung z. B. rechts unten, so tritt der Effekt der Beschneidung an der linken und oberen Bildseite auf. Der Ursprungsmodus hängt vom Programm bzw. von den Einträgen im Grafik-Header ab.

9.6.2 Prüfung der Farbwiedergabe

Zur Feststellung der Farbtiefe und eventueller Störungen der Farbwiedergabe können einfache Farbgrafiken sehr nützlich sein. Es handelt sich dabei in der Hauptsache um Verläufe zwischen maximalen Farbpegeln und Nullpegel (Schwarz). Unstetigkeiten der D/A-Wandler und der LUTs können auf diese Weise schnell und einfach aufgedeckt werden. In Abb. 9.19 ist die Anordnung solcher Farbmuster einer Testgrafik zu sehen.

Das Testmuster besteht aus vier Streifen mit den Übergängen Weiß/Schwarz, Rot/Schwarz, Grün/Schwarz und Blau/Schwarz, welche mit einem Pegel von 100 % beginnen, also nicht den Konventionen für FBAS-Signalkonvertierung entsprechen. Sie ermöglichen die Untersuchung der Kennlinien aller Farben. Kleinste senkrechte Farbstreifen im Grauübergang lassen auf Bit-Fehler oder krasse Unstetigkeiten schließen. Ist dagegen eine Abstufung mit 32 Treppen sichtbar, so liegt eine Wiedergabe-Farbtiefe von nur 16 Bit vor. Darunter befindet sich ein Streifen mit den Normfarbbalken, der ein Signal am FBAS-Ausgang nach Abb. 7.24 ergeben sollte.

Die senkrechten Farbstreifen auf der rechten Seite sollen die durchgängige Farbmodulation darstellen. Erscheinen hier (vor allem die oberen) Bereiche grau mit einer gewissen Regelmäßigkeit grau, so ist meist ein Fehler in der Burst-Generierung vorhanden, oder der Genlock-Modus ruft einen Farbsynchronisationsfehler hervor.

Das oben befindliche Streifenraster setzt bei jeweils einer Farbe nur ein Bit; Ausfälle in den D/A-Wandlern können somit leicht ausgemacht werden. Bei einer Farbtiefe von 16 Bit

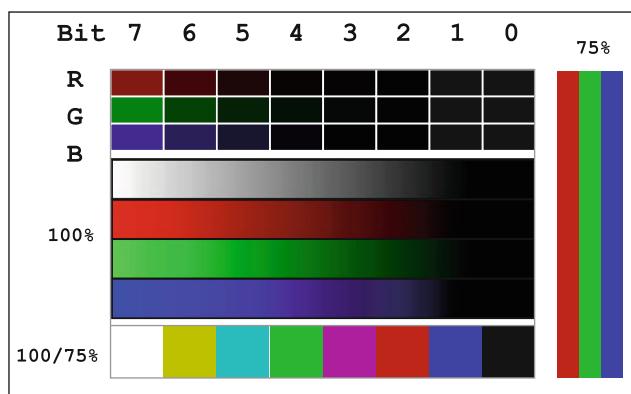


Abb. 9.19 Testgrafik zur Prüfung der Farbwiedergabe

bleiben die Felder der Bits 0...2 ständig schwarz, was jedoch nur deutlich wird bei voll aufgedrehter Helligkeit am Monitor.

Die möglichen Prüfungen erweitern sich natürlich auch auf die Funktionalität der LUTs – das oben Gesagte bezieht sich nämlich auf linear eingestellte LUTs. Invertiert man die Kennlinie aller LUTs, so ergeben sich komplementäre Verläufe (vgl. Abschn. 9.3.3). Auf diese Weise lassen sich auch die unteren Bits der sich anschließenden DACs testen, die ja sonst ein sehr dunkles Bildsignal erzeugen.

Vor jeder Produktion sollte ein derartiges Testbild geladen werden, weil die beschriebenen Fehler bzw. Einstellungen sonst nicht bei jedem beliebigen Bild auffallen.

9.6.3 Prüfung der Helligkeitsauflösung

Die Helligkeitsauflösung ist normalerweise nur ein Kriterium für die Signalqualitäten von FBAS und Y/C. RGB-Signale sollten in dieser Hinsicht keine spürbaren Mängel aufweisen.

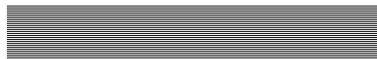
Im Vordergrund steht hier nicht die Quantisierung, die ja bereits mit einem Testbild nach Abb. 9.19 geprüft werden konnte, sondern die Wiedergabe bei hohen Frequenzen der Helligkeitsmodulation. Ein einfaches Muster nach Abb. 9.20 kann natürlich auch in andere Testbilder mit eingebunden werden. Da es sich nur um Schwarzweiß-Muster handelt, ist der Frequenzgang des Y-Signals von Bedeutung. Ein Verschwinden des Musters deutet auf einen Einbruch bei den entsprechenden Frequenzen hin.

Bei der Herstellung der Muster ist darauf zu achten, dass die Muster die entsprechende Signalfrequenz erreichen. Am besten eignet sich ein Grafikprogramm, welches im Rastermodus arbeitet. Die Muster sind dann lediglich pixelweise zu erstellen. Die senkrechten Strichkonturen ergeben bei der gewünschten Signalfrequenz f_1 und bei einer Pixelfrequenz von f_a pro Periode folgende Pixelzahl:

$$n = \frac{f_a}{f_1} \quad (9.7)$$

Bei einer Pixelabtastfrequenz von 13,5 MHz wäre so mit der feinsten Struktur (zwei Pixel, eines schwarz, eines weiß) theoretisch eine Frequenz von 6,75 MHz realisierbar.

Abb. 9.20 Prüfung des Y-Frequenzgangs und des Interlaced-Flimmers



Die waagerechten Linien, welche jeweils ein Pixel dick und ein Pixel Abstand voneinander haben, dienen der Prüfung von Interlaced-Flimmereffekten.

Literatur

1. Holtorf, K.: Handbuch der Grafikformate. München: Franzis-Verlag 1994.
2. Stotz, D.: Audio- und Videogeräte richtig einmessen und justieren.
München: Franzis-Verlag 1994.



Einfrieren von Videobildern

10

10.1 Besonderheiten der Signalaufbereitung beim Live-Modus

Wir wollen uns zunächst die grundsätzliche Signalaufbereitung in Abb. 10.1 betrachten. Die Darstellung geht davon aus, dass die Videografikkarte ein FBAS-Signal verarbeiten kann. Der TBC am Eingang ist bei weitem nicht bei allen Grabber-Karten enthalten.

Der Live-Modus stellt denjenigen Zustand dar, bei dem das eingehende Videosignal in irgendeiner Weise durch die Grabber-Karte wiedergegeben wird und noch kein Einfrieren stattgefunden hat. In diesem Stadium müssen sich daher alle Komponenten am Eingang der Karte auf das Eingangssignal einstellen (was ja nach dem Einfrieren nur sehr eingeschränkt gilt).

10.1.1 Timebase-Corrector

Der Timebase-Corrector (TBC) spielt in der Videotechnik bereits seit langem eine wichtige Rolle. Die von Videorecordern abgegebenen Signale sind aufgrund ihrer mechanischen Gleichlaufeigenschaften keinesfalls zeitlich stabil. Man kann zwar von einer langzeitstabilen Wiedergabe ausgehen, jedoch unterliegt die Abtastung der Kopftrommel zwangsläufig einem gewissen Jitter.

Der aus dem Eingangssignal genau wie im Genlock-Modus gewonnene Takt stammt von einer PLL-Schaltung (vgl. Abschn. 9.4). Relativ schnellen Änderungen kann die PLL nicht problemlos folgen, so dass es meist zu Bildverzügen kommt, wenn ein Videorecorder-Signal am Eingang der Karte liegt. Dieser Effekt verstärkt sich, wenn der Recorder selbst einen großen Regelbereich seiner Servo abverlangt (z. B. bei Bandzugproblemen usw.).

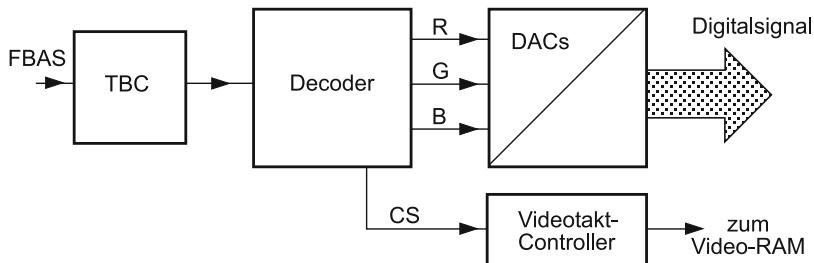
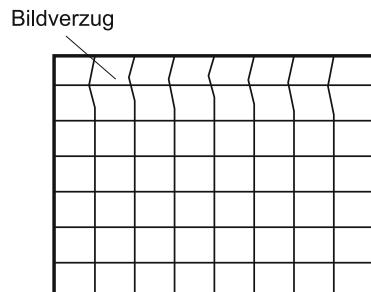


Abb. 10.1 Blockdiagramm der eingangsseitigen Signalaufbereitung

Abb. 10.2 Bildverzug beim angezeigten Live-Signal infolge der Kopfumschaltung eines Videorecorders



Der Effekt tritt meist besonders im oberen Bildbereich auf und äußert sich durch Geometriestörungen nach Abb. 10.2. Dies hängt hauptsächlich damit zusammen, dass kurz davor die Kopfumschaltung des Recorders stattgefunden hat und dort eine besonders große Diskontinuität der Synchronimpulse besteht.

Eine solche Unregelmäßigkeit kann ein TBC kompensieren; wie dies prinzipiell geschieht, ist in Abb. 10.3 zu sehen, wobei es natürlich auch hier Abwandlungen gibt. Das Signal gelangt zu einem Zwischenspeicher bzw. einem Schieberegister. Im Prinzip sind auch analoge Ausführungen denkbar, damit das FBAS-Signal direkt zwischengespeichert werden kann. Arbeitet der TBC jedoch mit digitalen Zwischenspeichern, so kann er in Abb. 10.1 auch in den DACs integriert sein. Der Schreibtakt dieses Speichers ist abhängig vom Takt des eingehenden Videosignals. Der Lesetakt hingegen ist hochstabil und entstammt einem Quarzoszillator. Das Schreiben und Auslesen geschieht bei diesem Speicher völlig unabhängig, was z. T. auch durch gesplittete Speicher mit alternierendem Schreibe- bzw. Lesevorgang realisiert wird.

Die Kapazität des Speichers kann für den oben genannten Zweck auf eine Größe von wenigen Zeileninhalten reduziert sein (ein TBC als separates Gerät kann meist ein gesamtes Bild zwischenspeichern). Da die nötige Aufbereitung im TBC verständlicherweise auch Qualitätseinbußen des Signalinhalts mit sich bringt, ist bei Grabber-Karten mit TBC-Einheit letztere auch (über Software) abschaltbar. Beim Anschluss von Kameras ist z. B. kein TBC notwendig, weil das Bildsignal hier ohnehin ein quarzstabiles Zeitverhalten aufweist. Betacam-Maschinen beinhalten einen TBC, so dass auch beim Einsatz solcher Geräte der TBC der Grabber-Karte ausgeschaltet sein sollte.

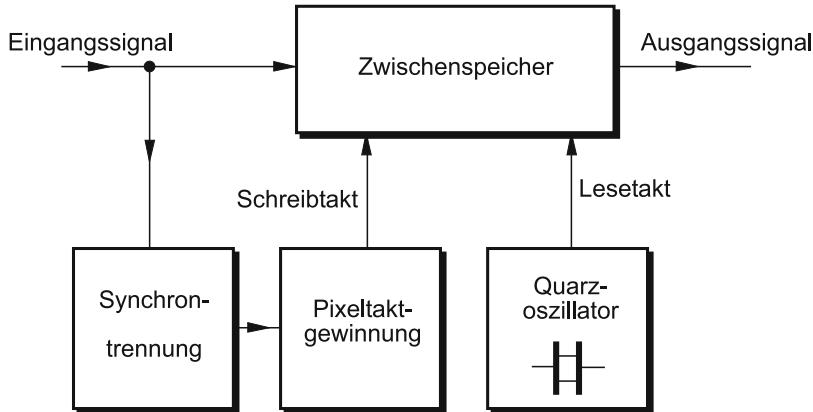


Abb. 10.3 Wirkungsweise eines TBC

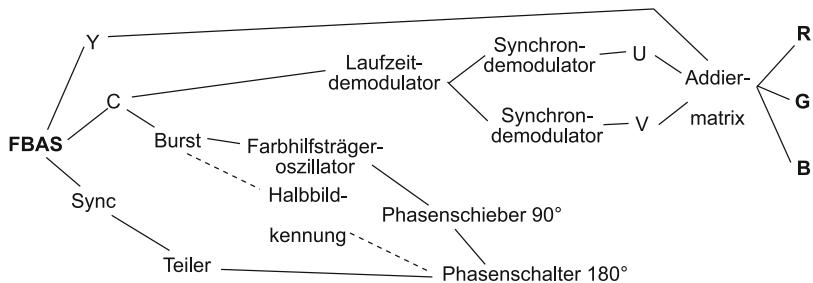


Abb. 10.4 Schematische Signalaufbereitung beim PAL-Decoder

10.1.2 Decodierung

Ein Decoder dient der Umwandlung von FBAS- oder Y/C-Signalen in die RGB-Form. Prinzipiell ist dabei der umgekehrte Weg als bei der Synthese zu beschreiben. In Abb. 10.4 ist die Aufbereitung schematisch dargestellt. S. 335 [1,2]

Die meisten Baugruppen sind bei modernen Decodern bereits integriert; als separate Bauelemente erscheinen bei Analog-Decodern vor allem die PAL-Laufzeitleitung und der Quarz für den Farbhilfsträgeroszillator.

Da analog arbeitende Decoder teilweise mit Temperaturdrift-Problemen behaftet sind, setzen Hersteller moderner Grabber-Karten gerne Digitallösungen ein, bei denen gleich am Videoeingang Digitalwandler erscheinen und der nachfolgende Aufbereitungsweg ausschließlich auf digitaler Ebene vollzogen wird. Auf diese Weise ist dann selbstverständlich auch die Realisierung des Zeilenpuffers für die Decodierung des Chromasignals möglich. (Wie wir wissen, erfolgt die Decodierung einer Zeile mit Hilfe des Chromasignals der vergangenen Zeile – dadurch erst kommt die Kompensation eines Farbfehlers bei PAL zustande. Vgl. auch Abschn. 7.2.6).

Der Decoder gestattet – unabhängig davon, ob er analog oder digital arbeitet – auch Einstellungen des Farbkontrastes, welcher häufig neben den Parametern Helligkeit und Luminanz-Kontrast mittels Kontrollprogramm beeinflusst werden kann. Mit den gewählten Einstellungen gelangt beim Einfrieren das digitalisierte Bildsignal in den Videospeicher.

10.2 Anzeige des Live-Bildes

Zur Kontrolle muss das Bild des eingehenden Videosignals zur Anzeige kommen. Prinzipiell sind hierbei zwei Möglichkeiten denkbar: die Vollbildanzeige und die Window-Anzeige. Beide haben ihre Vorteile und auch Probleme, was wir in den folgenden Abschnitten etwas erörtern wollen.

10.2.1 Vollbildanzeige

Die Vollbildanzeige ist i. Allg. statisch, also nicht skalier- und verschiebbar. Das digitale Bildsignal wird entsprechend Abb. 10.1 nach dem Passieren des sog. Videocrosspoint bzw. Mixer wieder rückgewandelt und gelangt als Videosignal zu einem Videomonitor.

Manche Videografikkarten, die auch über Videoausgänge verfügen, erlauben sogar während der Live-Wiedergabe (also vor dem Einfrieren) eine Non-Interlaced-Darstellung des Ausgangssignals. Der Videospeicher ist bei solchen Ausführungen fortwährend beschäftigt, da ja eine Zwischenspeicherung der Halbbilder erforderlich ist.

Bei einfacheren Videografikkarten gelangen jedoch während des Live-Modus keine Bildinformation zum Speicher, und das Wiedergabesignal entsteht wie oben beschrieben lediglich durch Rückwandlung. Somit entspricht es weitgehend dem Eingangssignal, was Inhalt und Zeitverhalten betrifft. Diese Art der Aufbereitung ist schematisch in Abb. 10.5 dargestellt.

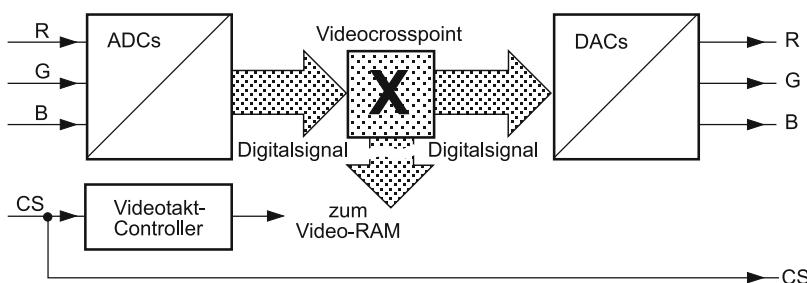


Abb. 10.5 Aufbereitung beim direkt angezeigten Live-Signal (Darstellung ohne Decoder und Encoder)

Das hier dargestellte Synchronsignal CS (Composite Sync) entspricht bei einer möglichen H-Phaseneinstellung nicht mehr dem Original, sondern hat bereits eine Aufbereitung nach Abb. 9.12 erfahren.

Eine Nutzung der Vollbildanzeige für Videozwecke ist besonders dann sinnvoll, wenn es um Stanz- und Overlay-Techniken geht. Wir kommen auf derartige Verfahren und Anwendungen in Kap. 11 ausführlicher zu sprechen. Hierbei geht es um die gemischte Wiedergabe von Live-Bild und Speicherbild.

10.2.2 Window-Anzeige

Eine etwas neuere Methode zur Anzeige des Live-Bildes ist die der Window-Anzeige, die jedoch nicht von allen Grabber-Karten unterstützt wird. Vor allem Betriebssysteme wie Windows sind für das Handling solcher skalierter Bilder geeignet, jedoch sind auch manche unabhängige Programme für eine derartige Darstellung ausgestattet.

Da das Zeitverhalten in diesem Modus nicht dem des Eingangssignals entspricht, ist ein einfaches Rückwandeln des Signals nicht ausreichend – in den Videospeicher sind ständig neue Bilder abzulegen. Es sind also kontinuierliche Einfriervorgänge notwendig, um das Bild im Window (Video-In-A-Window) in Echtzeit zu aktualisieren. Abb. 10.6 verdeutlicht die Signalaufbereitung bei der Window-Wiedergabe. Das Eingliedern des Videospeichers in das Window erfolgt über Mapping-Vorgänge, die eine spezifische Hardware der Karte voraussetzen.

Da beide Vorgänge – der Bildaufbau des Live-Bildes und der Bildaufbau des Window-Bildes – i. Allg. nicht innerhalb gleicher Zeiträume ablaufen, treten vor allem bei Bewegtbildern mit bestimmten Mustern Störeffekte zutage, denen man nur mit Kompromissen begegnen kann. Abb. 10.7 zeigt Gründe für das Entstehen solcher Effekte und deren Auswirkungen auf.

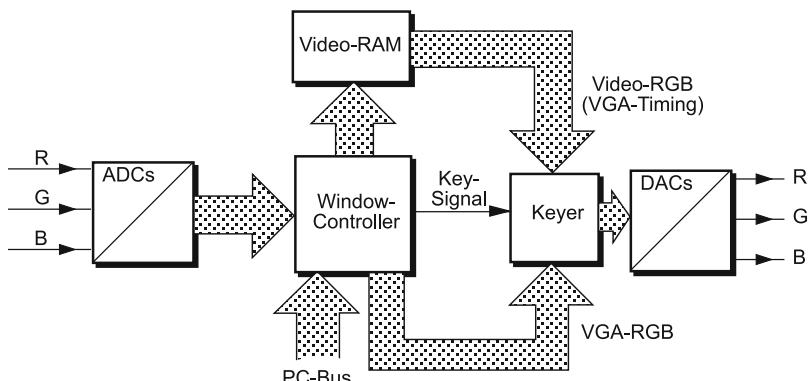


Abb. 10.6 Signalfluss bei Grabber-Karten mit Window-Wiedergabe (Video-In-A-Window)

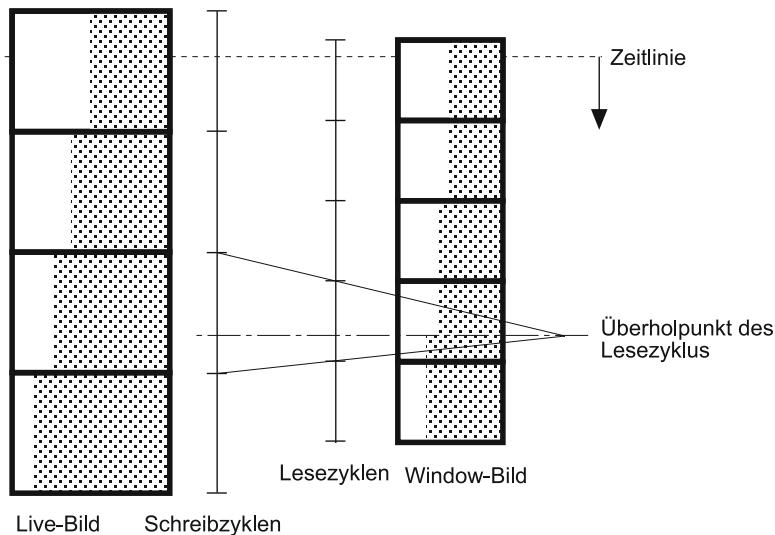
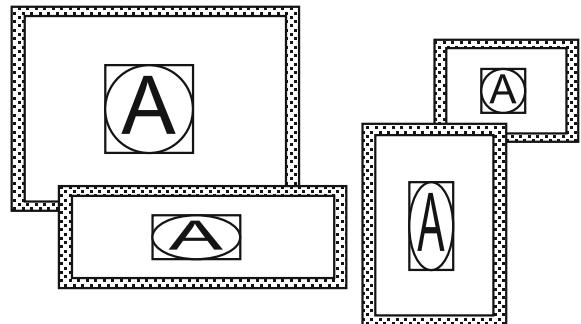


Abb. 10.7 Abtastphänomene und deren Ursache bei Window-Wiedergabe. Während die gedachte Zeitlinie über die Schreib- und Lesezyklen streicht, überholt an bestimmten Stellen die Leseposition des Window-Speichers die Schreibposition – die Folge ist das Schreiben des temporären Bildes. Die Darstellung orientiert sich an Vollbildern, obwohl ja genau genommen Halbbilder während der Abtastung in den Speicher gelangen. Findet eine schnelle horizontale Bewegung vertikaler Kanten statt, so äußern sich die zeitlichen Diskontinuitäten als besonders eklatante physische Unstetigkeiten

Abb. 10.8 Skalieren und verzerrten von Window-Live-Bildern in Echtzeit



Um das Window-Bild herum befindet sich die restliche Bedieneroberfläche des Programms oder des Betriebssystems. Die Eckpunkte sind verschiebbar, so dass das Bild einerseits frei skalierbar und andererseits sogar mit verzerrtem Maßstab wiederzugeben ist (siehe Abb. 10.8). Ein Skalier- bzw. Verzerrungsmechanismus kann über eine Abfrage-Interpolation der Speicherplätze für die Pixel geschehen. Die hierzu notwendigen Inkrementalwerte in Spalten- bzw. Zeilenrichtung werden mit Registern vom Host übergeben. Die Abfrage-Interpolation selbst wird vom Controller in Abb. 10.6 übernommen.

Es ist naheliegend, Grabber-Karten mit Window-Wiedergabe so zu konstruieren, dass mehrere Karten kaskadierbar sind. Auf diese Weise können dann auch mehrere Live-Bilder in unterschiedlichen Windows zu zeigen. Die Kaskadierung erfolgt jedoch dabei meist über interne Verbindungen, so dass man immer eine Master-Karte und mehrere Slaves hat.

Die erscheinenden Window-Bilder lassen sich natürlich auch übereinanderlegen, so dass auch eine bildmäßige Kaskadierung möglich ist.

10.2.3 Durchschleifen des Live-Bildes

Die mit Abstand einfachste Methode, an anliegendes Videosignal zu zeigen, besteht im einfachen Durchschleifen, was grundsätzlich keine markanten Signaländerungen mit sich bringt. Von dieser Option machen meist nur billige Grabber-Karten Gebrauch, weil keinerlei Aufbereitung notwendig ist. Teilweise lassen sich auch z. B. Parameter wie Helligkeit und Farbsättigung einstellen.

10.3 Vorgänge beim Einfrieren

Wir wollen bei Videografikkarten die Vorgänge zwischen Live- und Capture-Modus etwas näher betrachten. Für Karten, die Video-In-A-Window zeigen, finden ohnehin laufend Einfriervorgänge statt; dagegen gerät bei Videografikkarten mit Vollbild-Live-Video erst nach einem speziellen Befehl der Einfriermechanismus in Aktion.

10.4 Signalfluss

Signalechnische Vorgänge sind bereits in den obigen Ausführungen grob deutlich geworden. Wir wollen dennoch die Zustände am Videocrosspoint bzw. Mixer vor und nach dem Capture-Zeitpunkt noch detaillierter darstellen.

Betrachten hierzu Abb. 10.9, die beide Stati festhält. Der Videocrosspoint (VXP) kann als Datenweiche betrachtet werden. Sie führt im Live-Modus die Digitaldaten der Farbsignale einmal zu den DACs für die Rückwandlung und an das die Datenanschlüsse des Videospeichers, der jedoch noch nicht zum Schreiben bereit ist. Erst nach dem Erscheinen des Capture-Kommandos werden die Daten auch tatsächlich vom Videospeicher aufgenommen.

Unmittelbar mit dem Absetzen dieses Befehls schaltet die Weichenrichtung im VXP um; von nun an werden nur noch Daten des Videospeichers gelesen und zur Anzeige gebracht. Je nach Wahl bestimmter Parameter kann bei dieser Wiedergabe auch der Non-Interlaced-Modus aktiviert werden sowie eine Window-Wiedergabe erfolgen, was allerdings mehr oder weniger große Verzögerungszeiten für den Bildaufbau bedeuten.

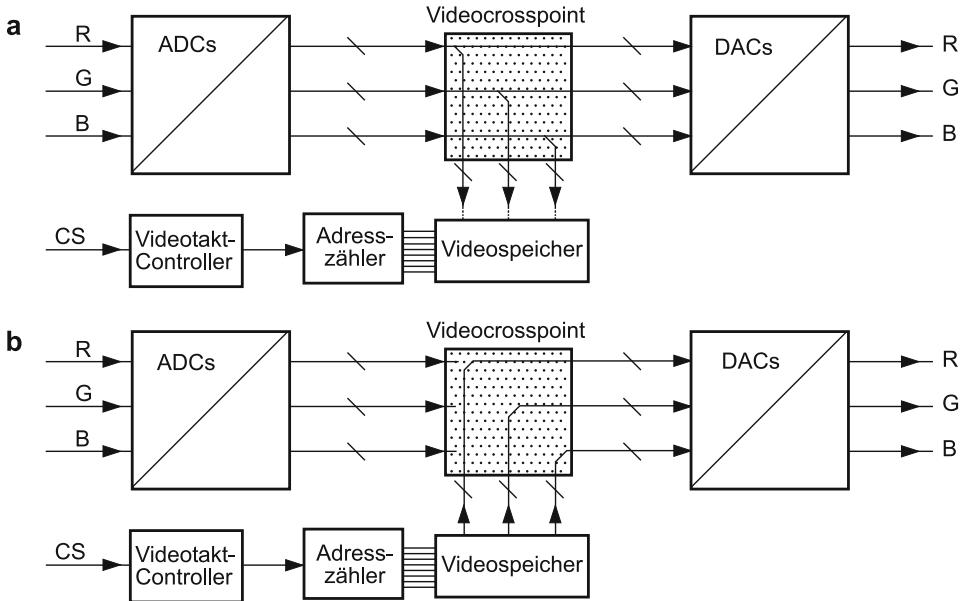


Abb. 10.9 Signalwege vor (a) und nach dem Einfrieren (b)

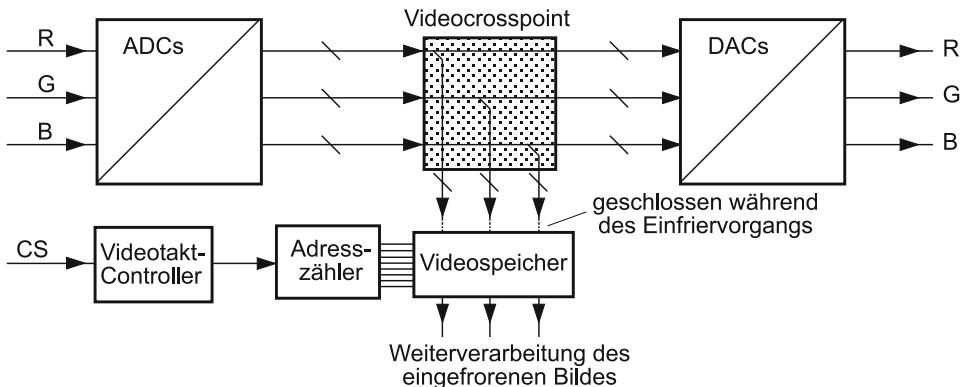


Abb. 10.10 Permanente Live-Bild-Darstellung bei verborgemem Einfrieren

Das jetzt im Videospeicher befindlichen Bilddaten können auf Festplatte gespeichert werden oder aber einer grafischen Weiterverarbeitung zugeführt werden.

Latenter Capture-Vorgang In speziellen Anwendungen kann es erwünscht sein, dass das Live-Bild zwar laufend gezeigt wird, im Hintergrund jedoch Bilddaten ständig oder nach programmiertem Muster im Videospeicher aktualisiert werden. Der Signalfluss ist in Abb. 10.10 zu sehen.

Der Durchgang von links nach rechts ist permanent vorhanden – bei jedem Einfrieren erlangt der Videospeicher für die Zeit eines Vollbildes die Bereitschaft zum Datenaufnehmen.

10.4.1 Einfrierzeitpunkt

Wir gingen bis jetzt immer vom willkürlichen Auslösen des Einfriervorgangs aus; in Wirklichkeit beginnt das Schreiben in den Videospeicher erst nach einem Halbbildwechsel.

Besteht das Live-Signal aus geschnittenen Szenen, so kann es beim Einfrieren zur Verschmelzung von Halbbildern kommen, die zwei unterschiedlichen Vollbildern angehören und zwei völlig differente Bildinhalte aufweisen. Außerdem stellt sich dabei zwangsläufig ein starkes Halbbildflimmern ein. Abb. 10.11 zeigt eine solche ungünstige Situation.

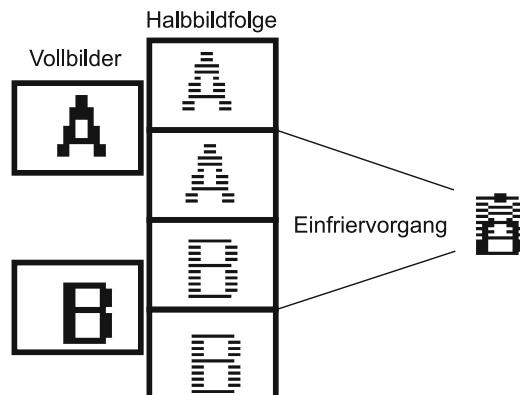
Um dies zu verhindern, kann der Einfrierzeitpunkt an den Anfang des ungeraden oder des geraden Halbbildes gelegt werden. Man geht zwar i. Allg. davon aus, dass Vollbilder keine Spaltung erfahren und somit das ungerade und das gerade Halbbild jeweils immer zusammen bleiben – somit müsste der Einfrierzeitpunkt genau am Beginn des ungeraden Halbbilds liegen. Manche Schnittsysteme erlauben jedoch auch halbbildweise Schnittvorgänge, so dass auch Trennungen der Vollbilder vorkommen können.

Die Wahl des Zeitpunkts wird über die Software bestimmt. Die signaltechnische Analyse der Vertikal-Synchronimpulse lässt eine eindeutige Identifizierung der Halbbilder zu.

10.4.2 Unterschiede zwischen Live- und Capture-Bild

Dass Live- und Capture-Bild aus ihrer zeitlichen Ausdehnung noch weitere Unterschiede zeigen, ist naheliegend. Einige dieser Effekte sind wieder systembedingt und lassen

Abb. 10.11 Einfrierzeitpunkt mit Verschmelzung von Halbbildern unterschiedlicher Szenen



sich lediglich minimieren, aber nicht vollständig beseitigen. Andere Eigenschaften sind wiederum modalbedingt und lassen sich meist individuell anpassen.

Überlagerung eines eingefrorenen Farbhilfsträgers Gelangt ein FBAS-Signal zum Einfrieren, so zeigt das Vollbildsignal, das in RGB-Form im Videospeicher liegt, Reste vom Farbhilfsträger infolge unzulänglicher Filterung und Aufbereitung in den Synchrongenmodulatoren. Wird das Speicherbild wieder als FBAS-Signal ausgegeben, so addieren sich reguläre Farbhilfsträgerschwingungen und parasitäre dergestalt, dass es zu Schwebungen mit einer Frequenz von 6,25 Hz kommt, weil der reguläre Hilfsträger nach Festlegung der 8er-Sequenz folgt, der parasitäre jedoch eine Periodizität der Dauer eines Vollbildes zeigt.

Beim Test dieses Phänomens mit Farbbalken entstehen auf dem Oszilloskop besagte Fluktuationen der Hilfsträgeramplituden, auf dem Vektorschreiber stellt sich eine mehr oder weniger schwache Rotation der Vektorendpunkte um ihre vorherigen Punkte ein. Mit dem Auge ist mitunter ein leichtes Farbflackern zu registrieren. Abb. 10.12 verdeutlicht das Entstehen dieser Schwebung.

Der oben geschilderte Effekt tritt natürlich nur bei Videografikkarten auf, die über FBAS-Eingang und -Ausgang verfügen. Das Oszilloskop bzw. Vektogramm wird unmittelbar nach dem Einfrieren unruhig. Das Ausmaß dieser Störung wird bestimmt von der Qualität und dem Abgleich des Decoders; Messungen in dieser Richtung können daher als Kriterium für den Abgleich gelten.

Reduzierte der Farbtiefe nach dem Einfrieren Videografikkarten, die im Live-Modus das Signal nach Abb. 10.5 bzw. Abb. 10.9 durchschleifen, zeigen dieses i. Allg. mit einer

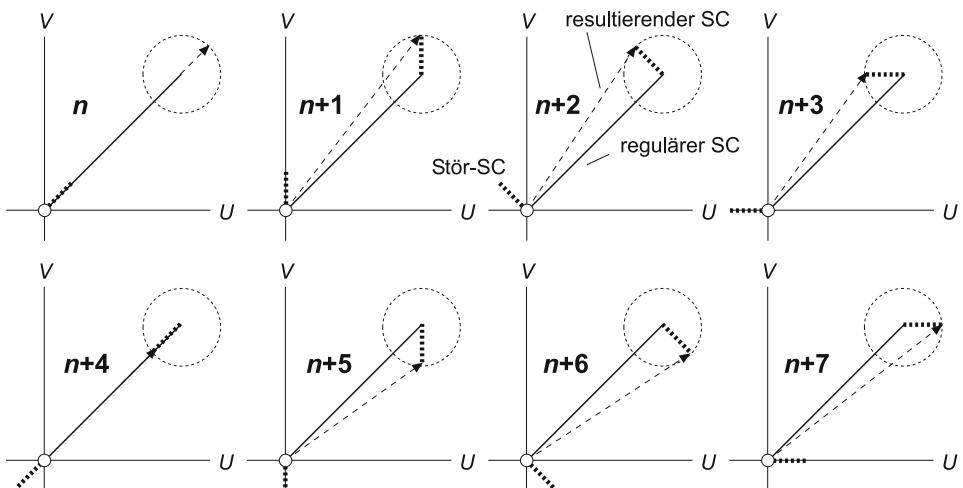


Abb. 10.12 Addition der Vektoren zwischen regulärem und parasitärem Farbhilfsträger (Restanteil vom decodierten FBAS-Signal)

Farbtiefe von 24 Bit (Truecolor). Nach dem Capture-Vorgang jedoch ist die Farbtiefe nicht mehr nur von der Auflösung der Wandler abhängig, sondern auch und vor allem vom eingestellten Farbtiefenmodus, der ja das Beschreiben des Videospeichers bestimmt.

Das eingefrorene Bild kann daher plötzlich stärkere Abstufungen und zwangsläufig auch vom vorherigen Live-Bild abweichende Farbtöne aufweisen. Beim Farbtiefen mit Palettencodierung sind Umrechnungen nötig, was sich durch deutlichen Zeitbedarf zwischen Einfrierzeitpunkt und Bilddarstellung äußert.

Sonstige Wiedergabeparameter Bei manchen Videografikkarten lässt sich für die Wiedergabe des Capture-Bildes eine automatische Umschaltung in den Non-Interlaced-Modus erzwingen. Ferner ist auch das Platzieren in ein Window möglich, von wo aus auch unmittelbar Ausschnitte wählbar und andere Bearbeitungen möglich sind.

An der Bildauflösung bzw. -größe hingegen ändert meist nichts, weil die Bildschirmtreiber neben der Auflösung selbst auch die Blanking-Zeiten festlegen, die dann auch für das Live-Bild gültig sind.

Literatur

1. Stotz, D.: Audio- und Videogeräte richtig einmessen und justieren. München: Franzis-Verlag 1994.
2. Häberle, H.: Elektronik, Teil 3: Nachrichtenelektronik. Wuppertal: Europa-Lehrmittel-Verlag 1980.



11.1 Overlay-Technik

Grafikkarten mit Videoeingängen und -ausgängen haben mitunter außerdem dem üblichen Bildspeicher auch noch einen Overlay-Speicher oder Alpha-Speicher. Dieser ist bei einer 32-Bit-Farbtiefe den Farbspeichern völlig äquivalent: es stehen 8 Bit pro Pixel für Attribut-Zuordnungen zur Verfügung. Doch auch bei einer Farbtiefe von 16 Bit steht 1 Bit pro Pixel für Overlay-Zwecke zur Verfügung.

Eines dieser Attribute kann die Überlagerung einer Computergrafik sein. Bei der Abfrage des Videospeichers entscheidet der Alpha-Wert darüber, ob das betreffende Pixel das Live-Bild oder die Computergrafik wiedergeben soll. Unter bestimmten Voraussetzungen können sogar „weiche“ Übergänge erzielt werden.

11.1.1 Overlay bei einer Farbtiefe von 16 Bit

Wie bereits erwähnt, verbleibt bei einer Farbtiefe von 16 Bit nur 1 Bit für den Overlay-Speicher. Demnach können die Pixel in puncto Overlay-Zustand jeweils nur zwei Zustände annehmen: Grafik- oder Live-Bild-Wiedergabe an der betreffenden Stelle des Pixels. Jedem Grafikpixel haftet also ein logisches Attribut an, das darüber entscheidet, ob es selbst oder aber ein Bildpunkt des Live-Video zur weiteren Verarbeitung und schließlich zur Anzeige kommen soll. Das Ergebnis dieser Methode geht aus Abb. 11.1 hervor. Die Videografikkarte befindet sich verständlicherweise im Interlaced-Modus, da ohnehin nur in diesem Modus eine videoteknische Nutzung des Overlay-Effekts sinnvoll ist.

Für alle drei Farbkanäle sind somit digitale Schalter vorhanden, die zwischen Grafik- und Live-Bild umschalten, je nachdem, ob das Alpha-Bit gesetzt ist oder nicht. Diese Schalter müssen jedoch in diesem Falle 3×5 Bit schalten können. Außerdem ist natürlich

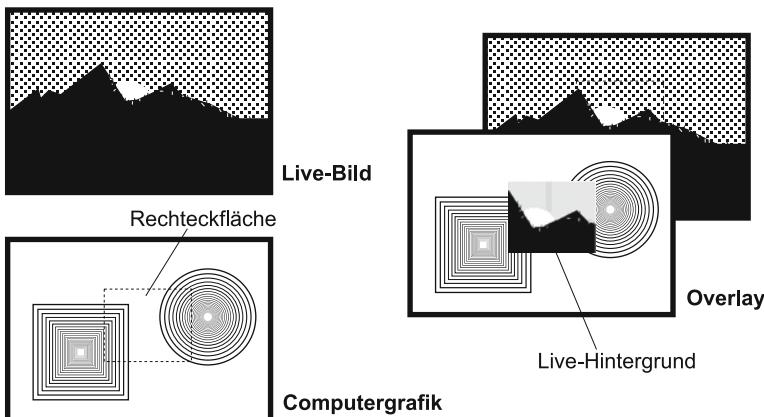


Abb. 11.1 Wirkungsweise beim Overlay

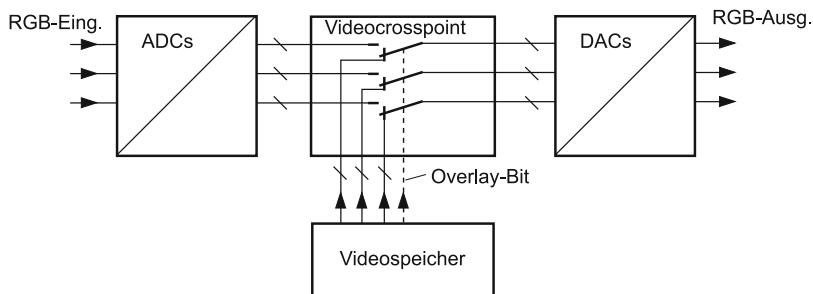


Abb. 11.2 Schematische Arbeitsweise der 1-Bit-Overlay-Einrichtung

eine enorme Schaltgeschwindigkeit vonnöten. Abb. 11.2 zeigt eine solche Anordnung schematisch.

Als Schalter dient meist der Videocrosspoint selbst, der über den Status des Overlay-Bit seine Schalterstellungen ändert.

Manche Karten verfügen über Mischer-Schaltkreise, die selbstverständlich auch für die oben genannten, trivialen Schaltfunktionen geeignet sind und darüber hinaus in Bezug auf Overlay noch weitere Fähigkeiten besitzen, wie wir weiter unten noch sehen werden.

Das 1-Bit-Overlay ermöglicht ein „hartes“ Schalten, das auf ein einzelnes Pixel beschränkt sein kann. Derartige scharfe Schaltkanten sind in der Videotechnik nicht erwünscht, weil es dadurch oft zu Aliasing-Störungen oder auch Schattenbildungen kommen kann.

