



Digitale Bildcodierung

Kurseinheit 1:
Grundlagen der digitalen Bildtechnik

Autoren:
Prof. Dr.-Ing. Thomas Bonse
Prof. Dr.-Ing. Firoz Kaderali

unter Mitarbeit von Dr.-Ing. Michael Stepping

Gesamt-Inhaltsverzeichnis

Kurseinheit 1

1	Einführung in die Bildtechnik	1-1
1.1	Übersicht - Warum Bilder „codieren“	1-1
1.2	Physikalische und psychophysische Grundlagen - Was ist sichtbar?	1-2
1.2.1	Sichtbarer Bereich der optischen Strahlung	1-2
1.2.2	Die Helligkeitsempfindung.....	1-4
1.2.3	Die Ortsauflösung und Sehschärfe.....	1-5
1.2.4	Die Zeitauflösung.....	1-7
1.3	Einführung in die Farbmétriek - Was ist „Farbe“	1-11
1.4	Einführung in die analoge Videotechnik - Wie es anfang	1-20
1.4.1	Analoge Fernsehabtastung	1-21
1.4.2	Das FBAS-Signal.....	1-24
2	Digitale Bilddarstellung und -übertragung	1-28
2.1	Abtastung und Quantisierung	1-28
2.2	Digitale Bildformate, Farbdarstellungen und Standards	1-33
2.2.1	Farbabastraster	1-33
2.2.2	Digitale Studionorm ITU-R BT.601	1-35
	Lösungshinweise	1-47

Kurseinheit 2

2.3	Digitale Bildübertragungssignale und Schnittstellen in der professionellen Bildtechnik.....	2-1
2.3.1	Die parallele Bildübertragung.....	2-1
2.3.2	Die serielle Bildübertragung	2-1
2.3.3	Hochauflösende Bildübertragung	2-5
2.3.4	Bildformate in der Computerbilddarstellung.....	2-7
2.4	Daten- und Kompressionsraten – Motivation für die Bilddatenkompression.....	2-9
3	Digitale Quellencodierung für Einzelbilder	2-13
3.1	Transformation als Basis der digitalen Quellencodierung	2-14
3.1.1	Motivation	2-14
3.1.2	Eigenschaften der Transformationscodierung.....	2-14
3.1.3	Auswahlkriterien für Basisfunktionen	2-15
3.1.4	Von der diskreten Fourier Transformation zur diskreten Cosinus Transformation	2-16
3.1.5	Bildfelderlegung und DCT-Basisbilder	2-19
3.1.6	Beispiele für DCT-Blöcke	2-20
3.2	Quantisierung der DCT-Koeffizienten	2-23
3.2.1	Übersicht	2-23
3.2.2	Visuelle Anpassung der Quantisierung	2-24

3.2.3	Einstellung des Kompressionsgrades	2-25
3.2.4	Beispiel	2-26
3.3	Entropiecodierung der quantisierten Koeffizienten	2-29
3.3.1	Zick-Zack-Scan	2-29
3.3.2	Laufängencodierung (Run Length Coding - RLC)	2-30
3.3.3	Variable Laufängencodierung (Variable Length Coding - VLC)	2-30
3.3.4	Beispiele	2-33
	Lösungshinweise	2-37

Kurseinheit 3

3.4	Wavelet-Codierung und JPEG 2000	3-1
3.4.1	Eigenschaften der DCT-Basisfunktionen	3-1
3.4.2	Blockschaltbild eines verbesserten Transformations-Encoders	3-2
3.4.3	Wavelets als Basisfunktionen	3-3
3.4.4	Beispiel einer Wavelet Transformation - das Haar-Wavelet	3-8
3.4.5	Vergleich verschiedener Wavelets	3-9
3.4.6	Zusammenfassen und Codieren der Koeffizienten	3-9
3.4.7	Der Standard JPEG 2000	3-12
3.5	Einführung in die fraktale Bildcodierung	3-16
3.5.1	Die Idee der Selbstähnlichkeit	3-16
3.5.2	Rekonstruktion durch Iteration	3-16
3.5.3	Die Partitionierung (Bildfeldzerlegung)	3-19
3.5.4	Messung der Codiereffizienz	3-22
4	Prinzipien der digitalen Quellencodierung für Bewegtbilder	3-27
4.1	Bewegtbildcodierung auf der Basis von Einzelbildcodierung	3-27
4.1.1	Von JPEG nach Motion-JPEG	3-27
4.1.2	Die DV-Codierung	3-28
4.2	Prädiktive Bildcodierung mit DPCM	3-38
4.2.1	Feed-Forward DPCM - D*PCM	3-39
4.2.2	Übergang zur Feed-Backward DPCM	3-39
4.2.3	Feed-Backward DPCM mit Bildspeicher als Prädiktor ..	3-41
	Lösungshinweise	3-43

Kurseinheit 4

4.3	Bewegungsgesteuerte, prädiktive Bildcodierung	4-1
4.3.1	Das Prinzip der Bewegungskompensation	4-1
4.3.2	Verfahren zur Bewegungsschätzung	4-3
4.3.3	Der Blockmatching Algorithmus im Detail	4-5

4.3.4	Minimierung des Rechenaufwandes beim Blockmatching Algorithmus	4-7
4.3.5	Rekursiv hierarchisches Blockmatching	4-9
4.3.6	Anwendung des Blockmatching Algorithmus mit der DPCM	4-9
4.3.7	Hybride DCT als Kombination aus DPCM und DCT ...	4-10
5	Standards der digitalen Videocodierung	4-12
5.1	Einfache Videocodierkonzepte der ITU-T - H.120 und H.261	4-13
5.1.1	H.120	4-13
5.1.2	H.261	4-14
5.2	MPEG-1	4-18
5.2.1	Einführung	4-18
5.2.2	Hierarchieebenen	4-20
5.2.3	MPEG Group-of-Pictures - Übersicht	4-21
5.2.4	Intraframe-Codierung – I-Blöcke und	4-22
5.2.5	Interframe-Codierung – P-Blöcke und P-Bilder	4-24
5.2.6	Bidirektionale Interframe-Codierung – B-Blöcke und B-Bilder	4-25
5.2.7	Regelung der Datenrate	4-29
5.2.8	Blockschaltbild des MPEG-1 Encoders	4-30
	Lösungshinweise	4-33
Kurseinheit 5		
5.3	MPEG-2 / H.262	5-1
5.3.1	Einführung	5-1
5.3.2	Profiles und Levels	5-4
5.3.3	Codierung von Halbbildern	5-8
5.3.4	MPEG-Datenstrukturen und -Multiplex	5-10
6	Videostandards in Anwendungen	5-16
6.1	Videostandards für die CD	5-16
6.1.1	Entwicklung der Video-CD	5-16
6.1.2	Parameter der Video-CD 2.0 und der Super Video-CD..	5-17
6.1.3	Beschreibung der Video-CD Standards im Detail	5-18
6.2	Der DVD-Video Standard	5-20
6.2.1	Parameter der	5-20
6.2.2	Besonderheiten der DVD-Video	5-21
6.2.3	Datenstruktur der DVD-Video	5-23
6.2.4	Logische Navigationsstruktur der DVD-Video	5-25
6.2.5	Navigation und Interaktivität.....	5-27
6.3	Multiplexstrukturen für Broadcastanwendungen	5-29
6.3.1	Einführung	5-29
6.3.2	PSI-Tabellen beim MPEG-2-TS	5-29
6.3.3	Zusätzliche Service Information bei DVB	5-31

6.3.4	Statischer und dynamischer Transportmultiplex	5-32
6.3.5	Beispiel	5-32
Lösungshinweise		5-35
 Kurseinheit 6		
6.4	Kanalcodierung für die digitale TV-Übertragung	6-1
6.4.1	Energieverwischung	6-1
6.4.2	Bitfehlerrate.....	6-3
6.4.3	Bitfehlermuster	6-3
6.4.4	Code-Klassen	6-4
6.4.5	Codeverkettung	6-5
6.4.6	Faltungsinterleaver	6-5
6.4.7	Reed-Solomon Code bei DVB	6-6
6.4.8	Faltungscode bei DVB	6-8
6.4.9	Punktierung von Faltungscode.....	6-9
6.5	DVB-Übertragungstechniken	6-11
6.5.1	Signalvorbereitung	6-11
6.5.2	Digitale Modulation für Satellitenkanäle	6-13
6.5.3	DVB-S Übertragungskonzept	6-14
6.5.4	Digitale Modulation für Breitbandkabelkanäle	6-16
6.5.5	DVB-C Übertragungskonzept.....	6-18
6.5.6	Digitale Modulation für die terrestrische TV- Übertragung.....	6-20
6.5.7	DVB-T Übertragungskonzept.....	6-23
7	Verbesserte Codiersysteme für die Bildkommunikation	6-27
7.1	ITU-T H.263	6-27
7.1.1	Übersicht	6-27
7.1.2	Struktur des ITU-T H.263 Standards in der Grundversion.....	6-28
7.1.3	ITU-T H.263, Version 1 – Codieroptionen	6-31
7.1.4	ITU-T H.263, Version 2 (H.263+) – Codieroptionen	6-32
7.1.5	ITU-T H.263, Version 3 (H.263++) – Codieroptionen ..	6-35
Lösungshinweise		6-39
 Kurseinheit 7		
7.2	ISO/IEC 14496-2 (MPEG-4 Video)	7-1
7.2.1	Konzept des MPEG-4 Standards.....	7-1
7.2.2	Leistungsmerkmale von MPEG-4 Video	7-4
7.2.3	Die bei MPEG-4 Video.....	7-4
7.2.4	Strukturebenen bei MPEG-4 Video.....	7-6
7.2.5	Funktionsdiagramm des MPEG-4 Videoencoders	7-7
7.2.6	Visuelle Profiles und Levels von MPEG-4	7-9

7.2.7	Beispiel für synthetisch generierte visuelle Objekte bei MPEG-4.....	7-15
7.2.8	Abschließende Bemerkungen	7-18
7.3	Advanced Video Coding - AVC.....	7-21
7.3.1	Überblick.....	7-21
7.3.2	Eigenschaften	7-22
7.3.3	Profiles und Levels	7-24
7.3.4	Blockbildung für die Bewegungskompensation	7-25
7.3.5	Örtliche Intra-Prädiktion	7-26
7.3.6	Integer-Transformation	7-32
7.3.7	Quantisierung und Auslesen der Koeffizienten	7-35
7.3.8	Weitere Elemente	7-37
7.3.9	Bewertung des Standards	7-37
7.4	Abschließende Bemerkungen	7-39
7.4.1	Entwicklung der digitalen Bildcodierung	7-39
7.4.2	Motion-JPEG 2000	7-40
7.4.3	Abschluss	7-40
	Lösungshinweise	7-43