Zeigegeräte wie Maus usw. können die Ebene des Alpha-Speichers beeinflussen, ohne dass dadurch am restlichen Bild etwas verändert wird. Dieses „Malen“ im Alpha-Speicher zeigt in Echtzeit Wirkung, indem sich die Bereiche der Transparenz variieren.

11.1.2 Overlay bei einer Farbtiefe von 32 Bit

Bei der höheren Farbtiefe von 32 Bit fallen dem Alpha-Speicher die gleiche Anzahl Bits zu wie den drei Farben, nämlich 8 Bit. Damit können ebenfalls 256 Werte realisiert werden.

Es liegt nun der Schluss nahe, man könne durch diese Dynamik des Alpha-Kanals nun kontinuierliche Übergänge zwischen Grafik und Live-Bild erzeugen, was ja in der Videotechnik meist sehr erwünscht ist, weil dadurch Aliasing-Störungen reduziert werden. Doch mit einem VXP nach Abb. 11.2 sind lediglich Schaltfunktionen realisierbar, so dass ein internes Overlay mit Antialiasing nicht möglich ist.

Externes Keying Eine 8-Bit-Tiefe für den Overlay-Kanal ist dennoch sinnvoll, wenn der Inhalt dieses Speichers als Analogausgangssignal zur Verfügung steht. Wie eine solche externe Steuerung des Overlay-Effekts aussehen kann, ist in Abb. 11.3 zu sehen.

Auffällig ist bei dieser Konstellation das Fehlen des Live-Signals im Grafikkartenbereich. Bei Studioanwendungen versucht man auch, solche Signalwege zu vermeiden. Statt dessen kommt von der Grafikkarte lediglich das Signal der Grafik sowie das Maskensignal (Key), mit dem die Grafik auf das Live-Signal gelegt werden soll. Die zeitliche Deckung zwischen beiden Signalen muss natürlich gewährleistet sein, weshalb das Keying mit dem RGB-Signal der Karte gegenüber einem Keying mit dem FBAS-Signal zu bevorzugen ist.

Nach welchem Algorithmus die Flanken des Key-Signals zu formen sind, legt meist eine Software fest, die dann den Alpha-Speicher einer entsprechenden Interpolation unterzieht.

Anwendungsgebiete für das oben genannte Verfahren sind z. B. Schriftgeneratoren, mit deren Hilfe Videotitel als Laufschrift einem bestehenden Videosignal zumischbar sind. Wir kommen auf solche Techniken noch in Abschn. 12.6 zu sprechen.

Trotz der hier beschriebenen Möglichkeiten, die natürlich wiederum von der Ausstattung der Videografikkarte abhängen, kann oft zusätzlich ein 1-Bit-Overlay mit internem Keying realisiert werden; als relevantes Bit gilt hier meist das MSB des Alpha-Speichers.

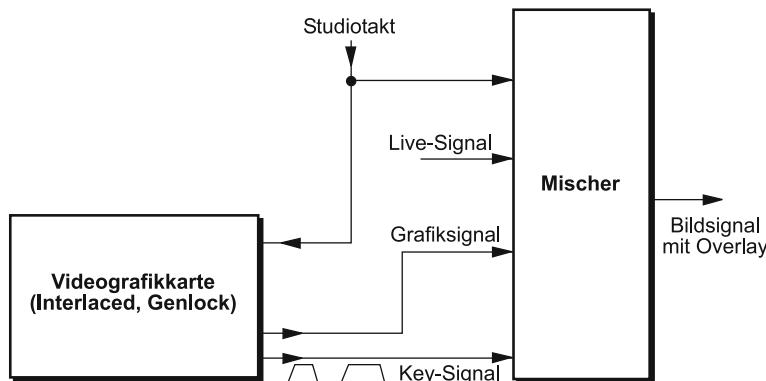
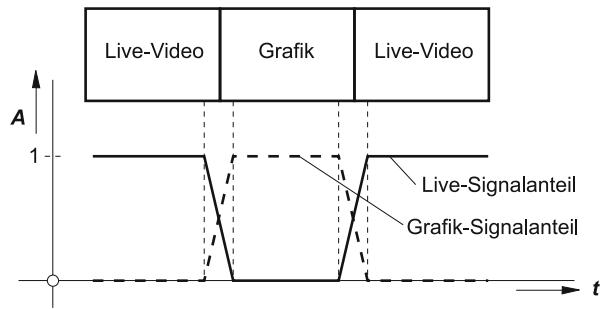


Abb. 11.3 Overlay mit externem, analogen Keying zur Erzeugung von „weichen“ Kanten

Abb. 11.4 Verlauf der Signalanteile beim Überstreichen von Overlay-Grenzbereichen



Internes Keying mit Mischer-Bausteinen Der stetige Übergang zwischen Live- und Grafiksignal war in der Konstellation nach Abb. 11.3 nicht möglich, da der Videocrosspoint nur als Vielfachschalter fungieren kann, nicht aber ein Mischen der digitalen Signale bewerkstelligen kann. Der Mischvorgang an den Kanten bewirkt eine Anteilsverschiebung der Signale nach Abb. 11.4.

Normalerweise korrespondiert ein hoher Alpha-Wert mit der Transparenz für das Live-Video, während ein Wert von null nur noch die Grafik erscheinen lässt. Der Alpha-Wert ist also ein direkt proportionales Maß zur Intensität des Live-Signals.

Für beide Signale findet an den Übergangszenen also ein schnelles Überblenden statt. Die Steuerung dieses Überblendens übernimmt wieder der Alpha-Speicher, der ja simultan zu den Farbspeichern abgefragt wird.

Für jedes Pixel erfolgen mehrere arithmetische Operationen (einfache, logische Operationen genügen hier nicht), die wie folgt aussehen können:

$$R = F \cdot \bar{\alpha} + L \cdot \alpha \quad (11.1)$$

R resultierender Pixelfarbwert, F Pixelwert der Farbe, α Pixelwert des Alpha-Speichers, L Live-Signalwert.

Natürlich muss für R noch eine Anpassung erfolgen, die jedoch über ein Auffangregister und einfache Bit-Verschiebung geschehen kann. Diese Operation, die für jedes Pixel und jede der drei Farben getrennt durchzuführen ist, erfordert hochintegrierte Bausteine.

Das Live-Signal muss in die Videografikkarte eingespeist werden, was für Studioanwendungen nicht sehr günstig ist; leider ist dies jedoch der einzige Weg, wenn das Key-Signal des Alpha-Speichers nicht am Ausgang verfügbar ist.

11.1.3 Kriterien für das Keying

Wir haben gesehen, dass das Keying durch den Status der einzelnen Pixelwerte des Alpha-Kanals bestimmt wird. Wodurch jedoch die Alpha-Komponenten definiert sind, ist noch

völlig offen. Man unterscheidet dabei vor allem zwei Kategorien: das örtliche und das inhaltliche Kriterium.

Örtliches Keying In Abb. 11.1 ist ein Rechteckfenster definiert, in welchem Live-Video erscheinen soll. Die Fläche, die auch wesentlich komplexer sein darf, ist eine örtliche Festlegung für die sich ergebende Transparenz. Die Vorgabe der Flächen kann dabei über Koordinaten oder auch durch ein Zeigegerät (z. B. Maus) erfolgen. Man kann Titelgeneratoren auch zu Anwendungen des örtlichen Keying zählen, wo der Alpha-Speicher laufend den geometrischen Begrenzungen des Textes angepasst wird.

Inhaltliches Keying Es ist ebenfalls möglich, den Alpha-Speicher aufgrund des Inhalts des Farbspeichers zu verändern. So kann z. B. ein bestimmter Farbwert der Grafik festgelegt werden, für den die betreffenden Pixel nur noch Live-Video zeigen. Daneben ist natürlich auch die Definition von Farbwertbereichen möglich. Die Manipulation des Alpha-Speichers aufgrund der Key-Vorgabe geschieht i. Allg. auf Software-Ebene, deshalb ist eine kurze Berechnungszeit unvermeidbar. Je nach Komplexität der Grafik entsteht ein mehr oder weniger bizarres Gebilde, in dem nun das Live-Bild erscheint.

11.2 Chromakeying

Ein komplementärer Effekt zum Overlay ist das Chromakeying, welches in der modernen Videotechnik nicht mehr wegzudenken ist. Vereinfacht ausgedrückt befindet sich hier die Computergrafik im Hintergrund, während im Vordergrund das Live-Signal überdeckend wirkt. Manche Videografikkarten sind ebenfalls in der Lage, Chromakeying durchzuführen; eine solche Anordnung zeigt Abb. 11.5.

Das Live-Signal, das – um dem Namen auch gerecht zu werden – in der Studiotechnik meist von einer Kamera stammt, nimmt Objekte auf wie bisher. Hinter den Objekten befindet sich jedoch eine monochrome (meist blaue) Fläche, die auch Stanzebene genannt wird. Die Ausmaße dieser Stanzebene kann auf kleinere Flächen begrenzt sein; sind große Kulissenbilder erwünscht, kann diese Fläche jedoch auch den gesamten Hintergrund des Studios ausfüllen.

Die Videografikkarte registriert das Stanzsignal dadurch, dass sie die Farbe mit einem vorgegebenem Farbwert vergleicht. Ist die Abweichung nur geringfügig, schaltet sie an dieser Stelle der Bildabtastung unmittelbar das Grafikbild ein. Objekte, die sich vor dieser (blauen) Stanzfläche befinden und keine Stanzfarbe aufweisen, gelangen als Live-Signal ungehindert durch die Karte hindurch.

Damit das Chromakeying ordnungsgemäß arbeiten kann, sind einige beleuchtungstechnische Voraussetzungen zu erfüllen. Die Stanzfläche muss den richtigen Farnton besitzen und muss gleichmäßig und kräftig ausgeleuchtet sein, so dass keinerlei Schattenbildung besteht. Objekte, die sich von der Stanzfläche abheben sollen (Sprecher usw.) dürfen keine Farben beinhalten, die der Stanzfarbe nahekommen. Schließlich sind gewisse Justagen

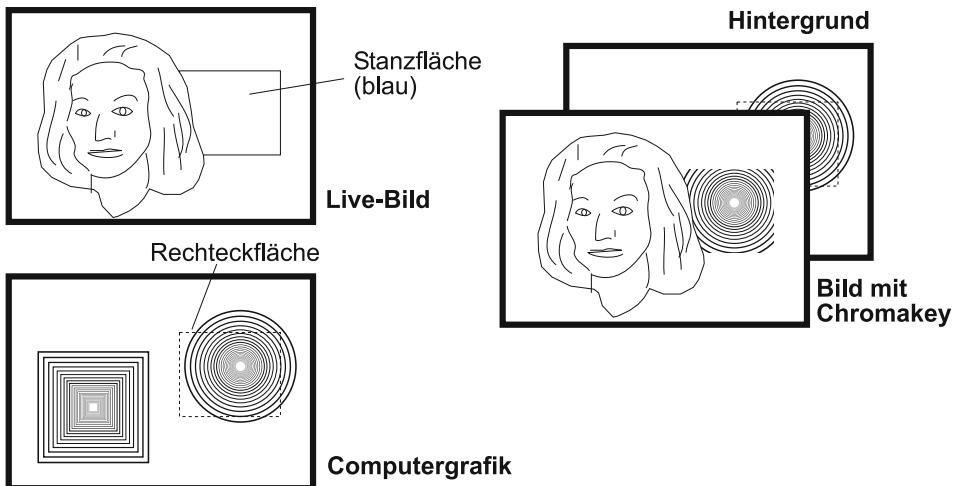


Abb. 11.5 Wirkungsweise des Chromakeyings

für die Karte zu absolvieren, die aber meist allesamt auf Software-Ebene ablaufen. Für Kontrollzwecke kann oftmals das in ein Schwarzweißbild konvertierte Stanzsignal zur Anzeige gebracht werden. Die Blaufläche muss darin gut weiß sein, während andere Bildbereiche möglichst dunkel sein sollten. Irgendwo knapp unterhalb dieses Weißwerts ist ein Schwellenwert einzustellen, auf den der Chromakeyer ansprechen soll.

Das Chromakeyings ist z. B. mit einem Mischer-Baustein oder einem Videocrosspoint-Baustein nach Abb. 11.2 durchzuführen, wo aber im Gegensatz zum Overlay-Effekt der Schaltvorgang nicht vom Alpha-Speicher, sondern von einem Farbdetektor im Live-Signalkreis initiiert wird. Auch hier ist wegen des Echtzeitprozesses ein schneller Videoschalter (für jede Farbe 8 Leitungen) erforderlich, wobei natürlich ebenfalls weiche Übergänge erwünscht sind (sonst entstehen bei senkrechten Begrenzungen eines Objekts im Live-Raum ausgefranste Stellen).

Als Stanzfarbe wird meist Blau gewählt, da sie in Körperfarben kaum vorkommt. Manche Karten mit dieser Einrichtung lassen jedoch auch eine alternative Farbwahl zu.

Anwendungen für Chromakeyings sind z. B. Wetterberichte, wo der Sprecher vor einer blauen Fläche steht, in die die Bilder der Wetterkarte eingestanzt und über eine Computeranimation oftmals auch in Bewegung gebracht werden. Doch auch beliebige Videotricks oder auch nur Kulissenbilder sind über das Chromakeyings realisierbar.

11.3 LUT-Effekte

Bereits in Abschn. 9.3.3 haben wir festgestellt, dass man die Farbwiedergabe einer Videografikkarte über LUTs beeinflussen kann, indem auf digitaler Seite einfach eine Manipulation der Farbkennlinien stattfindet.

Dieser Effekt ist je nach Ausführung der Karte sowohl für eingehende Live-Signale, als auch für wiedergegebene Grafiken möglich. LUT-Schaltkreise können daher sowohl nach den eingangsseitigen A/D-Wandlern (ADCs), als auch vor den ausgangsseitigen D/A-Wandlern (DACs) vorkommen.

Während Input-LUTs lediglich die Wiedergabe des Live-Bildes und das eingefrorene Bild verändern, beeinflussen sie die bloße Grafikwiedergabe nicht. Andererseits nehmen Output-LUTs bei der Wiedergabe immer einen Einfluss, der sich jedoch nicht auf die Daten eines gespeicherten Bildes niederschlägt.

LUTs stellen kleine Speicher (256 Bit) dar, deren Adressanschlüsse den Eingängen und deren Datenanschlüsse den Ausgängen entsprechen. Zur Veränderung von Kennlinien muss in diese Speicher geschrieben werden. Bei linearen LUTs entspricht der Ausgangswert immer dem Eingangswert; in einem solchen Falle kann die LUT auch mit einem Bypass überbrückt werden, ohne dass sich etwas ändert.

Für den Ausgleich von Kamera- und Monitorkennlinien sind LUTs i. Allg. nur einmal zu initialisieren. Sollen dagegen Effekte realisiert werden, so sind LUTs auch für einen dynamischen Einsatz geeignet, der ein ständiges Manipulieren des Speicherinhalts erforderlich macht. Als Zeitpunkt für das Beschreiben des LUT-Speichers dient üblicherweise die vertikale Bildaustastlücke, so dass eine Kennlinienveränderung sich erst beim neuen Halbbild bemerkbar macht.

Im Gegensatz zu Software-Algorithmen können LUTs in Echtzeit agieren – wenn auch mit sehr einfachen Möglichkeiten. Helligkeitsänderungen können beispielsweise so gestaltet werden, dass nicht der gesamte Bereich des Bildes einen gleichmäßigen Wechsel erfährt, sondern z. B. nur die hellen Flächen. Eine solche Manipulation würde Krümmungen oder Knicke in den Kennlinien erfordern. Andere Effekte, z. B. der der Solarisation sind ebenfalls möglich, indem bestimmte Peaks in den LUTs zu setzen sind.

Eine gezielte Veränderung der Bildwiedergabe durch Manipulation der LUT-Kennlinien erfordert jedoch einige Erfahrung, sonst stellen sich lediglich unkontrollierte Effekte ein.

11.4 Mischen und Überblenden

Die Effekte Mischen und Überblenden sind in der Videotechnik ebenfalls außerordentlich verbreitet – zum großen Teil fallen sie schon überhaupt nicht mehr auf. Um diese Dinge auf Videografikkarten zu bewerkstelligen, sind i. Allg. spezielle Mischer-Bausteine erforderlich, die wir bereits im Zusammenhang mit Overlay und Chromakeyung weiter oben erwähnt haben.

11.4.1 Homogene Blenden

Üblich sind Blenden zwischen Live-Bild und Grafik; das Mischen zwischen mehreren Eingangssignalen würde u. U. einen höheren Aufwand bedeuten, weil für FBAS bzw. Y/C separate Decoder erforderlich wären. Können dagegen in einem Grafikmodus zwei Seiten adressiert werden, so erlauben Videografikkarten mit Mischer-Bausteinen meist auch ein Überblenden zwischen zwei Grafiken.

Der Mischer-Baustein führt arithmetische Operationen für jedes laufende Pixel durch. Während diese Operationen beim Overlay-Effekt (siehe Abschn. 11.1) durch die einzelnen Pixelwerte des Alpha-Speichers gesteuert werden, geschieht dies beim der Verwendung als Überblendeinheit durch den globalen Wert k , die für das gesamte Bild gilt und der bei der Blende allmählich variiert. Gl. 11.1 ändert sich demnach zu:

$$R = F \cdot \bar{k} + L \cdot k \quad (11.2)$$

Bei konstantem k ergeben sich einheitliche Komponentenanteile über das gesamte Bild. Üblicherweise gibt es – äquivalent zur Overlay-Steuerung – 256 Stufen beim Überblenden, was zu einem akzeptablen Ergebnis führt, das sich wenig von kontinuierlichen Blenden unterscheidet.

11.4.2 Strukturblenden

Neben solchen bildhomogenen Blenden sind natürlich auch oft Strukturblenden erwünscht, die geometrische Begrenzungen zwischen den beiden zu mischenden Bildern erzeugt.

Formen Es gibt sehr viele unterschiedliche Formen von Strukturblenden; recht bekannte Gruppen führen die Bezeichnungen Wipe, Slide, Tile, Strip und Random. Abb. 11.6 zeigt einige Beispiele dieser Überblendformen.

In dieser Aufstellung bildet der Slide-Effekt eine Ausnahme, weil nur bei ihm der Inhalt von Bild **B** einer Bewegungsoperation unterliegt. Üblicherweise bedient man sich jedoch nicht einer Bewegung der Speicherinhalte, sondern gestaltet den Slide-Effekt durch eine Umlegung der Abfragegrenzen des Speichers, was natürlich voraussetzt, dass Bild **B** ein Speicher-Bild ist.

Arbeitsweise Alle anderen Strukturblenden reagieren über einen Key-Vorgang, der von einem separaten Effektbild (erzeugt von speziellen Effekt-Bausteinen im Zusammenspiel mit einem digitalen Videoschalter wie z. B. Crosspoint oder Mischer) gesteuert wird. Zu ersetzen ist hier also das digitale Alpha-Signal durch ein Effektsignal, das in Echtzeit konstruiert wird.

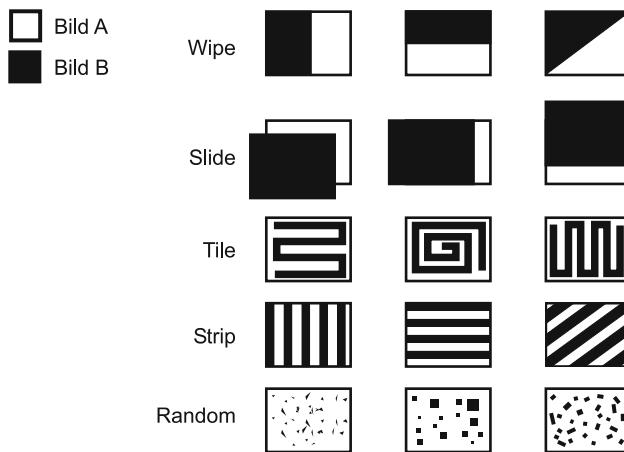


Abb. 11.6 Formen von Strukturblanden

Denkbar ist selbstverständlich auch eine Software-Überblendung, die mit einem Algorithmus die spezifischen Formen entwickeln und diese als Rasterbild dem Alpha-Kanal übergeben kann. Von dort aus betrachtet handelt es sich dann nur noch um ein Keying, das am Anfang des Kapitels erörtert wurde. Das Problem bei dieser Methode liegt jedoch häufig in der mangelhaften Geschwindigkeit. Eine zufriedenstellende Steigerung gelingt meist erst beim Einsatz spezieller Grafikprozessoren auf der Videografikkarte selbst, so dass die Rechner-CPU entlastet ist und sich anderen Dingen widmen kann.

11.5 Filtertechniken

Unterwirft man Pixelfarbwerte bestimmten Algorithmen, die keine Translation, Rotation oder sonstige Deformationen des Bildinhalts bewirken, so spricht man allgemein von Filterung.

11.5.1 Punktoperationen

Einfache Filteralgorithmen, die sich auf jedes einzelne Pixel gleichartig auswirken, haben wir mit den LUTs in Abschn. 11.3 und bei der Farbpegeleinstellung in Abschn. 9.2.3 kennengelernt. Solche punktdiskreten Vorgänge nennt man auch Punktoperationen (obwohl natürlich alle Bildpunkte behandelt werden). Mit LUT-Manipulationen lassen sich durchaus vielfältige Effekte realisieren, die optischen Farbfilterungen gleichkommen, wobei letztere sich sogar auf lineare Kennlinien der Primärfarben zurückführen lassen.

11.5.2 Ortsbereichsoperationen

Wesentlich interessanter und ergiebiger sind jedoch Filtertechniken, die das Bild als ein- oder zweidimensionales Signal betrachten, welches dann letztlich der spezifischen Filterung unterworfen wird. Es handelt sich bei derartigen Vorgängen um ein Filtern im Ortsbereich. Wir werden hier ein paar Filteralgorithmen und deren Auswirkungen betrachten; für tiefergehende Ausführungen sei auf spezielle Literatur verwiesen. S. 350 [1]

Zu unterscheiden sind Filtervorgänge auf Software- und auf Hardware-Basis. Grundsätzlich sind zwar alle Algorithmen auf Hardware-Basis und somit in Echtzeit (während des Schreibens eines Bildes) realisierbar. Bei der Vielzahl von unterschiedlichen Varianten kommt dies aus Aufwandsgründen jedoch nicht generell in Frage.

Man kann jedoch universelle FIR- oder IIR-Filter (Finite Impulse Response bzw. Infinite Impulse Response, siehe auch Abschn. 3.2) einsetzen, die immer von derselben Hardware-Verschaltung ausgehen und ihre Filtereigenschaften durch spezifische Filterkoeffizienten erreichen. Letztere sind nur vor der Operation einmal zu berechnen und verändern sich während des Filtervorgangs nicht mehr. Auf diese Weise sind jederzeit Hochpässe und Tiefpässe zu realisieren, ohne die Hardware verändern zu müssen.

Zweidimensionale Filterung Es ist einleuchtend, dass es nicht genügt, die zeilenweise abgetasteten Pixelwerte dem Filteralgorithmus zuzuführen, weil das Bild ja auch entlang einer Spalte ein Helligkeits- bzw. Farbprofil zeigt. Infolgedessen ist auch eine spaltenweise Filterung notwendig. Für eine FIR-Filterung ergibt sich der einfache Algorithmus zu:

$$\begin{aligned} P_{col} &= \sum_{i=0}^N \alpha_i \cdot P'_{col-i} \quad \text{für jede Zeile} \\ P_{lin} &= \sum_{i=0}^N \alpha_i \cdot P'_{lin-i} \quad \text{für jede Spalte} \end{aligned} \tag{11.3}$$

P Pixelwert (Farb- oder Luminanzwert), P' ehemaliger Pixelwert, col Spaltenindex, lin Zeilenindex, α Filterkoeffizient, N Filterordnung.

Alle Punkte jeder Zeile bzw. alle Punkte jeder Spalte sind chronologisch hiernach zu bestimmen; außerhalb liegende, virtuelle Punkte (mit negativem Index) sind i. Allg. auf den Wert zwischen Weiß und Schwarz festzulegen (Grau mit Halbbereichswert).

Verstärkung von Bildkanten Über einen sehr einfachen Rechenvorgang lassen sich Hochpässeigenschaften realisieren, welcher nur Bildübergänge zum Vorschein bringt, gleichförmige Flächen hingegen schwarz erscheinen lässt. Die Pixelwerte errechnen sich hier aus:

$$P = \frac{P' - \alpha \cdot \bar{P}}{1 - \alpha} \quad (11.4)$$

P Pixelwert (Farb- oder Luminanzwert), P' ehemaliger Pixelwert, α Verstärkungsfaktor, \bar{P} Durchschnittswert aus umgebendem Feld (z. B. 3×3 oder 5×5).

Für jeden Punkt ist zunächst der Durchschnitt aus einem Raster zu ermitteln, das ihn konzentrisch umgibt. Vereinfachend ist hier ein quadratisches Feld gewählt; idealer wäre natürlich ein rundes Feld. Außerdem ist es ratsam, für die Ermittlung der neuen Werte ständig die Originaldaten zu holen und nicht etwa manipulierte Daten (im Gegensatz zum FIR-Filter).

Das oben genannte Schema ist als Software-Werkzeug häufig anzutreffen; für Karten mit integrierten Filtern oder Grafikprozessoren ist die Verwirklichung jedoch ebenfalls kein großes Problem.

11.6 Hardware-Zoom und -Panning

Fast jede Grafiksoftware hat die Möglichkeit, durch eine Zoom-Funktion Bildausschnitte zu vergrößern oder zu verkleinern. Verständlicherweise ist dieser Vorgang sehr langsam, weil eine umfangreiche Kalkulation im Bildspeicher notwendig ist. Ähnlich viel Rechenaufwand entsteht durch das Verschieben des Bildausschnitts (Panning), was ja vor allem bei vergrößertem Bild oft gebraucht wird, wenn mittels Rollbalken oder direkt mit dem Zeigegerät eine der aktuelle Bildausschnitt zu bewegen ist.

Wesentlich günstiger wäre es, diese Funktionen mit in die Hardware der Videografikkarte zu integrieren. Es ist dann keine Manipulation des Bildspeichers mehr nötig, sondern nur noch dessen Abtastung bzw. zeitliche Zuordnung zum physischen Wiedergabebild.

Vorzustellen hat man sich die Zoom-Funktion durch eine Verminderung bei der Abtastung der Zeilen- und Kolumnenzähler, die für die Adressierung im Bildspeicher zuständig sind. Während also z. B. 4 physische Zeilen geschrieben werden, kommen lediglich die Informationen zweier Zeilen des Speichers zur Abfrage. Es leuchtet ein, dass immer eine Teilung durch ganze Zahlen möglich sind – vorzugsweise sogar nur Zweierpotenzen.

Gelangt das Zeigegerät an den Rand eines vergrößerten Bildes, so sollte unmittelbar die Panning-Funktion in Aktion treten. Verändert die Zoom-Funktion die Taktraten der Speicherabfrage, so entspricht das Panning einer Manipulation der Offsets von Zeilen- und Kolumnenzählern.

11.7 Hardware-Flip

Ein sehr einfach zu realisierender Effekt ist der Hardware-Flip, bei dem Zeilen in beliebiger Reihenfolge aus dem Speicher abgefragt werden. Der einfachste Fall ergibt eine Spiegelung an der horizontalen Mittellinie des Bildes (Kopfstand). Die sog. Rowtable (Reihentabelle) führt diese Umsortierung aus und besteht wie die LUT aus einem kleinen Speicher, dessen Adressierung vom Zeilenzähler vorgenommen wird (zuständig für Zeilenposition) und dessen Datenausgänge die tatsächlichen Zeilennummern des Speichers (zuständig für Zeileninhalt) entsprechen. Entsprechend einer solchen vertikalen Flip-Funktion sind natürlich auch horizontale denkbar, obwohl hier die Adressierung natürlich wesentlich schneller vonstatten geht und deshalb etwas kritischer ist.

11.8 Bild-im-Bild beim Fernsehen

Dieses Verfahren ist auch unter der Bezeichnung *Picture-in-Picture* (PiP) bekannt. Es handelt sich dabei um die Möglichkeit, ein Hauptbild und ein kleines Bild anderen Inhalts auf dem Wiedergabescreen gleichzeitig darzustellen. Bei einem zeilenmäßig abgetasteten Bild benötigt man hierzu lediglich einen Videoschalter, der beim Schreiben der Kleinbildfläche zwei Quellen umschaltet, die aber bezüglich Abtastung synchron sein müssen. Bei analogen Bildquellen ist dies i. Allg. nicht der Fall. Selbst wenn dies doch zuträfe, müsste man für das innere Bild eine modifizierte Ausgabe bewerkstelligen, weil andernfalls nur ein Bildausschnitt zur Darstellung käme. Soll es sich also beim Kleinbild um einen vollen Bildinhalt handeln, sind diese Bilddaten auf die entsprechende Auflösung umzurechnen und in einen digitalen Gesamtbildspeicher zu mappen. Wir sehen diesen Vorgang und das Ergebnis in Abb. 11.7.



Abb. 11.7 Bild-im-Bild. Das einzublendende kleine Bild mit der Rose hat zunächst gleiche Auflösung wie das Hauptbild. Vor dem Einblenden und Mappen in den Bildspeicher muss es in der Pixelzahl so reduziert werden, dass die Pixeldichte gleich groß ist wie beim Hauptbild

Ein Empfangsgerät kann beide Signale nur dann darstellen, wenn zwei separate Tuner vorhanden sind.

11.9 Videotext

Seit etwa den 80er-Jahren fand die Einführung von Videotext (allg. Teletext) bei den öffentlich rechtlichen Fernsehanstalten statt. Nur Empfangsgeräte mit entsprechendem Decoder und Digitalspeicher sind in der Lage, Videotextseiten wie beispielsweise in Abb. 11.8 darzustellen. In Abb. 11.9 ist der signaltechnische Aufbau für den Videotext zu sehen.

Bereits recht früh kam die Idee, während der vertikalen Austastlücke Informationen zu übertragen (siehe Abschn. 7.1). Wichtig dabei ist, dass die Vertikalsynchronisation (vor allem für ältere Geräte) nicht gestört wird. Es handelt sich um keine Bilddaten, sondern Binärdaten, die dann der Decoder erst zu Bilddaten konvertiert. Im Grunde benötigt das Gerät nicht unbedingt einen digitalen Bildspeicher – lediglich der Decoder benötigt einen Speicher, denn die Binärdaten liefern erst nach mehreren Halbbildern ein komplettes Textimage, welches dann ein Textgenerator als Bildinformation dem eigentlichen Hauptbild überlagert.

Jede Textseite kann aus 23 Zeilen à 40 Zeichen bestehen, außerdem stehen sechs Farben zur Verfügung. Wie die Zeichen aussehen, welche Schriftart erscheint, obliegt nicht den Sendeinformationen, sondern dies definiert sich durch die Hardware im Decoder auf Empfängerseite. Lediglich die Schriftgröße wird senderseitig gesteuert. Die Daten



Abb. 11.8 Videotext. Zu sehen ist hier die Startseite (stets Seite 100). Es ist bei den meisten Geräten möglich, den Hintergrund der Textseite transparent zu schalten, damit das laufende Programm erkennbar bleibt

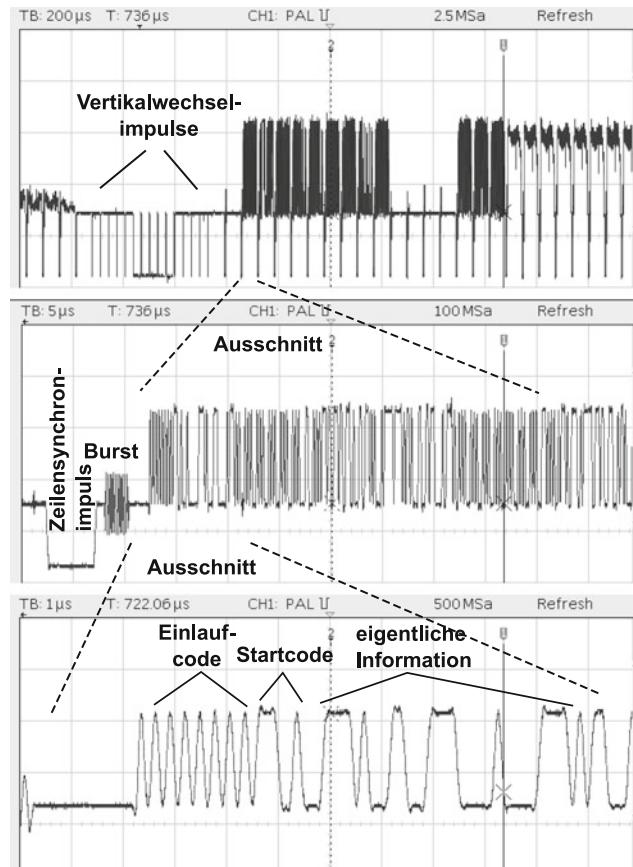


Abb. 11.9 Binärdaten während der Vertikalaustastlücke. Das Arrangement aus drei Oszilloskramen zeigt jeweils unterschiedliche Ausschnitte: Im oberen Oszilloskram sind Bildaustastlücke und mehreren Zeilen zu sehen, im mittleren eine einzige Zeile und im unteren die Anfangsdaten des Videotextes. Das Datenpaket beginnt mit Einlaufdaten für die Synchronisation des Videotextes, dann folgt eine Kennung, und erst im Anschluss kommen die binären Daten für den eigentlichen Text

von über 800 zur Verfügung stehenden Seiten erscheinen wie in einer Endlosschleife, allerdings sind wichtige Hauptseiten häufiger vertreten, damit die Wartezeit nicht zu groß wird.

Literatur

1. Steinbrecher, R.: Bildverarbeitung in der Praxis. München: Oldenbourg-Verlag 1993.



12.1 Bewegliche 2D-Modelle

Bei der Animation von zweidimensionalen Bildern erhält man lediglich eine Bewegung der Flächenobjekte; die räumliche Charakteristik haftet diesen Objekten weder in Form noch in Beleuchtung an. Folgerichtig sind geometrische Figuren auch nicht in Raumkoordinaten, sondern in Ebenenkoordinaten erfasst.

2D-Animationen kommen sehr häufig bei Zeichentrickfilmen vor, wo i. Allg. die Figuren und Objekte nur eine begrenzte Anzahl von Ansichten bereitstellen. Wir wollen hier aber nicht auf programmtechnische Möglichkeiten eingehen.

Der einfachste Fall in Bezug auf das Verfahren ist die Zusammenstellung mehrerer Einzelbilder, die die Animationssequenz beinhaltet. Dies bedeutet jedoch einen enormen Zeitaufwand bei der Erstellung der Bilder, so dass dies nur in Ausnahmefällen eine akzeptable Lösung ist.

Auch bei Programmen zur Erstellung von Zeichentrick-Animationen ist man dazu übergegangen, eine Vektorgrafik einzuführen. Damit lassen sich vor allem zwei wichtige Vorteile gewinnen: Einerseits ist eine nachträgliche Editierung der Szenen leicht durchzuführen, und andererseits sind oft nur Anfangs- und Endpunkte eines bewegten Objekts festzulegen – die Zwischenbilder generiert das Programm dann automatisch.

Derartige Trickanimationen benötigen oftmals nur eine sehr begrenzte Anzahl von Farben – eine Farbtiefe von 8 Bit stellt bereits 256 unterschiedliche bereit, die noch dazu über eine Palette adaptierbar sind. Weiterhin ist eine Begrenzung des Speicherbedarfs dadurch bedingt, dass oft viele gleichmäßige Farbflächen vorhanden sind, was wiederum einer komprimierte Aufbereitung der Bilddaten sehr entgegenkommt. Dies alles sind Faktoren, die ein Abspielen der Animation in Echtzeit begünstigen. Eine Wiedergabe im Videostudio oder das Aufzeichnen auf Videoband macht natürlich nach wie vor eine Videografikkarte erforderlich.

Der reibungslose Ablauf einer Sequenz hängt oft sehr stark von der Größe des Videospeichers und von der des Rechners ab. Häufig sind ganze Sequenzen im RAM ablegbar, von wo aus natürlich ein gleichförmiges Auslesen eher möglich ist als aus einer Festplatte. Wir kommen auf Überlegungen der Geschwindigkeit und der Auflösung noch am Ende des Kapitels zu sprechen.

12.2 Animation über Drahtgittermodelle

Besonders bei dreidimensionalen Animationen sind vektororientierte Objekte sehr vorteilhaft, weil sie über sog. Drahtgittermodelle eine geometrische Definition erhalten können. Es handelt sich dabei um Begrenzungslinien, über die der jeweilige Körper ausreichend definiert und auch darstellbar ist. Am Beispiel einer Kugel sehen wir das in Abb. 12.1.

Aus diesen und vielen weiteren geometrischen Grundgebilden lassen sich nun auch komplexere Drahtgittermodelle (Wireframes) formen, die zunächst als solche bestehen bleiben, da sie wenig Speicherplatz benötigen und daher auch während der Konstruktionsphase leicht beweglich sind.

12.2.1 Konstruktion eines Drahtgittermodells

Animationsprogramme bieten meist mannigfaltige Möglichkeiten, Drahtgittermodelle aus Einzelementen zusammenzusetzen. Da es sich um eine räumliche Konstruktion handelt, sind verschiedene Ansichten gleichzeitig darstellbar oder zumindest umschaltbar. Das Setzen von Elementen kann also immer aus einer günstigen Sichtposition heraus geschehen. Um einfache geometrische Gebilde wie Kugeln, Quader o. Ä. zu erzeugen, stehen grundsätzlich das direkte „Aufziehen“ über ein Zeigegerät (Maus usw.) oder aber die numerische Eingabe aller markanten Punkte als Möglichkeiten zur Verfügung. Insbesondere zur Generierung regelmäßiger Körper bedarf es i. Allg. nur eines Handgriffs mit dem Zeigegerät.

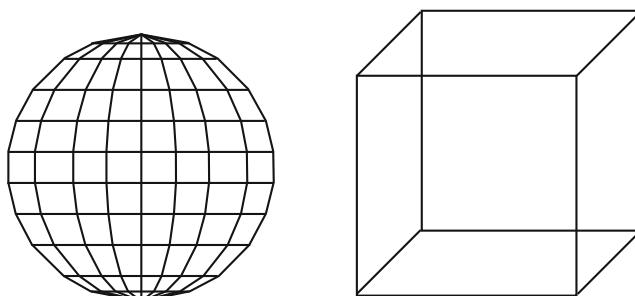


Abb. 12.1 Drahtgittermodell einer Kugel und eines Würfels

Sind Positionsänderungen notwendig, so sind diese jederzeit während der Modellierungsphase durchführbar. Das Drahtgittermodell stellt die unterste Stufe der Animation dar – es hat noch keinerlei Bewegung, sondern bildet lediglich die skelettierten Objekte und die Stati der zeitlichen Eckphasen.

Komplexere Figuren sind über die Konstruktion von Polygonen zu generieren. Der Verlauf der gedachten Oberflächenhaut stellt sich durch Approximation und ein Spline-Operationen ein.

12.2.2 Anfangs- und Endpunkt der Bewegung

Wir betrachten zunächst ein Quadrat als zweidimensionales Modell, welches über die Definition von Phasenzuständen einen späteren Bewegungsprozess ermöglichen soll (Abb. 12.2).

Es handelt sich dabei um eine Rotation um 90°, die in 25 Bildern (Frames) stattfinden soll. Durch die Festlegung der beiden Wireframes mit den Key-Nummern 1 und 2 und die dazugehörenden Frame-Nummern 0 und 25 (bzw. 0) hat der Keyframer erfasst, um welche Bewegung es sich handeln soll.

Die Generierung der Zwischenphasen (gestrichelt gezeichnet) erfolgt in unserem Beispiel linear; oft ist jedoch ein weicher Bewegungsablauf gefordert, der sich in der Nähe der Keyframes verzögert. Dies ist vor allem bei Bewegungen aus dem Ruhezustand heraus notwendig. Bei der Definition des Keyframing ist also die Kennlinie des Zeitzusammenhangs manipulierbar, wobei eine „geschmeidige“ Bewegung meist vorausgesetzt wird und nur davon abweichende frei zu definieren sind.

Startet man nun den Ablauf der Wireframe-Animation, so ist bereits eine grobe Beurteilung der Bewegung möglich. Die Wiedergabe einer solchen Animation in Echtzeit benötigt i. Allg. keine besondere Hardware, weil keine großen Datenmengen zu verarbeiten sind.

Die meisten Animationsprogramme verfügen weiterhin über eine Preview-Funktion (Vorschau), die aus dem Wireframe-Ablauf eine grobe „körperliche“ Animation macht, die ebenfalls noch wenig Speicherplatz benötigt und deshalb spielend mit jeder üblichen Grafikkarte abspielbar ist.

Die Wireframe-Datei mit Festlegung der Keyframes stellt die Basis dar für die endgültige Animation, bei der die Oberflächen der Objekte durch gewählte Muster oder

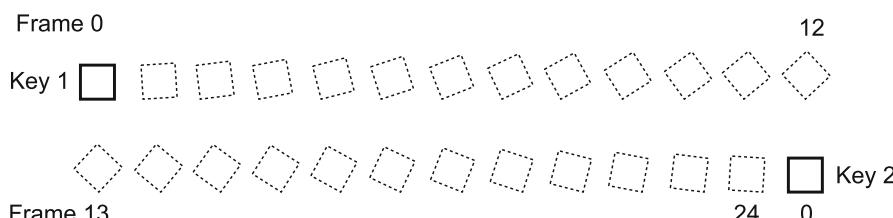


Abb. 12.2 Kontinuierliche Bewegung durch Definition von Eckphasen

Farben bedeckt werden. Solange diese Datei vorhanden ist, kann eine Reproduktion der endgültigen Animation und eine Änderung des Ablaufes herbeigerufen werden.

12.2.3 Rendering

Ein fertiges Bild entsteht aus einem Wireframe erst, wenn an die aufgespannten Flächen-elementen Oberflächenstrukturen projiziert werden. Dieser Prozess heißt Rendering (engl. *Übersetzung*) und ist für jedes einzelne Wireframe – nicht nur für die Keyframes einer Bewegung – durchzuführen.

Das Endresultat ist ein pixelorientiertes Bild, welches – bedingt durch die Auflösung und die Farbtiefe – einen entsprechend hohen Speicherbedarf zeigt. Es verhält sich grundsätzlich anders als das Wireframe, und die einzelnen Objekte (z. B. Kugel, Quader usw.) lassen sich nicht mehr als solche gezielt bearbeiten, weil ja deren Pixel nur dem Raster angehören und somit gleichberechtigt zu allen anderen Pixeln sind.