Gliederung

Kurseinheit 1

1	Einführung in die Bildtechnik	1-1
1.1	Übersicht - Warum Bilder „codieren“	1-1
1.2	Physikalische und psychophysische Grundlagen - Was ist sichtbar?	1-2
1.2.1	Sichtbarer Bereich der optischen Strahlung	1-2
1.2.2	Die Helligkeitsempfindung	1-4
1.2.3	Die Ortsauflösung und Sehschärfe	1-5
1.2.4	Die Zeitauflösung	1-7
1.3	Einführung in die Farbmeterik - Was ist „Farbe“	1-11
1.4	Einführung in die analoge Videotechnik - Wie es anfing	1-20
1.4.1	Analoge Fernsehabtastung	1-21
1.4.2	Das FBAS-Signal	1-24
2	Digitale Bilddarstellung und -übertragung	1-28
2.1	Abtastung und Quantisierung	1-28
2.2	Digitale Bildformate, Farbdarstellungen und Standards	1-33
2.2.1	Farbabastraster	1-33
2.2.2	Digitale Studionorm ITU-R BT.601	1-35
	Lösungshinweise	1-47

1 Einführung in die Bildtechnik

1.1 Übersicht - Warum Bilder „codieren“

Etwa mit Beginn der 90er Jahre beeinflussten vor allem zwei Entwicklungen den Umgang mit Bewegtbildern entscheidend: Es sind dies zum einen die rasanten Veränderungen im Bereich der digitalen Bildverarbeitung (*digitale Bildcodierung* und *Übertragung*) und zum anderen die sich hierauf aufbauenden zahlreichen neuen Anwendungsplattformen, wie beispielsweise Multimedia (MM), der Freizeitsektor, zahlreiche neue Kommunikations- und Datendienste sowie nicht zuletzt die Computer Animation (CA).

Eine entsprechende Bildqualität vorausgesetzt, benötigen die Signale bzw. die Daten von analogen bzw. digitalen Bewegtbildern eine relativ hohe Kanalbandbreite bei der Übertragung (TV-Kanalabstand: 8 MHz) oder eine hohe Kapazität bei der Speicherung auf entsprechenden Bild- und Datenträgern (z. B. *magnetische* oder *optische Speicherung*). Um eine entsprechende Ökonomie zu erreichen, ist die optimale Anpassung der Signalströme für den jeweiligen Anwendungsfall erforderlich. Insbesondere im digitalen Umfeld wird eine solche Anpassung im allgemeinen „Codierung“ genannt. Die Bildcodierung ist aber nicht prinzipiell auf digitale Bildsignale beschränkt. Auch beim analogen Fernsehen spricht man beispielsweise von Codierung, wenn es um die optimale Nutzung des TV-Übertragungskanals durch die Einführung der Zeilensprungtechnik geht. Auch die Einbettung des analogen *Farbartsignals* in das *Helligkeitssignals* stellt in diesem Zusammenhang eine Codierung dar. Es hat sich jedoch gezeigt, dass sich optimale Codierungen im analogen Umfeld oft nur mit erheblichem Aufwand realisieren lassen. Im digitalen Umfeld hingegen lassen sich erheblich vielfältigere Algorithmen entwickeln und in entsprechenden Computerumgebungen direkt verwirklichen. In sehr vielen Fällen geht es daher bei der *digitalen Bildcodierung* um eine Bewegtbildrepräsentation mit erheblich reduziertem Datenaufkommen, ohne dass es dabei zu einer sichtbaren Bildqualitätsverschlechterung kommt.

Der vorliegende Kurs „*Digitale Bildcodierung*“ beschäftigt sich mit den Grundlagen der digitalen Bildverarbeitung, wie sie heutzutage in zahlreichen Anwendungen genutzt werden. Als Anwendungsbeispiele sind hier zu nennen die *Digitale Fotografie*, die digitale Videodisk - *Digital Versatile Disk (DVD)* oder das digitale Fernsehen - *Digital Video Broadcasting (DVB)*.

Viele der aktuell in der digitalen Bildcodierung verwendeten Algorithmen sind das Ergebnis jahrzehnte langer intensiver, interdisziplinärer Forschungs- und Entwicklungsarbeiten. Zum vollständigen Verständnis der heutigen eingesetzten Verfahren zur Bildcodierung ist daher auch die Grundlagenkenntnis der Theorie der umfangreichen *analogen Bildtechnik* notwendig. Es würde jedoch den Rahmen dieses Kurses bei weitem sprengen, wollte man diese Grundlagen vollständig behandeln. So beschränkt sich dieser Kurs mit der Darstellung einiger weniger theoretischer Zusammenhänge, sofern diese dringend für das Verständnis der digitalen Bildco-

dierung erforderlich sind. Für eine Vertiefung sei auf die einschlägige Literatur verwiesen (z. B. [1.7] und [1.8]).

Inhaltsübersicht

Kapitel 1 befasst sich mit der *visuellen Physiologie*, der *Farbmetrik* und den *Grundlagen zur analogen Videotechnik*. **Kapitel 2** legt die *Grundlagen zur digitalen Bildverarbeitung*; der Übergang von der analogen zur digitalen Bildtechnik mit den üblichen *Formaten* und *Standards* wird dargestellt. In **Kapitel 3** wird die *digitale Quellencodierung für Stillbilder* behandelt. Zentraler Bestandteil dieses Kapitels ist der *JPEG-Standard*, der als quasi Ur-Vater vieler aktueller Codieralgorithmen der Bildtechnik gilt. Die weiter entwickelten Systeme, die auf der Wavelettransformation basieren (*JPEG 2000*) und die fraktale Bildcodierung sind ebenfalls Gegenstand dieses Kapitels. Die Prinzipien der digitalen Quellencodierung für Bewegtbilder bilden den Schwerpunkt des **Kapitel 4**. Als zentrale Verfahren werden *Motion-JPEG*, die *DV-Codierung* und die *prädiktive Bildcodierung* mit und ohne Bewegungskompensation vorgestellt. Die wichtigen Standards der digitalen Videocodierung sind Gegenstand des **Kapitel 5**. Vorgestellt werden unter anderem *ITU-T H.261*, *MPEG-1* und *MPEG-2*. Schwerpunkt des **Kapitel 6** ist die Beschreibung von Videostandards in ausgewählten Anwendungen. Hierzu werden gerechnet die Videocodierverfahren für die optischen Medien CD und DVD sowie die digitale Bildübertragung über TV-Broadcastnetze (*DVB-Verfahren*). *Verbesserte Codiersysteme für die Bildkommunikation* finden sich im abschließenden **Kapitel 7**. Hier werden die Grundideen der komplexen Standards *H.263*, *MPEG-4* und *H.264/Advanced Video Coding (AVC)* vorgestellt.

1.2 Physikalische und psychophysische Grundlagen - Was ist sichtbar?

Vorwiegend den Forschungsaktivitäten in der Physiologie, in der Psychologie und in der Medizin ist es zu verdanken, dass auf dem komplexen Gebiet der *visuellen Wahrnehmung* in den letzten 40 Jahren umfangreiche Erkenntnisse gesammelt werden konnten. Die interdisziplinäre Nutzung dieser Ergebnisse für die digitale Bildcodierung führte zu entscheidenden Fortschritten, insbesondere auf dem Gebiet der *Datenkompression*. Nachfolgend werden die physikalischen Grundlagen und die wichtigsten Erkenntnisse aus der visuellen Physiologie zusammengefasst dargestellt.

1.2.1 Sichtbarer Bereich der optischen Strahlung

Licht als
elektromagnetische
Strahlung

Im nachfolgenden soll unter *Licht* der Teil einer *optischen Strahlung* verstanden werden, der für das menschliche visuelle System wahrnehmbar ist. Die optische Strahlung wiederum kann physikalisch als elektromagnetische Welle beschrieben werden, in der sich die elektrische und magnetische Feldstärken periodisch ändern. Die Ausbreitungsgeschwindigkeit c einer solchen Welle ist im Vakuum gleich der Lichtgeschwindigkeit c_o :

$$c = c_o = 2,998 \cdot 10^8 \frac{m}{s}. \quad 1.2-1$$

Die Welle wiederholt sich periodisch in Raum und Zeit, so dass eine Periodendauer T definiert werden kann, welche die Anzahl der zeitlichen Perioden pro Sekunde als Frequenz f bzw. als Kreisfrequenz ω angibt:

$$f = \frac{1}{T} = \frac{\omega}{2\pi} \quad 1.2-2$$

Die räumliche Periodizität ist durch die Wellenlänge λ vorgegeben. Der formale Zusammenhang zwischen Wellenlänge λ und Frequenz f ist gegeben durch die Ausbreitungsgeschwindigkeit c :

$$c = \lambda \cdot f. \quad 1.2-3$$

Der für das menschliche Auge wahrnehmbare Bereich des elektromagnetischen Wellenspektrums belegt nur einen relativ geringen Teil des Frequenzspektrums. Die zugehörigen Wellenlängen erstrecken sich etwa über den Bereich zwischen 380 und 780 nm. Das Farbtonempfinden des Auges hängt ausschließlich von der Wellenlänge ab. So erscheint Licht mit der Wellenlänge von 400 nm blau-violett, 700 nm rot, alle weiteren Regenbogenfarben liegen dazwischen. Abb. 1.2-1 zeigt das Spektrum der elektromagnetischer Wellen. Besonders hervorgehoben ist der Bereich des sichtbaren Lichts.

Sichtbares Licht

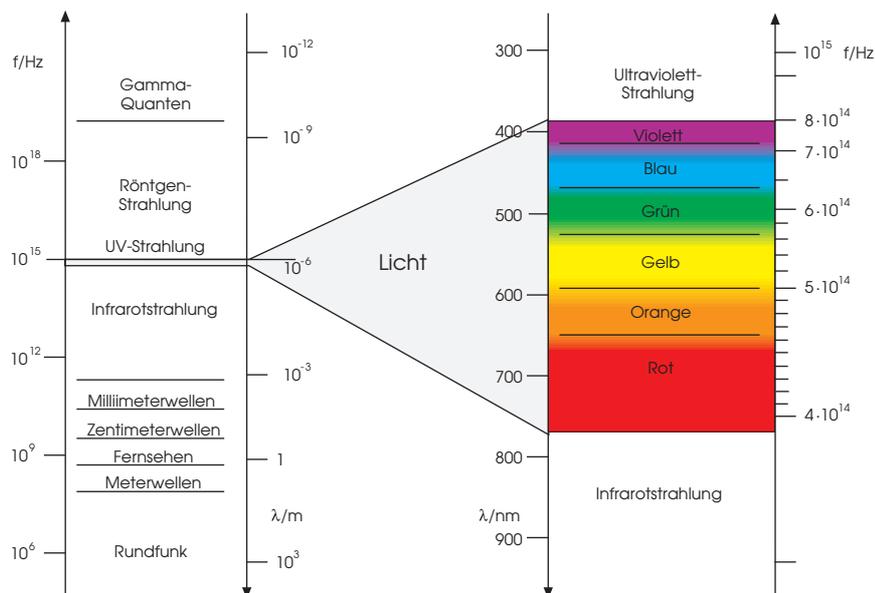


Abb. 1.2-1: Spektrum elektromagnetischer Wellen und sichtbarer Wellenlängenbereich

Hellempfindlichkeits-
grad
 $V(\lambda)$

1.2.2 Die Helligkeitsempfindung

Während die Frequenz einer elektromagnetischen Wellen eindeutig den *Farbton* des Lichts bestimmt, dienen die so genannten *fotometrische Größen* zur Messung der *Lichtintensität* und der *Helligkeit*. Dazu werden *lichttechnische Größen*, welche die energetischen Anteile des Lichts repräsentieren mit dem *spektralen Hellempfindlichkeitsgrad* $V(\lambda)$ des Auges gewichtet (vgl. [1.2]). Der relative spektrale Hellempfindlichkeitsgrad gibt die Frequenzabhängigkeit bezüglich der Augenempfindlichkeit an. Ein hell adaptiertes Auge („*Tagesehen*“) ist für den spektralen Bereich um 555 nm (*grün*) maximal empfindlich. Bei einem dunkel adaptierten Auge („*Nachtsehen*“) verschiebt sich dieses Maximum Richtung *blau* auf etwa 505 nm. Abb. 1.2-2. zeigt die spektrale Hellempfindlichkeit des menschlichen Auges.

Sehzellentypen

Der Grund für diese Verschiebung liegt in der unterschiedlichen Aktivität der beiden Sehzellentypen, der *Stäbchen* und der *Zapfen*, für dunkle und für helle Lichtumgebungen. Die Stäbchen besitzen eine größere Lichtempfindlichkeit und sind daher in dunkleren Umgebungen mehr aktiv. Für sie ist die Kurve für das dunkel adaptierte Auge charakteristisch (*grün*). Die Zapfen sind in den hellen Beleuchtungssituationen aktiv und bestimmen somit diese spektrale Hellempfindlichkeitskurve (*rot*). Zudem sind sie auch für die Farbunterscheidungsfähigkeit verantwortlich.

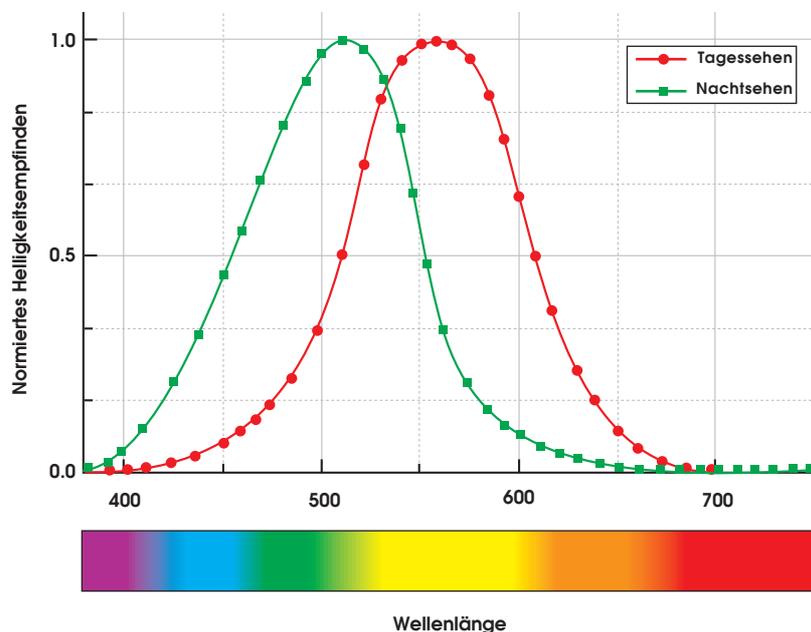


Abb. 1.2-2: Spektrale Hellempfindlichkeit des menschlichen Auges

Helligkeitsadaption

Die Hellempfindung ist also vom *Adaptionszustand* des Auges und der *Umfeldbeleuchtung* bestimmt. Die besondere Leistung des Auges besteht darin sich an Leuchtdichten anzupassen, die einen Bereich von 11 Zehnerpotenzen umfassen. Allerdings ist die Adaption mit verschiedenen Zeitkonstanten bis hin zu 30 Minuten verbunden [1.2]. Bei adaptiertem Auge können immerhin noch etwa 200 Helligkeitswerte (*Graustufen* oder *Luminanzwerte*) unterschieden werden.

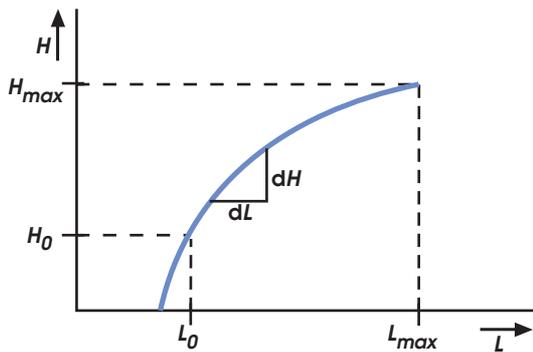


Abb. 1.2-3: Weber-Fechnerisches Gesetz

Der zuvor genannte Zusammenhang zwischen subjektiv wahrgenommener Helligkeit H und der relativer Leuchtdichte L ist im mittleren Leuchtdichtebereich *logarithmisch* (vgl. Abb. 1.2-3).

Weber-Fechnersches Gesetz

Dieser als *Weber-Fechnersches Gesetz* bekannte Zusammenhang lautet formal:

$$H = \log(a \cdot L), \quad 1.2-4$$

wobei a eine Konstante ist.

1.2.3 Die Ortsauflösung und Sehschärfe

Der Gesichtssinn umfasst horizontal zwar insgesamt einen Winkel von etwa 180° , dennoch können innerhalb dieses Bereiches nicht alle Lichtreize gleichzeitig und in gleicher Auflösung vom Auge wahrgenommen werden. Tatsächlich ist der Bereich, der „auf einen Blick“ erfassbar und scharf wahrnehmbar ist, nur etwa 3° groß. Dieser Bereich korrespondiert mit dem überwiegenden Teil des Sehzellentyps „Zapfen“ auf der Netzhaut. Soll ein größerer Bildbereich vom Auge betrachtet werden, muss dieser mit zeitlich sequenziellen Augenbewegungen überstrichen und kognitiv zu einem Gesamtbild zusammengesetzt werden (vgl. [1.9]).

Bereich der größten Schärfewahrnehmung

Die Trennbarkeit von Linien bzw. die Kontrastempfindlichkeit des Auges hängt vom *Modulationsgrad* und von der *Beleuchtungsstärke* ab. Die Empfindlichkeit ist bei 10 Linien pro Winkelgrad am höchsten und sinkt sowohl zu niedrigen als auch zu hohen Ortsfrequenzen hin ab. Insgesamt besitzt das Auge bezüglich seiner Ortsauflösung eine *Bandpasscharakteristik*. Diese Eigenschaft ist u. a. auf die so genannte *laterale Hemmung* zurückzuführen: Wird eine Sehzelle durch einen Lichtreiz aktiv, hemmt diese Aktivität gleichzeitig die Aktivität unmittelbar benachbarter Sinneszellen. Abb. 1.2-4 skizziert diesen Zusammenhang.

Laterale Hemmung

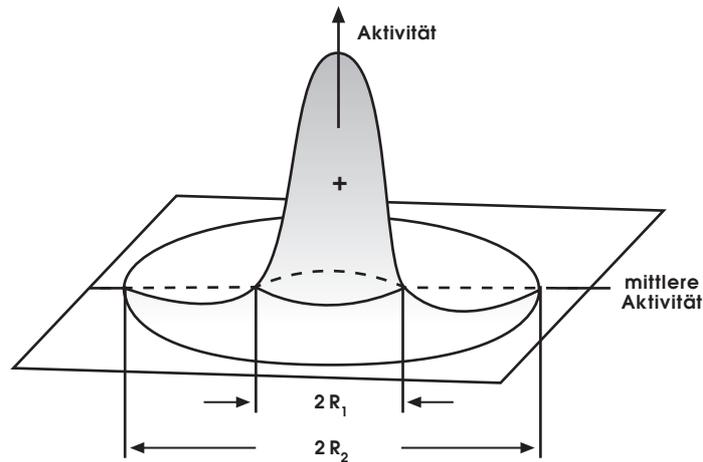
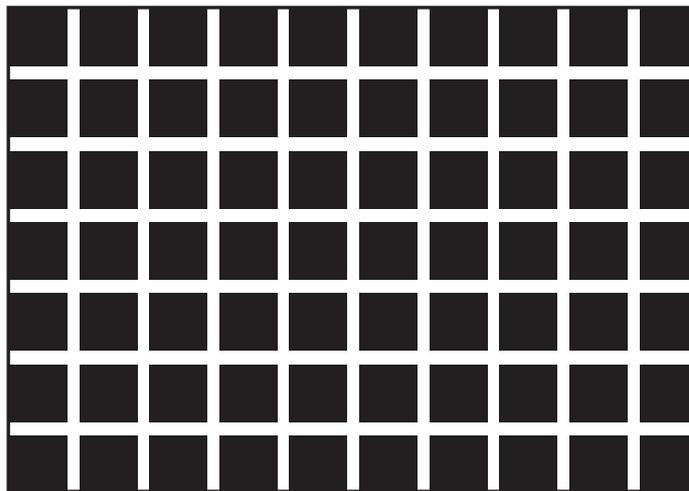


Abb. 1.2-4: Reizung im rezeptiven Feld einer Nervenfaser (nach [1.2])

Mach-Effekt Die laterale Hemmung hat interessante und gleichzeitig wichtige Auswirkungen auf die menschliche Sehleistung. So verursacht sie den *Mach-Effekt*. Dieser Effekt ist besonders gut am so genannten *Hermanngitter* zu beobachten (vgl. Abb. 1.2-5).