Der Vorgang des Rendering ist sehr zeitintensiv, weil jedes einzelne Pixel über spezielle Algorithmen zu berechnen ist, welche auch die eventuelle Verwendung von Oberflächenstrukturen (Texturen) berücksichtigen müssen. Die Projektion einer Textur auf eine gekrümmte Oberfläche beispielsweise geschieht über umfangreiche Rechengänge. Abb. 12.3 zeigt die Gegenüberstellung eines Drahtgittermodells und dem entsprechenden Bild nach dem Rendering-Prozess.

Als Hilfsmittel stellt das Programm meist eine große Zahl von wählbaren Oberflächen und Texturen in einer Bibliothek zur Verfügung, so dass viele künstliche Objekte mit natürlichem Aussehen geschaffen werden können. Diese Oberflächenattribute sind jedem Drahtgitterobjekt einzeln zuzuordnen, so dass natürlich eine Individualität der Objekte gewahrt bleibt.

Das Rendering sollte für videotechnische (jedoch auch für rein visuelle) Anwendungen ein einstellbares Antialiasing aufweisen, damit Übergänge beliebig weich zu gestalten sind. Außerdem sind auch hier die Bedingungen für die Farbpegel einzuhalten, die ggf. mit dem Programm als Grenzen einstellbar sind (vgl. Abschn. 9.2.3).

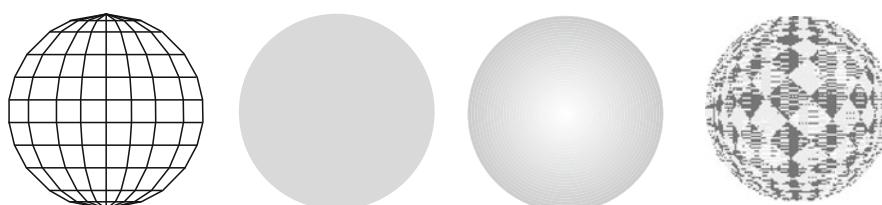


Abb. 12.3 Rendering-Prozess für eine Kugel mit unterschiedlichen Oberflächen

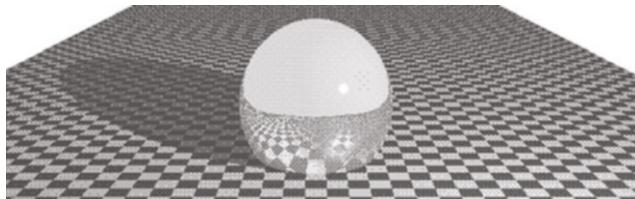


Abb. 12.4 Bildberechnung mit Raytracing und einer Lichtquelle

12.2.4 Raytracing

Da die konstruierten Körper wie in der Realität meist keine Selbststrahler sind, sondern nur ein auftreffendes Licht reflektieren, spielt die Verarbeitung von Lichtquellen eine außerordentlich große Rolle beim Rendering. Dieser Vorgang heißt Raytracing (engl. *Strahlverfolgung*) und ist eigentlich Bestandteil des Rendering.

Neben Reflexionen, die von einem Körper ausgehen, sind natürlich auch noch Schatten zu berücksichtigen. Einzelne Bewegungen der Objekte bedeuten verständlicherweise auch Veränderungen der Reflexionen und der Schatten, so dass wiederum jedes Bild einzeln berechnet werden muss. Abb. 12.4 stellt ein Beispiel für ein künstliches Bild dar, welches mit einer Lichtquelle bestrahlt wird. Die Nuancen der beleuchteten Oberfläche sowie die Schattenbildungen sind ein besonderes Merkmal beim Raytracing.

Die Reflexionen entstehen natürlich nicht nur durch die Lichtquelle selbst, sondern auch durch die bestrahlten Flächen, wie z. B. dem Boden, so dass auch die Oberflächenstruktur – hier ein Schachbrettmuster – auf andere Körper einwirkt und in entsprechender Form berechnet werden muss. Ebenfalls hat ein Schatten ja auch noch sekundäre Wirkungen. Bedenkt man weiterhin, dass auch mehrere Lichtquellen einsetzbar sind, so wird klar, dass die Rechenzeit dadurch sprunghaft ansteigt.

Bei den Objekten und den Lichtquellen sind viele verschiedene Parameter wie Oberflächenrauigkeit, Reflexionsfaktor, Farbe, Strahldivergenz, Strahlrichtung usw. zu definieren, die ebenfalls alle mit in die Berechnung eingehen. Ferner sind natürlich auch optische Eigenschaften und Effekte wie Brechung und Beugung zu berücksichtigen. Ein gutes Raytracing-Programm fragt daher bei Verwendung von Glaskörpern neben Form und Lage auch einige Eigenschaften ab. Jedem Programm zur Erstellung von dreidimensionalen Animationen muss die Vielfalt des Raytracing zur Verfügung stehen, da sonst keine natürlich wirkenden Lichtverhältnisse geschaffen werden können.

12.3 Morphing

In der modernen Tricktechnik hat sich besonders ein Effekt als besonders interessant und wirkungsvoll herausgestellt: das Morphing (eng. *Verwandlung*). Man kann durch dieses

sehr aufwendige Verfahren mit Computerunterstützung z. B. ein menschliches Gesicht in einen Tierkopf verwandeln und umgekehrt.

Ähnlich wie bei der Animation sind hier ebenfalls Eckphasen zu setzen, zwischen denen dann alle bildlichen Zwischenphasen errechnet werden. Da es nichts mit einem trivialen Überblenden zu tun hat, sondern Bildinhalte erkennen und verschieben kann, ist ein enormer Rechenaufwand zu treiben, der zwar durch entsprechende Hardware beschleunigt, aber dennoch nicht in Echtzeit zu vollziehen ist.

Das Morphing setzt Bildmustererkennung voraus, weil es bei oben genanntem Beispiel einzelne Partien des Gesichts identifizieren muss. Es ist jedoch verständlich, dass ein Morphing-Programm viele Details offen lassen muss, die zur Optimierung einzugeben bzw. anzupassen sind.

Die Bildmustererkennung kann (und wird) in den meisten Fällen über ein sog. Verzerrungsnetz ergänzt oder gar substituiert werden. Dabei hat der Programmbenutzer auf markante Stellen des ersten Bildes (z. B. Mund- und Augenpartien) Netzpunkte zu setzen, die also das Bild bereits in groben Zügen definiert. Dieses Punktenetz ist jetzt zu kopieren und auf das Endbild zu legen. Verständlicherweise müssen jetzt viele der Punkte auf die entsprechenden Stellen des Endbildes gelegt werden. Auf diese Weise erfolgt eine Mitteilung dem Computer darüber, welche Bildbereiche des Anfangsbildes mit denen des Endbildes korrespondieren.

Der Start des Morphing bringt nun die Verzerrungen der Bildbereiche in Gang, wobei viele Glättungsalgorithmen visuelle Brüche unterbinden.

Sollen nur Einzelheiten eines Bildes verändert werden (z. B. der Mundwinkel oder die Augenposition), so vereinfacht sich der Ablauf dahingehend, dass die Netzpunkte des Zielbildes ausgehend vom Erstbild zu manipulieren sind. Die errechneten Zwischenbilder speichert das Morphing-Programm ab, wobei die Anzahl vorher festzulegen ist. Zur Kontrolle bieten die meisten Programme eine Preview-Funktion, die einen groben Überblick des Ergebnisses nach wesentlich kürzerer Zeit geben kann.

12.4 Aufzeichnungstechniken

Liegen alle Bilder einer Animationssequenz vor, so werden diese üblicherweise auf ein Speichermedium (Band oder Laserdisc) aufgezeichnet. In dieser Form lässt sich die Sequenz jederzeit ohne Umstände in ein Videoprogramm einbinden.

12.4.1 Echtzeit-Animation

Ob die Animation in Echtzeit zur Aufzeichnung gelangen kann, ist sehr einfach feststellbar, indem sie mittels Animationsprogramm gestartet wird (falls es dafür eine Option gibt). Nach dem derzeitigen Stand der Technik sind Videoauflösungen von 768×576 (Vollauflösung) nicht als Echtzeitanimation abspielbar. Liegt sie in Form von Einzelbildern

unkomprimiert vor, so benötigt jedes einzelne Bild eine Speichergröße von ca. 1,3 MByte (bei einer Farbtiefe von 24 Bit). Auch wenn die Festplatte noch so groß ist, wäre der damit zusammenhängende Datenfluss von ca. 33 MByte/s nicht zu bewerkstelligen. Sicherlich könnte man Rechner und Festplatten konstruieren, die über Parallelbetrieb solchen Forderungen gerecht würden, aber eine wirtschaftliche Technologie wäre dies bestimmt nicht.

Als Alternative kommen entweder Teilbildanimationen oder Teilauflösungen im Vollbildformat in Frage. Ersteres ist für professionelle Videoanwendungen nicht akzeptabel; die zweite Möglichkeit bietet bereits passable Auflösungen, die jedoch ohne spezielle Hardware-Kompression bzw. -Dekompression nicht machbar sind.

12.4.2 Einzelbild-Aufzeichnungen

Aufzeichnungen von Animationen bei voller Videoauflösung können dennoch bewerkstelligt werden, wenn jedes einzelne Bild aufnimmt und große zeitliche Zwischenräume zulässt. Somit bleibt dem Computer genügend Zeit zum Aufbau jedes einzelnen Bildes.

Steuerung des Videorecorders Damit nun eine Sequenz von Einzelbildern chronologisch und vor allem automatisch auf Band aufgezeichnet wird, ist die Steuerung der Videobandmaschine erforderlich. Dies geschieht meist über ein Computerboard, dem sog. VTR-Controller. In Abb. 12.5 spiegelt sich das Zusammenspiel aller Komponenten wider.

Das die Grafikkarte verlassende Videosignal kann unterschiedliche Formate aufweisen – eine Anpassung zwischen Videorecorder und Grafikkarte ist daher manchmal erforderlich.

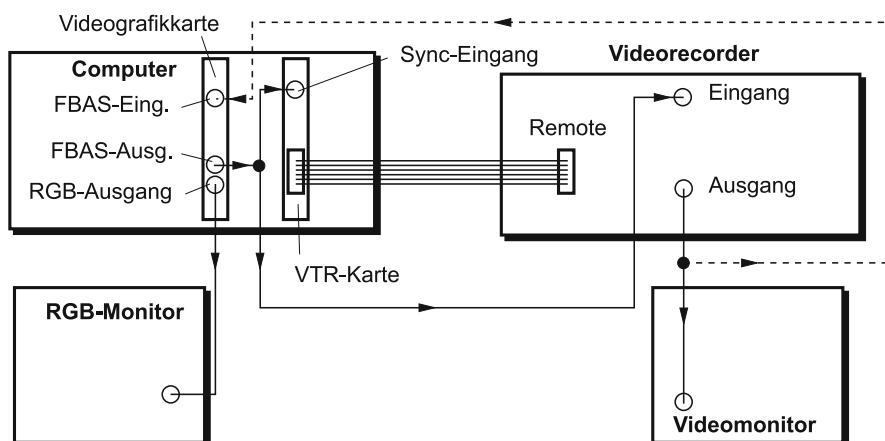


Abb. 12.5 Blockschaltbild für die Einzelbildaufzeichnung

Die VTR-Karte steuert zusammen mit dem Animationsprogramm den Videorecorder, der über eine entsprechende Schnittstelle verfügen muss (hier gibt es unterschiedliche Standards, wobei sich allerdings bei modernen Maschinen die serielle Schnittstelle RS422 allgemein durchgesetzt hat). Der Ablauf erfolgt nach folgendem Muster:

- Laden des (nächsten) aufzuzeichnenden Bildes in den Videospeicher der Videografikkarte.
- Zurückspulen des Bandes an eine definierte Stelle.
- Auf Wiedergabe schalten.
- An vorgegebene Bandposition aktuelles Bild aufzeichnen.
- Kurzer Bandauslauf, dann weiter bei erstem Schritt.

Damit die Aufzeichnung ordnungsgemäß vollzogen wird, schaltet die VTR-Karte den Recorder genau an derjenigen Bandstelle auf Record, die für das betreffende Bild reserviert ist. Damit genau nur an dieser Stelle für die Dauer einer Kopfradumdrehung aufgezeichnet wird, muss die VTR-Karte ständig die aktuelle Bandposition erfahren. Die Übermittlung solcher Informationen, die nur durch einen vorher auf Band aufgezeichneten Timecode möglich ist, erfolgt wie die Steuerung der Maschine über die Remote-Schnittstelle.

Der laufende Timecode kann zwar die absolute Position ständig aktualisieren, für die Übermittlung und Decodierung dieser Daten verstreicht jedoch eine nicht zu vernachlässigende Zeitspanne, so dass es hierdurch zu ungenauen Steuerungen und Aufzeichnungen käme. Aus diesem Grunde benötigt die VTR-Karte auch noch einen Syncrontakt, der mit dem der Bandmaschine und der Videografikkarte übereinstimmt. Die daraus gewonnenen Vertikalimpulse dienen dann letztlich dem Auslösen des Aufzeichnungskommandos.

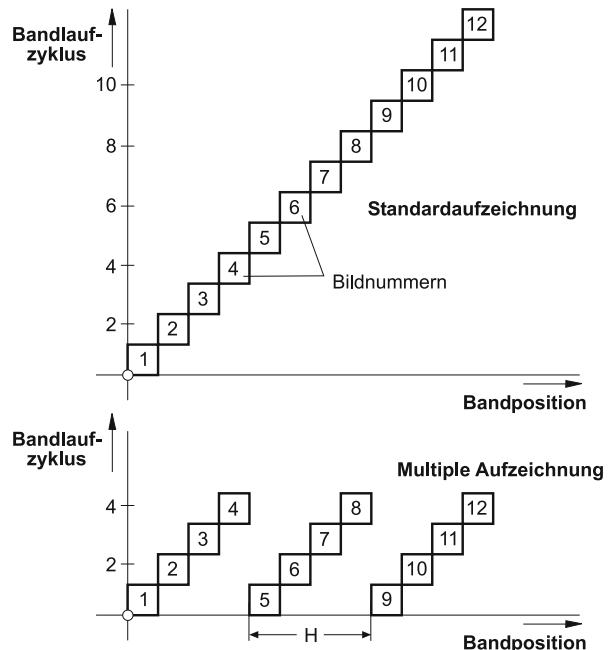
Kommen FBAS-Signale zur Aufzeichnung, so muss die Videografikkarte über einen FBAS-Eingang ein Genlock-Signal erhalten, das von der Video-Bandmaschinen stammt (gestrichelte Linie). Genaugenommen würde das Burst-Signal ausreichen. Nur so kann der FBAS-Encoder der Karte die Phase seines Farbhilfsträgers an diejenige der zuletzt aufgezeichneten Bilder anpassen.

Rendering der Einzelbilder bei der Aufzeichnung Das geschilderte Verfahren könnte man so abwandeln, dass nicht die fertigen Bilder in das RAM der Videografikkarte zu laden sind, sondern dass das Animationsprogramm jedes fertige Bild erst über den Rendering-Prozess vom Wireframe erzeugen muss. Dies dauert zwar länger, man darf allerdings nicht vergessen, dass die Zeiten für den Bandtransport ebenfalls nutzbar sind. Der große Vorteil, der mit diesem Verfahren gewonnen wird, besteht aus der enormen Ersparnis von Speicherplatz. Letztlich ist sowieso die Zeit für das Rendering einzukalkulieren, gleichgültig, ob es vor oder während der Aufzeichnung erledigt wird. Wenn es sich allerdings um ein leistungsfähiges Rechnersystem handelt, könnte die Zeit für den Bandtransport ausreichen, um das Rendering durchzuführen (was allerdings nicht für jedes Bild gleich lange dauert).

Multiple Aufzeichnung Das Zurückspulen und Abspielen sind Zeitkomponenten, die für jedes aufzuzeichnende Bild zählen. Außerdem wirken sich die häufigen Bandtransporte natürlich verschleißend auf die Mechanik der Maschine aus.

Um diesen Nachteilen zu begegnen, hat man eine naheliegende Idee verwirklicht: Während eines Bandzyklus können mehrere, zeitlich auseinander liegende Bilder an die entsprechenden Stellen des Bandes aufgezeichnet werden. Das zwischen den Bildern liegende Zeitintervall muss dabei so groß sein, dass das System das nächste Bild in den Videospeicher laden kann. Abb. 12.6 zeigt den Zusammenhang zwischen Bandposition und Bandlaufzyklen. Die Singleframe-Aufzeichnung benötigt wesentlich mehr Zyklen (nämlich genau so viele, wie Bilder aufzuzeichnen sind) als die multiple Aufzeichnung. H definiert den Abstand zweier Bilder, die in einem Zyklus zur Aufzeichnung kommen; dieser Abstand richtet sich nach der vom Computersystem benötigten Zeit zum Laden des nächsten Bildes. Das resultierende Zeitintervall muss vor dem Start der Aufzeichnung durch das System ermittelt werden, dabei ist natürlich ein Sicherheitszuschlag zu wählen. Kommt es dennoch zu Zeithorizonten, muss das Computersystem korrigierend eingreifen, so dass im Endresultat der Aufzeichnung nicht etwa Bilder fehlen oder an falschen Positionen liegen. Die Programmsteuerung muss zu jeder Zeit über eine genaue und vollständige Registrierung des Aufnahmestatus verfügen. So können mehrere, zeitlich auseinander liegende Bilder an die entsprechenden Stellen des Bandes aufgezeichnet werden. Das zwischen den Bildern liegende Zeitintervall muss dabei so groß sein, dass das System

Abb. 12.6 Schematischer Ablauf von Einzelbildaufzeichnungen und Zahl der notwendigen Zyklen



das nächste Bild in den Videospeicher laden kann. Abb. 12.6 zeigt den Zusammenhang zwischen Bandposition und Bandlaufzyklen. Die Singleframe-Aufzeichnung benötigt wesentlich mehr Zyklen (nämlich genau so viele, wie Bilder aufzuzeichnen sind) als die multiple Aufzeichnung. H definiert den Abstand zweier Bilder, die in einem Zyklus zur Aufzeichnung kommen; dieser Abstand richtet sich nach der vom Computersystem benötigten Zeit zum Laden des nächsten Bildes. Das resultierende Zeitintervall muss vor dem Start der Aufzeichnung durch das System ermittelt werden, dabei ist natürlich ein Sicherheitszuschlag zu wählen. Kommt es dennoch zu Zeitkonflikten, muss das Computersystem korrigierend eingreifen, so dass im Endresultat der Aufzeichnung nicht etwa Bilder fehlen oder an falschen Positionen liegen. Die Programmsteuerung muss zu jeder Zeit über eine genaue und vollständige Registrierung des Aufnahmestatus verfügen.

Das hier beschriebene Verfahren ist vor allem eine Ergänzung der Steuerungs-Software. Es bringt vor allem bei längeren Animationssequenzen zeitliche Vorteile und steigert die Standzeit der Maschinen. Eine Aufzeichnung mit integriertem Rendering allerdings ist nicht immer vorteilhaft, weil sonst die Länge H möglicherweise zu groß wird und das Verfahren unwirtschaftlich.

Zeitkalkulation bei der Einzelbildaufnahme Der Zeitfaktor ist ein wichtiges Kriterium bei der Einschätzung der Einzelbildaufzeichnung. Es ergibt sich für die Zeit der gesamten Aufzeichnung:

$$t = t_0 \cdot f_B \cdot T \quad (12.1)$$

t_0 endgültige Dauer der Videosequenz, f_B Vollbildfrequenz, T Zeit für Bandlaufzyklus.

Für eine gewünschte Sequenzdauer von 1 min und einer angenommenen Zykluszeit von 20 s ergibt sich also bei PAL ($f_B = 25$ Hz) eine Aufnahmezeit von:

$$t = 60 \text{ s} \cdot 25 \text{ Hz} \cdot 20 \text{ s} = 30,000 \text{ s} \approx 8,3 \text{ h}$$

Das setzt natürlich voraus, dass pro Bandlaufzyklus nur ein Bild aufgezeichnet wird; beim Einsatz der multiplen Aufzeichnung oder auch bei vereinzelten Szenen ohne Bildveränderungen (solche können durch Aufzeichnung in unmittelbarer Abfolge realisiert werden) verkürzt sich die Aufnahmezeit entsprechend. Allerdings verlängert sich dafür die Zykluszeit etwas.

Das Produkt $f_B \cdot T$ ist also der Faktor, um den die Zeit der Erstellung größer ist als die Wiedergabezeit des Endprodukts; in unserem Beispiel beträgt er 500.

12.5 Rotoscoping

Speziell bei Einzelbildaufzeichnungen ist ein besonderer Effekt realisierbar: das Rotoscoping. Es verbindet zwei unterschiedliche und unabhängige Geschehnisse zu einem zeitgenauen Ablauf.

12.5.1 Effektbeschreibung

Wir wollen uns hierzu ein einfaches Animationsmodell eines sich bewegenden und rotierenden Würfels anhand von Abb. 12.7 betrachten. Zunächst ist ein sich bewegender Körper wie bei einer normalen Animation Bestandteil des Rotoscoping, was hier von einem Würfel repräsentiert wird. Im allgemeinen Fall kann es sich jedoch um jede beliebige Fläche handeln, die in ihrem räumlichen Verlauf definiert ist (also auch gekrümmte Flächen). Diese Fläche wird nun über das Rotoscoping für eine Projektion eines zweiten Geschehens (hier ein sich bewegendes Gesicht) benötigt. Abhängig vom Verlauf dieser Fläche erfolgt laufend und für jedes einzelne Bild der Animation und des Projektionsablaufes (Gesicht) ein Anschmiegen des letzteren an diese Fläche. Dieser Vorgang lässt sich vergleichen mit dem Aufkleben eines Bildes auf eine gewölbte Fläche, wobei jedoch auch Projektionsalgorithmen denkbar sind, die auch Faltenbildungen vermeiden (Bild auf einer Gummihaut).

12.5.2 Arbeitsweise

Die Möglichkeit des Rotoscoping ist geknüpft an einige Fähigkeiten der Komponenten des Systems, deren Zusammenwirken aus Abb. 12.8 hervorgeht. Wir wollen davon ausgehen, dass die zu projizierende Bildsequenz von einem separaten Videorecorder zugespielt wird, weil sich bei Verwendung von nur einem Recorder Synchronisationsprobleme einstellen

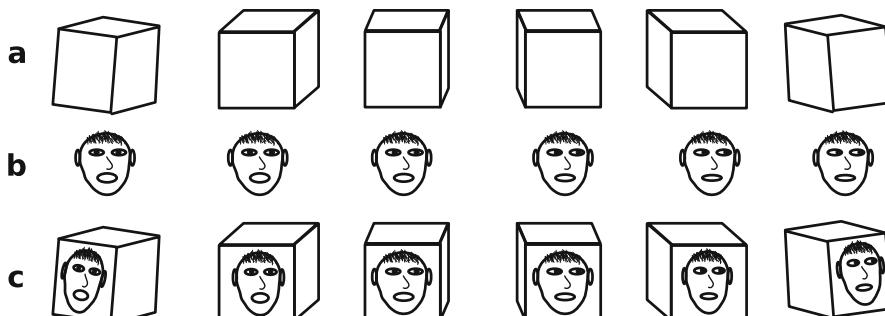


Abb. 12.7 Wirkungsweise des Rotoscoping mit (a) Animationskörper, (b) das Projektionsbild, (c) das fertige Rotoscoping

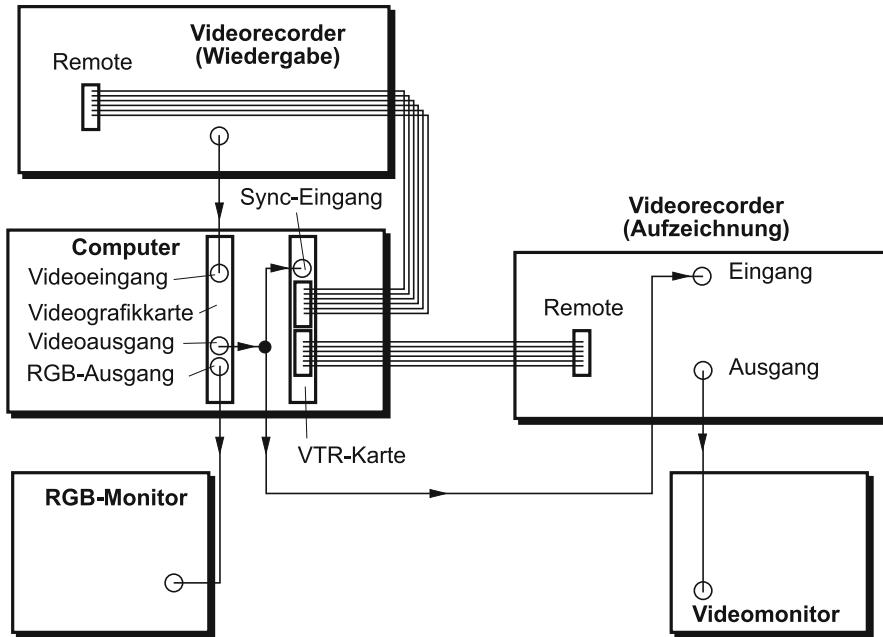


Abb. 12.8 Blockdiagramm beim Rotoscoping

können (die sich ergebende Signalschleife für die Synchronisation müsste künstlich aufgetrennt werden, oder die Genlock-Funktion der Videografikkarte ist zu unterbinden). Demnach muss auch die VTR-Karte über zwei Remote-Schnittstellen verfügen, die sich durch die Software unabhängig ansteuern lassen.

Der Ablauf gliedert sich wie folgt:

- Anfahren und Wiedergabe des (nächsten) zu projizierenden Bildes **B**.
- Einfrieren des betreffenden Bildes **B** durch Capture-Board.
- Rendering des (nächsten) Animationsbildes **A** mit Projektion auf vorgegebener Fläche.
- Aufzeichnen des Bildes auf vorgegebene Bandstelle im Einzelbildverfahren.
- Falls Sequenzende noch nicht erreicht, weiter beim ersten Schritt.

Aus dieser Aufstellung und auch aus dem obigen Blockdiagramm ist zu entnehmen, dass die Videografikkarte auch über einen Videoeingang verfügen muss, über den Bilder einzufrieren sind.

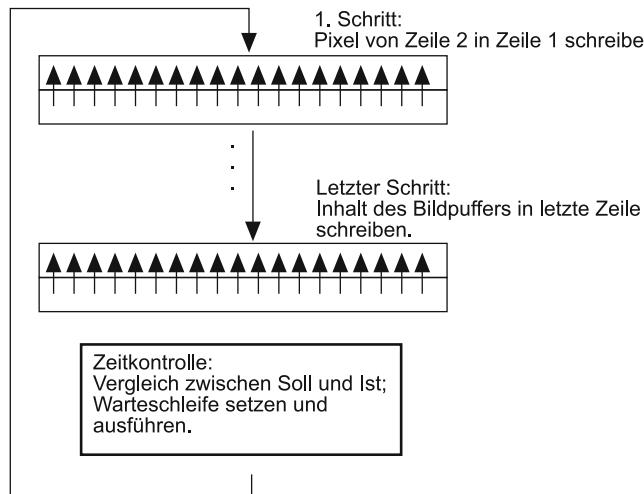


Abb. 12.9 Speichertransfer bei senkrechter Textbewegung

12.6 Schrift-Animation

Zur Erzeugung von Text-Overlays bei der Videovertitelung ist eine Animation der Schrift in Echtzeit notwendig. Dabei ist außer für den Bildinhalt auch ein Speichertransfer für den Alpha-Kanal erforderlich, weil er ja das Keying auslöst (vgl. Abschn. 11.1).

Obwohl man im ersten Moment dem zu bewegenden Text nicht sehr viel Komplexität und Speicherbedarf ansieht, müssen die Textvorlagen zunächst einen Rendering-Prozess durchlaufen, bevor ein Start der Animation möglich ist. Das hat jedoch auch den Hintergrund, dass die Textvorlage aus Gründen der Editierbarkeit objektorientiert sein muss, während die Animation nur Rasterbilder gebrauchen kann.

Bei der Realisierung geht es um das Hauptproblem der Datenbewegung innerhalb des Videospeichers. Dabei muss die Geschwindigkeit der Textbewegung unabhängig von der Komplexität des Textes und vom System überhaupt sein. Dies ist z. B. bei einer senkrechten Textbewegung nur dadurch zu gewährleisten, dass grundsätzlich der gesamte Bildspeicher zu bearbeiten und eine Zeitkontrolle für jeden Speicherzyklus durchzuführen ist. Abb. 12.9 zeigt das Schema eines solchen Speichertransfers.

Eine Echtzeitanimation ist hier jedoch vor allem deshalb leichter durchzuführen, weil i. Allg. kein Raytracing notwendig ist. Die Textobjekte sind also unabhängig von der Beleuchtung, und außerdem führen sie meist nur eine eindimensionale und höchstens eine zweidimensionale Bewegung aus, wodurch die Konturveränderung weitgehend vermieden wird.

Da der Zeitpunkt des Beginns der Laufschrift mit dem Videoprogramm eine feste Beziehung hat, sind qualitative Schriftengenerator-Programme in der Lage, auf einen spezifischen Steuerimpuls (Timecode) zu reagieren.

12.7 Kritische Parameter bei Echtzeit-Animationen

Speziell bei Echtzeit-Animationen kommt es auf die Geschwindigkeit des Systems an (daran entscheidet sich, ob sie überhaupt möglich ist oder nicht). Die technische Entwicklung läuft in eine Richtung, wo Einzelbild-Animationen für den Aufzeichnungsmodus nicht mehr nötig sein werden, auch wenn es sich um maximal komplexe Bilder bzw. Grafiken handelt.

Welche Faktoren dabei eine wesentliche Rolle spielen, wollen wir in diesem Abschnitt etwas näher untersuchen.

12.7.1 Anforderungen an das System

Ein Datentransfer von Bildern, deren Komplettdateien auf Festplatte liegen, zum Videospeicher, dauert in der Regel so lange, dass das Abspielen von Echtzeitvideos bzw. -animationen mit Einschränkungen an die Auflösung oder an die Farbtiefe verbunden sind. Diese Einschränkungen sind um so geringer, je schneller die dafür relevanten Einzelkomponenten arbeiten.

Festplatte Eine der langsamsten Komponenten beim Datentransfer ist die Festplatte des Computers. Dabei lässt sich für gewöhnlich zwar die Datenrate durch die Anzahl der Köpfe erhöhen, dennoch ist in Bezug auf die mittlere Zugriffszeit auf eine beliebige Spur nicht beliebig verkürzen. Die Unterschiede der Standards (z. B. AT-Bus, Local-Bus usw.) schlagen auch im Datentransfer nieder. Leider hängt jedoch die Zugriffszeit hauptsächlich von mechanischen Eigenschaften ab, weshalb eine individuelle Wahl der Festplatte erfolgen sollte.

Bus-Eigenschaften Der Bus eines Computers ist bekanntlich mit entscheidend für die mögliche Datenrate und damit für die Geschwindigkeit bzw. Ruckfreiheit ablaufender Prozesse verantwortlich. Um nicht die Abwärtskompatibilität zu verlieren, haben Rechnerhauptplatinen (Mainboards) meist immer auch Slots mit dem Standardbus. Daneben sollen jedoch schnellere Bus-Architekturen (bei IBM-kompatiblen Rechnern z. B. der Local Bus) die Möglichkeit bieten, mehr Leistungsfähigkeit zu ermöglichen. Die angeschlossene Karten müssen natürlich diesen schnelleren Standard unterstützen, sonst ergeben sich keine Vorteile bzw. arbeitet die Konstellation nicht ordnungsgemäß.

Grafikprozessor Manche hochwertigen Videografikkarten verfügen über einen speziellen Grafikprozessor (GSP), der viele Spezialaufgaben übernehmen kann. Er ist speziell an den Erfordernissen der Grafikdatenbearbeitung ausgerichtet und ist somit i. Allg. in der Lage, die geforderte Aufgabe schneller als der Zentralprozessor des Rechners auszuführen, zumal der Rechnerbus durch ihn umgangen wird. So liegt z. B. der Einsatz des GSP zur

Bewältigung von Rendering-Prozessen nahe; das Programm muss hierbei natürlich über spezielle Software-Schnittstellen und Instruktionen den GSP ansprechen.

Prozessortakt Für Vorgänge, die von der Rechengeschwindigkeit abhängen, trägt die Prozessortaktfrequenz unmittelbar zur Geschwindigkeit bei. Betrachtet man beispielsweise Rendering- oder Schiebevorgänge im Speicher, so ist ein hoher Prozessortakt grundsätzlich vorteilhaft. Zugriffe auf die Festplatte oder auf den Videospeicher der Grafikkarte jedoch unterliegen Begrenzungen anderer Art, auf die mit dem Prozessortakt nicht eingewirkt werden kann.

Speicherausbau Für sehr kurze Animationssequenzen kann die Möglichkeit bestehen, dass sie ins RAM des Rechners passen. Von dort aus ist natürlich ein schnellerer Zugriff möglich. Komplexe und längere Videoszenen jedoch finden in unkomprimierter Form keinen Platz im RAM, weshalb durch einen entsprechenden Ausbau auch keine Vorteile zu erwarten sind.

12.7.2 Auflösung

Die Auflösung der Einzelgrafiken wirken sich auf den Speicherbedarf aus. Beim Transport von Bilddaten ist daher für eine Grafik mit kleinerer Auflösung entsprechend weniger Zeit notwendig.

Veränderungsfaktoren für die Auflösung wirken sich verständlicherweise quadratisch aus, weil ein solcher Zusammenhang auch mit der Grafikfläche besteht, falls die Pixelgröße konstant bleibt. Verkleinert man die Auflösung von z. B. 640×480 auf 320×240 , so kann nach wie vor ein schirmfüllendes Bild (Full Screen) erreicht werden, wenn der Bildschirmtreiber dies unterstützt. Der Datendurchsatz pro Bild ist dagegen auf 1/4 gefallen, was einer Vervierfachung der Ladegeschwindigkeit gleichkommt.

Mit herkömmlichen Computersystemen und Grafikkarten ist noch keine Echtzeitanimation bei Full-Screen-Wiedergabe und voller Auflösung möglich. Spezielle Workstations dagegen können über Kompressionsverfahren (JPEG bzw. MPEG) eine solche Echtzeit-Wiedergabe schaffen. Allerdings kann man auch hier nicht sagen, dass bei der Wiedergabe kein Verlust erfolgte. Bereits die JPEG-Kompression generiert bei extremen Bildsituationen sehr wohl Veränderungen im Bild, die sich mitunter als sehr störend auswirken können.

12.7.3 Farbtiefe

Die Farbtiefe der Grafik geht linear in den Speicherbedarf ein, daher ist auch hier ein klarer Zusammenhang mit der Ladezeit eines Bildes zu sehen. Allerdings tritt meist allein schon deswegen eine Verlangsamung ein, wenn der Treiber für die Videografikkarte eine

andere Farbtiefe unterstützt als die aktuelle Grafik aufweist. Daher ist hier zunächst für Übereinstimmung zu sorgen, bevor Ladegeschwindigkeiten bemängelt werden.

Für eine Computergrafik mit angepasster Palette kann eine Farbtiefe von 8 Bit bereits völlig ausreichend sein; gegenüber 24 Bit spart man sich 2/3 des Speicherplatzes und erhält eine etwa dreifache Ladegeschwindigkeit.

12.7.4 Bildinhaltsveränderungen

Im allgemeinen Fall ergeben sich bei Animationssequenzen (wie übrigens auch bei üblichen Video- und Kameraszenen) von einem Bild zum nächsten keine solch drastischen Inhaltsveränderungen, die sich über die gesamte Bildfläche erstrecken. Diese Tatsache macht man sich beim MPEG-Verfahren zunutze, welches gleichbleibende Bildpartien bei der Speicherung und auch beim Laden ausklammert bzw. diese als Daten einfach im Video-Speicher stehenlässt. Der Datendurchsatz fällt dadurch auf ein erträgliches Maß ab.

Bei extremen Schnitten, die aber keine kontinuierliche Animation mehr darstellen, versagt dieses Verfahren vor allem dann, wenn es sich in der Hauptsache um solches Material handelt, das kaum noch gleichbleibende Bildflächen beinhaltet. Situationen solcher Art, die jedoch nur vorübergehend auftreten, können über eine Pufferung in ein RAM abgefangen werden, aus dem dann bei Bedarf auch ein neues Bild mit vollständig alternierendem Bildinhalt nachgeladen werden kann. Auch ist eine Pufferung in den Videospeicher denkbar, wenn die Videografikkarte mehrere Seiten aufnehmen kann. Im Gegensatz zu Videosequenzen sind bei Animationssequenzen i. Allg. alle Bildpartien, in denen eine Inhaltsänderung abläuft, von vornherein durch das Konstrukt der Animation bereits bekannt und definiert. Die Ausklammerung statischer Bildbereiche ist also recht einfach zu bewerkstelligen.

12.8 Problemlösungen zur Datenreduktion

Zur Steigerung der Leistungsfähigkeit bei der Wiedergabe von Animationen in Echtzeit sind einige Maßnahmen praktikabel, die als Paket sehr günstige Voraussetzungen schaffen. Es handelt sich dabei um Verfahren, die sich teilweise auf die Hardware und teilweise auf die Software verteilen.

12.8.1 JPEG-Verfahren

Um den Speicherplatzbedarf eines einzelnen Bildes zu reduzieren, sind vielerlei Kompressionsarten diskutiert worden; die erfolgreichste Methode heißt JPEG und wurde nach dem gleichnamigen Komitee benannt (Joint Photographic Experts Group). Die einzelnen

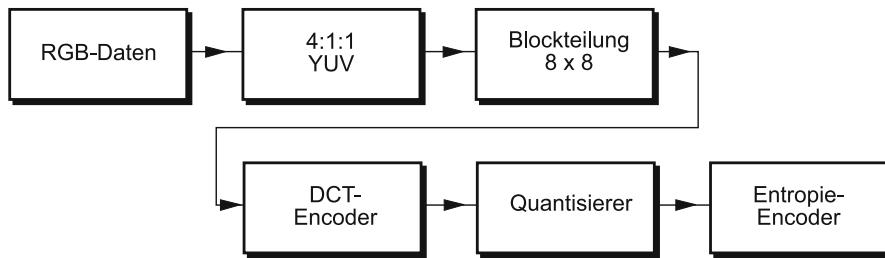


Abb. 12.10 Schema des JPEG-Encoders

Arbeitsschritte sind mit geeigneten Computerboards auf reiner Hardware-Basis in Echtzeit zu bewältigen; die hierzu nötigen Schritte müssen nicht über komplexe, elementare Befehle bewältigt werden, sondern sind mittels „verdrahteter“ Prozesse mit einem Minimalaufwand an Datenbewegungen durchführbar.

Wir wollen hier jedoch nur grob das Verfahren anschneiden, weil eine detaillierte Behandlung zu umfangreich wäre. Es handelt sich um eine Kombination aus Bilddatenkonversion und um eine mehrfache Kompression. Abb. 12.10 zeigt den Überblick der Wandlung in das JPEG-Datenformat. Die Rückwandlung geschieht ganz analog, nur in umgekehrter Richtung.

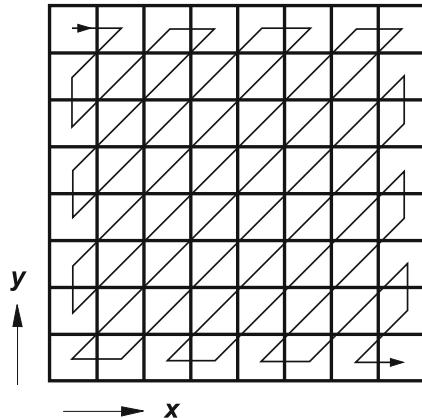
Umcodierung von RGB nach YCbCr RGB-Werte der Pixel unterliegen zunächst einer Wandlung in Luminanz und Chrominanz nach den Gleichungen:

$$\begin{aligned}
 Y &= 0,299 R + 0,587 G + 0,114 B \\
 C_b &= 0,1687 R - 0,3313 G + 0,5 B \\
 C_r &= 0,5 R - 0,4187 G - 0,0813 B
 \end{aligned} \tag{12.2}$$

Die Gleichung für die Luminanz (Y) ist inzwischen bereits mehrfach vorgekommen; lediglich die Farbdifferenzwerte unterliegen anderen arithmetischen Festlegungen als bei der Videosignaltechnik. Dieser Vorgang, der für sich betrachtet noch keinerlei Datenreduktion beinhaltet, bildet jedoch die Basis für eine erste Reduktion. Diese beruht auf der Tatsache, dass das menschliche Auge für Farbinformationen ein deutlich schlechteres Auflösungsvermögen aufweist. Für ein Bildelement von 2×2 Pixeln erfolgt nur noch eine Abtastung bzw. Wertangabe jeweils für C_b und C_r . Das Abtastverhältnis zwischen Luminanz und beider Chrominanzkanäle ist also 4:1:1, woraus sich die Namensgebung für diese Methode der Abtastreduktion ergibt. Für dieses Feld sind also keine 12 Werte (wie dies für RGB nötig ist), sondern nur noch 6 Werte relevant. Die Farbcodierung geschieht dabei nach den Mittelwerten der vier Pixel.

Nach diesem ersten Schritt findet bereits eine Datenreduktion auf 50 % statt. Die sich ergebenden Werte werden nun in Blöcke zu 8×8 Pixeln zusammengefasst.

Abb. 12.11 Zickzack-Abtastung des 8x8-Blocks bei der DCT



Diskrete Kosinus-Transformation (DCT) Der Umstrukturierung schließt sich die diskrete Kosinus-Transformation an – eine Sonderform der diskreten Fourier-Transformation. Die Transformationsgleichung lautet:

$$H(i, j) = \frac{1}{4} C(i) \cdot C(j) \left[\sum_{x=0}^7 \sum_{y=0}^7 A(x, y) \cdot \cos \frac{(2x+1)i\pi}{16} \cdot \cos \frac{(2y+1)j\pi}{16} \right] \quad (12.3)$$

$$\text{mit } C(i), C(j) = \begin{cases} \frac{1}{\sqrt{2}} & \text{für } i, j = 0 \\ 1 & \text{sonst} \end{cases}$$

Das Abtasten des 8x8-Datenblocks geschieht in Zickzacklinien nach Abb. 12.11, und zwar für alle Kanäle Y , C_b und C_r , die mit der Bildfunktion $A(x, y)$ korrespondieren. Es ergeben sich gleichfalls 64 Werte für H , wobei diese jetzt allerdings Frequenzkomponenten darstellen. Der unterste Wert nach der Transformation (für $i, j = 0$) stellt die DC-Komponente dar, also ein Offset, der sich allen Bildfunktionswerten überlagert. Die Komponenten von H mit Indizes, die auch für die Abtastung entlang der Zickzacklinie entstehen, repräsentieren immer höhere Frequenzen.