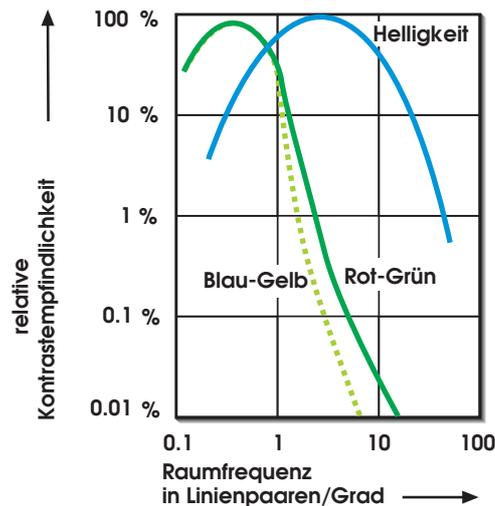


Hermanngitter

Abb. 1.2-5: Hermanngitter zur Verdeutlichung des Mach-Effekts

An den Kreuzungspunkten tauchen – quasi geisterhaft – scheinbare Verdunkelungen auf. Der Mach-Effekt ist auch an Objektkanten zu beobachten und führt in der Regel subjektiv zu einer *Kontrasterhöhung im Kantenbereich*. Das Ergebnis ist eine visuell *verbesserte Objektextraktion* auch bei wenig ausgeprägten Kontrastverhältnissen. Der für die Bildcodierung nutzbare Nebeneffekt ist die relative Unempfindlichkeit gegenüber leichten Verfälschungen im Kantenbereich selber. Leichte örtliche Verformungen sowie geringe Helligkeitsschwankungen werden hier maskiert oder zumindest wenig störend empfunden.

Anwendung Die hier beschriebenen örtlichen Auflösungseigenschaften des Gesichtssinns werden besonders effektiv in Codierverfahren genutzt, die mit einer Transformationscodierung arbeiten. In Kapitel 4 werden solche Transformationscodierungen vorgestellt. Der weit verbreitete *JPEG*-Algorithmus ist ein Beispiel hierfür.



Kontrastempfindlichkeit
Farbe und Helligkeit

Abb. 1.2-6: Kontrastempfindlichkeit des visuellen Systems für farbart- und helligkeitsmodulierte räumliche Gitter (nach [1.11])

Eine interessante visuelle Eigenschaft, die auch schon für eine Optimierung der analogen Farbbildcodierung genutzt wurde, ist die Tatsache, dass es unterschiedliche Kontrastempfindlichkeiten für farbartmodulierte und für rein helligkeitsmodulierte Gitter gibt. In [1.11] wurde dieser Sachverhalt intensiv untersucht.

In Abb. 1.2-6 ist das Ergebnis dargestellt. Es wird deutlich, dass die Kontrastempfindlichkeit für Farbartübergänge zu höheren Frequenzen wesentlich früher abfällt als für Helligkeitsübergänge. Somit ist die Auflösung des Gesichtsinns für Farbartübergänge (bei gleicher Helligkeit) wesentlich geringer als für Helligkeitsübergänge, oder anders ausgedrückt: Feine Objektdetails erkennt das menschliche Auge vorwiegend über einen Helligkeitssprung (*Kante*) und weniger über eine Änderung der Farbinformation.

Die unterschiedliche Empfindlichkeit des menschlichen visuellen Systems für Helligkeits- und Farbsprünge kann in der digitalen Bildcodierung vorteilhaft genutzt werden, in dem die Farbinformation örtlich 2 bis 4 mal niedriger aufgelöst codiert wird, als die Helligkeitsinformation. In Kapitel 2 werden diese Systeme vorgestellt, welche mit einer örtlichen Unterabtastung der Farbe arbeiten.

Anwendung

Abschließend sei bemerkt, dass die Auflösungsleistung des visuellen Systems nur für ruhende Bildvorlagen so groß ist, wie hier beschrieben. Soll das menschliche Auge unwillkürlich bewegten Objekten folgen, können diese meist in ihrer Struktur nur unscharf erkannt werden.

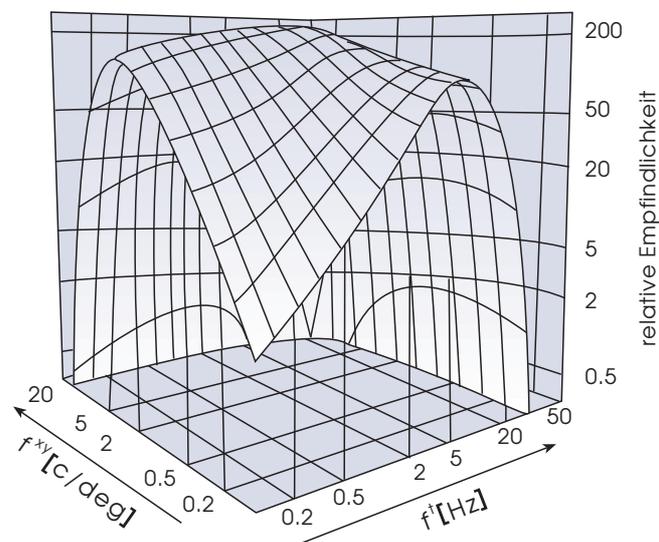
1.2.4 Die Zeitauflösung

Ähnlich wie in örtlicher Richtung weist das menschliche visuelle System auch in zeitlicher Richtung eine Bandpasscharakteristik auf. Konkret bedeutet das: Zeitlich sehr langsam wiederholte Sinnesreize werden weniger deutlich wahrgenommen als Sinnesreize, die in einem Frequenzbereich zwischen etwa 5 bis 8 Hz liegen. Höhere Frequenzen werden wieder weniger deutlich wahrgenommen. Dabei sind zeitliche

und örtliche Frequenzwahrnehmungen nicht unabhängig voneinander. Die Abb. 1.2-7 und Abb. 1.2-8 zeigen die empirisch aufgenommene örtlich-zeitliche *Schwellenempfindlichkeit* nach einer Messung aus [1.10].

Frequenz für eine Bewegungsverschmelzung

Für die Bewegtbildübertragung stellt sich weiterhin die wichtige Frage, wie viele Bilder pro Sekunde übertragen werden müssen, damit der Gesichtssinn einen in den Bildern dargestellten Bewegungsablauf als kontinuierlich zusammenhängend empfindet (*Bewegungsverschmelzung*). Untersuchungen haben gezeigt, dass ein gleichmäßig erscheinender Bewegungsablauf bereits mit einer Rate von etwa 20 Bildern pro Sekunde erzielbar ist [1.9]. Die Filmbildfrequenz beträgt bekanntermaßen 24 Hz. Trotzdem gibt es kein Bildwiedergabegerät, das mit einer derart niedrigen Frequenz eine Bewegtbilddarstellung realisiert.



Örtlich-zeitliche Schwellenempfindlichkeit

Abb. 1.2-7: Örtlich-zeitliche Schwellenempfindlichkeit - 3D-Darstellung (vgl. [1.9] und [1.10])

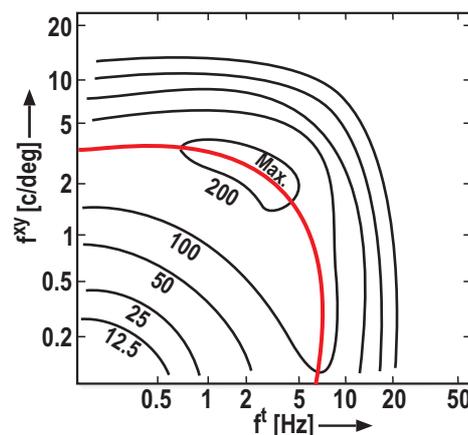


Abb. 1.2-8: Örtlich-zeitliche Schwellenempfindlichkeit - 2D-Querschnitt zur Verdeutlichung der Max-Kurve (vgl. [1.9])

Flimmer-Verschmelzungsfrequenz

Der mit Dunkelpausen verbundene Bildwechsel führt zu einer ständig wiederholten Erregung und Hemmung der Neuronen, was als sehr unangenehmes Flimmern empfunden wird. Dieses Flimmern wird als *Großflächenflimmern* bezeichnet, da es alle Bildpunkte zugleich betrifft. Die so genannte *Flimmer-Verschmelzungsfrequenz*, bei

der diese Flimmerempfindung verschwindet, liegt oberhalb von 50 Hz und steigt mit der Leuchtdichte der Bildarstellung und entscheidend auch mit dem Betrachtungswinkel: Große Bildarstellungen werden bei sonst gleichen Parametern eher als flimmernd wahrgenommen als kleine. Bei Wiedergabesysteme, bei denen das Bild einen großen Sehwinkel überstreicht (z. B. Computerdisplays), sollte die Bildwiederholfrequenz zwischen 70 Hz und 100 Hz liegen.

Dies ist der Grund, warum Kinofilmbilder, welche mit 24 Hz aufgenommen wurden bei der Wiedergabe zweimal (oder dreimal) projiziert werden, so dass sich eine Verdopplung (bzw. Verdreifachung) der Dunkelpausen bei einer Flimmerfrequenz von 48 Hz (bzw. 72 Hz) ergibt. Mit dieser Maßnahme reduziert sich zwar deutlich der Flimmereindruck, doch kann es bei Bewegungen (bewegte Objekte, Kameranachschwenk etc.) zu deutlich wahrnehmbaren Ruckelstörungen kommen, da die örtliche Position von bewegten Objekten bei dieser Darstellungsweise in den wiederholten Zwischenprojektionen der Erwartung des visuellen Systems in Bezug auf eine gleichmäßige Bewegung widerspricht.

Abb. 1.2-9 skizziert diesen Zusammenhang an Hand einer gleichmäßigen translatorischen Bewegung.

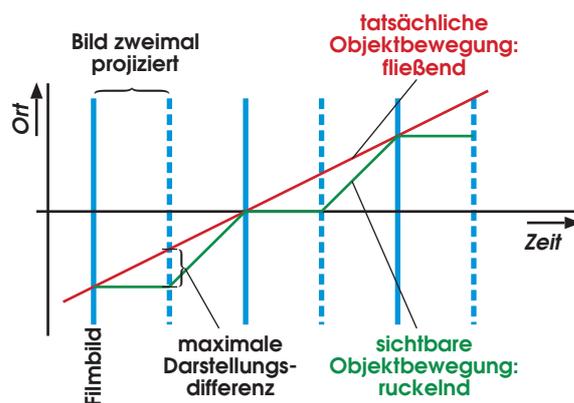


Abb. 1.2-9: Zur Entstehung von Ruckelstörungen bei der Kinoprojektion mit Bildwiederholung

Deutlich zu erkennen ist, dass die sichtbare Objektbewegung (grün) einen ungleichmäßigen Verlauf zeigt. Die maximale örtliche Differenz zwischen tatsächlicher Objektposition und dargestellter Position geschieht zum Zeitpunkt der Bildwiederholung durch eine erneute Projektion des Filmbildes.