Beim Vorgang der DCT entsteht also aus der Zeitfunktion des Bildes eine Frequenzfunktion. Prinzipiell erhält man zwar ebenfalls wieder 64 Werte, wobei allerdings bei den meisten Bildinformationen hauptsächlich die Komponenten niedriger Indizes einen von null verschiedenen Wert aufweisen.

Schnelle Grafikkarten mit moderner Prozessortechnologie sind heute in der Lage, manche Rendering-Prozesse in Echtzeit ablaufen zu lassen.

Diese Verteilung ist bei jedem Bild unterschiedlich – entsprechend sind natürlich später vom Bildinhalt abhängige Kompressionsfaktoren zu erwarten.

Quantisierung Die Unterteilung der sich ergebenden Werte in diskrete Stufen nennt man Quantisierung. Der sich später einstellende Kompressionsfaktor hängt maßgeblich mit der Feinheit der Unterteilung zusammen, wie dies auch aus Abb. 12.12 hervorgeht. Das gleiche Prinzip gilt auch bei der A/D-Wandlung, wo ebenfalls die Auflösung darüber entscheidet, wie viele Quantisierungsstufen existieren werden. Die Anzahl der diskreten Stufen ergibt die notwendige Informationsbreite. Durch die Quantisierung entsteht natürlich ein Datenverlust, der um so größer ist, je größer die Abstufung wird.

Entropie-Encoder Ein weiterer Schritt bei der Kompression basiert auf der Tatsache, dass die quantisierten Werte mit sehr unterschiedlichen Häufigkeiten vorkommen. Der mittlere Informationsgehalt aller Zeichen – auch als Entropie bezeichnet – verkleinert sich, wenn die Häufigkeitsverteilung ungleichmäßig verläuft. Eine Möglichkeit, dies auszunutzen, besteht in der sog. Huffman-Codierung. Hier setzt man eine oder mehrere Tabellen ein, die für häufig vorkommende Byte-Werte einfach einen Codes mit weniger als 8 Bits vorgibt. Nur seltene Werte müssen dann die ursprüngliche Codierung erhalten – allerdings eingeleitet von einem Kurzcode, der den Ausnahmefall ankündigt.

Die Tabellenwerte müssen bei der JPEG-Kompression mitgeliefert werden – zur Decodierung sind sie unbedingt notwendig. Die Anwendung kann statische Tabellen vorgeben, was allerdings den Kompressionsfaktor nicht besonders günstig ausfallen lässt. Werden mehrere, auf das jeweilige Bild angepasste Tabellen verwendet, so benötigt deren Erstellung i. Allg. einen separaten Durchlauf durch die Bilddaten.

Kompressionsergebnis Je nach Kompressionsfaktor kann der notwendige Speicherbedarf beim JPEG-Format ca. 0,25...2 Bit/Pixel betragen; entsprechend diesem Bereich ergeben sich auch Unterschiede der Qualität. Man muss sich aber immer vor Augen halten, dass der erreichte Kompressionsfaktor und somit der notwendige Speicherbedarf stark

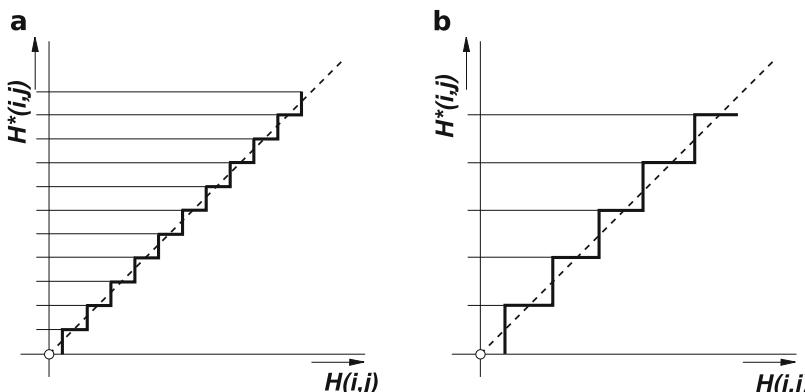


Abb. 12.12 Unterschiedliche Quantisierungen: (a) fein, (b) grob

vom Bildinhalt abhängen kann. Demnach kann alleine der erreichte Kompressionsfaktor nicht hinreichend die Bildqualität definieren.

Zum genauen Aufbau des JPEG-Formats sei auf Anhang [D](#) und auf Sekundärliteratur verwiesen. S. [372](#) [1,2]

12.8.2 MPEG-Verfahren

Zum Abspielen einer Animation in Echtzeit kann das JPEG-Verfahren u. U. nicht ausreichend sein, weil es vor allem auf die Kompression einzelner Bilder beschränkt und somit die Ähnlichkeiten aufeinanderfolgender Bilder unberücksichtigt lässt. Das MPEG-Verfahren (Motion Picture Expert Group) dagegen beinhaltet dieses Manko nicht.

Bildarten und Datenaufbereitung Auch hier kommt die diskrete Kosinus-Transformation zur Anwendung, allerdings nur bei komplett neu zu codierenden Bildern, nämlich den I-Bildern (Intra). Grundsätzlich erfolgt jedoch im Gegensatz zum JPEG-Verfahren eine Aufteilung des Bildes in 16×16 -Blöcke. Ein darauffolgendes Bild, welches keine wesentlichen Unterschiede zum laufenden aufweist, kann als P-Bild (Prediction) codiert werden, wobei nur die Abweichungen einer DCT unterliegen.

Daneben erfolgt auch die Codierung von B-Bildern (Bidirectional Prediction) durch Interpolation aus P- und I-Bildern. Die Bildbewegungen zwischen letzteren werden bei dieser Interpolation als gleichförmig vorausgesetzt. Das Originalbild dient also hier nicht als Grundlage – dennoch kommt es natürlich nur zu einer B-Bild-Codierung, wenn das Original keine großen Unterschiede zeigt, und dies ist gewährleistet, wenn die benachbarten P- und I-Bilder als Eckphasen genügen. Die Interpolation lässt sich vergleichen mit derjenigen zwischen Keyframes bei der Animation.

Immer jeweils bei Szenenwechseln oder Schnitten muss eine I-Codierung stattfinden, da keinerlei sonstigen Anhaltspunkte für das aktuelle Bild vorhanden sind.

Kompressionsergebnis Die individuelle Kompression ist auch hier abhängig vom Bildinhalt und von einzustellenden Parametern. Das MPEG-Verfahren kann mit Zusatzkarten auf Hardware-Basis bereits Video- bzw. Animationssequenzen mit voller Auflösung in Echtzeit verarbeiten. Der Größenfaktor beträgt hier ca. 1:100. Allerdings sind wie beim JPEG-Verfahren hohe Kompressionsfaktoren nur auf Kosten der Bildqualität zu erzielen; die Bildauflösung ist also kein hinreichendes Kriterium für die Übertragungs- oder Reproduktionsqualität, wenn bei den Umformungen Fehler entstehen, die sich im Bildeindruck niederschlagen.

12.8.3 QPEG-Verfahren

Nach den Standard-Verfahren JPEG und MPEG entstand ein weiteres Verfahren, welches von der Firma Q-Team Dr. Knabe entwickelt wurde – das QPEG-Verfahren (Quality Pictures Expert Group). Es führt die Datenreduktion auf abgewandelter Form durch und ist darauf ausgerichtet, dass die Wiedergabe der Sequenzen (und somit ggf. auch die Offline-Schnittbearbeitung) an einem Computer ohne Videografikkarte (z. B. lediglich mit einer VGA-Karte) erfolgen kann.

Ausgehend von Einzelbildern, die im Format TGA (siehe Anhang [D](#)) auf der Festplatte vorliegen, kann das QPEG-Verfahren Animationen oder auch Videosequenzen in Echtzeit abspielen, ohne dafür eine spezielle Dekompressions-Hardware oder eine spezielle Grafikkarte zu benötigen. Die Schritte bei der Kompression sollen hier etwas näher beleuchtet werden.

Reduktion der Farbtiefe Der erste Schritt bei der QPEG-Kompression besteht in der Reduktion der zur Verfügung stehenden Farben eines Bildes. Diese Überlegung geht von der i. Allg. gültigen Annahme aus, dass die meisten Bilder einen Farbwertvorrat von ca. 16,8 Millionen Farben bei einer Tiefe von 24 Bit nur zu einem sehr geringen Bruchteil ausschöpfen.

Das Ziel, Sequenzen auf herkömmlichen VGA-Karten abspielen zu können, setzt eine Reduktion auf die Farbtiefe von 8 Bit voraus. Von den zur Verfügung stehenden 256 Farben werden jedoch nur 236 genutzt; 16 weitere stehen als Hintergrundfarbe für die normale Darstellung im VGA-Modus zur Verfügung, und die vier restlichen Palettenplätze dienen Systemfunktionen.

Damit die Reduktion auf eine Farbtiefe von 8 Bit überhaupt möglich ist, muss – wie übrigens bei allen solchen codierten Farbtiefen – eine Palette angelegt werden. Nun genügt bei Bildsequenzen mit einer ursprünglichen Farbtiefe von 24 Bit keine einzelne Palette, sondern QPEG generiert Paletten, die für jede Szene (also mehrere Bilder) spezifisch sind und die Häufigkeit der im Original vorkommenden Farben berücksichtigt.

Berechnung von Delta-Frames Ähnlich wie bei MPEG wird auch bei QPEG der Hauptanteil der Kompression dadurch erreicht, dass Bilder einer Szene üblicherweise große redundante Flächen bilden. Man berücksichtigt also wiederum nur die Änderungen zwischen zwei Bildern. Im Gegensatz zu MPEG setzt QPEG dagegen nur an sehr vereinzelten Stellen ein I-Bild, also ein voll codiertes Bild, und B-Bilder kommen überhaupt nicht vor. Dieses Verfahren bringt keine Nachteile gegenüber dem kontinuierlichen Abspielen, jedoch gelingt eine beliebige Startpositionierung oder auch ein Rückwärtslauf verständlicherweise nicht.

Entropie-Encoder Nachdem die Daten aller Bilder nach Farbtiefenreduktion und Delta-Berechnung vorliegen, erfolgt wie bei JPEG und MPEG eine verlustfreie Kompression, die nach der Huffman-Codierung arbeitet (siehe Abschn. [12.8.2](#)).

Kompressionsergebnis Mit QPEG kann mit Größenfaktoren von 1:100 bis 1:250 gerechnet werden, was wiederum abhängig vom Bildinhalt ist.

12.9 Neue Animationstechniken

Die Animationstechnik hat sich drastisch verbessert in den letzten Jahren. Das Ziel war nicht nur, dargestellte Objekte oder Personen so realistisch wie nur möglich zu kreieren, sondern natürlich auch deren Bewegungsablauf. Wenn ein Mensch oder ein Tier geht oder läuft, so geschieht dies nach einem sehr komplexen Muster. Man hat hierfür unterschiedliche Methoden entwickelt.

12.9.1 Motion Capture

Beim Motion-Capture-Verfahren sind an realen Bewegungsobjekten Sensoren angebracht, um Daten für Bewegungsabläufe zu erhalten. Das Animationsprogramm verbindet diese Daten in ein Drahtgittermodell eines virtuellen Skeletts, woraus dann die gewünschten Personen durch Rendering entstehen. Alternativ hierzu sind auch ohne diese „Sampling-Daten“ manche Programme in der Lage, ein Computermodell einer Figur zu relativ natürlich anmutenden Bewegungen zu bringen. Der Ablauf orientiert sich an nach unterschiedlichen Kriterien (z. B. Energie) entwickelten Optimierungen. Dass solche Programme sehr komplex sind, lässt sich leicht erkennen.

Die Qualität bzw. Naturtreue der Animation hängt stark von der Auflösung des virtuellen Drahtgittermodells ab. Bei Bewegungen sind teilweise auch Gelenke relevant, die beim Design des virtuellen Skeletts nicht vorkommen. Demnach ist gut nachzuvollziehen, dass ohne diese Gelenke die Bewegung kantig oder roboterhaft aussieht.

12.9.2 Performance Capture

Neben Bewegungsabläufen eines ganzen Körpers ist natürlich auch die Gesichtsmimik von enormer Bedeutung. Die Vorgabe für diverse Stimmungen, die sich in spezifischen Gesichtsausdrücken äußert, entstammt meist aus entsprechenden „echten“ Bildern von Gesichtern. Wichtige Details sind dabei Mundwinkel, Augenliderstellung und Augenbrauen.

Literatur

1. Holtorf, K.: Handbuch der Grafikformate. München: Franzis-Verlag 1994.
2. Steinbrecher, R.: Bildverarbeitung in der Praxis. München: Oldenbourg-Verlag 1993.



13.1 Online-Schnittsysteme

Ein Schnittverfahren, bei dem ständig eine Steuerung der Zuspieler stattfindet und das Endprodukt – der endgültige Schnitt – direkt am Schluss der Arbeitsprozesse zur Verfügung steht, nennt man Online-Schnittverfahren.

13.1.1 Arbeitsweise

Nach Abb. 13.1 kann solch ein System mit nur zwei Zuspielern auskommen, wobei allerdings gewährleistet sein muss, dass die durch Schnitt aneinanderzufügenden Szenen immer auf beide Zuspieler verteilt sein muss. Diese Bedingung ist grundsätzlich erfüllt, wenn beide Zuspieler über alle Szenen verfügen. Um unnötige Spul- und Positionierzeiten zu vermeiden, ist jedoch eine grobe Einhaltung der zeitlichen Reihenfolge auf den Zuspielbändern von Vorteil. Beide Bänder müssen außerdem über einen durchgängigen Timecode verfügen, der aber nicht szeneneinheitlich sein muss.

Kennt man anhand eines Schnittplanes die Timecode-Werte der aneinanderzufügenden Szenen, so lassen sich diese entweder durch Anfahren der Bandstelle oder durch numerische Eingabe auf der Tastatur in den Computer übernehmen. Letzterer kann seinerseits alle drei Maschinen – die beiden Zuspieler und die MAZ für die Aufzeichnung – so steuern, dass diese sich in gewünschtem Synchronlauf befinden. Hierzu ist natürlich noch ein zentraler Takt notwendig, der jedoch auch vom Schnitt-Computer geliefert werden kann.

Nachdem nun beide Zuspieler im richtigen Zeitbezug zueinander ihre Bänder abspielen, kommt nach einigen Sekunden diejenige Stelle, die als Schnitt von Szene A/Band 1 und Szene B/Band 2 vorgesehen ist. In diesem Moment schickt der Computer an den Mischer

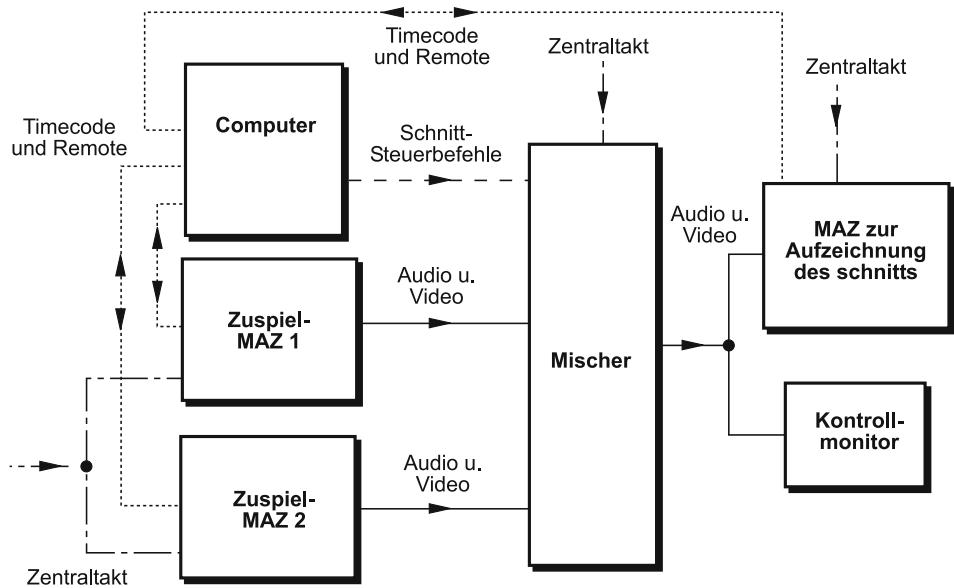


Abb. 13.1 Schematischer Aufbau beim Online-Schnitt

einen entsprechenden Steuercode, der eine Überblendung oder auch einen harten Schnitt auslöst. Der Kontrollmonitor ermöglicht eine Beurteilung des Schnitts. Falls nötig, lassen sich Schnittzeitpunkte beliebig verschieben und/oder die Schnitteffekte ändern.

Für die Computersteuerung wurden anfangs spezielle Systeme eingerichtet, die nur zur Schnittsteuerung geeignet waren. Heute sind jedoch zur Bewältigung dieser Aufgaben nur noch gängige Computer und eine Schnittstellen- und Genlock-Karte erforderlich. Teilweise sind sogar auch die Standard-Schnittstellen geeignet, weil sich die meisten Maschinen über RS-422 steuern lassen, was von Computern durch RS-232 substituiert oder auch direkt erfüllt werden kann.

13.1.2 Endgültiger Schnitt

Die MAZ für die Aufzeichnung läuft normalerweise erst am Schluss mit, wenn alle Schnittpositionen auf diese Art gefunden sind. Auch hier muss natürlich eine positionsgenaue und synchronisierte Steuerung stattfinden, da die Aufzeichnungs-MAZ bei sehr kurzen Szenen ja immer nur einen Schnitt bei kontinuierlichem Lauf aufzeichnen kann; der nächste Übergang macht dann einen weiteren Zyklus nötig, so dass zunächst Szene A und B, beim nächsten Durchlauf Szene B und C, dann Szene C und D usw. als Schnitt überspielt werden. An den Nahtstellen dieser Szenenpaare muss die Aufzeichnungs-MAZ zeitgenau gesteuert werden.

13.1.3 Nachteile des Verfahrens

Da aus dem Material der Zuspieler das endgültige Produkt durch Überspielen entsteht, sind folgende Bedingungen zu erfüllen:

- Die Aufnahmeequalität des Grundmaterials für die Zuspieler muss sehr hoch sein.
- Die Wiedergabequalität der Zuspieler muss ebenfalls hoch sein.
- Die Zuspielbänder müssen einen Timecode aufweisen.

Die ersten beiden Punkte implizieren, dass nur Kopien von Betacam- oder Digitalaufzeichnungen zu verwenden sind, bei denen das Qualitätsdefizit der ersten Generation zu vernachlässigen ist.

13.1.4 Besondere Schnittverfahren

Das oben beschriebene Schnittverfahren beinhaltet die allgemeinste Form des Schnittvorgangs; es erhält den Namen AB-Roll-Verfahren. Daneben können jedoch auch vereinfachte oder spezielle Verfahren angewandt werden, die bei entsprechender Ausstattung der Maschinen keinen Schnittcomputer benötigen. Diese Vorgänge beherrschen auch schon moderne Geräte für den nichtkommerziellen Gebrauch.

Assemble-Schnitt Sollen ausschließlich harte, übergangslose Schnitte vorgenommen werden, so genügen zwei Maschinen: ein Zuspieler und eine Aufzeichnungsmaschine, wo der Assemble-Schnitt (assemble = zusammenbauen) stattfindet. Ein solches Vorgehen kann ebenfalls nach Abb. 13.1 erfolgen, indem einfach der zweite Zuspieler entfällt. Abb. 13.2 verdeutlicht den Schnittvorgang. Bei besonderer Schnitteinrichtung beider Maschinen jedoch kann sogar auf den Schnitt-Computer verzichtet werden; in diesem Fall kommunizieren beide Maschinen über die Timecode- und Remote-Schnittstelle. Nach der Eingabe der Schnitt-Timecodes und dem Start des Schnittvorgangs fahren beide Maschinen eine Preroll-Position an, von der aus sie bei gemeinsamem Start bis zur Schnittposition zum erforderlichen Gleichlauf und Timecode-Bezug kommen. Hierzu ist i. Allg. außer der Kommunikation über den Timecode auch eine Synchronleitung erforderlich.

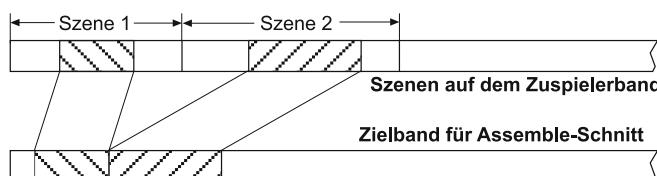


Abb. 13.2 Vorgänge beim Assemble-Schnitt

Insert-Schnitt Für das Einsetzen von Szenen in ein fertiges Programm kann der Insert-Schnitt zum Einsatz kommen. Auch hier genügen u. U. zwei Maschinen, die über die Timecode- und Remote-Schnittstelle miteinander kommunizieren. Allerdings ist hierbei zu beachten, dass das Material im Zielrecorder mit der einzufügenden Szene überschrieben wird – ein Vorrücken der der Schnittposition nachfolgenden Aufzeichnung findet bei dieser Konstellation verständlicherweise nicht statt. Zur Veranschaulichung diene Abb. 13.3.

13.2 Offline-Schnittsysteme

Beim Offline-Schnitt besteht eine vollkommene Trennung zwischen Schnittbearbeitung und endgültiger Schnittproduktion. Das Verfahren ist deshalb flexibler, und außerdem entfallen die Qualitätsrichtlinien für das Material der Zuspieler. Es geht während der Schnittbearbeitung in erster Linie um das Setzen der Schnittpositionen – dabei spielt die Qualität des Materials eine untergeordnete Rolle. Abb. 13.4 zeigt einen Überblick des Aufbaus beim Offline-Schnittverfahren. Das hier vorgestellte Verfahren arbeitet nicht nur offline – also getrennt vom Schneideprozess der endgültigen Produktion –, sondern gleichzeitig auch non-linear, was bedeutet, dass eine sequentielle Aufzeichnung des Arbeitsmaterials nicht vorliegen muss. Das zur Bearbeitung eingespielte Material erlaubt vielmehr sozusagen einen „wahlfreien Zugriff“.

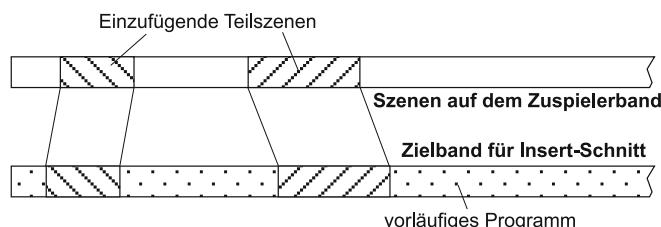


Abb. 13.3 Vorgänge beim Insert-Schnitt

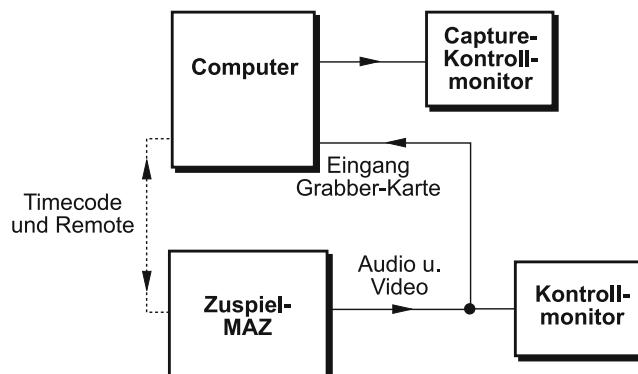


Abb. 13.4 Schematischer Aufbau beim Offline-Schnitt

13.2.1 Digitalisieren des Videomaterials

Zunächst muss das gesamte Material auf Festplatten digitalisiert werden, wobei für die Überspielung nicht zwingend ein Original hoher Güte notwendig ist. Zu beachten ist lediglich, dass die Kopie einen identischen und durchgängigen Timecode aufweist wie das Originalband.

Mit den Sequenzen, die nun in komprimierter und meist in schlechterer Auflösung auf der Festplatte vorliegen, wird der feste Bezug zum Timecode hergestellt. Er ist somit in die Einzeldateien eingebunden.

Nach dem Digitalisieren können die Einzelsequenzen abgespielt werden und auf dem Capture-Monitor (meist verkleinert in einem Fenster) betrachtet werden. Gleichzeitig erfolgt eine Anzeige des hierzu korrespondierenden Timecode. Zur besseren Identifizierung der einzelnen Szenen stellen die Programme unterschiedliche Hilfsmittel zur Verfügung: Zum Beispiel können Anfangs- und Endbild als kleines Bildsymbol auf der Benutzeroberfläche erscheinen, was dem Cutter meist genügt und ein Anspielen der Szene unnötig macht. Darüber hinaus sind meist die einzelnen Szenen mit einem Kommentar zu versehen, was ebenfalls die Identifizierung erleichtert und außerdem eine Unterscheidung von *Takes* gleicher Szenen möglich macht.

Der Transfer des Originaltones auf Festplatte bereitet normalerweise keine qualitäts-technischen Schwierigkeiten oder Beschränkungen. Die Qualität entspricht somit der des Zuspielers. Oftmals stehen sogar mehrere Tonkanäle zur Verfügung, die getrennt vom Bild editierbar sind.

Zur Speicherung dienen meist Wechselplatten (magnetisch oder magneto-optisch), da nach der Produktion das Material vorläufig nicht mehr benötigt wird und normalerweise zur Archivierung kommt. Da das Material ohnehin in einem komprimierten Format auf der Festplatte liegt, ergeben sich meist keine Möglichkeiten für eine weitere effektive Kompression. Allerdings können unbenutzte Teile des Materials gelöscht werden.

Während des Digitaliservorgangs benötigt das System ständig die aktuellen Timecode-Informationen des Zuspielers; damit Unterbrechungen dieses Datensignals keinen Totalausfall bewirken und auch kein Zeitversatz zwischen Timecode und Bildsignal vorkommt, ist oftmals noch eine zentrale Synchronisation erforderlich. Die Timecode- und Remote-Informationen können entweder direkt über eine serielle Schnittstelle oder aber über Konvertiergeräte ausgetauscht werden.

Das Digitalisieren der Sequenzen besorgt eine Grabber-Karte meist in Verbindung mit Kompressionskarten (JPEG o. Ä.). Für die Tonüberspielung ist eine Audiokarte notwendig, die u. a. die Zahl der Simultan-Kanäle festlegt. Die Bildauflösung sowie die Farbtiefe lässt sich auf Software-Ebene festlegen.

13.2.2 Durchführung des Offline-Schnitts

Nachdem das ganze Material auf Festplatte verfügbar ist, kann die eigentliche Schnittbearbeitung beginnen. Die Anfangs- und Endpunkte von Einzelszenen sind nach dem Anspielen von Gesamtszenen sehr einfach zu setzen. Das Aneinandersetzen solcher angepasster Szenen erfolgt ähnlich wie bei einem Textsystem durch Kopier-, Einsetz-, Überschreib- und Löschenfunktionen.

Unterstützt wird der gesamte Vorgang noch durch eine Timeline, die auch Blenden und andere Effekte darstellen kann. Außerdem lassen sich natürlich jederzeit Korrekturen der Schnittpositionen anbringen – ja sogar die gleichzeitige Erstellung mehrerer Schnittversionen ist möglich.

Das sich ergebende Storyboard gibt zu jeder Zeit Aufschluss über die aktuelle Zusammensetzung der einzelnen Parts.

Der Trimm-Modus erlaubt die Feinjustage der Schnittpunkte; jegliche Änderung lässt sich sofort im Preview-Modus beurteilen.

Ein wesentlicher Vorteil non-linearer Schnittsysteme ist der, dass eine beliebige Positionierung (Cueing) innerhalb kürzester Zeit im bestehenden Schnittprodukt möglich ist. Konventionelle Schnittverfahren benötigen dagegen eine nicht zu vernachlässigende Zeit für das Umspulen.

Die meisten Systeme für Computer-Offline-Schnitt erlauben auch das Einbinden von Grafiken und Animationen, die zuvor jedoch normalerweise eingeführt und auf das Format der Applikation gebracht werden müssen.

Die Darstellung der Clips kann entweder in Form von Symbolen oder nach vielen anderen Kriterien erfolgen. Eine Sortierungsmöglichkeit z. B. nach den Bezeichnungen ist ebenfalls üblich. Somit lassen sich auch Vorteile von Datenbanken für das Quellmaterial nutzen.

13.2.3 Rohschnittfassung

Viele Studios setzen bewusst Offline-Schnittsysteme ein, um das Originalmaterial und den Schnittcomputer zu schonen. Um diesen Vorteil noch zu steigern, sehen die meisten Offline-Systeme die Generierung eines Rohschnitts vor. Somit können verschiedene Schnittfassungen dem Auftraggeber vorgelegt werden, ohne den Online-Schnittcomputer benutzen zu müssen. Erst nach der Entscheidung für eine Fassung ist der endgültige Schnitt mit guter Qualität zu vollziehen.

Moderne Offline-Schnittsysteme schaffen heute bereits eine Rohschnittqualität von S-VHS – der endgültige Schnitt ist also nicht unbedingt auffällig besser.

Voraussetzung für die Erzeugung einer Rohschnittfassung vom Offline-System heraus ist das Vorhandensein einer Videografikkarte oder eines entsprechenden Scan-Konverters, damit beim Abspielen überhaupt erst ein Videosignal zur Aufnahme verfügbar ist.

Die Rohschnittfassung kann – wegen ihrer starken Kompression – auch über ISDN-Leitungen übertragen werden, was eine schnelle Beurteilung ebenfalls erleichtert.

13.2.4 Schnittliste und endgültiger Schnitt

Nach erledigter Schnittarbeit legt ein Offline-Schnittsystem eine Schnittliste (EDL), die in digitaler Form auch auf Diskette ausgegeben werden kann. Sie enthält alle Daten über timecode-orientierte Schnittpositionen, Blenden und Steuerung von Audiodaten.

Für die Daten der Schnittliste gibt es verschiedene Formate, die vor allem nach den Kriterien unterschiedlicher Hersteller für Schnittsysteme und Bandmaschinen entstanden sind. Normalerweise sind die Programme für Offline-Schnittsysteme in der Lage, die Schnittliste für unterschiedliche Formate zu erstellen. Man kann die Übersetzung des internen Codes etwa mit der Funktion eines Druckertreibers vergleichen, der für einen Text die nötigen Steuerzeichen und natürlich die Zeichen für den Text selbst in Form eines Signalcodes anlegt, der vom jeweiligen Drucker verstanden wird und ihn zum korrekten Arbeiten bringt.

Das Originalmaterial oder die Arbeitskopie zusammen mit dieser Schnittliste gelangt nun zum Online-Schnittcomputer, wo der endgültige Schnitt mit maximaler Qualität erledigt wird.

Das Offline-Schnittsystem kommt also nicht ohne Online-System aus (wohl aber umgekehrt), aber dennoch bringt seine Zwischenschaltung wesentliche arbeitstechnische und wirtschaftliche Vorteile.

13.3 Audionachbearbeitung

Obwohl man primär vermutet, der Originalton sei ja mit den Szenen bereits bei der Schnittproduktion berücksichtigt worden, so bedarf es im Allgemeinfall dennoch einer Nachbearbeitung des Tones. Die Originalgeräuschkulisse ist zwar beim Schnitt auf das Zielband gekommen, jedoch würde solch eine Produktion ohne musikalische oder geräuschmäßige Untermalung etwas fade klingen. Da diese Untermalung ja ihre eigene Kontinuität besitzt und oftmals über die Schnittpositionen hinausgeht, muss dieser Ton in einem separaten Arbeitsgang zugemischt werden, oder er ist zumindest beim Online-Schnitt synchron hinzuzusetzen (was aber genau aus Gründen der Diskontinuität in den meisten Situationen nicht machbar ist).

13.3.1 Asynchrone Mischung

Handelt es sich tatsächlich nur um eine Untermalung, die zu ergänzen ist, so ist oftmals keine absolute Synchronisierung erforderlich. Das vereinfacht die Nachbearbeitung

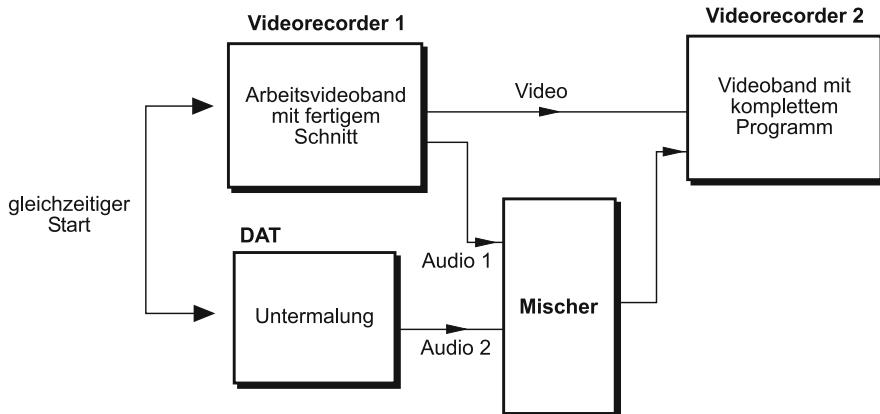


Abb. 13.5 Einfachste, asynchrone Zumischung von untermalendem Ton

dergestalt, dass der Ton bei einem Kopiervorgang zum bestehenden, den Szenen zugehörigen Ton zugemischt werden kann (Abb. 13.5). Beim gleichzeitigen Start von Videorecorder 1 und DAT kann man davon ausgehen, dass beide Bänder innerhalb einer Stunde maximal 1 s auseinanderlaufen. Dieser hohe Gleichlauf besteht jedoch nur dann, wenn es sich um Laufwerke mit Quarzsteuerung handelt, was bei Videorecordern und DATs der Fall ist.

Die grobe Zeitabstimmung muss anhand der Zeitzählerstände von Videorecorder und DAT erfolgen. Nach Beurteilung der zeitlichen Aufteilung des Videoprogramms kann der Zusatzton zusammengestellt und aufgezeichnet werden.

13.3.2 Synchrone Mischung

Eine synchrone Mischung kann durch Zuführen eines zentralen Taktes auf den Videorecorder 1 und den DAT oder im einfacheren Falle durch Synchronisation der DAT-Maschine oder sonstige Audiomaschine durch das Ausgangssignal und den Timecode von Videorecorder 1 erreicht werden. Das setzt natürlich eine entsprechende externe Synchronisierbarkeit des Audiozuspielers voraus. Außerdem muss über das Timecode-Signal ein definierter Start und eine Bezugsposition beider Zuspieler erreicht werden können. Wie die zu ergänzende Musikspur über MIDI synchronisierbar ist, geht aus Abschn. 4.8 hervor.

Direkte Tonspur vom Computer Eine ähnliche Situation ergibt sich für die Synchronisation von Tonspuren, die direkt aus dem Computer via Audiokarte kommen, die allerdings keine MIDI-Arrangements sind. Auch hier muss das Programm, das die Sampling-Dateien bearbeitet und ausgibt, durch den SMPTE-Timecode der Videobandmaschine synchronisiert werden können. Die zeitliche Anpassung während der Editierung des Tones

muss separat vorgenommen werden; hierzu ist ein Szenenplan mit genauen Timecode-Angaben nützlich. Die einzelnen Schritte könnten also wie folgt aussehen:

- Offline-Schnitt ausführen und Rohschnitt anfertigen.
- Nach Schnittliste im Online-Verfahren vorläufige Videoaufzeichnung mit Szenenton erstellen.
- Nach Rohschnitt Szenenplan für Zusatzton mit Timecode-Angaben aufstellen.
- Zusatzton zusammenstellen und nach Szenenplan ausrichten.
- Zusatzton aufzeichnen (DAT mit Timecode).
- Vorläufige Videoaufzeichnung und Zusatzton synchron mischen und endgültiges Programm aufzeichnen.

Generierung des Tons beim Offline-Schnitt Eine weitere Möglichkeit besteht darin, den Zusatzton bereits beim Offline-Schnitt auf die Originalkulisse der Videoszenen zu setzen. Je nach Systemausstattung können hierfür separate Spuren angelegt werden, die jedoch immer absolut synchron zum Originalton sind. Im Falle separater Spuren kann nach vollendetem Schnitt diese Tonspur auf DAT mit Timecode überspielt und nach dem Online-Schnitt synchron zugemischt werden (wieder über einen weiteren Kopievorgang). Die Zusammenstellung des Zusatztons ist beim Offline-Schnittverfahren besonders einfach und komfortabel, weil ja immer ein synchroner Szenenton und das synchrone visuelle Programm verfügbar sind. Toneinsätze lassen sich somit gleich mit dem richtigen Zeitbezug herstellen und trimmen. Die Parameter dieser Tonspur müssen dann ebenfalls mit auf die Schnittliste gelangen, wobei ein korrekter Timecode-Bezug des aufzeichnenden Mediums (DAT o. ä.) erforderlich ist.

Bei dieser Methode entfallen die Arbeitsgänge Rohschnitt anfertigen und Szenenplan aufstellen, was die Bearbeitung wesentlich vereinfacht und beschleunigt.

13.4 Praxisbeispiel mit einfachem Schnittsystem

Zum Schluss dieses Kapitels sei noch ein Beispiel für ein einfaches Schnittsystem genannt. Es handelt sich um ein recht preiswertes System mit dem Namen *Magix Video deluxe*, das jedoch schon an die Möglichkeiten eines professionellen Videoprogramms heranreicht.

Nachdem die Musikdatei als MP3 fertig war, galt diese als feste Spur, nach der sich alle Videoclips und Bildszenen auszurichten hatten. Diverse Gitarrenclips wurden synchron nachgespielt und als Videos an die entsprechenden Stellen in der Timeline positioniert. Das Fein-Trimmen ist in einem solchen Falle einfacher, wenn die nachgespielten Szenen mit Ton aufgezeichnet werden. Siehe Abb. 13.6, 13.7 und 13.8.

Abb. 13.6 Die *Timeline* (oben) repräsentiert alle Clips, Bilder und Audienteile in der gewünschten Reihenfolge. Die Bedienung ist denkbar einfach. Nach dem Import aller Komponenten erfolgt die Anordnung auf der Zeitachse. Dabei ist die grobe Platzierung per Maus nur der erste Schritt. Für frame-genaues Justieren sind passende Werkzeuge vorhanden. Das Abspielen einer Sequenz bei gleichzeitigem Trimmen ist natürlich ebenfalls möglich. Das *Storyboard* (unten) zeigt uns den chronologischen Verlauf aller Filmszenen (Quelle: Magix Video deluxe)



Abb. 13.7 Der kleine Monitor ist natürlich sehr wichtig zur Prüfung des momentanen Standes der Produktion (Quelle: Magix Video deluxe)



Abb. 13.8 Überblendungen von mannigfaltiger Art sowie Betitelung sind sehr wichtige Werkzeuge für ein modernes Schnittsystem. Die Titel sind ebenfalls verschiedenartig animierbar (Quelle: Magix Video deluxe)





14.1 Signalarten-Wandler

Signalarten-Wandler sind immer dann einzusetzen, wenn das externe Videosignal nicht mit der Signalart der Videografikkarte übereinstimmt. Das Funktionsprinzip solcher Wandler ist in Kap. 7 erläutert, so dass wir uns hier auf wesentliche Merkmale und Richtlinien beschränken können.

14.1.1 Wandlung von RGB nach Y/C bzw. FBAS

Diese Wandlung führt ein sog. Encoder bzw. PAL-Encoder durch. Die erreichte Signalqualität hängt maßgeblich von der Technik und vom Abgleich des Encoders ab. Außerdem hat das RGB-Signal Voraussetzungen zu erfüllen, die in den Abschn. 7.2.4 und 9.2.3 erläutert sind.

Liegt ein RGB-Signal z. B. von einer Kamera vor, die Videografikkarte hat jedoch nur einen FBAS- bzw. Y/C-Eingang, so ist das Zwischenschalten eines entsprechenden Konverters wenig sinnvoll, da man eine Doppelwandlung in und somit eine zusätzliche Verschlechterung des Signals Kauf nimmt. Die Rückwandlung in eine RGB-Signal findet nämlich auf der Karte statt. Günstiger wäre es in diesem Fall, eine Karte mit RGB-Eingang zu verwenden (die optimale Lösung bietet ohnehin eine Karte mit möglichst allen Signalarten).

Eine externe Wandlung von RGB nach FBAS bzw. Y/C ist dann zu empfehlen, wenn die Videografikkarte nur ein RGB-Signal ausgibt und das anzuschließende Gerät nur FBAS oder Y/C verarbeiten kann, was für viele Videorecorder zutreffend ist. In diesem Falle ist die Anzahl der Wandlungen auf ein Mindestmaß beschränkt.

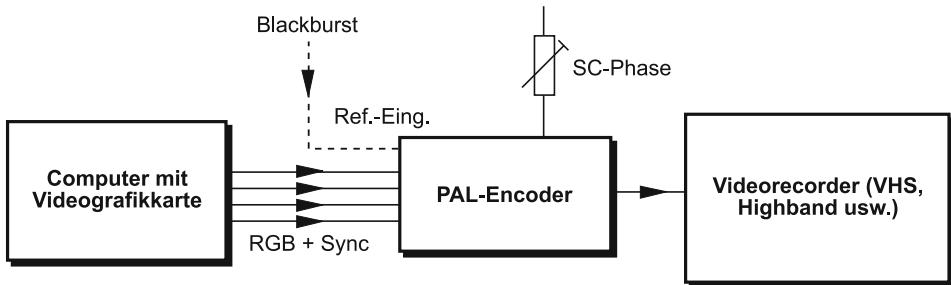


Abb. 14.1 Häufiger Einsatz von PAL-Encodern

Der einzusetzende Encoder sollte für den Einsatz im Studiobereich nicht nur beste Signalqualität aufweisen, sondern auch genlock-fähig sein. Das bedeutet im Detail, dass das Gerät einen Eingang für den Zentraltakt (Blackburst) besitzen muss, und der Phasenwinkel des ausgangsseitigen Farbsignals zeigt einen festen Bezug zum Referenz-Burst. Der Bezugswinkel sollte allerdings einstellbar sein, damit Phasenverschiebungen und somit Farbtonverfälschungen – hervorgerufen z. B. durch die Laufzeit von Leitungen – kompensiert werden können. Abb. 14.1 zeigt die wichtigste Applikation von PAL-Encodern.

Bei ausschließlich dezentraler Anwendung (also mit vom Studiotakt unabhängigen Geräten) ist die Synchronisierung durch ein Blackburst-Signal nicht notwendig. Dennoch gibt es Konstellationen, die auch dann besonders kritisch sind. In Kap. 7 wurde die feste Phasenbeziehung zwischen H- und SC-Phase erwähnt. Eine solche Verkopplung benötigen z. B. MII-Maschinen, obwohl diese als eigentliches Signal YUV aufzeichnen. Wird dieses Referenzsignal nicht unter diesen Bedingungen geliefert, unterbricht eine MII-Maschine i. Allg. den Aufnahmebetrieb. Auch eine 1-Zoll-MAZ stellt normalerweise derartige Anforderungen an ein eingehendes FBAS-Signal.

Bei der Anschaffung von PAL-Encodern sollten die oben genannten Aspekte berücksichtigt werden. Die Signalqualität verlangt von Zeit zu Zeit eine gewissenhafte Überprüfung mittels RGB-Testbild, Oszilloskop und Vektorschop; ggf. ist ein neuer Abgleich nötig, wenn Pegel und Phasenwinkel vom ordnungsgemäßen Vektorgramm abweichen. Mitunter kommt es auch vor, dass Videografikkarten RGB und FBAS am Ausgang anbieten; kommt es jedoch auf höchste Qualität des FBAS-Signals an, so könnte ein externer Encoder in der Anwendung nach Abb. 14.1 bessere Ergebnisse bringen als der auf der Karte integrierte.

Das Zwischensignal Y/C entsteht bei der Aufbereitung nach FBAS. Es bedeutet also keinen besonderen Schaltungsaufwand, auch Y/C als Ausgangssignal anzubieten. Der sich daraus ergebende Vorteil ist die Anschlussmöglichkeit von S-VHS-Recordern, die ja bekanntlich wegen fehlender Cross-Color und -Luminanz eine wesentlich bessere Bildqualität liefern als Recorder, die nur FBAS-Signale akzeptieren.

Ist ein RGB-Signal zu wandeln, welches das Synchronsignal auf Grün liefert, ist die spezielle Ankopplung nach Abb. 14.3 anzuwenden (siehe Abschn. 14.1.4).

14.1.2 Wandlung von FBAS bzw. Y/C nach RGB

Die Wandlung eines FBAS-Signals in ein RGB-Signal übernimmt ein Decoder. Beim PAL-Übertragungssystem ist ein sog. Laufzeitdemodulator nötig, der in der Analogtechnik als Ultraschall-Verzögerungsleitung ausgeführt ist. Auch bei dieser Aufbereitung ergeben sich system- und schaltungsbedingt Verluste, die lediglich auf ein gewisses Maß zu reduzieren, aber nicht ganz auszumerzen sind.

Modernere PAL-Decoder – vor allem auf Grabber-Karten integrierte – führen die Laufzeitdemodulation durch digitale Verzögerung (also Speicherung) einer Zeileninformation durch. Auch hierbei entstehen natürlich Verluste, allerdings ist die nachfolgende, nunmehr digitale Aufbereitung gänzlich reproduzierbar, d. h. ein Drift-Verhalten durch Temperaturschwankungen usw. ist auf dieser Ebene nicht mehr zu erwarten. Abb. 14.2 zeigt die schematische Funktionsweise eines analog arbeitenden PAL-Decoders.

Da in Videografikkarten die Signale meist als digitale RGB-Information benötigt werden, ist ein FBAS-Signal grundsätzlich mit einem Decoder zu wandeln. Besonders eine

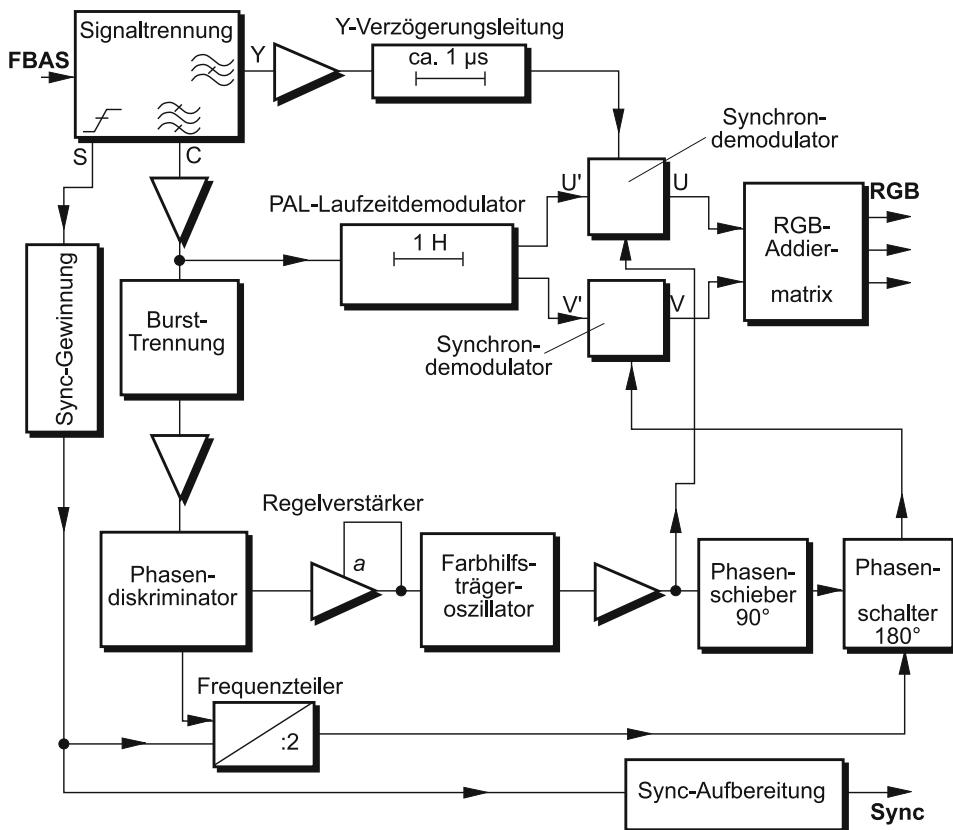


Abb. 14.2 Blockscheme eines PAL-Decoders

solche Einheit ist – bei analoger Arbeitsweise – sehr stark von der Qualität des Abgleichs abhängig. Mit Farbtonverfälschungen, PAL-Jalousie und Pegelfehler seien nur wenige Störungen genannt, die maßgeblich von einer falschen Justage des Decoders herrühren. S. 395 [1]

Auch hier ist bei einem vorhandenen RGB-Eingang die Möglichkeit in Betracht zu ziehen, einen externen, hochwertigeren Decoder zu verwenden.

Decoderqualitäten können nur mit Hilfe von speziellen Signal- bzw. Testbildgeneratoren exakt ermittelt werden, welche oftmals nicht zur Verfügung stehen. Dennoch ist auch mit einfacheren Mitteln schnell festzustellen, ob ein Decoder grobe Mängel aufweist. Folgende Punkte sind bei einem eingefrorenen Bild eines decodierten FBAS-Signals zu beurteilen:

- richtige Wiedergabe aller Farben im Spektrum (kann über histogrammatische Untersuchung verifiziert werden)
- Zeilenstruktur homogener Bereiche (PAL-Jalousie, Pegelausgleich zwischen verzögertem und unverzögertem Signal im Laufzeitdemodulator)
- Wiedergabe grauer Feinstruktur (Y-Frequenzgang)
- Schattenbildung (Signalreflexion)
- Deckung von Farb- und Helligkeitssignal (Y-Verzögerung)

Auch hier ist eine Konvertierung des Y/C-Signals vorzuziehen, falls die Signalquelle ein solches Signal anbietet und der Decoder einen entsprechenden Eingang aufweist.

14.1.3 Komponenten-Transcoder

Im Videostudio wird im Stadium der Produktionsbearbeitung hauptsächlich mit Komponentensignalen YUV gearbeitet. Die wichtigsten Aufzeichnungsgeräte, die dieses Signal verarbeiten sind Betacam und MII, obwohl letztere außer den Komponenten auch noch ein Referenzsignal (Blackburst) benötigt, das noch dazu kritische Bedingungen erfüllen muss.

Zur Übersetzung von YUV nach RGB und umgekehrt kommen sog. Komponenten-Transcoder zum Einsatz, die im Wesentlichen aus Addierstufen und Verstärkern bestehen. Signaltechnisch gesehen sind bei der Umsetzung folgende Gleichungen für die Wandlungen in beide Richtungen notwendig, die bereits in Kap. 7 genannt wurden.

$$\begin{aligned}
 Y &= 0,299 R + 0,587 G + 0,114 B \\
 U &= 0,493 \cdot (B - Y) \\
 V &= 0,877 \cdot (R - Y)
 \end{aligned} \tag{14.1}$$

$$\begin{aligned}
 R &= \frac{V}{0,877} + Y \\
 G &= Y - 0,394 \cdot U - 0,581 \cdot V \\
 B &= \frac{U}{0,493} + Y
 \end{aligned} \tag{14.2}$$

Außer dieser Realisierung mathematischer Sachverhalte durch elektronische Netzwerke ist natürlich noch der Synchronimpuls korrekt umzusetzen. Im RGB-Format liegt entweder eine separate Synchronisationsleitung vor, oder das Synchronsignal ist im Grünsignal integriert. Bei der Umsetzung nach YUV wird dem Leuchtdichtesignal Y das Synchronsignal überlagert.

Die Umsetzung in beide Richtungen bringt für die Bildqualität kaum Verluste, da keine Trägercodierung vorliegt und man somit keine Frequenzumsetzungen oder Filterungen vorzunehmen hat.

Einige Videografikkarten besitzen Anschlüsse für YUV, was vor allem in Studioumgebung recht günstig ist. Die interne Weiterverarbeitung kann mit der Signalart RGB oder auch direkt in YUV geschehen. In jedem Falle ist der elektronische Aufwand nicht hoch, und die zu erwartenden Qualitätseinbußen sind ebenfalls gering.

Bei Verwendung externer Transcoder ist darauf zu achten, dass die Signale U und V durch Abschwächung von B-Y und R-Y gewonnen werden. Bei manchen Geräten ist eine Umschaltung zwischen beiden Pegeln möglich.

14.1.4 Spezielle Sync-Einkopplungen

Soll ein RGB-Signal mit Synchronsignal auf Grün gewandelt werden, so ist hierzu auch ein Wandler mit separatem Sync-Eingang geeignet. Dabei ist die Syncrhoneinkopplung nach Abb. 14.3 vorzunehmen.

Das Grünsignal ist in diesem Falle also einfach auf beide Eingänge zu legen, wobei bei einem der Abschluss von $75\ \Omega$ ausgeschaltet werden muss. Lassen sich nur alle Abschlüsse gemeinsam ausschalten, so können über T-Verteiler externe Einzelabschlüsse angeschlossen werden.

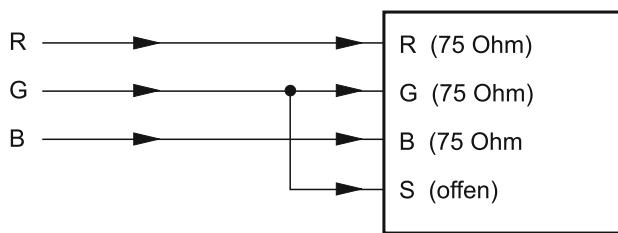


Abb. 14.3 Ankopplung von RGBS-Eingängen an RGB-Signale mit Sync auf Grün

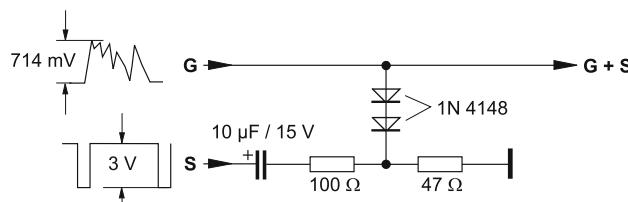


Abb. 14.4 Einkopplung eines TTL-Sync in das Grün-Signal

Der umgekehrte Weg – also die Verschmelzung von Grün- und Synchronsignal – ist normalerweise nicht ohne zusätzliche Bauelemente möglich. Eine entsprechende Konstellation ist aus Abb. 14.4 ersichtlich.

Der Aufwand hält sich in Grenzen, und es sind lediglich passive Bauelemente und keine Stromversorgung notwendig. Eine solche Schaltung lässt sich ohne weiteres in einem Steckergehäuse der Sub-D-Serie unterbringen. Bei anderen Sync-Pegeln muss einer der Widerstände geändert werden, oder beide lassen sich auch als Potentiometer ausführen.

14.2 Normenwandler

Um Bildsignale auf andere Ablenkfrequenzen zu konvertieren, benötigt man sog. Normenwandler, auch bekannt unter der Bezeichnung Scan-Konverter. Soll beispielsweise ein VGA-Signal in ein PAL-Signal überführt werden, sind die Pixelinformationen mit anderer Geschwindigkeit auszulesen – und dies bedeutet sowohl für die Zeilen- als auch die Bildablenkung und -abtastung eine Variation der Taktfrequenzen.

Im einfachsten Falle handelt es sich bei solchen Geräten um digitale Einzelbildspeicher, die getrenntes Ein- und Auslesen erlauben und somit auch unabhängige Taktfrequenzen zulassen. Mit Abb. 14.5 haben wir das Funktionsschema eines Normenwandlers vorliegen. Die eingehenden Signale sind hier bereits getrennt in Farb- und Sync-Signale. Während das Sync-Signal zur Erzeugung des Pixeltaktes und der Zeilen- und Bildumschaltimpulse dient, werden die Farbsignale der Analog-Digitalwandlung zugeführt (entweder getrennte Wandler für jeden Farbkanal oder ein gemeinsamer Wandler mit vorgesetztem Multiplexer). Der Bildspeicher nimmt laufend die digitalisierten Daten auf und gibt sie mit anderem Takt wieder aus, der von einem internen Generator stammt. Nach der Digital-Analogwandlung stehen die Bilddaten nun im gewünschten Takt zur Verfügung.

Unterschiedliche Bildfrequenzen machen jedoch Wiederholungen oder Auslassungen von Vollbildern erforderlich, damit der Bildwechsel nicht in die Zeit der Abtastung fällt und somit Diskontinuitäten sichtbar werden (vgl. auch Abschn. 10.2.2). Während bei NTSC beispielsweise 6 Vollbilder erscheinen, sind es bei PAL lediglich 5. Eine Wandlung in dieser Richtung würde demnach bedeuten, dass ein NTSC-Bild ignoriert werden muss. Bei einer Wandlung von PAL nach NTSC hingegen ist ein PAL-Bild im NTSC-Standard zu

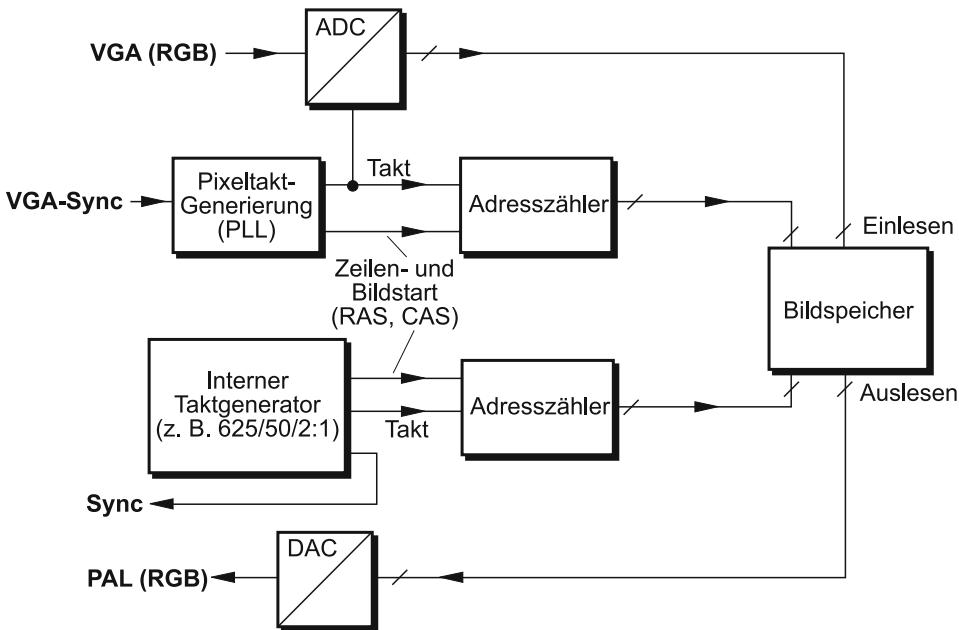


Abb. 14.5 Prinzip eines Normenwandlers

wiederholen. Noch bessere Resultate liefert jedoch ein Verfahren, das die Auswahl nicht nur unter Vollbildern, sondern sogar unter Halbbildern trifft.

Die Steigerung oder Minderung der Zeilenanzahl geschieht meist durch Interpolation, wodurch jedoch die effektive Vertikalauflösung verschlechtert wird. Die beste Qualität ist daher zu erreichen, wenn die Zeilenzahl konstant bleibt. Spezielle Videoprozessoren bewerkstelligen bzw. steuern alle oben genannten Operationen, so dass die Umwandlung selbstverständlich in Echtzeit abläuft. Es leuchtet ein, dass derart komplexe Vorgänge nur noch auf digitaler Ebene stattfinden können. Ist ferner eine Wandlung von bzw. zum FBAS-Signal notwendig, ist am Eingang ein Decoder und am Ausgang ein Encoder des gewünschten Standards zu schalten.

Bei der Ausstrahlung amerikanischer Videoserien sind teilweise die Folgen oben genannter Operation bei der Wandlung zu bemerken. Bei FBAS-Bildern (und um solche handelt es sich ja beim Fernsehbild) sind die Effekte allerdings nicht besonders auffällig. Dennoch ist ein konvertiertes Videobild mit geschultem Auge sofort zu erkennen.

Eine Wandlung, bei der die Auflösung konstant bleibt und nur die Zeilenfrequenz verdoppelt oder halbiert wird, kann mit recht geringem Qualitätsverlust geschehen, weil ja weder Interpolationen noch Bildausschüsse bzw. -verdopplungen erforderlich sind.

14.3 Timebase-Corrector

In Abschn. 10.1.1 kam bereits der Timebase-Corrector (TBC) zur Sprache, der ein mit Jitter behaftetes Videosignal derart aufbereitet, dass es mit quarzstabilem Zeitverhalten wieder ausgegeben wird.

Zeitliche Schwankungen während der Wiedergabe kommen vor allem bei Videorecordern vor, da durch die Abtastung durch ein mechanisches System wie dem Kopfrad samt Antrieb keine hohe Geschwindigkeitsstabilität zu gewährleisten ist. Die Langzeitstabilität ist zwar durch die Servo erreichbar, während einer Kopfradumdrehung jedoch ist durchaus mit nicht zu vernachlässigenden Gleichlaufschwankungen zu rechnen. Bei modernen Geräten mit Halbleiterspeichern sind keine Jitter-Effekte zu erwarten und somit ist auch kein TBC notwendig, da der Datenstrom bereits phasenstabil zur Verfügung gestellt werden kann.

Vor allem Video-Grabber-Karten macht solch ein Signal am Eingang zu schaffen, falls diese nicht über einen internen TBC verfügen. Schließlich darf man nicht vergessen, dass die für das Genlock-System notwendige PLL schnelle Schwankungen nicht unverzögert mitmachen kann. Die Folge davon sind Bildverzüge, die sich wie in Abb. 10.2 äußern.

Auch der TBC setzt – ähnlich wie der Normenwandler – einen Bildspeicher ein, in den mit zeitlichem Jitter eingelesen und aus dem mit Quarzreferenz wieder ausgelesen wird. Die Größe des Bildspeichers richtet sich nach den Anforderungen an den TBC: zur Unterdrückung von Bildverzügen genügt i. Allg. eine Speichergröße zur Aufnahme von wenigen Zeilen, während die Verschiebung bzw. Ausrichtung eines Bildes an einer Taktreferenz eine Speichergröße für ein gesamtes Bild erfordert. Externe Geräte der höheren Preisklasse setzen meist Speichergrößen für ein Vollbild ein, um den Einsatzgebiete möglichst universell zu halten.

Nun kann natürlich der TBC nicht einen Ausgangstakt wählen, der vollkommen unabhängig vom Eingangstakt ist. Ist der TBC nicht extern synchronisiert, so kann er zwar dennoch ein quarzgenaues Ausgangssignal erzeugen, welches ein vom Eingangssignal abweichendes Timing zeigt. Allerdings dreht die Videokopftrommel trotzdem mit einer Drehzahl von 25 Hz, die durch eine interne Quarzreferenz stabil gehalten wird. Der aus dem Videosignal gewonnene Bildwechseltakt ist langfristig gesehen also quarzgenau, er unterliegt nur kurzfristigen Schwankungen. Über einen trügen PLL-Regelkreis ist also daraus ein synchroner, quarzgenauer Bildtakt zu gewinnen, der auch kurzfristig keine stärkeren Schwankungen mehr zeigt.

Eine externe Synchronisation ist ebenfalls möglich; bei Studioeinbindung ist dies sicherlich ein Muss, da sich das Bildsignal einem einheitlichen Takt unterwerfen muss. Die horizontale Bildlage lässt sich meist am TBC ebenfalls einstellen. Abb. 14.6 zeigt die einfache Verschaltung bei interner und externer Synchronisation. Die prinzipielle Arbeitsweise eines TBC geht aus Abb. 10.3 hervor, wobei allerdings noch A/D- bzw. D/A-Wandlungen an den Bildspeicher anschließen.

Bei einer Betacam-Maschine (siehe Abschn. 7.3.4) ist kein externer TBC für die Kompensation von Zeitfehlern notwendig, weil dieser bereits im Gerät vorhanden ist.

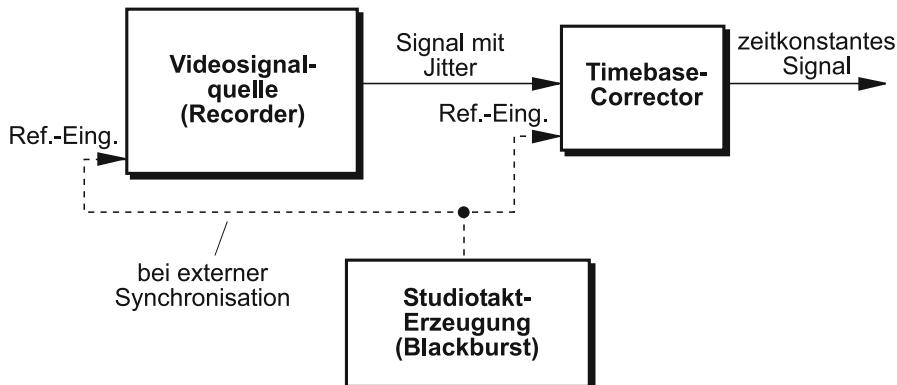


Abb. 14.6 Verschaltung und Synchronisation von TBCs

Doch grundsätzlich können auch bei quarzstabilen Signalen TBCs zur Anwendung kommen, wenn es um die horizontale Verschiebung des Bildes nach links geht, was ja einer virtuellen zeitlichen Vorverlegung gleichkommt. Für eine Verschiebung nach rechts hingegen genügen einfache Verzögerungsglieder, auf welche wir im nächsten Abschnitt noch zu sprechen kommen.

TBCs zeigen auch Unterschiede in Bezug auf die eingangs- bzw. ausgangsseitigen Signalarten. Neben FBAS- und Y/C-Signalen sollten aus Kompatibilitätsgründen auch ungeträgte Signale (RGB oder YUV) zur Verfügung stehen. Die Qualität von FBAS und Y/C hängt natürlich auch hier vom Niveau und vom Abgleich der internen Decoder und Encoder.

Natürlich lässt sich nicht verschweigen, dass die Wandlung von analog nach digital und zurück mit einem gewissen Verlust verbunden ist; das Zwischenschalten eines TBC korrigiert zwar das Zeitverhalten, führt aber zu einer geringfügigen Verschlechterung der Signalübertragung im Frequenzverhalten.

14.4 Verzögerungsglieder

In der Videotechnik kommt es häufig zu statischen zeitlichen Verschiebungen einzelner Signale zueinander. Die Gründe hierfür liegen in der unterschiedlichen Bandbreite oder Aufbereitung von Signalen. Doch auch gleichartige Signale unterliegen einer Zeitverschiebung, wenn sie Leitungen unterschiedlicher Länge passieren müssen.

Besonders auffällig äußert sich dieser Effekt, wenn der zeitliche Offset zwischen einem Key-Signal und dem Stanzsignal besteht. Das zu integrierende Signal wird dann nicht mehr an den richtigen Stellen aufgeschaltet.

Um dem zu begegnen, kann die Laufzeit voreilender Signale künstlich mit Verzögerungsgliedern dem zeitlich zuletzt eintreffenden Signal angeglichen werden. Solche Bausteine lassen sich digital oder analog aufbauen.

In der analogen Arbeitsweise sind sogar passive Glieder möglich, die lediglich in den Signalweg einzukoppeln sind und keinerlei Versorgungsspannung benötigen. Allerdings ist dann auch mit einer gewissen Dämpfung zu rechnen. Das verzögernde Glied besteht hierbei aus einem Allpass bzw. einer Kette aus LC-Gliedern nach Abb. 14.7.

Abhängig von der Anzahl der LC-Glieder ergibt sich eine bestimmte Laufzeit. Es gibt daher Verzögerungsglieder, die eine stufenweise Umschaltung der Laufzeit erlauben. Es sind aber auch Ausführungen erhältlich, die feste Laufzeiten aufweisen und einer Staffelung nach Zweierpotenzen unterliegen. Ein Sortiment hiervon erlaubt ebenfalls eine sehr genaue Annäherung an eine bestimmte Verzögerungszeit. Die oben angegebene Methode der Verzögerung findet bei der Verzögerung des Leuchtdichtesignals gegenüber dem Farbsignal in PAL-Decodern und -Encodern statt.

Anderen Technologien setzen analoge oder digitale Schieberegister ein, wobei die Länge des Schieberegisters und die Taktfrequenz die Verzögerungszeit bestimmen. Bei der digitalen Ausführung muss das Register die notwendige Tiefe besitzen, und außerdem sind natürlich A/D- und D/A-Wandlung notwendig, wie dies in Abb. 14.8 zu sehen ist.

Die Einstellung der Taktfrequenz ist natürlich nur in einem gewissen Rahmen möglich, da das Abtasttheorem immer erfüllt sein muss. Außerdem kann es auch zu Zwischenfrequenzen mit angeschlossenen, abtastenden Systemen kommen.

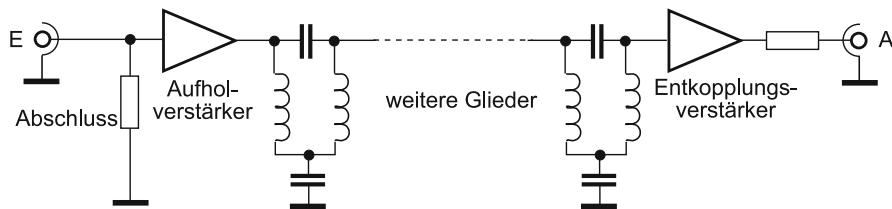


Abb. 14.7 Aufbau analoger Verzögerungsglieder

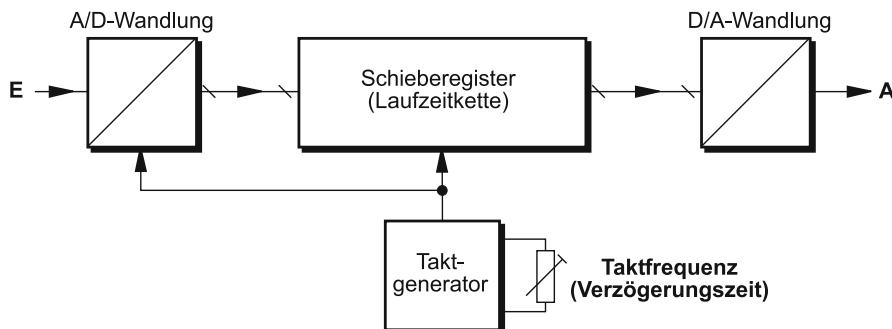


Abb. 14.8 Digitales Laufzeitglied

14.5 Synchrongenerator

Oft ist es vorteilhaft oder sogar zwingend notwendig, dass alle Videogeräte (so auch die Videografikkarte usw.) mit einem einheitlichen, phasenstarren Bild- und Zeilentakt arbeiten. Eine solche Synchronisation erfordert meist einen speziellen Generator, der nicht nur Horizontal- und Vertikalimpulse liefert, sondern auch die Burst-Impulse für die Farbsynchronisation bei FBAS- bzw. Y/C-Systemen.

Eine funktionsfähige Synchronisierung ergibt sich jedoch nur, wenn alle Geräte über einen entsprechenden Referenzeingang verfügen. Bei Video-Bandmaschinen ist dies meist der FBAS-Eingang. Videografikkarten müssen ebenfalls einen solchen Eingang vorweisen und außerdem die Möglichkeit des Genlock-Betriebs besitzen.

Im Offline-Betrieb, wo nur zwei Videogeräte zusammengeschaltet sind, ist ein zentraler Takt meist verzichtbar, doch bereits bei Verwendung von drei Geräten kann die Verkettung des Synchronsignals Probleme bereiten. Eine sternförmige Verbindung mit einem zentralen Synchrongenerator bietet hier eine Lösung – jedes Gerät wird dann mit dem Blackburst- oder einfachen Synchronsignal versorgt, und Durchlaufverzögerungen der Impulse wirken sich nicht aus. Dennoch ist zu bedenken, dass bereits unterschiedliche Leitungslängen eine merkliche Taktverschiebung verursachen können.

Ein Synchrongenerator bietet mehrere entkoppelte Ausgänge an, die zwar identische Signale liefern, aber dennoch unabhängig voneinander belastet werden können. Wichtig ist bei der Wahl eines solchen Gerätes auch die Anzahl der verfügbaren, entkoppelten Ausgänge. Weiterhin sollte das Burst-Signal eine feste H-Phasenverkopplung nach der 8er-Sequenz für PAL aufweisen. Sinnvoll sind auch Ausgänge ohne Burst-Signal (Composite Sync); mitunter sind bei den einzelnen Ausgängen einfach die Burst-Signale schaltbar.

Literatur

1. Stotz, D.: Audio- und Videogeräte richtig einmessen und justieren.
München: Franzis-Verlag 1994.

Anhang A: Audio-Datenformate

In diesem Teil des Anhangs sind Aufbau bzw. Arrangement von Audiodaten innerhalb der Dateien und den im Header geführten Informationen dargestellt.

Leider gibt es – wie auch übrigens bei Grafikdatenformaten – keinen einheitlichen Standard, obwohl es im Audiobereich weitaus weniger Unterschiede in Bezug auf Anwendungsbereich und Entstehung als im Videobereich gibt. So existiert z. B. bei Sampling-Dateien kein Format, das der Vektor-Orientierung bei Grafiken entspricht. Dagegen folgt aus dem Verhalten bei verlustfreier Kompression, dass z. B. das bloße Aneinanderreihen von 16-Bit-Werten und ggf. die Alternierung der Stereokanaldaten völlig ungeeignet ist.

Weitaus bessere Methoden gliedern die Daten in Pakete auf. Darüber hinaus ist die Speicherung der Direktdaten des PCM-Codes weniger vorteilhaft als vielmehr die Intervalldaten, also die Unterschiede zwischen den ursprünglichen PCM-Werten benachbarter Abtastpunkte.

Das zweite Handicap ist die fehlende Kompatibilität von Audiodateien mit gleicher Datei-Extension. Man muss also auch hier teilweise eine weiterführende Spezifikation angeben.

A.1 Format WAV

Für den Windows-Standard hat sich hauptsächlich das WAVE-Format durchgesetzt. Die Anordnung der Daten gestaltet sich recht einfach: Nach dem Header (meist aus 44 Bytes bestehend, siehe weiter unten Tab. A.1) schließen sich direkt die Datenbytes an. Hier ist zu berücksichtigen, dass auch im Datenbereich bei 16-Bit-Auflösung zunächst das LSB, dann das MSB erscheint. Bei Stereodateien alternieren ständig die Daten für den linken und den rechten Kanal, wobei mit dem linken Kanal begonnen wird. Bei 12-Bit-Samples

erfolgt ebenfalls i. Allg. eine Speicherung im 16-Bit-Format, wobei die oberen 4 Bits des MSB (Nibble) auf 0 gesetzt sind. Durch diese direkte Datenstruktur ermöglicht das WAVE-Format eine einfache Möglichkeit der Wiedergabe.

A.2 Format VOC

Dieses Format ist ein Standard der Fa. *Creative Labs*. Gegenüber dem WAVE-Format ist der den Daten vorangehende Header vergleichsweise einfach; dagegen ist die sich anschließende Datenstruktur etwas komplexer, u. a. weil es unterschiedliche Subblock-Arten gibt.

Nach dem Offset, der in hier Byte 14 und 15 definiert ist, beginnt der erste Subblock an Adresse 1A. Nach der unten folgenden Tabelle gibt es 8 verschiedene Subblöcke, wobei jedoch nicht jeder für sich alleine stehen oder in beliebiger Reihenfolge erscheinen kann. Daraus geht hervor, dass hier neben Audiodaten auch Text, Markierungen usw. enthalten sein können. Außerdem besteht die Möglichkeit einer Kompression, die jedoch nicht blockweise variieren kann, d. h. über die gesamte Datei konstant ist. Siehe Tab. A.2 und A.3.

A.3 Format SMP

Es handelt sich um das Format SampleVision der amerikanischen Firma *Turtle Beach Systems*. Es ist vorwiegend zum Einsatz in Sampler-Geräten gedacht und besitzt deshalb vorläufig nur einen Kanal (Mono) in 16-Bit-Auflösung. Nachfolgende Tab. A.4 zeigt den Datenaufbau, der sich hier nicht nur auf den Header beschränkt, sondern sozusagen auch einen Footer besitzt.

Vorzugsweise wird das SMP-Format in MIDI-Applikationen bzw. Sampler-Geräten eingesetzt; man erkennt dies vor allem am MIDI-Noten-Bezug, dem einstellbaren SMPTE-Offset und den setzbaren Schleifenpunkten.

A.4 Format SFI/SFD

Ebenfalls ein Format der amerikanischen Firma *Turtle Beach Systems* ist SoundStage – ein auf spezielle Workstations angepasstes Format für Harddisk-Recording. Es werden zwei Dateien angelegt, nämlich die Informationsdatei (SFI = Sound File Info) und die eigentliche Sampling-Datei (SFD = Sound File Data); letztere beinhaltet lediglich die Abtastwerte im Intel-Format (zuerst LSB, dann MSB), und bei Stereo-Samples alternieren L- und R-Abtastwerte, wobei mit dem linken Kanal begonnen wird.

Bei Mono-Dateien entfällt natürlich die Alternierung von Kanälen, und jedes 16-Bit-Wort entspricht auch einer Abtastung. Den Aufbau der SFI-Datei, die alle Parameter der SFD-Datei enthält, entnehmen wir nachfolgender Tab. A.5.

Die SFI-Datei lässt sich eigentlich in zwei Bereiche unterteilen, nämlich in den Header und den Bereich der definierten Marken. Letzterer fehlt bei einfachen Dateien, in denen keine Marken definiert sind, sonst schließen sich deren Informationen direkt an den SFI-Header an. Über den Aufbau einer einzelnen Marke gibt der untere Teil von Tab. A.6 Aufschluss.

Die SFI-Datei beinhaltet zwei Informationen von jeweils Byte-Größe, die einerseits den Typus der SFD-Datei und andererseits die Art der Marken beschreibt. Während der SFD-Typ für eine Datei uniform ist, kann für jede Marke, die im Anschluss an den SFI-Header definiert ist, eine individuelle Spezifizierung erfolgen, was aus Tab. A.6 hervorgeht.

Im Gegensatz zum SMP-Format kann das SFD-Format alle denkbaren Kanalkonstellationen verarbeiten, selbst eine Kanalextraktion aus einer Stereodatei ist ohne weiteres möglich, ohne dass die eindeutige Zuordnung über ein Kennzeichen verlorengeht. Die Auflösung ist jedoch auf 16 Bit beschränkt; ein Export einer 8-Bit-Sampling-Datei bringt deswegen zwangsläufig etwa eine Verdopplung des Speicherbedarfs.

A.5 Format AIFF

Dieses Format entstammt einer relativ frühen Standardisierung und heißt Audio Interchange Fileformat. Es war ursprünglich gedacht für einen einfachen Austausch von Audiodateien zwischen unterschiedlichen Anwenderprogrammen und sogar verschiedenen Computer-Plattformen. Siehe nachfolgende Tab. A.7.

Diese Angaben und auch der Aufbau der Daten selbst sind nicht gesichert, da die Angaben hier sehr stark differieren. Es existieren bei diesem Format noch weitere Definitionen, die sich jedoch auf MIDI-Daten beziehen. Es erfolgt dann anstatt der SNDD-Kennung eine entsprechende Variante.

A.6 Format ADPCM

Ein Format, welches eine Kompression vornimmt, stellt ADPCM (Adaptive Delta Pulse Code Modulation) dar, das von Microsoft konzipiert wurde und ebenfalls wie das Wave-Format die Dateiendung **WAV** aufweist. Die primär bei der Abtastung entstehenden Werte kommen hier nicht direkt zur Abspeicherung; es werden vielmehr zunächst Delta-Werte (Differenzen benachbarter Abtastwerte) gebildet, die ja normalerweise kleiner sind als die Abtastwerte selbst. Danach findet eine nicht ganz verlustfreie Kompression statt.

Es handelt sich in der Tat um eine Abwandlung des Wave-Formats, was sich bei der Betrachtung des Header bestätigt. Dieser zeigt große Übereinstimmung mit dem in Tab. A.1, wobei allerdings die Länge des zweiten Blocks größer ist, der auch Daten für die Kompression festlegt. Ebenfalls ergibt sich eine Änderung für den Formattyp, dessen Wert von 1 auf 2 wechselt.

Die zusätzlichen Daten sowie der Algorithmus für die Kompression, mit der man einen Kompressionsfaktor von bis zu 1 : 4 erhält, ist allerdings nicht geklärt. Da es sich um eine verlustbehaftete Kompression handelt, beschränkt sich das Verfahren nicht auf eine Huffman-Codierung (Umcodierung der häufig vorkommenden Werte durch verkürzte Worte). Vielmehr ist eine Quantisierung anzunehmen, welche irreversible Verluste entstehen lässt (vgl. Abschn. 12.8.1, JPEG-Verfahren). Der Grad des Verlustes richtet sich weitgehend nach dem Inhalt der Originaldatei, da der Kompressionsgrad konstant bleibt.

Die Daten selbst werden in Blöcke von je 1024 Bytes zusammengefasst, die durch 10H getrennt sind. Diese Blocklänge erscheint ebenfalls im Header.

Die Daten können in Echtzeit allerdings nur durch den entsprechenden Codec geschrieben bzw. gelesen werden. Die schaffen dies im allgemeinen ohne separaten DSP (Digital Signal Processor).

A.7 Speicherbedarf der Dateien

Bereits in Kap. 2 war zu erfahren, wie groß die zu erwartende Datenrate sein wird (Gl. 2.7); danach ergibt sich eine Dateigröße von:

$$M = K \cdot F_A \cdot t \cdot D / 8 \quad [\text{Byte}] \quad (\text{A.1})$$

K Kanalanzahl, F_A Abtastrate in 1/s, t Zeit in s, D Auflösung in Bit.

Für eine Minute benötigt man also für ein Stereoformat mit 16-Bit-Auflösung und einer Abtastrate von 44 100 Hz einen Speicherplatz von:

$$M = 2 \cdot 44100 \cdot 60 \cdot 18 / 8 \text{ Byte} \approx 10 \text{ MByte}$$

Die oben durchgeführte Rechnung berücksichtigt nicht den Platzbedarf, den ein Header erfordert, was sich jedoch unerheblich auswirkt. Außerdem ist hier Vorbedingung, dass keine Kompression stattfindet.

Kompressionsalgorithmen, die für den Echtzeitbetrieb ausgerichtet sind, sollten eine dynamische Betriebsart erlauben, die bei leisen Stellen auch den Zugriffshäufigkeit auf die Festplatte reduziert.