Wird das Filmbild zweimal wiederholt, entsteht eine Bildfolgefrequenz von 72 Hz. Das korrespondierende Bildflimmern reduziert sich damit weiter. Allerdings erhöht sich auch die maximale örtliche Differenz zwischen tatsächlicher Objektposition und dargestellter Position zum Zeitpunkt der zweiten Filmbildwiederholung. Als Konsequenz daraus verstärkt sich der Ruckeleindruck bei Objekt- und Kamerabewegungen. In der Praxis wird daher als Kompromiss aus Ruckeln und Flimmern bei der Kinobildarstellung nur eine einfache Bildwiederholung angewendet.

Korrelation Ruckeln und Bildwiederholung

Selbsttestaufgabe 1.2-1:

Geben Sie den ungefähren Wellenlängenbereich für die sichtbare optische Strahlung an!

Selbsttestaufgabe 1.2-2:

Wie heißen die beiden Sensorentypen der menschlichen Netzhaut? Welche Aufgaben kommen welchen Sensoren zu?

Selbsttestaufgabe 1.2-3:

Beschreibung und erklären Sie kurz den Mach-Effekt!

Selbsttestaufgabe 1.2-4:

Erklären Sie den Unterschied zwischen Bewegungsverschmelzungsfrequenz und Flimmerverschmelzungsfrequenz!

Selbsttestaufgabe 1.2-5:

Mit welcher Maßnahme lässt sich eine flimmerfreie Bilddarstellung beim Kinofilm realisieren? Welchen gravierenden Nachteil muss man dabei in Kauf genommen werden?

Selbsttestaufgabe 1.2-6:

Wie verhält sich das örtliche Auflösungsvermögen des menschlichen Auges in Bezug auf Helligkeitsübergänge im Vergleich zum Auflösungsvermögen bei Farbübergängen?

1.3 Einführung in die Farbmeterik - Was ist „Farbe“

In den nachfolgenden Abschnitten wird auf der Basis der farbmeterischen Eigenschaften des menschlichen Sehorgans die allgemeine *Farbtheorie* und *Farbenlehre* hergeleitet.

Wie bereits in Abschnitt 1.2.2 ausgeführt wurde, besitzt der menschliche Gesichtssinn zwei Typen von lichtempfindlichen Organen: Die *Stäbchen* und die *Zapfen*.

Die *Zapfen* haben ihre größte Dichte auf der Netzhaut innerhalb eines zentralen Bereiches (*Fovea Centralis*) von etwa 3° (bezogen auf den horizontalen und vertikalen Blickwinkel des Sehorgans). Diese Sinnesorgane sind zum einen verantwortlich für die Hellempfindlichkeitskurve „*Tagesehen*“ (vgl. Abb. 1.2-2) und zum anderen ermöglichen sie die Wahrnehmung von Farben. Dazu stellte bereits 1807 der englische Mediziner und Physiker Thomas Young in seinen „*Lectures of Natural Philosophy*“ die Hypothese auf, dass es im Auge drei Typen von Farbsehzellen geben muss. In der zweiten Hälfte des 19. Jahrhunderts baute der Physiologe und Physiker Hermann von Helmholtz die Theorie in seiner *Dreikomponententheorie* aus. Tatsächlich gibt es drei unterschiedliche Zapfenarten, die sich hinsichtlich der Zapfensehstoffe unterscheiden. So reagieren diese Zapfenarten auf die einfallende Lichtstrahlung im Spektralbereich von *blau* (etwa 400 bis 500 nm), *grün* (etwa 500 bis 600 nm) und *rot* (etwa 600 bis 700 nm) jeweils unterschiedlich. Über eine Mischung der drei Teilempfindungen entsteht der Farbeindruck. Abb. 1.3-1 zeigt die Absorptionsspektren der drei Zapfenarten.

Zapfen

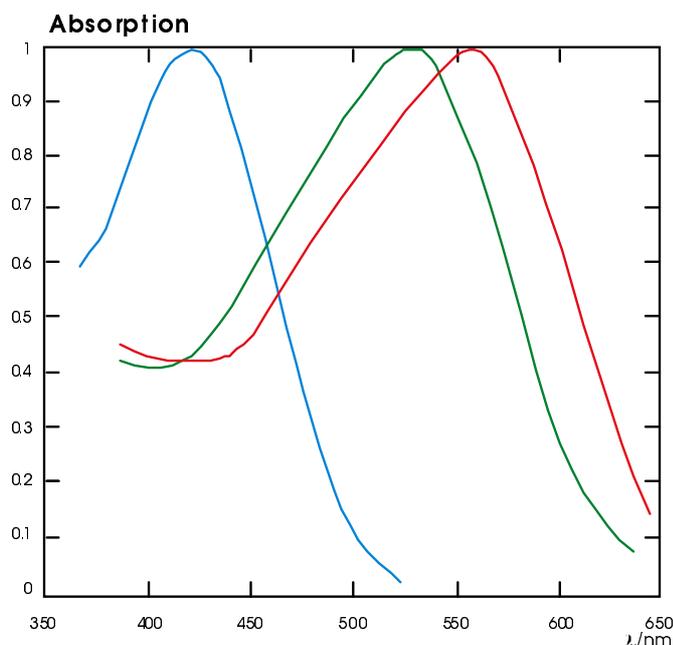


Abb. 1.3-1: Absorptionsspektren der drei Zapfenarten beim menschlichen Auge nach [1.5]

Im Gegensatz zu den Zapfen, gibt es von den *Stäbchen* nur eine Art. Diese Zellenart besitzt ebenfalls einen Sehstoff, nämlich *Rhodopsin*, der bei der Absorption von

Stäbchen

Licht ausgebleicht wird. So ist für diese Sehzellen eine Helladaption über einen großen Leuchtdichtebereich möglich. Insgesamt können diese Sehzellen noch bei wesentlich geringeren Leuchtdichteverhältnissen Helligkeiten differenzieren als die Zapfen, obschon sie dabei keine Farbunterscheidungsfähigkeit besitzen („Bei Nacht sind alle Katzen grau.“). Die spektrale Empfindlichkeitskurve für das „Nachtsehen“ aus Abb. 1.2-2 beruht im Wesentlichen auf den zuvor genannten Eigenschaften der Stäbchen.

Lange vor den Erkenntnissen über die Eigenschaften des menschlichen Auges war experimentell erprobt, dass sich die verschiedenen Farbeindrücke aus drei geeigneten Grundfarben (z. B. *rot*, *grün* und *blau*) ermischen lassen (z. B. bei der Malerei). Es zeigt sich aber, dass sich aus den Grundfarben nach zwei unterschiedlichen Mischtechniken jeweils sehr verschiedene Mischfarben erzielen lassen.

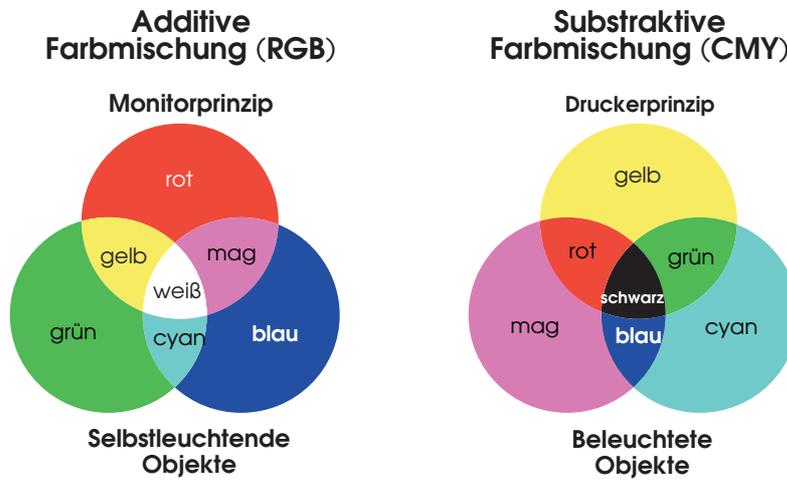
Subtraktive Farbmischung

Ein Beispiel für die so genannte *subtraktive Farbmischung* ist das Zusammenführen verschiedenfarbiger Farbstoffe, wie es bei der Malerei geschieht. Dieses Mischen von übereinander liegenden Farbstoffen kann man als das Zusammenfügen von Farbfiltern mit unterschiedlichen farblichen Absorptionseigenschaften verstehen. So erhält man beispielsweise *grünliches* Licht, wenn man vor eine *weiße* Lichtquelle ein *blaues* und dahinter ein *gelbes* Filter setzt. Der resultierende Farbeindruck entsteht also durch die nach der optischen Filterung verbleibenden Farbanteile einer ursprünglich weißen Lichtquelle. Dazu wird angenommen, dass die weiße Lichtquelle eine weitgehend kontinuierliche Energieverteilung über das ganze sichtbare Spektrum besitzt. Diese Annahme ist in der Regel für natürliche Lichtquellen gegeben. Die subtraktive Farbmischung wird vielfach auch bei Farbdruckern verwendet, die mit drei oder mehr Grundfarben arbeiten. Als Grundfarben werden dann aber oft *gelb*, *magenta*, *cyan* und ggf. weitere verwendet.

Additive Farbmischung

Dazu entsteht im Gegensatz bei der *additiven Farbmischung* der resultierende Farbeindruck, in dem kleine, dicht beieinander liegende, selbst leuchtende Farbpunkte aus so großer Entfernung betrachtet werden, dass die einzelnen Farbpunkte nicht mehr separat aufgelöst werden können, sondern ineinander *verschmelzen*. Ebenso ist die additive Mischung auch durch die Projektion von einem roten, grünen und blauen Bildauszug auf eine gemeinsame Bildwand möglich. Das Wesentliche der additiven Mischung liegt darin, dass die Farbanteile der verschiedenen Farben additiv von den Sehzellen des menschlichen Auges wahrgenommen werden. Die additive Mischung funktioniert, solange eine Trennung der einzelnen Farbanteile vom menschlichen Auge nicht in örtlicher oder zeitlicher Richtung möglich ist. In dem zuvor genannten Beispiel mit den kleinen Farbpunkten ist das dann der Fall, wenn ein hinreichend großer Betrachtungsabstand gegeben ist. Die zeitliche Trägheit des menschlichen Auges (vgl. auch Abb. 1.2-7 und Abb. 1.2-8) ermöglicht auch eine additive Farbmischung ohne dass die Farbanteile gleichzeitig vorhanden sein müssen. Vielmehr kann die Farbdarstellung auch gelingen, wenn zeitlich sequenziell gezeigte Farbblitze so schnell auftreten, dass sie vom Auge hinreichend gut zu einem Gesamtfarbeindruck integriert werden. Eine Reihe von Farbbildprojektoren arbeiten nach diesem Prinzip (*Farbkreisel*).