Tab. A.1 Header des Audiodatenformats WAV

Byte Nr. Hex	Festwert Hex	Bedeutung	Bemerkung
00–03	52, 49, 46, 46	‘RIFF’	WAV-Kennung
4	–	Dateilänge –8	Byte 0
5	–	Dateilänge –8	Byte 1
6	–	Dateilänge –8	Byte 2
7	–	Dateilänge –8	Byte 3
08–0E	57, 41, 56, 45, 66, 6D, 74	‘WAVEfmt’	WAV-Kennung
0F	20	Blank	–
10	10	Länge des nächsten Blocks	Byte 0
11	0		Byte 1
12	0		Byte 2
13	0		Byte 3
14	1	Formattyp	LSB
15	0	Formattyp	MSB
16	–	Anzahl der Kanäle	LSB
17	–	Anzahl der Kanäle	MSB
18	–	Einfache Sampleclock	Byte 0
19	–	Einfache Sampleclock	Byte 1
1A	–	Einfache Sampleclock	Byte 2
1B	–	Einfache Sampleclock	Byte 3
1C	–	Datenbytes/sec	Byte 0
1D	–	Datenbytes/sec	Byte 1
1E	–	Datenbytes/sec	Byte 2
1F	–	Datenbytes/sec	Byte 3
20	–	Gesamtdatenbytes/Sample	LSB
21	–	Gesamtdatenbytes/Sample	MSB
22	–	Auflösung	LSB
23	–	Auflösung	MSB
24–27	64, 61, 74, 61	‘data’	Kennung für Datenanfang
28	–	Gesamtanz. Datenbytes	Byte 0
29	–	Gesamtanz. Datenbytes	Byte 1
2A	–	Gesamtanz. Datenbytes	Byte 2
2B	–	Gesamtanz. Datenbytes	Byte 3
2C	–	Erstes Datenbyte	LSB bei 16-Bit-Auflösung

Tab. A.2 Header des Audiodatenformats VOC

Byte Nr. Hex	Festwert Hex	Bedeutung	Bemerkung
00–12	43, 72, 65, 61, 74, 69, 76, 65, 20, 56, 6F, 69, 63, 65, 20, 46, 69, 6C, 65	‘Creative Voice File’	VOC-Kennung
13	1A	Abschluss der Kennung	–
14–15	1A, 00	Offset des ersten Subblocks	LSB, MSB
16	0A	Formattyp	LSB
17	1	Formattyp	MSB
18	29	Prüfsumme	LSB
19	11	Prüfsumme	MSB
1A	–	Subblock-Anfang	

Tab. A.3 Aufbau der Subblöcke beim VOC-Format

Byte Nr. Hex	Festwert Hex	Bedeutung	Bemerkung
	Terminator	Symbol für Dateiende	
1A	0	Kennzeichen Terminator	
	Voice data	Erster Audiodatenblock	
1A	1	Kennzeichen Voice Data	
1B	–	Länge des Datenbereichs	Byte 0
1C	–	dto.	Byte 1
1D	–	dto.	Byte 2
1E	–	codierte Sampling-Rate	
1F	–	Kompressionsart	
20–...	–	Audiodaten ...	
	Voice continuation	Folgender Audiodatenblock	
1A	2	Kennzeichen Voice Continuation	
1B	–	Länge des Datenbereichs	Byte 0
1C	–	dto.	Byte 1
1D	–	dto.	Byte 2
1E–...	–	Audiodaten ...	
	Silence	Sequenz mit Nulldaten (Ruhe)	
1A	3	Kennzeichen Silence	
1B	–	Länge des Datenbereichs	Byte 0
1C	–	dto.	Byte 1
1D	–	dto.	Byte 2
1E	–	Pausenlänge in Samples – 1	Byte 0
1F	–	dto.	Byte 1
20	–	codierte Sampling-Rate	

(Fortsetzung)

Tab. A.3 Fortsetzung

Byte Nr. Hex	Festwert Hex	Bedeutung	Bemerkung
	Marker	Bereichsmarkierung	
1A	4	Kennzeichen Marker	
1B	2	Länge des Datenbereichs	
1C	0	dto.	
1D	0	dto.	
1E	–	Markierungsbereich	Byte 0
1F	–	dto.	Byte 1
	ASCII-Text	Textbereich	
1A	5	Kennzeichen ASCII-Text	
1B	2	Länge des Datenbereichs	Byte 0
1C	0	dto.	Byte 1
1D	0	dto.	Byte 2
1E–...	–	Zeichenkette ASCII	
	Repeat loop	Beginn einer Schleife	
1A	6	Kennzeichen Repeat Loop	
1B	2	Länge des Datenbereichs	Byte 0
1C	0	dto.	Byte 1
1D	0	dto.	Byte 2
1E	–	Anzahl der Wiederholungen –1	Byte 0
1F	–	dto.	Byte 1
	End repeat loop	Ende einer Schleife	
1A	7	Kennzeichen End Repeat Loop	
1B	0	Länge des Datenbereichs	Byte 0
1C	0	dto.	Byte 1
1D	0	dto.	Byte 2

Tab. A.4 Aufbau des Audio-Datenformats SampleVision (SMP)

Byte Nr. Hex	Festwert Hex	Bedeutung	Bemerkung
00–11	53, 4F, 55, 4E, 44, 20, 53, 41, 4D, 50, 4C, 45, 20, 44, 41, 54, 41, 20	‘SOUND SAMPLE DATA’	SMP-Kennung
12–15	32, 2E, 31, 20	‘2.1’	Version
16–51	–	Anwender-Kommentar	60 Zeichen
52–6F	–	Sample-Name	30 Zeichen
70	s	Sample-Größe	Byte 0
71	s	dto.	Byte 1
72	s	dto.	Byte 2

(Fortsetzung)

Tab. A.4 Fortsetzung

Byte Nr. Hex	Festwert Hex	Bedeutung	Bemerkung
73	s	dto.	Byte 3
74	–	Audiodaten (1. Wert)	LSB
75	–	Audiodaten (1. Wert)	MSB
:	:	:	:
72 + 2s	–	Audiodaten (letzter Wert)	LSB
73 + 2s	–	Audiodaten (letzter Wert)	MSB
74 + 2s	–	reserviert	
75 + 2s	–	reserviert	
76 + 2s	–	Anfang Schleife 1	Byte 0
77 + 2s	–	dto.	Byte 1
78 + 2s	–	dto.	Byte 2
79 + 2s	–	dto.	Byte 3
7A + 2s	–	Ende Schleife 1	Byte 0
7B + 2s	–	dto.	Byte 1
7C + 2s	–	dto.	Byte 2
7D + 2s	–	dto.	Byte 3
7E + 2s	0 = aus, 1 = vorwärts, 2 = vorwärts/rückwärts	Typ Schleife 1	
7F + 2s	–	Wiederholungen Schleife 1	LSB
80 + 2s	–	dto.	MSB
81 + 2s ... 8A + 2s	–	Name von Marke 1	10 Zeichen
8B + 2s	–	Position Marke 1	Byte 0
8C + 2s	–	dto.	Byte 1
8D + 2s	–	dto.	Byte 2
8E + 2s	–	dto.	Byte 3
8F + 2s	–	MIDI-Referenz-Note	
90 + 2s	–	Sampling-Frequenz	Byte 0
91 + 2s	–	dto.	Byte 1
92 + 2s	–	dto.	Byte 2
93 + 2s	–	dto.	Byte 3
94 + 2s	–	SMPTE-Offset Subframes	Byte 0
95 + 2s	–	dto.	Byte 1
96 + 2s	–	dto.	Byte 2
97 + 2s	–	dto.	Byte 3
98 + 2s	–	Zyklus-Größe in Samples	Byte 0
99 + 2s	–	dto.	Byte 1
9A + 2s	–	dto.	Byte 2
9B + 2s	–	dto.	Byte 3

Tab. A.5 Aufbau der SFI-Datei

Byte Nr. Hex	Festwert Hex	Bedeutung	Bemerkung
	Header		
0	4	Zahl der Dateneinheiten für Marken (Marker Records)	
01–16	53, 6F, 75, 6E, 64, 20, 46, 69, 6C, 65, 20, 49, 6E, 66, 6F, 72, 6D, 61, 74, 69, 6F, 6E	‘Sound File Information’	SFI-Kennung
17–1B	20, 31, 2E, 30, 20	‘1.0’	Version
1C–57		Kommentar	60 Zeichen
58		Sampling-Frequenz	Byte 0
59		dto.	Byte 1
5A		dto.	Byte 2
5B		dto.	Byte 3
5C		SFD-Länge	Byte 0
5D		dto.	Byte 1
5E		dto.	Byte 2
5F		dto.	Byte 3
60		SMPTE-Offset (Stunden)	
61		SMPTE-Offset (Minuten)	
62		SMPTE-Offset (Sekunden)	
63		SMPTE-Offset (Bilder)	
64–67		reserviert	
68		MIDI-Referenz-Note	
69	0 = aus, 1 = ein	Emphasis	
6A	0 = aus, 1 = ein	Kopierschutz	
6B		SFD-Typ	s. nächste Tabelle
6C		Anzahl definierter Marken	
6D		Tempo	
6E–8B		reserviert für spätere Erweiterungen	
	Marken		
8C	–	Markentyp	s. nächste Tabelle
8D–98	–	Markenname	12 Zeichen
99	–	Startbild	Byte 0
9A	–	dto.	Byte 1
9B	–	dto.	Byte 2
9C	–	dto.	Byte 3
9D	–	Anzahl der Bilder	Byte 0
9E	–	dto.	Byte 1
9F	–	dto.	Byte 2
A0	–	dto.	Byte 3
	ggf. weitere Marken		

Tab. A.6 SFD- und Markentypen

Parameter	Bedeutung	Wert (Hex)
Audiodaten (SFD-Typ)		
linker Kanal	0	
rechter Kanal	1	
Stereo	40	
Mono	41	
Marken-Typ		
Schleife	0	
Marke	1	
Zone	2	
Text	3	
sonstiges	FE	
nicht verwendet	FF	

Tab. A.7 Header des Audio-Datenformats AIFF

Offset Hex	Festwert Hex	Bedeutung	Bemerkung
	Datei-Header		
00–03	43, 4F, 4D, 4D	COMM'	Kennung
04–07	–	Blocklänge in Byte	
08–09	–	Anzahl der Kanäle	
0A–0D	–	Anzahl der Rahmen (Frames)	
0E–0F	–	Auflösung/Sample	
10–19	–	Frames/Sekunde	
	Erster Datenblock-Header		
00–03 (rel.)	53, 4E, 44, 44	‘SSND’	SoundData-Kennung
04–07 (rel.)	–	Blocklänge in Byte	
08–0B (rel.)	–	Offset zum Rahmen	
0C–0F (rel.)	–	Größe des Rahmen	
	Daten		
	Letzter Datenblock-Header		
00–03 (rel.)	53, 4E, 44, 44	‘SSND’	SoundData-Kennung
04–07 (rel.)	–	Blocklänge in Byte	
08–0B (rel.)	–	Offset zum Rahmen	
0C–0F (rel.)	–	Größe des Rahmen	
	Daten		
	:	:	:

Anhang B: Tabellen und Diagramme

Abb. B.1 Kurven gleicher Lautstärke

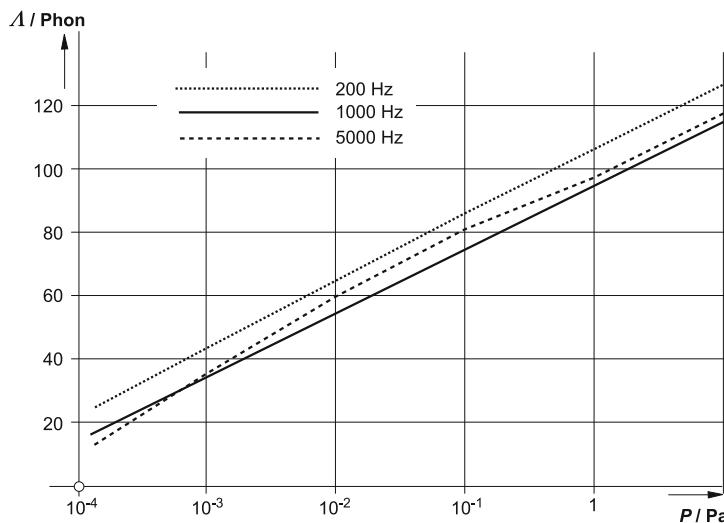
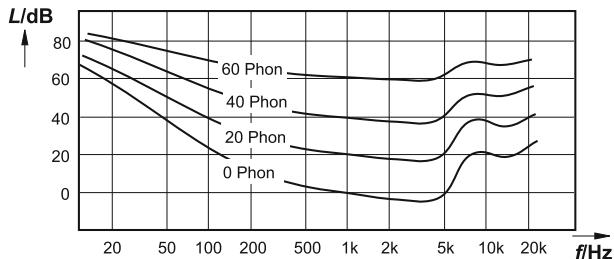


Abb. B.2 Abhängigkeit der Lautstärke vom Schalldruck bei verschiedenen Frequenzen

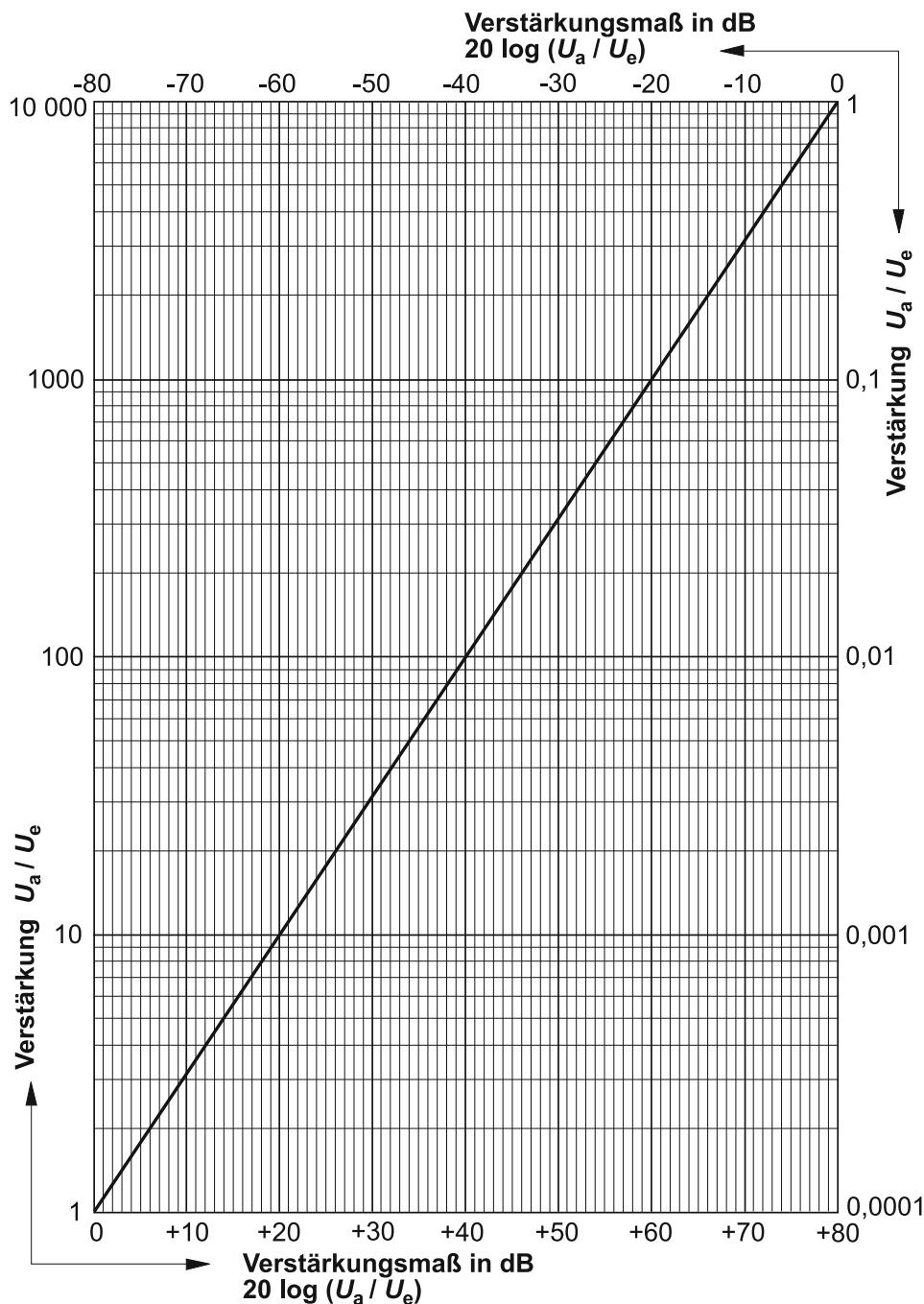


Abb. B.3 Verstärkung und logarithmisches Verstärkungsmaß

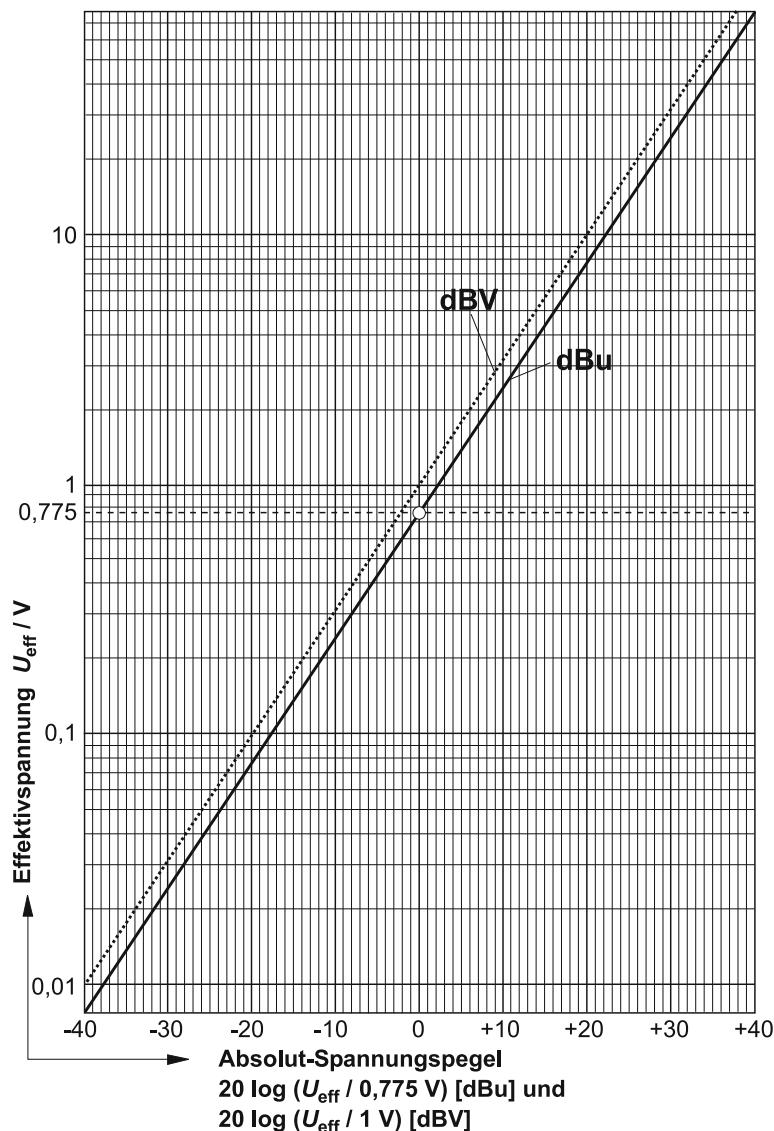


Abb. B.4 Effektivspannung und absoluter Spannungspegel

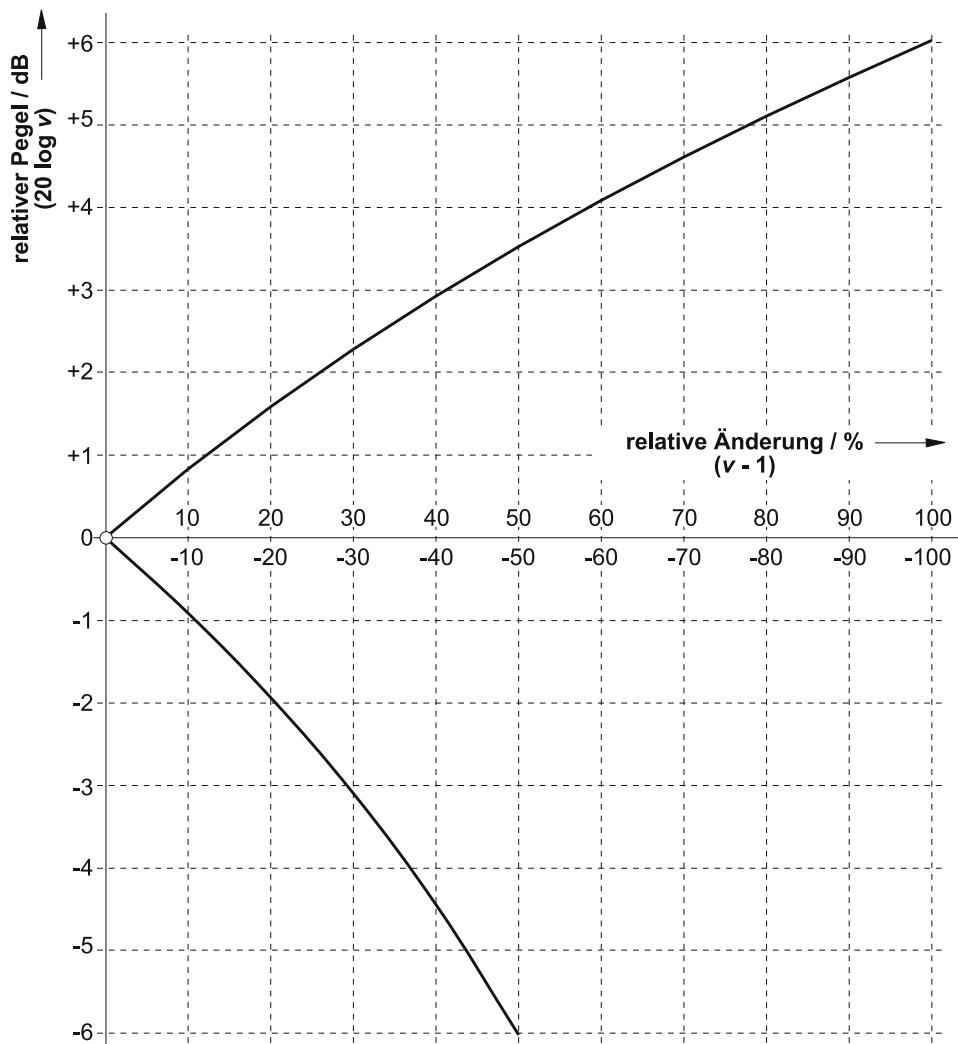


Abb. B.5 Relativpegel in Abhängigkeit von der prozentualen Änderung der Spannung

Tab. B.1 Gegenüberstellung der Systeme PAL und NTSC

Größe/Funktion	Einheit	PAL (B/G)	NTSC (M)
Frequenz der Halbbilder	Hz	50	59,94
Zeilenfrequenz	Hz	15625	15734,26
Gesamtzeilenzahl pro Vollbild	–	625	525
Abtastfrequenz für Digitalanwendungen			
Luminanz	MHz	13,5	13,5
Chrominanz ($R - Y, B - Y$)	MHz	6,75	6,75
Zahl der Abtastungen pro Gesamtzeile bei der Luminanz	–	864	858
Mindestzahl der Abtastungen pro aktive Zeilendarstellung bei der Luminanz (nach CCIR 601)	–	720	720
Zeilenauastastzeit	μs	$12 \pm 0,3$	$10,8 \pm 0,2$
Beziehung zwischen Horizontalfrequenz f_h und Abtastfrequenz f_L bei der Luminanz	–	$f_h \cdot 144 = f_L/6 = 2,25$	$f_h \cdot 143 = f_L/6 = 2,25$
Beziehung zwischen Farbhilfsträgerfrequenz f_{sc} und Horizontalfrequenz f_h	–	$f_{sc} = \frac{1135}{4} + (\frac{1}{625}) \cdot f_h$	$f_{sc} = \frac{455}{2} \cdot f_h$
Farbhilfsträgerfrequenz	Hz	$4\,433\,618,75 \pm 5$	$3\,579\,545 \pm 10$
Burst-Schwingungen	–	10 ± 1	≥ 8
Burst-Phasenlage gegenüber U -Achse	°	± 135	180
Automat. Farbtonfehler-Kompensation	–	ja	nein
Decodierung	–	Laufzeit-Demodulation über Ultraschall-Verzögerungsleitung	direkt
Videobandbreite	MHz	5	4,2
Bildaustastdauer	Zeilen	25	$19 \dots 21$
Farbhilfsachsen	–	keine	Q bei 33°
			I bei 123°

Tab. B.2 Gegenüberstellung der Systeme PAL und HDTV Eureka 95 S. 432 [1]

Größe/Funktion	Einheit	PAL (B/G)	HDTV
Bildformat	–	4:3	16:9
Zeilenfrequenz	Hz	15625	62 500 ^a
Gesamtzeilenzahl pro Vollbild	–	625	1250
Sichtbare Zeilen	–	≤ 576	1150
Zahl der Abtastungen pro Gesamtzeile	–	864 ^b	–
Gültige Bildpunkte pro Zeile	–	768 (720 ^b)	2035
Abtastfrequenz für Digitalanwendungen			
Luminanz	MHz	13,5	
Chrominanz ($R - Y$, $B - Y$)	MHz	6,75	

^aProduktionsstandard mit Abtastung 1:1 (Non-Interlaced)^bNach CCIR 601 für digitale Studios**Tab. B.3** Konstanten in der Audiotechnik und Akustik

Größe	Einheit	Wert	Bemerkung
Mindestschalldruck p_0	Pa	$2 \cdot 10^{-5}$	bei 1000 Hz (Bezugsleistung 0 Phon)
Schmerzgrenze der Lautstärke	Phon	ca. 130	
Kammerton A	Hz	440	Tonhöhenreferenz
Spannung bei			
Absolutpegel 0 dBu	V	0,775	

Tab. B.4 Konstanten in der Audiotechnik und Akustik

Größe	Einheit	Wert	Bemerkung
Gesamtspannungshub zwischen Sync-Niveau und Weißwert	V	1	
Weißpegel gegenüber Schwarzniveau	V	1	
Wellenwiderstand der verwendeten Leitungen	Ω	75	
Maximale Farbsättigung	%	100	bei PAL (B/G)
Maximaler Reinfarbpegel	%	75	bei PAL (B/G)

Anhang C: Beschaltungen

In diesem Teil des Anhangs sind einige Beschaltungsbeispiele und -festlegungen erläutert. Dabei sollen nicht nur Belegungen von Steckverbindungen zur Sprache kommen, sondern auch die Erkennung bzw. Vermeidung von Problemen im Zusammenhang mit der Zusammenschaltung von Geräten.

Abb. C.1 Entstehung und

Vermeidung von
Brummschleifen.

(a) Bezugspunkte der
Signalquelle sowie des
angeschlossenen Gerätes sind
z. B. über den Netzzschluss an
Erdmasse geführt. (b) Dadurch
entsteht ein geschlossener
Stromkreis, wobei ein Teil der
Brummspannung U_{Br} am
Kabelwiderstand R_{K2} abfällt.
Dieser Teil addiert sich zur
Nutzsignalspannung. U_{Br} kann
durch Induktion oder
Störströme durch R_{K1}
entstehen. Im unteren Bereich
der Abbildung sind
Möglichkeiten zur Vermeidung
von Brummschleifen gezeigt

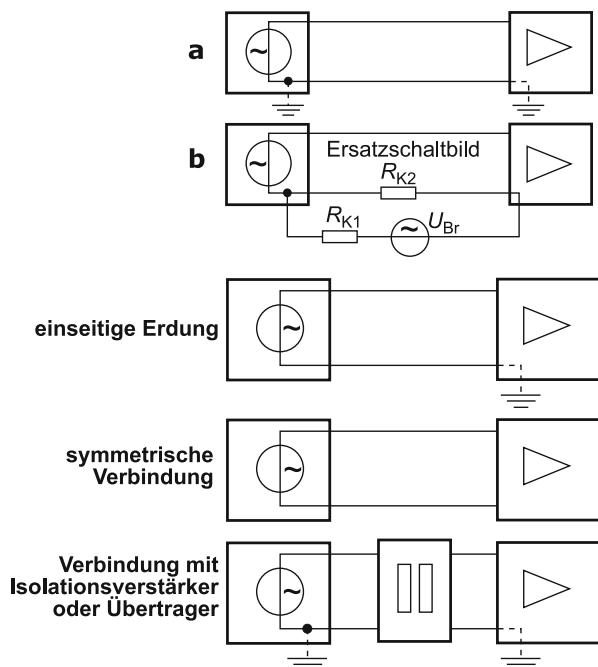
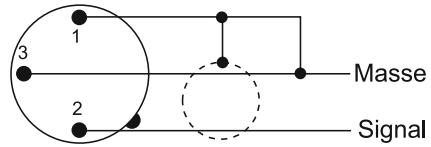
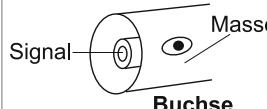
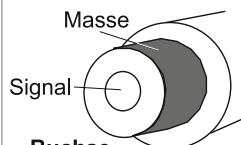
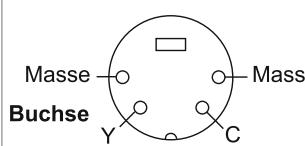
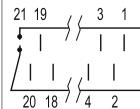
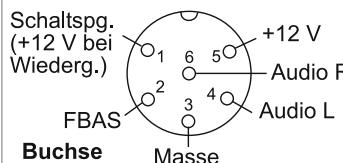
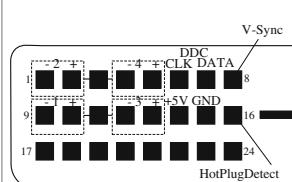


Abb. C.2 Symmetrischer Anschluss XLR auf asymmetrischen Anschluss



Tab. C.1 Belegung verschiedener Steckverbindungen im Videobereich

Norm	Polzahl	Zweck	Anschlussbild (Draufsicht auf Kontaktseite)																																																
BNC (koaxial)	2	Verbindung eines Einzelsignals (z. B. FBAS)	 Buchse																																																
Cinch	2	Verbindung eines Einzelsignals (z. B. FBAS)	 Buchse																																																
Mini-DIN	4	Verbindung des Y/C-Signals (S-VHS-Geräte)	 Buchse																																																
Scart	21	Kombination sverbindung mit Audio und Einschleifmöglichkeit	 <table> <tbody> <tr><td>21</td><td>19</td><td>3</td><td>1</td></tr> <tr><td>20</td><td>18</td><td>4</td><td>2</td></tr> </tbody> </table> <table> <tbody> <tr><td>1</td><td>Audio-Out R</td><td>12</td><td>Datenleitung 1</td></tr> <tr><td>2</td><td>Audio-In R</td><td>13</td><td>Rot-Masse</td></tr> <tr><td>3</td><td>Audio-Out L</td><td>14</td><td>Reserviert</td></tr> <tr><td>4</td><td>Audio-Masse</td><td>15</td><td>Rot</td></tr> <tr><td>5</td><td>Blau-Masse</td><td>16</td><td>Sync</td></tr> <tr><td>6</td><td>Audio-In L</td><td>17</td><td>FBAS-Masse</td></tr> <tr><td>7</td><td>Blau</td><td>18</td><td>Sync-Masse</td></tr> <tr><td>8</td><td>Schaltspg.</td><td>19</td><td>FBAS-Out</td></tr> <tr><td>9</td><td>Grün-Masse</td><td>20</td><td>FBAS-In</td></tr> <tr><td>10</td><td>Grün</td><td>21</td><td>Abschirmung</td></tr> </tbody> </table>	21	19	3	1	20	18	4	2	1	Audio-Out R	12	Datenleitung 1	2	Audio-In R	13	Rot-Masse	3	Audio-Out L	14	Reserviert	4	Audio-Masse	15	Rot	5	Blau-Masse	16	Sync	6	Audio-In L	17	FBAS-Masse	7	Blau	18	Sync-Masse	8	Schaltspg.	19	FBAS-Out	9	Grün-Masse	20	FBAS-In	10	Grün	21	Abschirmung
21	19	3	1																																																
20	18	4	2																																																
1	Audio-Out R	12	Datenleitung 1																																																
2	Audio-In R	13	Rot-Masse																																																
3	Audio-Out L	14	Reserviert																																																
4	Audio-Masse	15	Rot																																																
5	Blau-Masse	16	Sync																																																
6	Audio-In L	17	FBAS-Masse																																																
7	Blau	18	Sync-Masse																																																
8	Schaltspg.	19	FBAS-Out																																																
9	Grün-Masse	20	FBAS-In																																																
10	Grün	21	Abschirmung																																																
DIN 45322	6	Kombination sverbindung mit Audio und FBAS für Heimgeräte																																																	
DVI-D-Anschluss	24	Übertragung digitaler Bilddaten zu einem Wiedergabegerät	 <table> <tbody> <tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td></tr> <tr><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td></tr> <tr><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td></tr> <tr><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td></tr> </tbody> </table> <p>V-Sync DDC CLK DATA 5V GND HotPlugDetect</p>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24																								
1	2	3	4	5	6																																														
7	8	9	10	11	12																																														
13	14	15	16	17	18																																														
19	20	21	22	23	24																																														

(Fortsetzung)

Tab. C.1 Fortsetzung

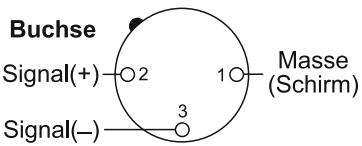
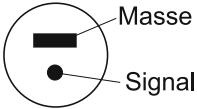
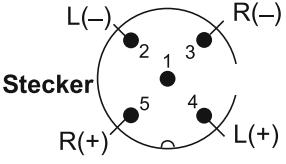
Norm	Polzahl	Zweck	Anschlussbild (Draufsicht auf Kontaktseite)
HDMI-Type-A-Anschluss	19	Übertragung digitaler Audio- und Videodaten	

Tab. C.2 Belegung verschiedener Steckverbindungen im Audiobereich

Norm	Polzahl	Zweck	Anschlussbild (Draufsicht auf Kontaktseite)
DIN 41524 'Diode'	5	Stereo Input/Output Kleinsignalverbindung asymmetrisch	
DIN 41524 'Diode'	3	Mono Input/Output Kleinsignalverbindung asymmetrisch	
Cinch	2	Mono Input oder Output Kleinsignalverbindung asymmetrisch	
Klinke 6,35 mm	2	Mono Input oder Output Kleinsignal- oder Lautsprecherverbindung asymmetrisch	
Klinke 6,35 mm	3	Stereo Input oder Output Kleinsignal- oder Kopfhörerverbindung asymmetrisch	

(Fortsetzung)

Tab. C.2 Fortsetzung

Norm	Polzahl	Zweck	Anschlussbild (Draufsicht auf Kontaktseite)
Klinke 3,5 mm	2	Mono Input oder Output Kleinsignalverbindung asymmetrisch	siehe Klinke 6,35 mm Mono
Klinke 3,5 mm	3	Stereo Input oder Output Kleinsignal- oder Kopfhörerverbindung asymmetrisch	siehe Klinke 6,35 mm Stereo
XLR 'Cannon'	3	Mono Input oder Output Kleinsignalverbindung symmetrisch	Buchse 
DIN 41529	2	Mono Output Lautsprecherverbindung	
DIN 45327 'Würfel 5'	5	Stereo Output Kopfhörerverbindung	Stecker 

Tab. C.3 Einige wichtige Mikrofon-Beschaltungen

Verbindung	Beschreibung	Anschlussbild
DIN 41524 'Diode'	Asymmetrisch	
DIN 41524 'Diode'	Symmetrische Beschaltung nach DIN 45594	
DIN 41624 'Tuchel'	Symmetrisch	
XLR 'Cannon'	Symmetrisch	
versch.	Phantomspeisung nach DIN 45595 eines Kondensatormikrofons	

Anhang D: Grafikformate

Dieses Kapitel kann nicht als Referenz für Grafikformate dienen; dafür sind letztere zu umfangreich, und auch bestimmte Formate sind viel zu komplex, um sie vollständig zu beschreiben. Dennoch sollen die wichtigsten zur Sprache kommen, wobei die Besonderheiten bereits aus dem Header hervorgehen. Für weitergehende Informationen und Details sei auf spezielle Literatur verwiesen. S. 432 [2].

Die Datenchronologie ist – falls nicht ausdrücklich erwähnt – nach dem Intel-Format ausgerichtet, d. h. bei Word-Daten erscheint zunächst das LSB, dann das MSB.

D.1 Kurzer Überblick

In den folgenden Abschnitten finden wir Informationen über den Aufbau der Datei-Header und somit über den Aufbau der Datei selbst.

Es sollen einige sehr häufig anzutreffende Raster-Grafikformate behandelt werden, die auch in vielen Programmen anzutreffen sind: TIFF, BMP, TGA, PCX und PICT. Während TIFF und TGA völlig unabhängig von Betriebssystemen sind, gehören BMP und PCX dem Betriebssystem MS-Windows an. Mit PICT soll dann noch ein Apple-Format zur Sprache kommen.

Abschließend werfen wir dann noch einen Blick auf das JPEG-Format, obwohl es kein eigentliches Grafikformat, sondern ein Verfahren zur Aufbereitung und Kompression von Bilddaten ist.

Mit dem Wissen über den Aufbau der Formate erhalten teilweise sicherlich auch Entwickler von Programmen Unterstützung. Dennoch muss bedacht werden, dass manche Formate bei verschiedenen Applikationen auch geringfügige Abweichungen vom Standard aufweisen. Ferner sind natürlich auch Weiterentwicklungen der Formate möglich, die nicht immer völlig abwärtskompatibel sind.

D.2 Format TIFF

Das Tag Image File Format TIFF stammt von der Aldus Corporation und ist eines der ältesten und am meisten verbreiteten Formate im Grafikbereich. Es erlaubt alle Farbtiefen bis 32 Bit. Ferner erlaubt TIFF die Verarbeitung in unterschiedlichen Farbmodellen, also nicht nur in RGB, was im Druckgrafikbereich sehr nützlich ist. Außerdem unterstützt TIFF zahlreiche Kompressionsarten und kann in einer Datei mehrere Grafiken speichern.

Ein einzelnes Bild ist in Felder oder Streifen aufgeteilt, und jeder Streifen erhält ein Etikett, ein sog. Tag, das aber nicht im Block der eigentlichen Bilddaten platziert wird, sondern direkt nach dem festen Teil des Header. Die Anzahl der Streifen und somit der Tags muss also festgelegt sein, und zwar am Anfang jeder sich anschließenden Image-File-Directory (IFD). Jeder IFD-Eintrag entspricht einem Tag, dessen Identifikation vorangestellt ist. Aus Abb. D.1 geht der Aufbau hervor, und Tab. D.1 definiert Details.

Wie man sieht, können mehrere IFDs existieren – es kann jedoch auch eine einzige ausreichend sein. Jede IFD beschreibt genau ein Bild und ist unterteilt in Tag-Einträge, und jeder Tag-Eintrag ist wiederum (fest) unterteilt in 12 Bytes. Existiert nur eine IFD, so nimmt der Zeiger im unteren Teil den Wert 0 an.

Die in der IFD verschlüsselten Tag-Einträge müssen mit aufsteigenden Tag-IDs erscheinen. Dabei sind natürlich nur die notwendigen Tags zu beschreiben, allerdings sind manche hiervon zwingend notwendig. Tab. D.2 zeigt eine Liste der derzeit möglichen und definierten Tag-Einträge eines IFD-Eintrags.

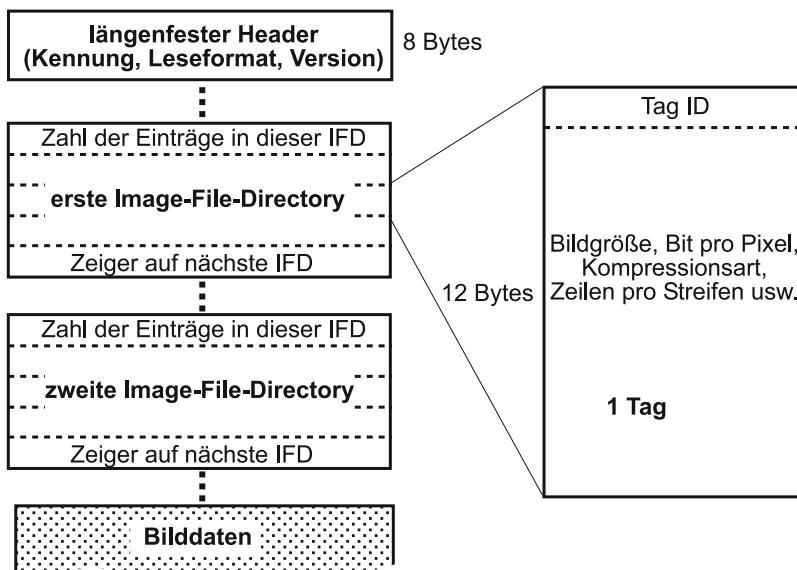


Abb. D.1 Aufbau einer Grafik im TIFF-Format

Tab. D.1 Header des Grafikdatenformats TIFF

Offset Hex	Länge	Festwert Hex	Bedeutung
00	Word	<i>längenkonstanter Header</i> Intel: 4949, Motorola: 4D4D	Leseformat (Byte-Wertigk.)
02	Word	002A	Identifikation
04	DWord		Offset der ersten IFD
00 + m	Word	<i>Image-File-Directory</i>	Anzahl der IFD-Einträge
02 + m	Word		Tag-ID
04 + m	Word	1 = Byte, 2 = ASCII, 3 = SHORT (16 Bit), 4 = LONG (32 Bit), 5 = RATIONAL (2 × LONG), 6 = SBYTE (Byte mit Vorzeichen), 8 = SSHORT (16 Bit mit Vorz.), 9 = SLONG (32 Bit mit Vorz.), 10 = SRATIONAL (wie 5 mit Vorz.), 11 = FLOAT (IEEE einfach), 12 = DOUBLE (IEEE doppelt)	Feldtyp (für Berechnungen weiter unten: f_1 = Feldlänge in Byte)
06 + m	DWord		Anzahl der Werte n_1
0A + m	DWord		Wenn $n_1 \cdot f_1 > 4$, dann Offset zum Wert, sonst Wert selbst
0E + m	Word		Tag-ID
10 + m	Word	s. o.	Feldtyp
12 + m	DWord		Anzahl der Werte n_2
16 + m	DWord		Wenn $n_2 \cdot f_2 > 4$, dann Offset zum Wert, sonst Wert selbst
		<i>weitere, insges. t IFD-Einträge (Tags)</i>	
⋮	⋮	⋮	⋮
$t \cdot 12 + 2 + m$	DWord	0: keine weitere IFD	Zeiger auf nächste IFD
		<i>Bilddaten</i>	
⋮	⋮	⋮	⋮

Tab. D.2 Tags des Grafikdatenformats TIFF

Tag-ID Hex	Anzahl Werte	Feldtyp ^a	Bedeutung
FE	1	4	Neuer Unterdatei-Typ
FF	1	3	Unterdatei-Typ
100	1	3 od. 4	Bildbreite in Pixel (Spaltenzahl)
101	1	3 od. 4	Bildhöhe in Pixel (Zeilenzahl)
102	var. ^b	3	Anzahl Bits pro Pixel (bpp)
103	1	3	Kompressionsart (Werte z. B.: 1: keine, 5: LZW, 6: JPEG)
106	1	3	Photometrische Interpretation (Werte: 0: Weiß = Min, 1: Schwarz = Min, 3: RGB-Palette, 4: Transparenz, 5: CMYK-Modell, 6: YCbCr-Modell)
107	1	3	Schwellenmodus (Threshholding) (Werte: 1: Bilevel, 2: Halftone, 3: Error Diffused)
108	1	3	Zellenbreite in Pixel bei Dithering
109	1	3	Zellenhöhe in Pixel bei Dithering
10A	1	3	Füllordnung (Werte: 1 (Standard): MSBit auf erstes Pixel, 2: LSBit auf erstes Pixel)
10D	undef.	2	Name des Dokuments
10E	undef.	2	Bildbeschreibung (verbal)
10F	undef.	2	Scanner-Hersteller
110	undef.	2	Modell (des verwendeten Gerätes usw.)
111	var. ^c	3 od. 4	Streifen-Offset (relativ zum Dateianfang)
112	1	3	Orientierung des Ursprungs (Werte: formatrichtig: 1: links oben, 2: rechts oben, 3: rechts unten, 4: links unten; mit Formatwechsel: 5: links oben, 6: rechts oben, 7: rechts unten, 8: links unten)
115	1	3	Anzahl der Werte pro Pixel (Werte: 1 für Schwarzweiß, 3 für RGB) = n/p
116	1	3 od. 4	Zeilen pro Streifen
117	var. ^c	3 od. 4	Byte-Anzahl pro Streifen
118	n/p	3	Min-Wert für photometrische Interpretation
119	n/p	3	Max-Wert für photometrische Interpretation
11A	1	5	Auflösung in X-Richtung (Pixel pro Einheit)
11B	1	5	Auflösung in Y-Richtung (Pixel pro Einheit)
11C	1	3	Planar-Konfiguration der Farbinformation (Werte: 1 farbalternativ, 2: zeilenalternativ)
11D	undef.	2	Seitenname
11E	1	5	X-Position (Abstand/Inch des Bildes von der linken Seite)
11F	1	5	Y-Position (Abstand/Inch des Bildes vom oben)
120	undef.	4	Freier Offset (Adressen freier Blöcke)

(Fortsetzung)

Tab. D.2 Fortsetzung

Tag-ID Hex	Anzahl Werte	Feldtyp ^a	Bedeutung
121	undef.	4	Freie Bytes
122	1	3	Grauverlauf-Einheit (Werte: x: 10-x für x 1...5)
123	2bpp	3	Grauverlauf-Kurve (Wert: Zeiger auf Tabelle)
124	1	4	Optionen für Gruppe 3
125	1	4	Optionen für Gruppe 4
128	1	3	Einheit der Auflösung (Werte: 1: undef., 2: Inch, 3: cm)
129	2	3	Anzahl der Seiten (1. Wert: Seite, 2. Wert: Seitenanzahl)
12D	undef.	3	Transfer-Funktion
131	undef.	2	verwendete Software
132	20	2	Datum
13B	undef.	2	Künstler
13C	undef.	2	Computer
13D	1	3	Predictor (Werte: 1: keine Bearbeitung, 2: zeilenweise Differenzbildung bei LZW)
13E	2	5	Weißpunkt
13F	6	5	Primärfarb-Eigenschaften
140	32bpp	3	Farbtabelle (Colormap)
141	2	3	Halbtön-Zusatz
142	1	3 od. 4	TileWidth
143	1	3 od. 4	TileLength
144	var. ^d	4	TileOffset
145	var. ^d	3 od. 4	TileByteCounts
14C	1	3	Tintensatz
14D	undef.	2	Tintenbezeichnung
14E	1	3	Anzahl der Tinten
150	2	1 od. 3	Punktbereich
151	undef.	2	Zieldrucker
152	var.	1	Sonderabtastungen (Zeiger auf Tabelle)
153	var. ^e	3	Sample-Format
154	var. ^e	var.	SMinSampleValue
155	var. ^e	var.	SMaxSampleValue
156	6	3	Transfer-Bereich
200	1	3	JPEG-Prozess (Werte: 1: sequentiell DCT, 2: Huffman)
201	1	4	JPEGInterchangeFormat (für den JPEG-Datenaustausch)
202	1	4	Länge des JPEG-Datenfeldes
203	1	3	JPEGRestartInterval (Zahl der Minimumcode-Einheiten)

(Fortsetzung)

Tab. D.2 Fortsetzung

Tag-ID Hex	Anzahl Werte	Feldtyp ^a	Bedeutung
205	var. ^e	3	JPEGLosslessPredictor
206	var. ^e	3	JPEGPointTransforms
207	var. ^e	4	JPEG-Q-Tabellen
208	var. ^e	4	JPEG-DC-Tabellen
209	var. ^e	4	JPEG-AC-Tabellen
211	3	5	YCbCr-Koeffizienten
212	2	3	YCbCr-Unterabtastung
213	1	3	YCbCr-Positionierung
214	var. ^f	4	Referenz Schwarz/Weiß

^aSiehe Tab. D.1^bIst bei Monochrombildern 1, bei Mehrfarbbildern mit unterschiedlichen Bit-Tiefen ein Zeiger-Offset, der auf ein Feld mit den entsprechenden Angaben gerichtet ist^cFür nur einen Datenblock (Anzahl Werte = 1) ist der Offset direkt im Wert enthalten, ansonsten richtet sich der angegebene Zeiger auf ein Feld mit entsprechenden Angaben^dTiles per Image^eSamples pro Pixel^f3. Samples pro Pixel**Tab. D.3** Fileheader des Grafikdatenformats BMP

Offset Hex	Länge	Festwert Hex	Bedeutung
00	Word	4D42	Dateikennung 'BM'
02	DWord		Größe der Bilddatei in Byte
06	Word		Reserviert
08	Word		Reserviert
0A	DWord		Byte-Offset des Fileheader zu den Bilddaten

D.3 Format BMP

Hier handelt es sich um das Bitmap-Format BMP unter dem Betriebssystem MS-Windows (für das IBM-Betriebssystem OS/2 gibt es hierfür eine gesonderte Definition). Es wurde von Microsoft entwickelt. Mit BMP wurde beabsichtigt, ein geräteunabhängiges Format zu schaffen (Device Independent Bitmap = DIB). S. 432 [3].