Animation 1.3-2 zeigt die beiden Farbmischverfahren auf der Basis von *rot*, *grün* und *blau* (*Additive Farbmischung*) und *gelb*, *magenta* und *cyan* (*Subtraktive Farbmischung*).



Animation 1.3-2: Additive und subtraktive Farbmischung

Die additive Farbmischung findet bei fast allen elektronischen Bild Darstellungen in der analogen und digitalen Bildtechnik Anwendung. Auf dieser Basis werden auch die nachfolgenden technischen Farbbild Darstellungen vorgestellt. In Anlehnung an die Farbphysiologie des menschlichen Auges kann man für den Farbmonitor (hier: Kathodenstrahlmonitor) primäre Farbphosphore *rot* (*R*), *grün* (*G*) und *blau* (*B*) definieren. Bei der Herstellung solcher Bildschirmphosphore und bei der Farbmischung ist man jedoch einigen Beschränkungen unterworfen. So lassen sich nicht in jedem Fall zur natürlichen Bildvorlage spektral identische Farbdarstellungen mit dem Monitor erzielen. Eine mögliche Lösung dieses Problems besteht darin, so genannte *metamere Farben* auf dem Monitor zu generieren. Diese sind Farbmischungen, die für das menschliche Auge zwar gleich aussehend sind, aber eventuell eine andere spektrale Zusammensetzung besitzen.

Dieses führt uns zu den drei *Graßmannschen Gesetzender Farbmimetrik*:

Graßmannsche Gesetze

1. *Für das Ergebnis einer additiven Mischung ist nur das Aussehen nicht die spektrale Zusammensetzung der Komponenten maßgeblich.*

Der Satz drückt aus, dass in additiven Farbmischungen die Komponenten durch ihre *Valenzen* bestimmt sind während ihre spektrale Zusammensetzung keine Rolle spielen. Man kann jede Mischungskomponente durch eine metamere Komponente gleicher Valenzen ersetzen, ohne dass sich die Valenz der gesamten Mischung ändert. Mischungen aus gleich aussehenden Farben ergeben dabei wieder gleich aussehende Farben.

2. *Alle Farbmischungen verlaufen stetig. Die Stetigkeit von Mischungen drückt sich in der Stetigkeit der Spektralwertkurven aus.*

3. Zur Festlegung einer Farbe sind 3 Bestimmungsstücke notwendig und hinreichend.

Diese Feststellung drückt die Erfahrungstatsache aus, dass drei Farbkomponenten *rot (R)*, *grün (G)* und *blau (B)* notwendig sind (aber auch ausreichen), um mit einer additiven (äußeren) Mischung alle übrigen Farben zu ermischen. Als alternative Bestimmungsstücke bietet sich auch das Tripel aus *Helligkeit*, *Farbton* und *Farbsättigung* an.

Bei der Einführung eines technischen Farbkoordinatensystems gibt es verschiedene Möglichkeiten. Sie alle müssen grundsätzlich den Graßmannschen Gesetzen genügen. Von daher sind viele Farbkoordinatensysteme dreidimensional (z. B. der *RGB*-Farbwürfel-Raum, vgl. [1.5]). Leider sind die 3D-Koordinatensysteme recht unübersichtlich und wenig anschaulich. Um Aussagen über die „Farbe“ – bestehend aus Farbton und Farbsättigung – machen zu können reicht oft auch eine 2D-Darstellung, die dann das dritte Bestimmungsstück – die Helligkeit – unberücksichtigt lässt. Dies ist der Grund, warum die meisten standardisierten Farbkoordinatensysteme nur zweidimensional ausgelegt sind.

Ein einfaches Beispiel aus der Bildtechnik ist das so genannte *Farbdreieck* oder die *Farbtafel*. Hierin sind folgende Begriffe definiert:

- Als *Farbwerte* sind die Bildschirmprimärvalenzen *R*, *G* und *B* definiert.
- Eine *Farbvalenz (F)* ist festgelegt durch: $(F) = R(R) + G(G) + B(B)$.
- *Helligkeit (F)_{grau}* = $a [1(R) + 1(G) + 1(B)]$.
- Die *Farbart* ist definiert durch den *Farbton* und die *Farbsättigung*.
- Die *Farbwertanteile* sind definiert durch

Begriffe
Farbdreieck, Farbtafel

$$r = \frac{R}{R + G + B}; \quad g = \frac{G}{R + G + B}; \quad b = \frac{B}{R + G + B} \quad 1.3-1$$

Unbuntpunkt Für das Farbdreieck wird ein gleichseitiges Dreieck mit den Bildschirmprimärvalenzen *R*, *G* und *B* als Eckpunkte angesetzt (vgl. Abb. 1.3-3). Der Mittelpunkt des Dreiecks stellt den so genannten *Unbuntpunkt* dar, ein Punkt, an dem Farbton und Farbsättigung gleich Null sind.

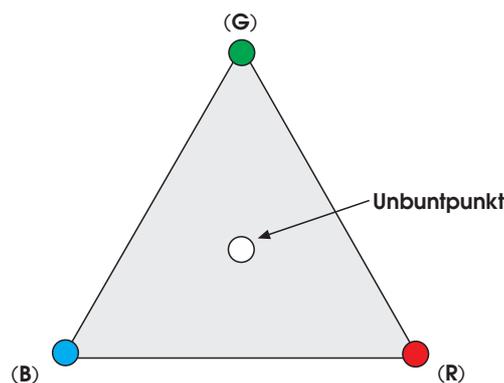


Abb. 1.3-3: Farbtafel

In der Entwicklungsphase des analogen Farbfemsehens wurde die Definition der Bildschirmprimärvalenzen auf der Basis von herstellbaren Farbphosphoren für die Kathodenstrahlröhre definiert. Für Europa wurden die so genannten *EBU-Phosphore* (EBU Tech. 3213 [1.12]), für die USA die *Phosphore* nach der *SMPTE-Norm RP 145* [1.13] festgelegt. Diese beiden Farbsysteme unterscheiden sich geringfügig voneinander. Beiden ist aber eines gemeinsam: Es sind grundsätzlich nicht alle in der Natur vorkommenden Farbarten (Farbton, Farbsättigung) exakt darstellbar. Die mangelnde Darstellungsfähigkeit beruht insbesondere auf der beschränkten Möglichkeit der zugrunde gelegten Bildschirmphosphore, hohe Farbsättigungen für alle Farbtonbereiche gleichermaßen gut wiedergeben zu können.

Mit Hilfe der Gesetze der *metameren Farbmimetrik* lassen sich allerdings dreidimensionale Farbkoordinatensysteme konstruieren, mit Hilfe dessen sich jeder in der Natur vorkommende spektrale Verteilungsverlauf gleich aussehend – metamer – aus einer Linearkombination der drei spektralen Verteilungskurven nachbilden lässt, die dieses Farbkoordinatensystem aufspannen. Die so genannte *CIE-Normfarbtafel* wird zunächst durch drei spektrale Verteilungsdichten $X(\lambda)$, $Y(\lambda)$ und $Z(\lambda)$ gebildet.

Metamere Farbmimetrik

Die Besonderheiten dieses Farbkoordinatensystems sind:

- Die so genannten *Normspektralwertkurven* $X(\lambda)$, $Y(\lambda)$ und $Z(\lambda)$ sind überall positiv.
- Die $Y(\lambda)$ -Kurve stimmt mit der Hellempfindlichkeitsverteilung des menschlichen Gesichtssinns für das Tagsehen überein (vgl. Abb. 1.2-2). Damit ist die $Y(\lambda)$ -Kurve bei einer Wellenlänge von 555 nm exakt auf den Wert 1 normiert.

Als weitere Besonderheit muss beachtet werden, dass die $X(\lambda)$ -Kurve zwei Maxima besitzt. Abb. 1.3-4 zeigt die Normspektralwertkurven $X(\lambda)$, $Y(\lambda)$ und $Z(\lambda)$ der *CIE-Normfarbtafel* von 1931 für einen 2°-Normalbeobachter.

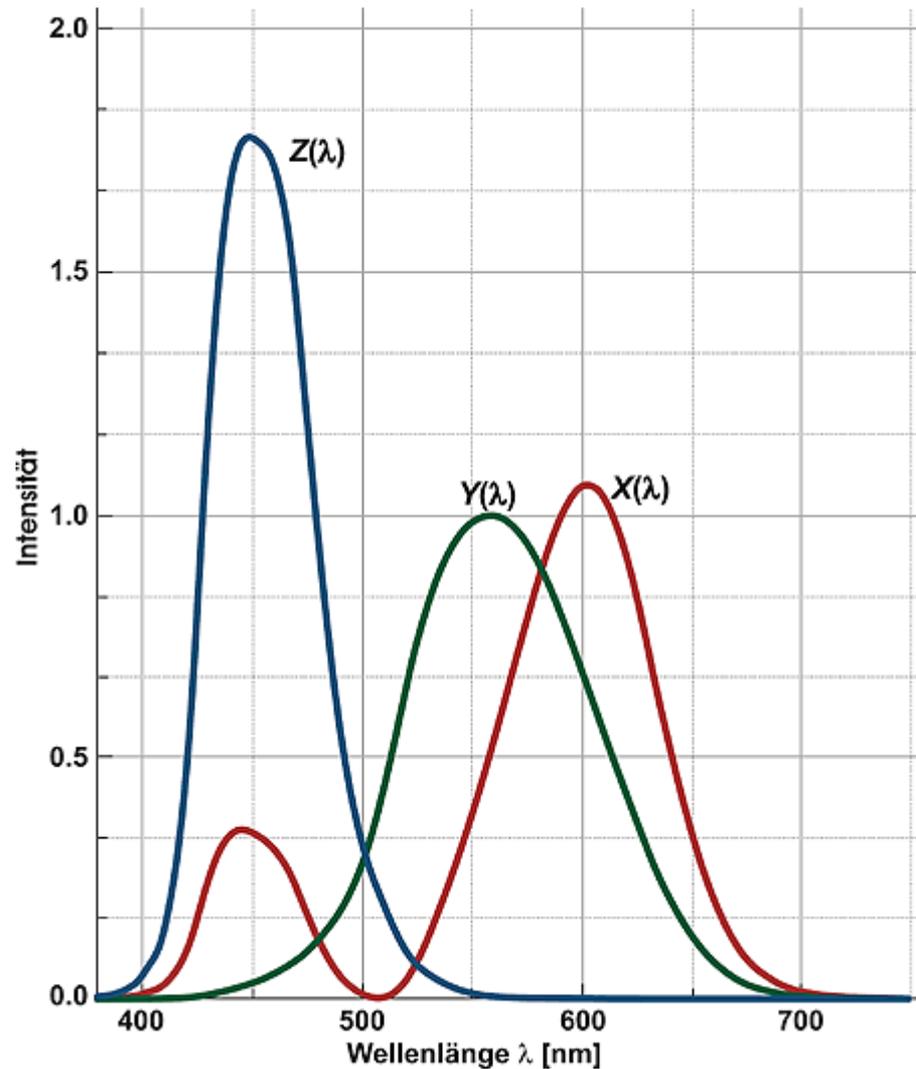


Abb. 1.3-4: Normenspektralwertkurven für den farbmetrischen Normalbeobachter (2°)