Die Beschreibung über den Informationsaufbau der Datei ist unterteilt in Bitmap-Fileheader, -Infoheader und RGBQuad. In den folgenden Tabellen sind diese Datenfelder entschlüsselt. Das Leseformat entspricht hier durchweg dem Intel-Format, also zuerst LSB, dann MSB. Siehe Tab. D.3, Tab. D.4 und Tab. D.5.

Die Farbtabelle schließt sich normalerweise unmittelbar an die beiden Header-Teile an, allerdings lässt sich mit dem Wert für die Größe des Infoheader eine Verschiebung des

Tab. D.4 Infoheader des Grafikdatenformats BMP

Offset Hex	Länge	Festwert Hex	Bedeutung
0E	DWord		Größe des Infoheader in Byte
12	DWord		Bildbreite in Pixel (Kolumnen)
16	DWord		Bildhöhe in Pixel (Zeilen)
1A	Word	1	Anzahl der Bildebenen
1C	Word	1, 4, 8 oder 18	Bit pro Pixel (Farbtiefe)
1E	DWord	0: ohne Kompression 1: Lauflängencodierung RLE8 2: Lauflängencodierung RLE4	Kompressionsart
22	DWord		Anzahl der Bilddaten in Byte
26	DWord		Hor. Auflösung in Pixel pro Meter
2A	DWord		Vert. Auflösung in Pixel pro Meter
2E	DWord		Anzahl der benutzten Farbwerte in der Farbtabelle
32	DWord	0: alle	Anzahl der für die Wiedergabe wichtigen Farben

Tab. D.5 RGBQUAD-Farbtabelle des Grafikdatenformats BMP

Offset Hex	Länge	Farbvektornummer	Bedeutung
36	Byte	0	Blauwert
37	Byte		Grünwert
38	Byte		Rotwert
39	Byte		Reserviert (auf 0 gesetzt)
⋮	⋮	⋮	⋮
$36 + m^a \cdot 4$	Byte	$m - 1$	Blauwert
$37 + m^a \cdot 4$	Byte		Grünwert
$38 + m^a \cdot 4$	Byte		Rotwert
$39 + m^a \cdot 4$			Reserviert (auf 0 gesetzt)

^a $m = \text{Anzahl der benutzten Farben}$

Zugriffs bewirken. Bei einer Farbtiefe von 24 Bit entfällt die Farbtabelle, weil hier die Bilddaten direkt für die Farbinformationen stehen.

Der Ursprung einer BMP-Grafik befindet sich links unten. Die Daten einer Zeile sind am Format LONG ausgerichtet, d. h. die Anzahl der Bytes pro Zeile ist ein ganzzahliges Vielfaches von 4. Jede Zeile beginnt mit einem auf null gesetzten Byte, welches zur oberen Zählung mitgerechnet wird.

Tab. D.6 Farbtiefen und Farbcodierung beim Grafikdatenformats BMP

Farbtiefe in Bit	Anzahl der Farben	Codierung	Beschreibung
1	2	über Farbtabelle (indiziert)	Jedes Bit im Bilddatenbereich bestimmt die Auswahl der Farbdefinition in der Farbtabelle: Ist das Bit = 0, ist der erste Farbeintrag zu wählen, ist das Bit = 1, so ist der zweite Eintrag der Tabelle zu wählen
4	16	über Farbtabelle (indiziert)	Jedes Nibble (Halbbyte) im Bilddatenbereich bestimmt die Auswahl aus der Farbtabelle: Der Nibble-Wert entspricht der Farbvektornummer, die wiederum korrespondiert mit derjenigen Stelle in der Farbtabelle, die die Farbzusammensetzung bestimmt. Die Adressierung erfolgt mithin prinzipiell so wie oben
8	256	über Farbtabelle (indiziert)	Jedes Byte im Bilddatenbereich bestimmt die Auswahl aus der Farbtabelle. Die Adressierung erfolgt mithin prinzipiell so wie oben
24	16 777 216	direkt	Es wird keine Farbtabelle benötigt. Die Kennzeichnung einer Farbe erfolgt auf direktem Wege: Für jedes Pixel werden 4 Bytes bereitgestellt, welche in der Reihenfolge Blau, Grün, Rot, Reserviert die Zusammensetzung der jeweiligen Farbe definieren

Bei einer indizierten Farbcodierung (bei Farbtiefen <8 Bit) entsprechen die Bilddaten der Farbvektornummer der Farbtabelle (z. B. bei 16 benutzten Farben, also bei einer Farbtiefe von 4 Bit, gibt jedes Nibble die Nummer des Vektors der Farbtabelle an, der die im Pixel verwendete Farbe beschreibt). Tab. D.6 erläutert die Farbcodierung für die einzelnen Farbtiefen. Die verwendete Farbtiefe in einer BMP-Grafik ist als Wert im Infoheader niedergeschrieben.

Für die Komprimierung von Bilddaten sind zur Zeit nur zwei Modi definiert, die 4-Bit- bzw. 8-Bit-Daten mit einer Lauflängencodierung behandelt, welche im Encoded Mode aus zwei Bytes im Datenbereich besteht, wovon das erste die Anzahl der nachfolgenden Pixel angibt und das zweite den Index in der Tabelle bestimmt. Eine Unterbrechung dieser Codierung wird durch eine Null für das erste Byte erzwungen, und die Bedeutung für

Tab. D.7 Header des Grafikdatenformats TGA

Offset Hex	Festwert Hex	Bedeutung
00	–	Länge des ID-Feldes
01	0: ohne Farbtabelle 1: mit Farbtabelle	Art der Farbtabelle
02	0: keine Bilddaten 1: mit Farbtabelle 2: True-Color-Bild 3: Schwarzweißbild 9: mit Farbtabelle u. RLE ^a A: True-Color-Bild mit RLE ^a B: Schwarzweißbild mit RLE ^a	Bild-Typ
03,04		Index des ersten Eintrags
05,06		Zahl der Farbtabelleneinträge
07	0F, 10, 18 oder 20	Bit pro Eintrag
08,09		X-Offset der Ecke unten links
0A, 0B		Y-Offset der Ecke unten links
0C, 0D		Hor. Pixelzahl (Kolumnen)
0E, 0F		Vertikale Pixelzahl (Zeilen)
10		Bit pro Pixel (Farbtiefe)
11	Bit 0 – 3: Attribute pro Pixel Bit 4: Ursprungposition hor. Bit 5: Ursprungposition vert. Bit 6 – 7: reserviert	Bildbeschreibung 0: links, 1: rechts 0: unten, 1: oben
12		Erste Daten (ID-, Farbtabellen- oder Bilddaten)

^aRLE = Run Length Encoding (Lauflängencodierung für Kompression)

den Abbruch bestimmt das zweite Byte (0: Zeilenende, 1: BMP-Ende, 2: Deltafunktion mit zwei folgenden Bytes, die den X- bzw. Y-Offset für das nächste Pixel zur aktuellen Position beschreiben). Im Absolute Mode ist das erste Steuerbyte ebenfalls null und das zweite im Bereich 03...FF; es repräsentiert die Anzahl der nachfolgenden Farbeinträge eines einzelnen Pixels. Jeder Durchlauf ist Word-ausgerichtet. S. [432](#) [3].

D.4 Format TGA

Das Rasterformat TGA (= Targa) wurde von der amerikanischen Firma *Truevision* entwickelt. Es hat einen Header, der 18 Bytes lang ist, siehe Tab. [D.7](#). Aus ihm ist bereits ersichtlich, dass auch mit in der Grafik eingebundenen Farbtabellen gearbeitet werden kann.

D.5 Format PCX

Das PCX-Format stammt von der Firma ZSoft Corporation. Es wurde anfänglich für das Grafikprogramm Paintbrush entwickelt, ist jedoch heute zu einem wichtigen Format für viele Programme geworden. Tab. D.8 soll Aufschluss über den Header geben.

Bei PCX-Grafiken wird eine Lauflängencodierung sehr häufig verwendet. Die Lauflänge beträgt dabei maximal eine Zeile. Sind in den Daten die Bits 6 und 7 gesetzt, so geben die 6 unteren Bits die Anzahl der Pixel und das nachfolgende Byte deren Farbwert an. Daraus folgt, dass Farben mit einem Byte-Wert > 127 immer eine Lauflängencodierung benötigen. Der Wert Bit pro Pixel und Farbebenen entspricht nicht zwangsläufig der Farbtiefe, da teilweise mehrere Farbebenen benutzt werden. In diesem Falle erfolgt eine bitweise Verteilung der Bilddaten nach der Vorgabe der Variablen Bit pro Pixel und Farbebenen. S. 432 [4].

Tab. D.8 Header des Grafikdatenformats PCX

Offset Hex	Länge	Festwert Hex	Bedeutung
00	Byte	0A	Dateikennung
01	Byte	0: Version 2.5 2: Version 2.8 mit Palette 3: dto. ohne Palette 5: Version 3.0 mit Palette	Version
02	Byte	0: ohne Kompression 1: mit Kompression	Kennzeichen für Kompression
03	Byte	1: bei Farbtiefen 1 Bit, 4 Bit 8: bei Farbtiefen 8 Bit, 24 Bit	Bit pro Pixel und Farbebenen
04	Word		Xmin (Begrenzung links)
06	Word		Ymin (Begrenzung oben)
08	Word		Xmax (Begrenzung rechts)
0A	Word		Ymax (Begrenzung unten)
0C	Word		Horizontalauflösung der Grafik in dpi
0E	Word		Vertikalauflösung der Grafik in dpi
10	48 Byte		Farbpalette mit 16 Farbeinträgen
40	Byte		Reserviert
41	Byte	1: bei Farbtiefen 1 Bit, 8 Bit 3: bei Farbtiefe 24 Bit 4: bei Farbtiefe 4 Bit	Anzahl der Farbebenen
42	Word		Byte pro Zeile
44	Word	1: für Farbe oder Schwarzweiß 2: für Graustufen	Palettenart
46	58 Byte	00	Auffüllfeld auf 128 Byte
		<i>Bilddaten</i>	

Tab. D.9 Header des Grafikdatenformats PICT

Offset Hex	Länge	Festwert Hex	Bedeutung
200	Word		Dateigröße – 200h
202	DWord		unklar
206	Word		Bildbreite in Pixel (Kolumnenzahl)
208	Word		Bildhöhe in Pixel (Zeilenzahl)
20A	45 Byte		unklar
237	Byte	98 oder anders	Identifikation
238	Byte		Format des Steuer-Byte
		<i>für Identifikation = 98</i>	
256	Word		Anzahl der Bildebenen
258	Word		Bit pro Pixel
25A	12 Byte		unklar
266	n Byte		Text mit n Zeichen
266 + n	Word	0	Textabschluss
268 + n	Word		Anzahl der Farben – 1
		<i>für Identifikation ≠ 98</i>	
25A	Word		Anzahl der Bildebenen
:	:		:

Das Feld für die Farbpalette ist grundsätzlich für 16 Farben ausgelegt; sind 256 Farben definiert, so werden die fehlenden Einträge ans Ende der Datei angehängt. Für jede Farbe ist in der Palette ein Byte vorgesehen, insgesamt also 3 Byte pro Eintrag, und zwar in der Reihenfolge Rot, Grün und Blau. Die Farbtiefe von 24 Bit benötigt keine Palette – die Farbcodierung erfolgt dann direkt.

Da die Grafikauflösung mit abgespeichert wird, ist die Aspect Ratio der Grafik am Zeitpunkt der Erzeugung immer verfügbar.

D.6 Format PICT

Der Vollständigkeit halber soll noch ein Format von Apple-Macintosh beschrieben werden, wobei allerdings einige Angaben offen bleiben müssen. Die Leserichtung erfolgt nach dem Motorola-Format, d. h. bei Word und größeren Längen erscheinen die Bytes in absteigender Wertigkeit.

Nach diesen Header-Daten folgt die Farbtabelle, die aus 8 Byte besteht. Ein einzelner solcher Eintrag geht aus Tab. D.10 hervor. Es sind Farbtiefe bis 24 Bit zulässig, wobei hier – ähnlich zum PCX-Format – drei Ebenen verwendet werden, was den Eintrag für Bit pro Pixel auf 8 reduziert.

D.7 Format JPEG

Das nach Abschn. 12.8.1 erläuterte Verfahren der JPEG-Kompression muss ebenfalls einen Header erhalten, nach dem die komprimierte Datei (mit der Endung JPG) wieder decodiert werden kann. Diese Definitionen sind in Blöcke unterteilt, die jeweils durch eine Kennung eingeleitet werden. Tab. D.11 zeigt diese Blockeinheiten.

Allumfassende Beschreibungen und Definitionen des JPEG-Standards kann nur spezielle Literatur bieten. S. 432 [5, 6].

Tab. D.10 Eintrag der Farbtabelle beim Grafikdatenformat PICT

Offset Hex	Länge	Bedeutung
00 + m	Word	Nummer des Farbtabelleneintrags (MSB, LSB)
02 + m	Byte	Rotwert
03 + m	Byte	Reserviert
04 + m	Byte	Grünwert
05 + m	Byte	Reserviert
06 + m	Byte	Blauwert
07 + m	Byte	Reserviert

Tab. D.11 Header des Grafikdatenformats JPG

Offset Hex	Länge	Festwert Hex	Bedeutung
0	Word	<i>Bildstart</i> FFD8	Kennung für Bildstart
00 + m	Word	<i>Bildparameter</i> FFE0	Block-Kennung für Bildparameter
02 + m	Word		Anzahl der folgenden Bytes
04 + m	5 Byte	4A, 46, 49, 46, 00	Identifizierung 'JFIF'
09 + m	Word		Versionsnummer
0B + m	Byte	0: ohne Einheit, 1: dpi, 2: Dots pro cm	Auflösungseinheit in beiden Richtungen
0C + m	Word		Horizontalauflösung in Pixel
0E + m	Word		Vertikalaufösung in Pixel
10 + m	Byte		Breite in Pixel der Vorschau = b
11 + m	Byte		Höhe in Pixel der Vorschau = h
12 + m	$3 \cdot b \cdot h$ Byte		Bilddaten der Vorschau (RGB)
00 + n	Word	<i>Huffman-Tabelle</i> FFC4	Kennung für Huffman-Tabelle
02 + n	Word		Anzahl der folgenden Bytes
04 + n	Byte		Zeiger auf Farbkomponente
05 + n	16 Byte		Einzellängen der Einträge (Summe = s)
15 + n	s Byte		Tabelleneinträge

(Fortsetzung)

Tab. D.11 Fortsetzung

Offset Hex	Länge	Festwert Hex	Bedeutung
$00 + p$	Word	<i>Arithmetik-Codierung</i> FFCC	Kennung für Arithmetik-Codierung
$02 + p$	Word		Anzahl der folgenden Bytes
$04 + p$	Byte		Zeiger auf Farbkomponente
$05 + p$	Byte		Wert eintrag
$00 + q$	Word	<i>Quantisierungstabelle</i> FFDB	Kennung der Q-Tabelle
$02 + q$	Word		Anzahl der folgenden Bytes
$04 + q$	Byte	Bit 0–3: Index der Tab. Bit 4–7: 0: Byte, 1: Word	Index und Verschlüsselung
$05 + q$	Byte/Word		Tabellenwerte
		<i>Restart-Intervall</i>	
$00 + r$	Word		Kennung für Restart
$02 + r$	Word		Anzahl der folgenden Bytes
$04 + r$	Word		Wert für Restart-Intervall
		<i>Bilddaten-Parameter</i>	
$00 + s$	Word	FFCx ^a	Kennung für Bilddaten-Parameter
$02 + s$	Word		Anzahl der folgenden Bytes
$04 + s$	Byte		Daten-Präzision
$05 + s$	Word		Bildhöhe in Pixel
$07 + s$	Word		Bildbreite in Pixel
$09 + s$	Byte		Anzahl der Farbkanäle
$0A + s$	Byte		Nummer des ersten Farbkanals
$0B + s$	Byte	Bit 0–3: Vertikalwert Bit 4–7: Horizontalwert	Unterabtastungsfaktor für Farbe
$0C + s$	Byte		Index der zugehörigen Q-Tabelle
$0A + ss$	Byte		Nummer des letzten Farbkanals
$0B + ss$	Byte	Bit 0–3: Vertikalwert Bit 4–7: Horizontalwert	Unterabtastungsfaktor für Farbe
$0C + ss$	Byte		Index der zugehörigen Q-Tabelle
⋮	⋮	⋮	⋮
		<i>Start der Abtastung</i>	
$00 + t$	Word		Kennung
$02 + t$	Word		Anzahl der folgenden Bytes
$04 + t$	Byte		Anzahl der Komponenten
$05 + t$	Byte		Nummer der Komponente
$06 + t$	Byte	Byte 0–3: AC-Tabelle Byte 4–7: DC-Tabelle	Nummern der ausgewählten Arithmetik-Tabellen

^ax legt die Codierungsart fest (Arithmetik- oder Huffman-Codierung)

Literatur

1. Hübscher, H.: Elektrotechnik Fachbildung Kommunikationselektronik 2
Radio-/Fernseh-/Funktechnik. Braunschweig: Westermann Schulbuchverlag 1989.
2. Holtorf, K.: Handbuch der Grafikformate. München: Franzis-Verlag 1994.
3. Microsoft Windows Multimedia: Programmer's Reference.
Washington: Microsoft Press 1991.
4. Steinbrecher, R.: Bildverarbeitung in der Praxis. München: Oldenbourg-Verlag 1993.
5. Wallace, G. K.: The JPEG Still Picture Compression. IEEE-Computer, 1991.
6. C-Cube Microsystems: JPEG File Interchange Format. Mailbox-Textdatei, 1992.

Glossar

2er-Komplement Das 2er-K. ist eine besondere Darstellung von Digitaldaten. Die positiven Werte bleiben dabei unverändert, während bei negativen Daten das MSB 1 ist. Bei 8 Bit ist demnach ein Bereich von 10000000 (= −128) bis 01111111 (= +127) gültig. Operationen mit negativen Zahlen führt beim 2er-K. zum richtigen Ergebnis. Im Gegensatz dazu wird teilweise auch das Format →Offset-Dual (auch Offset Binary genannt) eingesetzt. 4-2-2-Verfahren. Bei digitalen Systemen liegt hier für die →Leuchtdichte Y eine doppelt so häufige Abtastung vor als für U und V , weil das menschliche Auge ebenfalls die Leuchtdichte besser aufzulösen vermag als Farbinformationen.

8er-Sequenz Der Burst weist bei der Zeile n bezüglich des Horizontalsynchronimpulses eine bestimmte Phasenlage auf. Nach einer Sequenz von 8 Halbbildern erreicht der Burst wieder dieselbe Phasenlage bei dieser Zeile (Übereinstimmung besteht natürlich auch für die restlichen Zeilen). Im Schnittbetrieb ist es manchmal erforderlich, die 8er-Sequenz einzuhalten, da sonst die Phasenlage des Burst „springt“ und zu Instabilitäten der Farbsynchronisation führt. Man spricht häufig vom Viertelzeilen-Offset-Verfahren, weil auch Phasengleichheit nach Ablauf von jeweils 4 Zeilen besteht. Die 8er-Sequenz kann nur bei einer strengen und festen Phasenverkopplung zwischen Burst und Horizontalfrequenz eingehalten werden.

Abschlusswiderstand In der Videotechnik verwendet man häufig einen A, um Leitungsreflexionen und somit Schattenbildungen zu vermeiden. Der A. hat dabei einen Wert, der dem →Wellenwiderstand der Leitung entspricht. Für jede Quelle ist am Ende der Leitung, die bei der Verschaltung mehrerer Geräte in linienförmig (und nicht sternförmig) sein muss, ein einziger A. anzukoppeln. Auch beim Anschluss offener Leitungen kommt es beim Fehlen des A.s zu Reflexionen.

Aliasing Das Shannon-Theorem schreibt vor, dass die höchste Eingangsfrequenz höchstens dem halben Wert der Abtastfrequenz entsprechen darf. Bei Nichteinhaltung des Shannon-Theorems entstehen bei der Abtastung eines analogen Signals Störfrequenzen, die sich aus der Differenz zwischen Abtastfrequenz und Eingangsfrequenz ergeben. Die auftretende Störung nennt man A.

Amplitude Eine periodische Schwingung zeigt Auslenkungen um ihren Nullpunkt; das Maximum der Auslenkungen entspricht der Amplitude und somit der Hälfte des Spitzespitze-Werts.

Aspect Ratio Unter A. versteht man das Seitenverhältnis eines Pixels, welches sich aus dem Ausdruck (Zeilenzahl/Spaltenzahl) → 4/3 ergibt. Eine A. von 1 (quadratische Pixel) ergibt sich z. B. bei den Auflösungen 768 × 576 (PAL) und 640 × 480 (NTSC).

Attack Erster Parameter von → Hüllkurven. Beschreibt den Einklingvorgang in Form und Zeit. → Decay, → Sustain, → Release.

Bandbreite Ein Übertragungssystem verarbeitet Signalfrequenzen, die innerhalb eines Intervalls liegen, das durch die untere und obere → Grenzfrequenz definiert ist. Bei diesen Frequenzen ist die Spannung jedoch auf ca. 71 % (−3 dB) gegenüber der Maximalspannung abgefallen. Die B. kennzeichnet jedoch nur die Differenz beider Frequenzen, weshalb die Angabe der Grenzfrequenzen günstiger ist.

Bark Maß für wahrgenommene Tonhöhe. Das Spektrum bis 16 kHz wird in 24 Bark-Einheiten eingeteilt. Es besteht kein linearer Zusammenhang zwischen Frequenz und empfundener Tonhöhe, jedoch entspricht 1 Bark der Frequenz von ca. 100 Hz.

Bildaustastung Im Bereich zwischen zwei Halbbildern liegen Impulse zur Vertikalsynchronisation. Da diese somit in die Zeit des Strahlrücklaufs (Bildrücklauf) fallen, muss eine B. dafür sorgen, dass keine Störungen auf dem Bildschirm sichtbar werden. → Zeilenaustastung

Bildformat Das B. definiert das Seitenverhältnis eines Bildes oder einer Grafik. Handelt es sich um ein schirmfüllendes Bild, so ist das B. bei herkömmlicher Technik 4:3, während bei HDTV das B. 16:9 ist.

Bildschirmauflösung Die B. legt i. Allg. die Spalten- und Zeilenzahl eines Bildes (Pixel/Zeile × Zeilen) fest. Diese Angabe sagt nicht nur etwas über die Qualität des Bildes aus, sondern auch darüber, ob es sich um → Overscan oder → Underscan handelt und welches Pixelformat (→ Aspect Ratio) vorliegt. Die B. ist nicht zu verwechseln mit der → Grafikauflösung.

Bildschirmseite → Page

Blackburst Zur Synchronisation von Videogeräten verwendet man oft eine Signal, welches zwar keinen Bildinhalt (Schwarzbild), jedoch Synchronimpulse und Farbsynchronsignal (→ Burst) aufweist. In Studios sind alle Geräte über einen solchen B. gekoppelt und somit synchronisiert.

Blu-ray Disc Weiterentwicklung der → DVD mit mehr Speicherkapazität. Dient vor allem als Datenträger für Videofilme.

Burst Zur Synchronisation des Farbdecoders setzt das FBAS- und das Y/C-Signal den B. ein, der sich (beim FBAS-Signal) auf der hinteren Schwarzschanze befindet. Der B. besteht aus mehreren Schwingungen des Farbhilfsträgers (4,43 MHz), welche einen Oszillator im Empfänger (Decoder) anstoßen bzw. synchronisieren, der zur Entschlüsselung der Farbinformationen innerhalb des darauffolgenden Zeile dient. Auf die Phasenlage des Burst beziehen sich nämlich die Phasenlagen der Farbartsignale der Zeile.

Chroma-Keying In der Fernsehtechnik (vor allem bei Sendungen innerhalb des Studios) werden oft Stanzebenen eingesetzt, in die ein Fremdbild elektronisch eingetastet werden kann. Da die Stanzebene eine markante Farbe (meist Blau) aufweist, wird das Verfahren C. genannt. Es kommt u. a. bei Wettervorhersagen zum Einsatz, wo in eine Projektionstafel eine Animation (z. B. vom Computer) gesetzt wird. Doch das C. dient auch Speialeffekten und Kulissenbildern.

Chrominanz Die C. definiert die Farbinformation. Dabei kann sie im Sinne eines Signals oder im Sinne von einzelnen Pixelwerten verstanden werden. Durch die C. besteht eine eindeutige Aussage über Farbton und →Farbsättigung der wiederzugebenden Farbe. Dagegen kann von der C. nicht auf die →Leuchtdichte geschlossen werden.

Containerformat Das C. ist kein eigentliches Datenformat, sondern lediglich eine Anweisung, wie verschiedene Datenformate in diesem Container aufzubewahren sind. Beispiele für solche Formate sind AVI, PDF usw.

Cross-Color Ist das Helligkeitssignal mit dem Farbsignal verschmolzen, wie das bei →FBAS der Fall ist, können sich Frequenzen im Helligkeitssignal, die in der Nähe des Farbhilfsträger sind, auf den Farbkanal des Decoders gelangen und somit als Farbe interpretiert werden. Bekanntes Beispiel kann beim Fernsehen beobachtet werden, wenn karierte Graumuster (Anzüge usw.) in schillernden Farben erscheinen. Dieser Effekt heißt C, welcher systembedingt nicht unterdrückt werden kann.

Cross-Luminanz Entsprechend der →Cross-Color können bei der Übertragung mit dem →FBAS-Signal C.-Störungen entstehen, wenn die Farbschwingungen in den Luminanzkanal des Decoders gelangen und als Helligkeitsschwankungen interpretiert werden. Ist auf dem TV-Monitor ein sog. Perl schnurmuster zu sehen, so liegt eine C.-Störung vor. Eine weitgehende Unterdrückung dieses Effektes ist möglich durch Fernhalten der →Farbhilfsträgerschwingungen vom Luminanzkanal; dies kann jedoch nur auf Kosten des Luminanzfrequenzgangs geschehen.

Decay Zweiter Parameter von →Hüllkurven. Beschreibt den Rückfallvorgang der Amplitude in Form und Zeit nach der Phase →Attack, →Sustain, →Release.

Dezibel (dB) Unter diesem logarithmischen Maß gibt man Verstärkung bzw. Dämpfung eines Übertragungssystems an. Aber auch Pegelunterschiede können hierdurch spezifiziert werden. Ist das Verhältnis zweier maßgebenden Spannungen bekannt (z. B. Eingangs- und Ausgangsspannung), so multipliziert man den Zehnerlogarithmus dieses Verhältnisses mit 20 und erhält den Pegelunterschied.

Differenztonfaktor →Intermodulation

Dirac-Kamm Periodische Stoßfolge als mathematische Funktion, wobei ein Stoß unendlich kurz, aber unendlich hoch ist, sodass als Flächeninhalt die Einheitsfläche entsteht. Den D. ist von grundlegender Bedeutung in der Signaltheorie, nämlich beim idealen Abtasten, bei welchem das abzutastende Signal mit der Funktion des D.s multipliziert wird.

Dropout Vor allem ein Bandspeichermedium kann Störungen verursachen, die auf einen D. zurückzuführen sind. Auslöser hierfür können Knickstellen oder auch Schichtlöcher des Bandes sein. Analog wiedergebende Systeme beantworten einen D. mit

vorübergehendem Pegelabfall, während Digitalgeräte bis zu einem gewissen Grade mittels Fehlerkorrektur den Datenverlust kompensieren oder wenigstens interpolieren können, bei größeren D.s jedoch kommt es meist zum Abschalten des Geräts.

Druckempfänger Eine Mikrofonkapsel, die eine dem Schalldruck entsprechende Spannung abgibt, ist ein D. Die Membran ist nur von einer Seite für den Schall zugänglich, somit ist nur der Schalldruck für die Membranbewegung maßgebend. Die →Richtcharakteristik solcher Kapseln ist meist kugelförmig.

Druckgradientenempfänger Ein D. ist eine Mikrofonkapsel, die auf Druckunterschiede zwischen Membranober- und -unterseite reagiert bzw. eine dem Druckunterschied entsprechende Spannung abgibt. Die Membran ist für den Schall für beide Seiten zugänglich, und sie ist ungedämpft aufgehängt; gelangen jedoch Schallwellen gleicher Phasenlagen auf beide Seiten, so bedeutet dies auch zu jeder Zeit gleiche Drücke zu beiden Seiten, und die Membran bewegt sich (im Idealfall) nicht. Bei konstanter Schalldruckamplitude steigt mit der Frequenz der Druckgradient und somit die Mikrofonspannung. Die →Richtcharakteristik ist im Normalfall achtförmig.

DVD Die D. (Digital Versatile Disc) dient als Speichermedium. Die Kapazität kann bei Single-Layer-Typen bis etwa 4,7 GByte betragen, bei Dual-Layer-Typen fast das Doppelte. Normalerweise handelt es sich um ein einmal beschreibbares Medium. Es gibt jedoch Sonderformen, die mehrfach wiederbeschreibbar sind.

Dynamik Die D. definiert den möglichen Pegelunterschied zwischen leisester und lautester Stelle einer Audioübertragung. Digitale Systeme gestatten meist eine Bestimmung der D. durch die Kenntnis der Auflösung bzw. der Bit-Breite der →Quantisierung.

Eckfrequenz Beschreibt man den Frequenzgang durch zwei Tangenten (z. B. durch eine waagerechte und eine schräge Tangente bei einem Tiefpass), so schneiden sich diese bei der E. (bei einem 6-dB-Tiefpass entspricht dies der →Grenzfrequenz).

Effektivwert Der E. beschreibt einen Spannungswert einer Wechselspannung. Legt man an einen Wirkwiderstand eine Gleichspannung von gleichem Wert, so wird die gleiche Leistung umgesetzt wie durch die Wechselspannung. Das Verhältnis zwischen E. und arithmetischem Mittelwert entspricht dem sog. Formfaktor, der spezifisch ist für jede Signalform. Bei der reinen Sinusform beträgt der Formfaktor ca. 1,11, und der E. ist ca. 71 % vom Scheitelwert (→Amplitude).

Elongation Unter der E. versteht man die momentane Auslenkung einer Schwingung (im Gegensatz zur →Amplitude, die der maximalen E. entspricht).

Equalizer Der E. ist ein Gerät zur Beeinflussung des →Frequenzgangs. Somit ist eine Korrektur bzw. Kompensation von Frequenzgangfehlern und auch die Unterdrückung von Störsignalen möglich. Ferner können E. zur Erzeugung von Spezialeffekten eingesetzt werden (Telefonstimme usw.).

Ersatzlautstärke Die E. ist ein veralteter Begriff zur Spezifikation von Störpegeln bei Mikrofonen. Sie entspricht demjenigen Schallpegel, der am Mikrofon die gleiche Signalspannung hervorrufen würde wie die abgegebene Störspannung. Heute verwendet man nur noch den Begriff des →Geräuschspannungsabstandes a_r , welcher

bei Mikrofonen bezogen wird auf einen Bezugsschalldruck von 1 Pa, was einem Schallpegel von 94 dB entspricht. Somit ergibt sich die E. aus der Differenz 94 dB a_r .

Farbdifferenzsignal Bei der Wandlung von →RGB nach →FBAS und umgekehrt ist als Zwischenstufe die Erzeugung zweier F.e notwendig, die sich vor der Modulation durch den Farbhilfsträger aus der Differenz von $B - Y$ (Blau minus →Leuchtdichte) und $R - Y$ (Rot minus Leuchtdichte) ergeben. Im Studio werden in modifizierter Form (Reduzierung der Amplituden) als Signale U und V zusammen mit dem Leuchtdichtesignal Y weiterverarbeitet und aufgezeichnet (Betacam). Die Farbe Grün wird nicht codiert, sondern kann über Y zurückgewonnen werden.

Farbhilfsträger Um in einem sequentiell zu übertragenden Bildsignal (z. B. bei der Fernsehübertragung) die Farbinformation codieren zu können, ist ein F. notwendig, der bei →PAL eine Frequenz von 4,43 MHz hat. Jede zu übertragende Farbe hat eine bestimmte Zusammensetzung der beiden Farbdifferenzsignale, und entsprechend dieser entstehen zwei F-Schwingungen, die um 90° zueinander phasenverschoben sind. Die Addition dieser beiden Schwingungen ergibt die Farbartschwingung, welche innerhalb der Zeile übertragen und mit dem außerhalb liegenden →Burst wieder decodiert wird.

Farbpalette Bei kleinen →Farbtiefen (8 Bit und weniger) stellt ein System normalerweise eine Auswahl von Farben zur Verfügung, deren Zusammensetzung auch frei gewählt werden kann. Das heißt die Anzahl der Paletteneinträge ist zwar begrenzt, jedoch lassen sich die Einträge ändern und mit höherer Farbtiefe definieren (18 Bit bis 24 Bit).

Farbsättigung Nach der Reinheit einer Farbe ergibt sich die F; bei Verwendung von nur einer oder zwei der drei Primärfarben besteht eine F. von 100 %. Werden zu einer Farbe jedoch die fehlenden Primärfarben beigemischt, so reduziert sich die F, weil man sich sozusagen dadurch Weiß nähert. Die F. wirkt sich auf die Länge der Vektoren des →Farbdifferenzsignals aus (dagegen ist ein Schluss von der Vektorlänge auf die Farbsättigung nicht möglich, da auch der Helligkeitspegel Einfluss auf die Vektorlänge nimmt).

Farbtiefe Unter F. versteht man i. Allg. die Bit-Breite, mit der die Farbinformation durch ein digitales System (z. B. A/D-Wandler einer Framegrabber-Karte) verarbeitet werden kann. Auch gespeicherte Bilder werden durch die F. spezifiziert. Eine F. von 24 Bit ermöglicht beispielsweise ca. 16,8 Millionen unterschiedliche Kombinationen der Primärfarbwerte, was jedoch nicht gleichbedeutend ist mit genauso viel unterschiedlichen Farbtönen. Bei einer F. von 16 Bit sind nur noch 32 768 Werte möglich, und 8 Bit gar hat nur noch 256 verschiedene Farben (hier wird meist mit einer →Farbpalette gearbeitet, die eine Zusammenstellung der wichtigsten Farben enthalten sollte).

FBAS F. ist ein Kürzel für Farbe-, Bild-, Austast- und Synchronsignal und steht somit für eine Art der Bildsignalübertragung, bei der jegliche Information und Hilfssignale auf einer Leitung geführt werden. Dadurch können sich →Cross-Color- und →Cross-Luminanz-Störungen ergeben, weil sich Luminanz- und Farbsignal gegenseitig beeinflussen, bzw. weil beide nur mangelhaft voneinander zu unterscheiden sind.

FM-Synthese Bei der F. handelt es sich um ein Verfahren, bei dem mehrere VCOs, die einer Frequenzmodulation unterliegen, den Klang eines Instruments nachbilden. Auch

Audiokarten bedienen sich teilweise dieser Technik. Modernere Methoden arbeiten jedoch mit den Klängen digitalisierter Instrumente (→ Sampling-Synthesizer).

Fremdspannungsabstand Der F. ist ein Maß für den Pegelunterschied zwischen Nutzsignal (eigentlich Gesamtsignal) und Störsignal. Es finden hier alle Störgeräusche Berücksichtigung. → Geräuschspannungsabstand

Frequenzgang Ein Audiogerät sollte alle hörbaren Frequenzen möglichst mit gleichem Pegel wiedergeben bzw. konstanter Verstärkung verarbeiten. Ein Maß für diese Linearität ist der F, welcher als Bereich und mit der zugehörenden Abweichung angegeben wird.

Gammakorrektur Die G. dient der Kompensation von Krümmungen der Kennlinien an Kameras und Monitoren. Es besteht i. Allg. nämlich kein linearer Zusammenhang zwischen Leuchtdichte und Steuer- bzw. Signalspannung. Manche Programme zum Einfrieren von Bildern können daher optional eine G. vornehmen.

Genlock Um ein Bildsignal einer Grafikkarte in einem Studio verwenden zu können, muss es normalerweise durch einen zentralen Takt synchronisiert werden können. Ein solches G. erfordert einen Eingang an der Grafikkarte, an den der Studiotakt (→ Blackburst) anzuschließen ist. In der Karte muss dann der Pixeltakt mit entsprechend höherer Frequenz (jedoch phasenstarr zur Zeilenfrequenz) arbeiten.

Geräuschspannungsabstand Wie der → Fremdspannungsabstand ist auch der G. ein Maß für den Pegelunterschied zwischen Nutz- und Störsignal. Allerdings werden hier nur Rauschsignale gemessen. Oftmals erfolgt eine Bewertung durch ein Filter, das den Frequenzgang des Ohres nachbilden soll (Ohrkurvenfilter). Auf diese Weise sagt der G. mehr über seine physiologische Empfindung aus. Bei Mikrofonen bezieht der G. meist auf einen Schalldruck von 1 Pa (→ Ersatzlautstärke).

Grafikauflösung Die G. definiert die Kolumnen- und die Zeilenzahl einer Grafik. Bedeckt diese gerade vollständig den Bildschirm, so entspricht die G. genau der → Bildschirmauflösung.

Grafikformat Die spezifische Datenanordnung einer Grafik wird durch das G. bestimmt. Darunter fallen z. B. TIFF, EPS, BMP, TGA usw. Das G. ist nicht zu verwechseln mit der → Grafikauflösung, dem → Bildformat oder der → Bildschirmauflösung.

Grenzfrequenz Die G. ist definitionsgemäß diejenige Frequenz, bei der ein Pegel um 3 dB (ca. 71 % der Spannung) gegenüber dem Maximalpegel abgefallen ist. Damit ist die G. ein wichtiger Parameter eines Filters. → Eckfrequenz

Halbbild Bei der Wiedergabe in der Betriebsart → Interlaced setzt sich ein ganzes Bild (Vollbild) aus zwei zeitlich hintereinander folgenden, zeilenweise ineinander geschriebenen H.ern zusammen. Dadurch erreicht man eine recht flimmerfreie Wiedergabe, weil die Vertikalablenkrate hoch ist und dennoch der Informationsgehalt niedrig gehalten werden kann.

Harmonische H. sind Signale, die ein ganzzahliges Vielfaches der Frequenz der Grundschwingung aufweisen. Reine Sinussignale zeigen keine H., sie bestehen daher nur aus ihrer Grundschwingung. Jede andere periodische Signalform jedoch besteht aus

der Grundschwingung und H.n, deren Amplitude und Frequenz von der Form des Gesamtsignals abhängen.

H-Phase Die H. bestimmt die horizontale Lage eines Bildes. Ob bei einer Veränderung der H. der Bildinhalt oder nur der Synchronimpuls zeitlich verschoben wird, hängt von den Möglichkeiten und auch von der übrigen Situation ab, die eine solche Justage erforderlich macht (beispielsweise zur Optimierung der Kongruenz von Stanzmasken). Im Genlock-Betrieb sind manchen Video-Grafikkarten in der Lage, die Computergrafik gegenüber dem Genlock-Takt zu verschieben.

Hüllkurven Synthesizer steuern über H. die Lautstärke, also die Charakteristik und die Zeiten der Vorgänge → Attack, → Decay, → Sustain und → Release. Zur Erzeugung der h.-gesteuerten Verstärkung ist ein → VCA notwendig.

Impedanz Ein- und Ausgänge von Geräten haben einen Scheinwiderstand, der sich durch vektorielle Addition aus dem Wirk- und dem Blindwiderstand ergibt. Dieser Wert entspricht der I, die wegen ihrer Frequenzabhängigkeit meist mit einer entsprechenden Frequenzangabe versehen ist. Wichtig sind in diesem Zusammenhang vor allem Mikrofon- und Lautsprecherimpedanzen sowie die I. von Geräten in der Videotechnik, die wegen ihrer Verschaltung untereinander meist mit dem → Wellenwiderstand der Leitungen übereinstimmt.

Interferenz Treffen zwei Wellen gleicher Art (z. B. Schallwellen) aufeinander, so kann es – je nach Phasenlage – zur Verstärkung oder auch zur Auslöschung kommen. Als Ergebnis dieser Überlagerung (Superposition) ergibt sich die I., wobei der Effekt bei Wellen gleicher Wellenlänge bzw. konstanter Phasenlage zu stehenden I.-Bildern führen kann.

Interlaced I. ist der englische Ausdruck für einen Modus der Bildwiedergabe, bei dem immer nur die Zeilen eines Halbbildes (also nur jede zweite Zeile des Gesamtbildes) geschrieben bzw. abgetastet werden. Die Zeilen des darauffolgendes Halbbildes werden dann in die verbleibenden Lücken des ersten Halbbildes geschrieben. Auf diese Weise erfolgt die (geometrische) Bildablenkung in der halben Zeit als bei der normalen Ablenkung bei → Noninterlaced, was eine relativ flimmerfreie Wiedergabe bei nicht allzu hoher Informationsbandbreite ermöglicht. I. wird vor allem in der Videotechnik und Noninterlaced in der Computertechnik eingesetzt.

Intermodulation Unter I. versteht man denjenigen Vorgang, der in einem Übertragungssystem stattfindet, wenn es nichtlineare Verzerrungen zeigt und wenn mindestens zwei Sinustöne zu verarbeiten sind. Als Ergebnis entstehen I.s-Produkte, die sich aus der Differenzfrequenz beider Töne ergeben.

Interpolation Bei der I. ein Wert durch zwei benachbarte Werte berechnet; hierfür gibt es unterschiedliche Rechenarten, die am Ziel der I. ausgerichtet sind. In der Digitaltechnik tritt I. zum Beispiel dann auf, wenn bei einem Audiosignal ein Wert durch Beschädigung des Datenträgers verloren ging. Auch beim Vorgang des Oversampling errechnet ein digitales Filter Zwischenwerte durch I., die im Original überhaupt nicht vorhanden sind.

Jitter Zeigt ein Signal regelmäßige oder unregelmäßige Phasensprünge, so spricht man von J. Als Folgeerscheinung können u. a. Synchronisationsfehler entstehen.

Klirrfaktor Der K. definiert den Grad der harmonischen Verzerrungen, die sich aus der Nichtlinearität einer Kennlinie ergeben können. Ein reines Sinussignal, das keinerlei Obertöne aufweist, hat einen K. von 0; passiert dieses Signal ein nichtlinear verzerrndes System, so entstehen zusätzlich → Harmonische, deren Gesamteffektivspannung im Verhältnis zur effektiven Gesamtspannung den K. ergibt. Eine weitere Möglichkeit zu Spezifizierung nichtlinearer Verzerrungen ist die → Intermodulation.

Komponentensignal → YUV-Signal.

Leuchtdichte In der Photometrie (Lichtmesstechnik) versteht man unter L. die Lichtstärke eines Körpers, bezogen auf seine Fläche. Mit welcher Helligkeit ein strahlender Körper empfunden wird, hängt von der L. ab, jedoch nicht von der Entfernung. Im Bereich Videotechnik versteht man unter L. meist das in dem Zusammenhang stehende Signal, welches die Helligkeit der Bildinformation darstellt und somit die Helligkeit der Bildpunkte bestimmt.

LUT Look Up Table. Eine L. ist sozusagen eine Hardware-Tabelle, mit der einem 8-Bit-Muster (Adresse) willkürlich ein beliebiges 8-Bit-Muster (Speicherwert) zugeordnet werden kann. Auf diese Weise kann man z. B. digitale Werte einer Primärfarbe gemäß einer Kennlinie manipulieren. L. dienen der → Gammakorrektur, Komplementärdarstellung und anderen Effekten.

Luminanz → Leuchtdichte.

Mixdown Das M. ist der Abmischvorgang, bei dem die Aufzeichnungen der separaten Spuren auf die endgültigen Stereospuren der Master-Aufzeichnung reduziert werden. Pegel, Panorama, Effekte usw. ist mittels Mischpult für jede einzelne Spur festzulegen bzw. während des M. zu verändern.

Nibble Das N. enthält 4 binäre Stellen eines Bytes und entspricht somit einer Stelle des Hexadezimalwerts. Ein Byte ist also aus zwei N.s zusammengesetzt, dem unterem und dem oberen N. In der Informatik kommt der Begriff des N. sehr häufig vor, wenn z. B. Bytes als Informationseinheiten zu groß sind und deshalb unterschiedliche Größen speichern.

Noninterlaced N. steht für einen Wiedergabe- oder Abtastmodus bei der Bildübertragung, bei dem die Bilder gleichförmig Zeile für Zeile geschrieben werden. Da für die Flimmerfreiheit eine gewisse minimale Bildablenkfrequenz erforderlich ist, macht N. eine hohe Übertragungsbandbreite notwendig, was in der Video- bzw. Fernsehtechnik nicht erfüllt werden kann. Im Gegensatz zu N. steht der Modus → Interlaced.

Offset-Dual Bei der D/A-Wandlung wird neben dem → 2er-Komplement auch das Format O. verwendet. Dabei wird der Nullpunkt im Analogbereich einem Offset des einfach codierten Digitalformats zugewiesen (meist in der Mitte des Bereiches). Somit erhält man immer positive Binärwerte.

Overlaying Im Gegensatz zum → Chromakeying stanzt das O. in ein speziell präpariertes Computerbild das Live-Bild einer Kamera oder anderer Quellen. Das Computerbild wird also über das Live-Bild „gelegt“. Zur Verwirklichung des Effekts sind Grafikkarten

notwendig, die über einen sog. Alpha-Speicher verfügen – ein separater Bildspeicher für das Hinterlegen von Attributen, die für jedes Pixel anders geartet sein können.

Page Unter P. versteht man die Einheit einer Bildschirmseite, die von der Grafikkarte dargestellt werden kann. Dabei besteht i. Allg. eine definierte Aufteilung des Videospeichers in P.s, wobei die momentan eingestellte → Bildschirmauflösung innerhalb den physischen Page-Grenzen liegen muss, wenn zwischen den P.s umgeschaltet werden soll. Werden diese Grenzen überschritten, können normalerweise die bedeckten P.s nicht mehr separat angesprochen werden.

PAL Phase Alternation Line. In den westeuropäischen Ländern hat sich mit P. ein Standard für die Farübertragung durchgesetzt, der gegenüber anderen Systemen (NTSC und SECAM) entscheidende Vorteile aufweist. Nach jeder Zeile wird der Farbvektor und der → Burst an der U-Achse gespiegelt (bzw. die Modulationsachse umgeschaltet). Dadurch kompensieren sich Farbtonfehler, die durch Phasenfehler zwischen Burst und Farbartsignal entstanden sind (der Nachteil der geringfügigen Farbentsättigung kann in Kauf genommen werden). Bei P. muss die Farbinformation jeder Zeile zwischengespeichert werden, da durch sie immer die jeweils darauffolgende Zeile demoduliert wird (durch den sog. Laufzeitdemodulator).

Panorama Das P. beschreibt den Horizont des Stereoraums. Zur Platzierung von Stimmen oder Instrumenten verändert man bei Mischpulten die Stellung des P.-Reglers.

PCM Pulse Code Modulation. Bei der Wandlung ins digitale Format gibt es verschiedene Methoden. Eine davon stellt das P. dar, bei dem die Digitaldaten als direkte, äquivalente Werte zum Analogsignal darstellen. Bei der Wandlung entsteht meist das Format mit sog. → 2er-Komplement.

Phantomschallquelle Eine nur virtuell existierende Schallquelle, deren Zentrum aus der subjektiven Wahrnehmung heraus lokalisiert wird.

Phon Das P. ist die logarithmische Einheit der Größe Lautstärke. Letztere ergibt sich ähnlich wie das → Dezibel, wobei hier der Quotient aus Schalldruck und Bezugsschalldruck (Schwellwert 2 → 10–5 bei 1000 Hz) gebildet wird.

Pitch Das P. ist derjenige Parameter eines Klanges, der seine Tonhöhenlage definiert. Es kann daher sowohl eine elektronische Einrichtung in Form eines P.-Wheel oder auch ein Software-Datum zur Festlegung der Tonhöhe eines Tones.

Polarisation Wenn es eine Vorzugsrichtung der Schwingungsrichtung gibt, spricht man von P. In der Optik kann lineare P. (also solche mit einer Schwingungsrichtung) durch doppelbrechendes Material wie Kalkspat erzeugt werden. Bei der Überlagerung frequenzgleicher, phasenverschobener Wellen entsteht elliptische bzw. bei gleichen Amplituden zirkulare P.

Quantisierung Bei der Wandlung eines analogen Signals erfolgt die Q, eine Aufteilung in (meist gleichmäßige) Intervalle des Analogbereiches. Jedem Intervall wird dann ein Binärkode zugeordnet.

Release Letzter Parameter von → Hüllkurven. Beschreibt den Ausklingvorgang in Form und Zeit nach der Phase → Sustain. → Attack, → Decay.

RGB Unter R. versteht man eine Signalform, bei dem das Bild über drei separate Signale der Primärfarben Rot, Grün und Blau übertragen wird. Die Qualität ist sehr hoch, da hier im Gegensatz zu →FBAS keine Trägercodierung der Farbinformation notwendig ist.

Richtcharakteristik Die R. beschreibt die Abhängigkeit der Mikrofonspannung vom Richtwinkel (bei konstantem Schalldruck). Die bekanntesten R.a sind Kugel, Acht und Niere. Die R. eines Mikrofon hängt von der akustischen Arbeitsweise ab (z. B. →Druckempfänger usw.). Durch Kombination zweier Mikrofone unterschiedlicher R. kann eine neue geschaffen werden; so ergibt z. B. die Kombination aus Kugel und Acht eine Nierencharakteristik. Jedoch kann dies ebenfalls durch eine gemischte Bauweise der Kapsel selbst erreicht werden.

Rosa Rauschen Im Gegensatz zum Signal →Weißes Rauschen zeigt das R. für Frequenzintervalle mit konstantem Verhältnis (z. B. Oktave) jeweils konstante Energien. In der Messtechnik wird R. häufig zur Messung des →Frequenzgangs herangezogen.

Sampling-Synthesizer Beim S. wird der Klang eines Instruments dadurch erreicht, dass ein echtes Instrument über einen Sampler digitalisiert wurde. Der so als Digitaldaten zur Verfügung stehende Klang wird nur noch über →Hüllkurven und das →Pitch gesteuert, um Dynamik sowie Tonhöhe zu gestalten.

SC-Phase Zwischen dem Horizontal-Synchronimpuls und dem (an dieser Stelle gedachten) →Burst-Signal besteht ein fester Phasenbezug, der sich für eine bestimmte Zeile erst nach 8 Halbbildern (→8er-Sequenz) bzw. nach 4 Zeilen (Viertelzeilen-Offset) wiederholt. Diese starre S. ist außerordentlich wichtig im Studiobetrieb, weil dort mit besonders kritischen Aufzeichnungsmaschinen wie z. B. der 1-MAZ gearbeitet wird. Gute Videografikkarten erfüllen solche Forderungen bezüglich H-SC-Verkopplung und erlauben im →Genlock-Betrieb auch eine Phasenveränderung zwischen eingehendem und ausgehendem →Farbhilfsträger. Somit ist eine Farbtonanpassung bei zentraler Farbsynchronisation möglich.

Schnelleempfänger Eine Mikrofonkapsel, deren Membran auf beiden Seiten für Schall zugänglich ist, reagiert auf die Schnelle des letzteren, sofern eine starke Bewegungsdämpfung (erreichbar durch sich ändernde Volumenverhältnisse bei Membranschwingungen) vorliegt. Trotz des mit der Frequenz steigenden Druckgradienten wird damit die Membran-Elongation konstant gehalten, und somit liegt eine Abhängigkeit mit der Schallschnelle vor, welche die Bewegungsgeschwindigkeit der Luftteilchen beschreibt. S. zeigen wie →Druckgradientenempfänger eine achtförmige →Richtcharakteristik.

Sustain Dritter Parameter von →Hüllkurven. S. beschreibt die zeitliche Länge konstanter Höhe von →Hüllkurven und somit die Dauer mit konstantem Pegel. →Attack, →Decay, →Release.

Synchronisation Bei der Bildwiedergabe ist eine S. für den Monitor wichtig, damit überhaupt erst ein stehendes Bild zustande kommt. Die Impulse für den Bild- und Zeilenanfang gewinnt man bereits bei der Videoquelle (also z. B. Kamera usw.); bei allen Aufzeichnungen auf Videoband werden diese Initialisierungsimpulse mitgeführt. Beim Mischen mehrerer Quellen im Studio erreicht man mit der S., dass die Bildsignale

jeweils zeitgleich beginnen. Zur S. von Farbinformationen ist der →Burst notwendig, der mit jeder Zeilen-S. erscheint. Damit Ton und Bild zeitlich abgestimmt sind und bleiben, setzt man hier zur S. (bei Verwendung von separaten Geräten für Bild und Ton) meist den →Timecode ein.

Timbre Das T. beschreibt die Klangfarbe einer Stimme. Im Zusammenhang mit dem MIDI-Standard spezifiziert das T. die Instrumentenklänge der Programm-Nummern; auch die Zahl der simultan ansteuerbaren Instrumente (nicht der Stimmen!) entspricht der Zahl der unabhängigen T.s (multitimbral).

Timecode Spezieller Code, der zur Synchronisation und Steuerung von Audio- und Videogeräten dient. Vor allem im Schnittbetrieb ist der T. unentbehrlich. Der T. kann unterteilt werden in verschiedene Codeformate und in die Aufzeichnungsarten.

VCA Voltage Controlled Amplifier. Hier handelt es sich um eine Funktionseinheit, mit der man den Verstärkungsfaktor über eine Spannung steuern kann. Der V. kommt bei Synthesizern sehr häufig vor, um →Hüllkurven zu erzeugen.

VCO Voltage Controlled Oscillator. Der V. stellt einen Oszillator dar, dessen Frequenz von einer Steuerspannung abhängt. Die ursprünglichen, monofonen Synthesizer arbeiteten mit V.s, die eine Steuerspannung von der Klaviatur erhalten, wobei ein Oktavsprung meist 1 V Spannungsunterschied bewirkte. Heute werden V.s ebenfalls eingesetzt, wenn es um künstliche Klangerzeugung geht. Allerdings handelt es sich um eine komplexere Anordnung mehrerer V.s, die auch eine →FM-Synthese bewerkstelligen.

Voxel Kleinste Bildeinheit des dreidimensionellen Bildes (im Gegensatz zum Pixel beim zweidimensionalen Bild).

Weiße Rauschen Beim W. handelt es sich um ein Signal, bei dem alle Frequenzen des hörbaren Spektrums mit gleicher Amplitude vorkommen. Eine spektrale Analyse ergibt daher im zeitlichen Mittel eine Waagerechte für die Spannungen. In der Messtechnik spielt das W. eine bedeutende Rolle. →Rosa Rauschen

Wellenwiderstand Der W. ist ein Parameter einer Übertragungsstrecke (meist Leitungen). Wird das Leitungsende mit dem W. abgeschlossen, so bilden sich keine Reflexionen und somit keine stehenden Wellen. Der W. wird auch als Widerstand der unendlich langen, verlustfreien Leitung bezeichnet. In einem gewissen Frequenzbereich hängt der W. nur von der Leitungsinduktivität (Induktivitätsbelag L_0) und der Leitungskapazität (Kapazitätsbelag C_0) ab. Der Wellenwiderstand kann theoretisch mit einer LC-Messbrücke gemessen werden; die Kapazität wird bei offener Leitung und die Induktivität bei am anderen Ende kurzgeschlossener Leitung ermittelt. Die Wurzel aus dem Verhältnis von L_0 und C_0 ergibt den W.

Wobbelsignal Zur Messung des Frequenzgangs bei Audiogeräten wird häufig ein W. eingesetzt, welches ein Sinussignal mit fortlaufender Frequenzvariation darstellt. Ist das zur Frequenz proportionale Steuersignal des W.s vorhanden, so lässt sich der Frequenzgang unmittelbar auf einem Oszilloskop darstellen. Man spricht heute zumeist vom *Sweepsignal*.

Y/C Es handelt sich bei Y. um eine Signalform, bei der das Luminanz- und das Chrominanzsignal separat übertragen und weiterverarbeitet wird. Im Gegensatz zu →FBAS entsteht wegen dieser Trennung weder →Cross-Color noch →Cross-Luminanz. Jedoch setzt man auch hier eine →Farbhilfsträger-Modulation ein, was die Farbauflösung z. B. gegenüber →RGB einschränkt.

YUV Beim Y. handelt es sich um ein Farübertragungssignal, welches ebenfalls wie →RGB ohne →Farbhilfsträger auskommt. Es wird u. a. bei Betacam-Maschinen verwendet und trägt daher auch den Namen Betacam-Signal oder auch Komponentensignal. Für die Übertragung ist ein dreidriges Kabel notwendig, welches das Signal der Leuchtdichte Y und der beiden Farbdifferenzsignale U und V führt. Letztere weisen jeweils nur die halbe Bandbreite von Y auf, was auch beim →4-2-2-Verfahren praktiziert wird.

Zeilenaustastung Nachdem eine Zeile am Monitor geschrieben wurde (Dauer ca. 52 µs bei 625/50-Systemen), erfolgt der Zeilenrücklauf, bei dem der Katodenstrahl also von rechts nach links abgelenkt wird (Ausgangsstellung für die nächste zu schreibende Zeile). Während dieser Rücklaufzeit erfolgt keine Übertragung von Bildinformationen, und der Strahl muss dunkelgetastet werden, damit sich der Rücklauf nicht störend auf das sichtbare Bild auswirkt. Diese Dunkeltastung während des Zeilenrücklaufs (Dauer ca. 12 µs) nennt man Z. →Bildaustastung

Zeilensprungverfahren →Interlaced

Zeitcode →Timecode

Stichwortverzeichnis

A

AAC (Advanced Audio Coding) 75
Abrieb 33
Abspielgeschwindigkeit 118
Abtastbedingung 120
Abtastfrequenz 302
Abtastspektrum 45
Abtastung, ideale 45
AC3 26, 78
Acht-Charakteristik 22
Active Pixel Sensor (APS) 233
Active Sensing 185
ADPCM 69, 399
Advanced Audio Coding (AAC) 75
Aftertouch 138, 143
AIFF 399
Akkommodation 278
Aliasing 302
-Störung 50, 241
Allpass 106, 114
Alpha-Signal 308
Alpha-Speicher 340
Amplitudenmodulation 95
Amplitudensprung 166
Animation, 2D-351
Animationssoftware 280
Anschlagsdynamik 174, 185
Anti-Flicker-Filter 318
Antialiasing-Filter 50
Aspect Ratio 290, 295, 300, 321
Assemble-Schnitt 260
Asymmetrie der Verstärkung 116
Audionachbearbeitung 379
Auflösung 56, 80, 287, 295, 365

Aufschaukeln 109
Aufzeichnung, multiple 359
Ausbreitungsgeschwindigkeit 273
Auslöschung 114
Ausschneiden 101
Aussteuerung 80
Aussteuerungsgrenze 65
Azimut 211

B

Balance 142
Bandbegrenzung 9
Bandpass 106
Bandrauschen 80
Bandsperrre 106
Bandzugproblem 325
Bank-Wechsel 156
Basisbreite 193
BAS-Signal 224
BBC (Binaural Cue Coding) 74
B-Bild 370
Benutzeroberfläche, Wiedergabe der 312
Begrenzung 6, 115, 116
Bel 4
Beschneidung 81
Betacam 258, 309
Betamax 259
Beugung 18
Beugungsmuster 276
Bewertung 12
Bezugspegel 3
Bildaustastung 221
Bildgeometrie 321

- Bild-im-Bild 348
 Bildkante 346
 Bildlagefehler 319
 Bildseitenformat 287
 Bildseitenverhältnis 219
 Bildstandstörung 320
 Bildverzug 392
 Binaural Cue Coding (BBC) 74
 Blauertsche Bänder 31
 Blaufläche 342
 Blenden
 - homogene 344
 - Struktur- 344
 Blooming-Effekt 233
 Blu-ray 281
 Bodenschallreflexion 82
 Bodenschwingung 84
 Boot-Record 39
 Braunsche Röhre 234
 Brechung 19
 Brechungssindizes 19
 Brille, 3D- 275
 Brummeinstreuung 87
 Brummschleife 413
 Burst 245
 Bus-Eigenschaften 364
- C**
 Capture-Zeitpunkt 331
 CCD-Wandler 232
 CCIR 601 309
 CD 36
 CD-ROM Betriebsarten 37
 Channel
 - Message 139
 - Mode 144
 - Voice 140
 Chromakey 341
 CIRC (Cross Interleaved Reed
 Solomon Code) 37
 Clipping 6, 65
 Clips 378
 Compact Disc (CD) 36
 Cross-Color 319
 - Störung 249
 Cross-Luminanz 319
 - Störung 250
 Cross-Mix 103
 Cueing 378
- D**
 Dämmstoff 18
 Dämpfungsmaß 3
 Dateigröße 400
 Datendurchsatz 66
 Datenhandschuh 279
 Datenreduktion 366
 dBV 4
 DBX 96, 124
 DCC (Digital Compact Cassette) 34
 Delta-Frames 371
 Delta-Sigma
 - Verfahren 58
 - Wandler 79, 80
 - Wandlung 60
 Dezibel 4
 Dezimation 52, 54
 Differenzfrequenz 49
 Differenztonfaktor 10
 Diffraktion 18
 Digital
 - Compact Cassette 34
 - Signal Processor – DSP 127
 - Versatile Disc 280
 Digital-Decoder 327
 Digitalisieren 377
 Dirac-Kamm 45
 Direktverfahren 58
 Display, holografisches 276
 Display-Modus 315
 Dithering 63
 Dither-Verfahren 285
 Dolby 96, 124
 - C 33
 - Digital 26
 - Pro Logic 27
 - Pro Logic II 28
 - Surround-Sound 25
 Doppelabtastung 46
 Doppler-Effekt 20, 93
 Double Super Twisted Nematic (DSTN) 237
 Down-Sampling 53
 Drahtgittermodell 352
 Dropout 68
 Dropout-Kompensator 260

- Druckempfänger 86
Druckgradientenempfänger 86
D1-Signal 309
DSP (Digital Signal Processor) 198
DSTN-Zelle 237
Dunkeltastung 221
Durabis 281
DVD (Digital Versatile Disc) 39, 280
Dynamik 12
Dynamikkompression 123
- E**
EBU-Farbbalken 240
Echo 18, 109, 209
Echogerät 92
Echtzeitanimation 356
Editierfunktion 176
Effektgerät 91
Effektzumischung 90
EFM (Eight-to-Fourteen-Modulation) 37
e-Funktion 45
Eimerketten-Speicher 93
Einfügen 100
Einfrierzeitpunkt 333
Eingangsimpedanz 88
Einmal-Echo 109
Einmessen 211
Einzelbild-Aufzeichnung 357
Einzelwert-Bearbeitung 126
Ellipsenkonstante 304
Emphasis 96, 125
Entropie-Encoder 369, 371
Envelope-Kurve 166
Equalizer 91, 207
Error Feedback 64
Ersatzlautstärke 88
Ersetzen 101
EUREKA 95 266
Export 306
Expression 142
- F**
Farbart 229
Farbbalken 299
Farbcodierung 426
Farbdetektor 342
Farbebene 428
- Farbensehen 228
Farbentsättigung 304
Farbfehlerkompensation 247
Farbhilfsträger 244
Störungen vom 334
Farbhilfsträgeroszillator 316
Farbmetrik 231
Farbort 229
Farbpalette 429
Farbpegel 298, 304
Farbreinheit 235
Farbsättigung 229, 246, 299
Farbstreifen 322
Farbsynchronisation 245, 254
Farbtabelle 424
Farbtiefe 283, 297, 303, 365
reduzierte 334
Farbton 246, 316
Farbtonfehler 253, 264
Farbtonverfälschung 388
Farbtripel 234
Farbwert 283
Fast-Fourier-Transformation (FFT) 121
FBAS-Signal 249, 309
Fehlerkorrektur 34
Fehlerverdeckung durch Interpolation 37
Feld-Leerlauf-Übertragungsfaktor 86
FFT (Fast-Fourier-Transformation) 103
Field-PAL-Sequence, 8- 251
Film, holografischer 276
Filter 345
Filtergüte 108
Filterkoeffizient 105
Filterordnung 105
Filterung 208
zweidimensionale 346
FIR-Filter 104, 127, 303, 346
Firewire 213
Flanger 93
Flanging 114
Flankensteilheit 108
Flimmereffekt 317
Flimmerfreiheit 220
Flimmen 235
Flip 348
Fluoreszenzschicht 234
Flystick 279
Force-Feedback 279

- Format
 BMP 424
 JPEG 430
 PCX 428
 PICT 429
 TGA 427
 TIFF 420
Fourier-Koeffizient 164
Fourier-Reihe 164
Frame 72
Fremdspannungsabstand 12
Frequenzgang 4, 211
Frequenzmodulation 21, 162, 192, 256
Frequenzverdopplung 223
Frequenzverschiebung 192
- G**
Gamma-Korrektur 310
Gate-Impuls 161
Genlock
 -Betrieb 313
 -Signal 358
Geometriestörung 326
Geräuschspannungsabstand 11, 96
Gerbernorm 219
Gewindegang einer Schraube 275
Gitarrengriffbrett 175
Gitarrenkonverter 175
Glättungsalgorithmus 356
Gleichgewichtsorgan 279
Gleichlauf 192
Gleichlaufeigenschaft 325
Gleichlaufschwankung 392
Gleitband 125
GM-Standard 156
Goniometer 194
Grafikprozessor 364
Granular-Rauschen 56
Grenzflächendämpfung 113
Grenzflächenmikrofon 81
Grenzwert-Interpolation 104
Gruppenlaufzeit 14, 107
GS-Standard 156
- H**
Halbbildlücke 221
Halbbildverdopplung 318
- Halbzeilenverschiebung 224
Hall 18, 112, 209
Halldauer 113
Hallplatte 92
Hallsimulation 112
Hamming-Fensterfunktion 105
Haptik 279
Hard-Disk-Recording 36, 212
Hardware-Kompression 357
Harmonische 7, 115
Harmonizer 92, 121
Harvard-Architektur 127
HDTV 287
Helligkeitsauflösung 323
High-Band U-matic 257
HighCom 96, 124
Hilbert-Transformation 200
Hilbert-Transformator 27
Hinterleuchtung 240
Hochpass 106
Holografie 276
H-Phasen-Einstellung 314
Hüllkurve 131
Hüllkurvenkontrolle 123
Hüllkurvensteuerung 163
Huffman-Codierung 73, 369
- I**
I-Bild 370
ID-Nummer 147
IIR-Filter 107
Im-Kopf-Lokalisation 25
Immersion 279
Instrument 137
Intel-Format 398
Intensitäts-Stereofonie
 22, 194
Interferenz 17, 81
Interferenzstörung 219
Interlaced 220, 274
 -Flimmern 317
Interleaving 261
Intermodulation 8
Interpolation 53, 102, 287, 300
Interpolationswert 119
Invers-Echo 111
IRE-Einheit 224

J

Jitter 325, 392
Joint Stereo 73
JPEG-Verfahren 366

K

Kanaltrennung
Kassette 33
Kennlinie 310
Keyframe 353, 370
Keying 363
 externes 339
 inhaltliches 341
 internes 340
 örtliches 341
Key-Velocity 174
Klangregelung 90
Klangsynthese
 analoge 159
 FM- 162
 Fourier- 164
 LA- 165
 Sampling- 165
 Waveshaping- 164
Klaviatur 173
Klemmschaltung 225
Klirrdämpfung 8
Klirrfaktor 7, 66, 81, 211
Klirrkomponente 7
Knackgeräusch 103, 125
Koinzidenz-Mikrofon 22
Kolumnenzahl 287
Komplementärfarbe 299
Komponenten-Transcoder 388
Komponentensignal 309
Kompression 67, 399
Komprimierung 426
Kontrastkorrektur 310
Konvergenz 234
Konversionszeit 60
Kopfbewegung 25
Kopfneigung 211, 273
Kopfumschaltung 326
Kopiereffekt 33
Kopieren 99
Korrelationsgrad 73, 193
Kosinustransformation 368
 eindimensionale 72

Kunstkopf-Stereofonie 24
Kurven gleicher Lautstärke 11
Kurzschluss, akustischer 19

L

$\lambda/4$ -Plättchen 273
Laserdisc 260
Latenz-Jitter 213
Latenz-Zeit 212
Laufband 279
Lauflängencodierung 426, 428
Laufzeitdemodulation 387
Laufzeitstereofonie 24, 82, 193
Lautstärke 147
LCD
 -Monitor 235
 -Projektor 238
LED-Bildschirm 239
Leslie 93
Letterbox 265
Licht-Beamer 239
Lichtquelle 355
Lied-Positionszeiger 169, 180
Linearcodierung 297
Linearinterpolation 120
Linearitätsfehler, differentielle 65
Live-Modus 315, 325
Lochblende 234
Löschen 102
Löschkopf, fliegender 260
Longitudinal-Tonspur 255
Look-Up-Tables 310
Low-Band U-matic 257
LTC 169
Luftströmung 175
LUT 310, 342

M

MAC-Operation 127
Magnetbandaufzeichnung 33
Maskensignal 339
Maskierung 71
Master 138
 -Keyboard 174
 -Mode 320
Matrixverfahren 25

- MDCT (Modified Diskrete Cosinus Transformation) 70
- ME-Band 259
- Mediumrauschen 123
- Mehrwegsystem 21
- Messfrequenz 8
- Methode, 4-1-1- 286
- Metronom 181
- MIDI 398
- Datenprotokoll 135
 - Expander 157
 - Gerät 137
 - In 133
 - Instrument 137
 - Interface 135
 - Kanal 155
 - Keymap 159
 - Mapping 157
 - Message 139
 - Notennummer 138
 - Out 133
 - Patchmaps 158
 - Programm 138
 - Setup 157
 - Steckverbindung 135
 - Steuergerät 137, 174
 - Synchronisation 169
 - Thru 133
 - Verschaltung 134
- Mikrofon 81
- MiniDisc 35
- Mischer-Baustein 342
- Mischprodukt 10
- Mischung 8
- asynchrone 379
 - synchrone 380
- Missing Codes 60, 65
- Mittelwertbildung 54
- Modell, psychoakustisches 72
- Modulationsfrequenz 163
- Modulationstiefe 163
- Moiré-Bildung 241
- Monitoring 212
- Mono-Kompatibilität 193
- Mono Mode 144
- Monotonie 65
- Morphing 355
- Motion
- Adaptive Colour Plus (MACP) 264
- Capture 372
- Sickness 279
- MPEG-Verfahren 366, 370
- MP3 Surround 74
- MP3-Tags 41
- MP3-Verfahren 70
- MS-Verfahren 22, 73
- MTC-Generator 179
- Multimode 147, 155
- Multiplizierer 95
- Multiply & Accumulate 127
- Multitimbralität 156
- Mute 102
- N**
- Nachbarstützstelle 120
- Nachleuchtdauer 235
- Nachleuchten 318
- Nachvertonung 168, 177
- Nachzieheffekt 232, 235
- Nahtstelle 100, 102
- Nebengeräusch 119
- Nesa-Film 231
- Nierencharakteristik 22
- Noise-Normalization 77
- Noise Shaping 64
- Non-Interlaced 222
- Modus 331
- Normally Black Mode (NBM) 237
- Normally White Mode (NWM) 237
- Normenwandler 312, 390
- Normfarbvalenzen 231
- Notationsprogramme 175
- Notennummer 185
- NTSC (National Television System Committee) 262
- Nullstellen 114
- Nullwert 46
- Nyquist 50
- O**
- Oberflächenrauigkeit 260
- Oberflächenstruktur 354
- Oberton 7
- Objektgrafik 300
- Offset-Dual-Codierung 55
- Ohrkurvenfilter 11

- Omni-Mode 144
Operation,
 Ortsbereichs- 346
 Punkt- 345
Optokoppler 134
Ordnung 108
Organic Light Emitting Diode (OLED) 239
Ortung einer Schallquelle 22
OSS-Technik 83
Overlay 337
Overlay-Technik 329
Oversampling 51
Overscan 288
- P**
Pages 307
PAL
 -Decoder 387
 -Encoder 386
 -Flipflop 247
 -Jalousie 388
 -Laufzeitdemodulator 247
 -Sequenz, 8er- 319
Palette 285
Palettencodierung 297
Panning 347
Panorama 89, 142, 147
Parameter, registrierter 142
PASC (Precision Adaptive Sub-Coding) 34
PASC-Verfahren 69
P-Bild 370
PCM (Pulsecode-Modulation) 55
Peaks 110, 125
Pegelabweichung 4
Pegelanzeige 90
Pegelbearbeitung 122
Pegeleinstellung 89
Pegelfehler 388
Performance Capture 372
Phantomschallquelle 31
Phasenfrequenzgang 14
Phasenschieberkette 114
Phasensprung 166
Phasenverkopplung 252
Phasenverzerrung 14
Phaser 93
Phasing 21, 114, 209
Picture-in-Picture 348
- Pitch 120
 -Bereich 143
 -Ratio 92, 121
Pixelgrafik 299
Pixel-Interpolation 292
Pixelrate 293
Pixeltakt 313
Plasma-Schirm 238
PLL 392
PLL-Schaltung 325
Polarisation
 elliptische 273
 lineare 271
 zirkulare 273
Polarisationsfilter 272
Polarisationswinkel 273
Poly Mode 145
Polyphonie 156
Poppfilter 83
Precision Adaptive Sub-Coding (PASC) 34, 69
Primärfarben 228
Primärvalenz 230
Prisma 233
Programmänderung 143
Programmnummer 153
Programmwechsel 156
Projektion 361
Pulsecode-Modulation 55
Pulsweitenverfahren 260
Pulsweitenmodulation 78
Pyramid Vector Quantisation 78
- Q**
QPEG-Verfahren 371
Quadrofonie 25
Quantisierung 369
Quantisierungsfehler 56, 62
Quantisierungsrauschen 62
Quantisierungsstufe 55, 56
- R**
r 316
Rauschen 6, 94
RAMDAC 310
Random 344
Raumakustik 16, 81, 200
Rauschanteil 164

- Rauschimpuls 6
Rauschunterdrückung 96
Rauschunterdrückungsverfahren 33
Raytracing 355
Rechenzeit 107
Red
 Book 38
 Pitaya 201
Referenzton 184
Referenzwelle 276
Reflexion 17, 355
Reflexionspunkt 112
Reflexionsschall 25
Refraktion 19
Reinfarben 298
Remote-Schnittstelle 375
Rendering 279, 354
RGB-Signal 240, 308
Richtmikrofon 82
Richtungshören 22
Ringmodulator 95
Röhrenkamera 231
Rotary Digital Audio Tape
 (R-DAT) 34
Rotoscoping 361
Row-Address 292
Rowtable 348
RS
 232 173
 422 173
Rückkopplung 109
Rückkopplungspfad 107
Rumpeln 12
- S**
Sample&Hold 44
Sampler 167
 -Dumps 148
Sampling
 -Datei 183
 -Frequenz 119
Scan-Konverter 378
Schallbündelung 19
Schatten 355
Schattenbildung 388
Schattenschwingung 241
Schattenwelle 31
Schlaginstrument 152, 154
Schleife 166
Schmiereffekt 318
Schneidetechnik 99
Schnelleempfänger 87
Schnitt
 Assemble- 375
 -Computer 373
 Insert- 376
 non-linearer 376
 Offline- 376
 Online- 373
 Roh- 378
 physischer 205
 virtueller 206
Schnitteffekt 374
Schnittfehler 206
Schnittgenauigkeit 206
Schnittkanten-Interpolation 206
Schnittliste 379
Schnittposition 205
Schriftengenerator-Programm 363
Schroeder-Plot 200
Schwarzwert 225
Schwebung 334
Schwindelgefühl 279
SC-Phasen-Einstellung 315
SECAM (séquentielle à mémoire) 264
Seite 307
Seitenschwingung 163
Seitenverhältnis 290
Sektorenplan 39
Selektivität 108
Sequenz, 8er- 334
Sequenzer-Programm 169, 175, 178
SFI/SFD 398
Shannon 50
 -Theorem 48, 78
Shutter-Brille 271
Signal, S-VHS- 244
Signalart 308
 -Wandler 385
Signalprozessor 104, 208
Signalverformung 6
 manuelle 210
Single-Image-Stereogram 269
Slave 138
 -Mode 320
Slewrate-Filter 125
Slide 344

- SMP 398
SMPTE 169
-Offset 398
-Timecode 178
Spannungsintervall 55
Spannungspiegel 3
Spannungssprung 102
Speichertransfer 363
Spektralempfindlichkeit 229
Spektrumanalyser 6
Sphärendisplay 277
Spiegel, dichroitischer 233
Spiegelfrequenz 10
Split-Image-Verfahren 274
Sprungantwort, symmetrische 14
Square-Polar-Mapping 76
Steuerung, lokale 147
Störabstand 11, 88
Störrauschanpassung 11
Störzone 288
Standard, 4-2-2- 243, 261
Stanzfarbe 342
Stanzfläche 341
Stanzsignal 341
Statusbyte 140, 185
Stereofonie 22
Stereoskopie 268
Steuercomputer 133
Steuerungsänderung 140
Stimme 138
Stimmenanzahl 154
Stimmenumfang 138
Stimmung 142
Storyboard 378, 382
Strahlrücklauf 221
Strahlumkehr 273
Streifenfilter 233
Streuung 19
Strip 344
Studiotakt 316
Stufensignal 45
Stufungshöhe 47
Stummschalten 102
Sustain-Effekt 116
Synchrongenerator 247
Synchrongenerator 395
Synchronimpulse 326
Synchronisation 168, 177, 223
Synchronisationsfehler 320
Synchronsignal auf Grün 389
Synchrontakt 358
Synchronzeitpunkt 179
System
 Common 149
 Exclusive 147, 171
 Exclusive / Non-Realtime 148
 Exclusive / Realtime 148
 Realtime 151
- T**
Tag 420
Taktwechsel 175
Tastendruck 143
Tempiwechsel 175
Terzband 6
Testbild 321
Text-Overlay 363
Textur 354
TFT-Monitor 235
THD (Total Harmonic Distortion) 8
Theorem von Shannon 48, 78
Tieppass 106
 digitaler 52
 Filterung 51, 224
Tile 344
Timebase-Corrector 259, 315, 325, 392
Timecode 169, 358, 373, 377
 -Daten 148
 -Spur 257
 -Wandlung 171
Timeline 378, 382
Timing-Fehler 134
Tonerzeugung, multitimbraler 147
Tonhöhenänderung 143
Tonhöhenlage 92
Tonlage 120
Tonspeicherung 33
Trägerfrequenz 163
Transparenz 340
Treibermodus 291
Trittschall 84
Trittschallfilter 84
Tschebyscheff-Polynom 164
- U**
Überblenden 340
Überblendform 344

Überdeckung 34
 Übernahmeverzerrung 116
 Übersprechdämpfung 13
 Übersteuerung 81
 Übertragungsqualität 5
 UHDTV 268
 Underscan 289
 Undo-Funktion 210
 Unstetigkeit 102
 Ursprung 425

V

VCA (Voltage Controlled Amplifier) 28, 161
 VCF (Voltage Controlled Filter) 161
 VCO (Voltage Controlled Oscillator) 161
 Vektorskop 316
 Velocity 138
 Verdeckungseffekt 71
 Verfahren, S-VHS- 256
 Vertikalaustastlücke 288
 Verzerrung 115, 291
 kubische 10
 nichtlineare 6
 quadratische 10
 Verzerrung, lineare 5
 Verzerrungsnetz 356
 Verzögerung 107
 Verzögerungsglieder 113, 393
 Verzögerungszeit 90, 93, 109
 VGA-Feature-Connector 311
 VHS-HiFi-System 255
 VHS-Verfahren 254
 Video 2000 259
 Hi-8 259
 Videocrosspoint 331, 338
 Videoformat 287
 Vidikon 232
 Viertelbild-Message 170
 Viertelzeilenoffset 251
 Visual Cube 277
 VITC 169
 VOC 398
 Vocoder 93
 Vollausssteuerung 191
 Vollbildanzeige 328
 Voll-Message 171
 Volumendisplay 277
 Vorbis 75

Vorgang, rekursiver 108
 Vorecho 72
 Vormagnetisierung 211
 Vorspann 184
 Vorverlegung, zeitliche 393
 VR-Brille 278
 VTR-Controller 357

W

Wägeverfahren 58
 Wandlungs-Linearität 65
 Wandlungskennlinie 58
 WAV 397
 Wavetable-Karte 183
 Weißpunkt 230
 Wellenfeld-Synthese 30
 Wellenlängen-Multiplex-Verfahren 274
 Window-Anzeige 329
 Window-Wiedergabe, Probleme bei 330
 Wipe 344
 Wireframe 352
 Wobbelsignal 114
 WOLED (White-OLED) 239
 Wow & Flutter 192

X

XA-Format 39
 XLR 414
 XY-Verfahren 23

Y

Y/C-Signal 244, 309
 Yellow Book 38
 YUV-Signal 242, 309
 Y-Verzögerung 388

Z

Zählverfahren 58, 78
 Zeilenfrequenz 221
 Zeilenrücklauf 221
 Zeilensprungverfahren 220
 Zeilenzähler 348
 Zeilenzahl 287
 Zeitdehnung 122, 210
 Zeitinversion 210

- Zeitverschiebung 93
Zero Latency Monitoring (ZLM) 212
ZLM (Zero Latency Monitoring) 212
Zoom 347
1-Zoll-MAZ 260
- Zugriffszeit 364
Zuspieler 373
Zweierkomplementformat 55
Zweitonverfahren 9
Zwischenablage 99