Als Koordinaten im rechtwinkligen Farbdreieck X, Y, Z lassen sich die Farbwertanteile x, y und z in Analogie zu Gl. 1.3-1 wie folgt berechnen:

$$x = \frac{X}{X + Y + Z}; \quad y = \frac{Y}{X + Y + Z}; \quad z = 1 - x - y \quad 1.3-2$$

Da alle Spektralwertkurven nur positive Anteile besitzen, die die so genannte *Spektralwertkurve*, die alle Farbtöne mit der in der Natur maximal möglichen Sättigung zeigt, überall mit positiven Anteilen beschreibbar. Innerhalb dieser Kurve liegen dann alle möglichen Farbarten, auch die, die ein Monitor darstellen kann. Für die Farb-Transformation zwischen dem *RGB*-Farbwürfel-Raum und der CIE-Normfarbtafel werden die folgenden Transformationsgleichungen verwendet:

EBU-Phosphore in der
CIE-Normfarbtafel

$$\begin{aligned} x &= 0.640R + 0.290G + 0.150B \\ y &= 0.330R + 0.600G + 0.060B \\ z &= 0.030R + 0.110G + 0.790B \end{aligned} \quad 1.3-3$$

$$x = 0.630R + 0.310G + 0.155B$$

$$y = 0.340R + 0.595G + 0.070B$$

$$z = 0.030R + 0.095G + 0.775B$$

1.3-4

SMPTE-Phosphore in
der CIE-Normfarbtafel

Abb. 1.3-5 zeigt die *CIE-Normfarbtafel* mit den eingezeichneten Koordinaten der Primärvalenzen nach *SMPTE*- und *EBU-Norm*. Die beiden ursprünglich für das analoge Farbfernsehen definierten Farbdreiecke der EBU (Europa) und der SMPTE (USA) sind relativ ähnlich. Weiterhin ist der Zug der Spektralfarben (*Spektralfarbenzug*) eingezeichnet. Er ist mit den Wellenlängen (in [nm]) parametrisiert. Der Spektralfarbenzug wird durch die Purpurgerade geschlossen und kennzeichnet so alle reellen Farben.

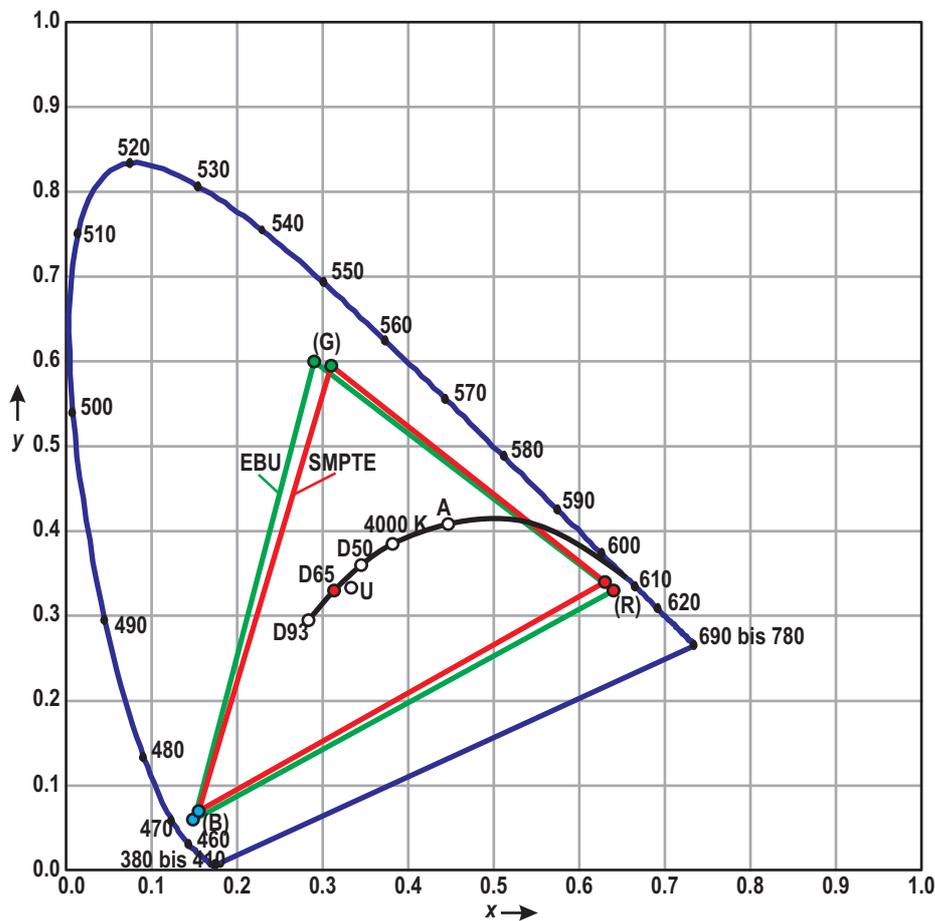


Abb. 1.3-5: CIE-Normfarbtafel

Die *CIE-Normfarbtafel* besitzt im farbmeterischen Sinne einen theoretischen Unbuntpunkt (Weißpunkt) *U* mit den Koordinaten $x = y = z = 1/3$. Jede Gerade, die von diesem Weißpunkt ausgeht, kennzeichnet eine Farbart mit gleichem Farbton aber verschiedener Farbsättigung.

Als *praktischen Unbuntpunkt* wird in der Regel nicht der Weißpunkt U , sondern ein Farbort gewählt, der auf der schwarz eingezeichneten Kurve liegt. Diese Kurve gibt die Farborte der so genannten *Planckschen Strahlung*¹ von wichtigen Lichtquellen an. Sie ist in Grad Kelvin [K] skaliert und gibt den Farbort an, den eine plancksche Strahlungsquelle mit der angegebenen Temperatur besitzt (*Farbtemperatur*). Einige wichtige Punkte sind mit Buchstaben und Zahlen gekennzeichnet. So wird für die beiden analogen Farbfernsehsysteme der EBU und SMPTE als Weißpunkt der Farbort $D65$

Weißpunkt D65

festgelegt

$$D65(x, y, z) = (0.3127, 0.3290, 0.3582) \quad 1.3-5$$

Er kennzeichnet eine plancksche Strahlungsquelle mit der Farbtemperatur von 6500 K. Dieses entspricht in etwa den Farbstimmungsverhältnissen für Tageslicht im Außenbereich (Sonnenlicht). Für Kunstlicht im Innenbereich (Glühbirne) liegt die Farbtemperatur etwa bei 2800 bis 3000 K und entspricht so näherungsweise dem Normlicht A.

So wie alle reellen Farben, das heißt Farben, die in der Natur vorkommen, nur innerhalb des vom Spektralfarbenzug umschlossenen Bereiches liegen, lassen sich mit den *EBU-* bzw. *SMPTE-Primärvalenzen* grundsätzlich auch nur Farben darstellen, die innerhalb des jeweiligen Farbdreiecks liegen.

Aus den zuvor genannten Gleichungen lässt sich schließlich auch der *Luminanzanteil* Y ermitteln. Die nachfolgende Gleichung gilt insbesondere auch für den digitalen Standards *ITU-R BT.601* (vgl. Kapitel 2):

$$Y = 0,299R + 0,587G + 0,114B. \quad 1.3-6$$

Selbsttestaufgabe 1.3-1:

Erklären Sie den Unterschied zwischen additiver und subtraktiver Farbmischung! Geben Sie jeweils ein technisches Beispiel an!

Selbsttestaufgabe 1.3-2:

Was sind metamere Farben?

1 Die Begriffe Plancksche Strahlung und Farbtemperatur soll hier nicht näher hergeleitet werden. Weiterführende Hinweise finden sich beispielsweise in [1.3]

Selbsttestaufgabe 1.3-3:

Worin besteht die Einschränkung eines EBU- oder SMPTE-Farbmonitors in Bezug auf die Farbdarstellung?

Selbsttestaufgabe 1.3-4:

Geben Sie die Farbwerte des Unbuntpunktes für EBU- und SMPTE-Farbfernsehsysteme in der CIE-Normfarbtafel an!

1.4 Einführung in die analoge Videotechnik - Wie es anfang

Für die Übertragung von bewegten Bildern („Video“) muss das sich ändernde zweidimensionale Abbild einer realen Szene in ein geeignetes elektrisches Signal für die Übertragung umgewandelt werden. Um diesen Vorgang vereinfacht beschreiben zu können sei im folgenden nur die Abbildung und Übertragung der Helligkeitsverteilung der Bildvorlage betrachtet.

Die Bildvorlage wird in horizontale, aneinander angrenzende Zeilen zerlegt. Diese Aufgabe übernimmt ein optisch-elektrischer Wandler, der auch das korrespondierende Signal zeitsequenziell an ein Übertragungssystem weiterreicht.

Empfängsseitig wird, nach entsprechender Aufbereitung, das elektrische Signal einem elektrisch-optischen Wandler, dem Monitor, zugeführt und wieder als ein Abbild der Helligkeitsverteilung der Bildvorlage wiedergegeben.

Bildfeldzerlegung Der Vorgang der *Bildfeldzerlegung in Zeilen* und die entsprechende Rückwandlung im Empfänger muss zeitsynchron geschehen. Dazu werden im Aufnahmesystem zusätzliche Synchronsignale generiert und ebenfalls zum Empfangssystem übertragen. Abb. 1.4-1 skizziert dieses Prinzip der analogen Fernsehübertragung.

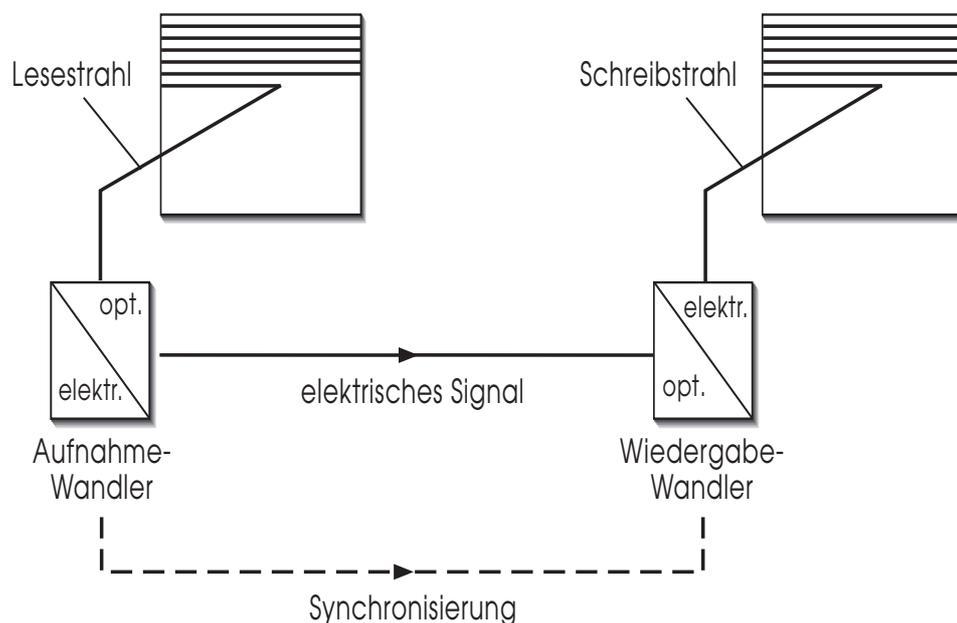


Abb. 1.4-1: Prinzip der Fernsehübertragung

In den Anfängen der analogen Bildübertragung gab es eine Reihe von technischen Randbedingungen, die eingehalten werden mussten. Die hieraus entwickelten Normen sind auch heute noch Standard, wenngleich viele technische Einschränkungen nicht mehr gelten. Aus Gründen der Kompatibilität ist eine Veränderung im Sinne des technischen Fortschritts innerhalb eines Standards nur schwer zu verwirklichen.

Die wichtigsten Eigenschaften der analogen Video- und Fernsehtechnik werden nachfolgend kurz zusammenfassend dargestellt.

1.4.1 Analoge Fernsehabtastung

Das in Abb. 1.4-1 skizzierte Prinzip der Fernsehbildübertragung geht von der Annahme aus, dass eine Art *Lesestrahle* zeilenweise die Bildvorlage abtastet. Tatsächlich hatte dieses Modell des Lesestrahls bei der Einführung des Fernsehens als technische Vorlage die Bildaufnahmeröhre. In modernen Bildwandlern stimmt die Abtastung mit dem Lesestrahle nur insofern, als dass beide Sensoren das gleiche zeitlich abgetastete Zeitsignal erzeugen. Physikalisch haben sich die Verfahren prinzipiell geändert. Trotzdem wollen wir nachfolgend bei diesem Bild des Abtastlesestrahls und des korrespondierenden Schreibstrahls bleiben, um den Abtastvorgang als solchen besser untersuchen zu können.

So tastet der Lesestrahle die Bildvorlage horizontal und vertikal von oben links nach unten rechts ab. Man kann diesen Vorgang mit der Bewegung des Gesichtsfeldes beim Lesen eines Textes vergleichen: Lesen innerhalb der Zeile von links nach rechts mit anschließenden raschen Zurückspringen zum Beginn der folgenden Textzeile. Für die Führung eines solchen Lesestrahls benötigt man eine periodische Ablenkung horizontal (*schneller Sägezahn*) und vertikal (*langsamer Sägezahn*).

Abb. 1.4-2 zeigt die beiden *Ablenksignale*: (horizontal – blau und vertikal – rot). Es ist auch deutlich zu sehen, dass die Rücksprungzeit vom Ende einer Zeile zur nächstfolgenden Zeile nicht beliebig kurz sein kann. Das gleiche gilt für den Rücksprung am Ende der letzten Zeile eines Bildes zur ersten Zeile des nachfolgenden Bildes. Es kann demnach nicht die komplette Zeit *aktiv* für das Lesen einer Bildvorlage genutzt werden; es wird unterschieden zwischen aktiver und nomineller Zeilenzahl, sowie zwischen aktiver und nomineller Zeilendauer.

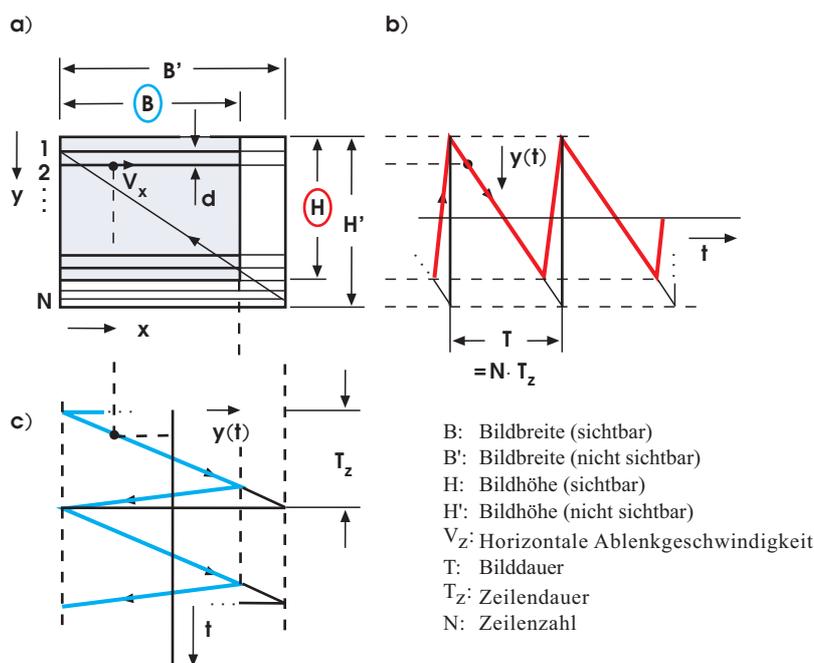


Abb. 1.4-2: Sichtbarer und nicht sichtbarer Bildbereich und die Ablenkung des Lese- bzw. Schreibstrahles

Die Qualität des reproduzierbaren Bildes hängt entscheidend von der Anzahl der

Ablenksignale für
H und V

Notwendige Zeilenzahl

verwendeten Zeilen ab. Zahlreichen Untersuchungen hatten das Ziel, eine allgemeine Vorschrift für die Anzahl der notwendigen Zeilen zu finden. Eine solche Näherungsformel lautet:

$$N_{aktiv} = \frac{2500}{E/H}, \quad 1.4-1$$

mit E als Betrachtungsabstand und H als sichtbare Bildhöhe (vgl. Abb. 1.4-2). Bei einem Betrachtungsabstand der dem fünffachen der Bildhöhe entspricht ergibt sich damit eine mindestens notwendige Anzahl von 500 aktiven Zeilen. Schließlich standardisiert wurden die folgende Werte:

$$\begin{aligned} 50 \text{ Hz-Länder} : N &= 625; N_{aktiv,analog} = 575; N_{aktiv,digital} = 576 \\ 60 \text{ Hz-Länder} : N &= 525; N_{aktiv,analog} = 485; N_{aktiv,digital} = 480 \end{aligned} \quad 1.4-2$$

Weitere wichtige Normen für die beiden Systeme betreffen das *Bild-Seitenverhältnis* sowie die *Bildwechselfrequenz*.

Bild-Seitenverhältnis Unter dem *Bild-Seitenverhältnis* versteht man das Verhältnis der Bildbreite B zur Bildhöhe H . Für beide Systeme galt ursprünglich:

$$\frac{B}{H} = \frac{4}{3}. \quad 1.4-3$$

Die vielen Breitbildformate beim heutigen Kino haben zu der Einführung eines weiteren Bild-Seitenverhältnisses (ursprünglich nur für *HDTV*, weltweit) geführt:

$$\frac{B}{H} = \frac{16}{9}. \quad 1.4-4$$

Beide Bild-Seitenverhältnisse sind heute üblich. In der Praxis wird das Format durch das Bildwiedergabesystem festgelegt. Dies kann zu schwarzen Balken oben und unten („*Letterbox*“) bei der Wiedergabe eines 16:9-Signals auf einem 4:3-Bildschirm bzw. zu schwarzen Balken links und rechts („*Side-Panel*“) bei der Wiedergabe eines 4:3-Signals auf einem 16:9-Bildschirm führen.

Bildwechselfrequenz Ähnlich wie beim Kino wird bei der analogen Bildtechnik die bewegte Bildvorlage als eine Folge von Einzelbildern übertragen. In Abschnitt 1.2.4 wurde bereits dargestellt, dass im Kino dazu 24 Bilder pro Sekunde verwendet werden, die aber mittels Flügelblende bei der Filmprojektion jeweils zweimal² gezeigt werden. Es entsteht somit eine effektive *Bildwechselfrequenz* von 48 Hz. Bei gut abgedunkelten Räumen und den damit verbunden geringen Leuchtdichten bewegt sich das wahrnehmbare Flimmern im vertretbaren Rahmen. In der analogen Bildtechnik ist eine solche empfängerseitige Bildwiederholung nicht so leicht zu realisieren. Um dennoch

2 In einigen wenigen Fällen ist sogar die dreimalige Wiederholung eines einzelnen Filmbildes üblich. Dann ergibt sich die effektive Bildwechselfrequenz zu 72 Hz.

eine weitgehend flimmerfreie Bilddarstellung zu erreichen, wird die Bildvorlage abweichend von Abb. 1.4-1 und Abb. 1.4-2 im so genannten Zeilensprungverfahren abgetastet und übertragen.

Zunächst verkoppelt man die Bildfrequenz mit der Netzfrequenz des jeweiligen Landes, um mögliche Interferenzstörungen mit Lichtquellen ausschließen zu können. So wird in Europa statt der Abtastung aller 625 Zeilen innerhalb von 40 ms (25 Hz) zunächst nur die Hälfte der Zeilen abgetastet (also 312,5), und zwar im ersten Schritt alle ungeraden Zeilen und danach alle geraden Zeilen. So entstehen zwei Halbbilder, welche mit der Halbbildfrequenz f_v von 50 Hz übertragen werden. Abb. 1.4-3 zeigt die so entstandenen abgetasteten Halbbilder.

Zeilensprungabtastung

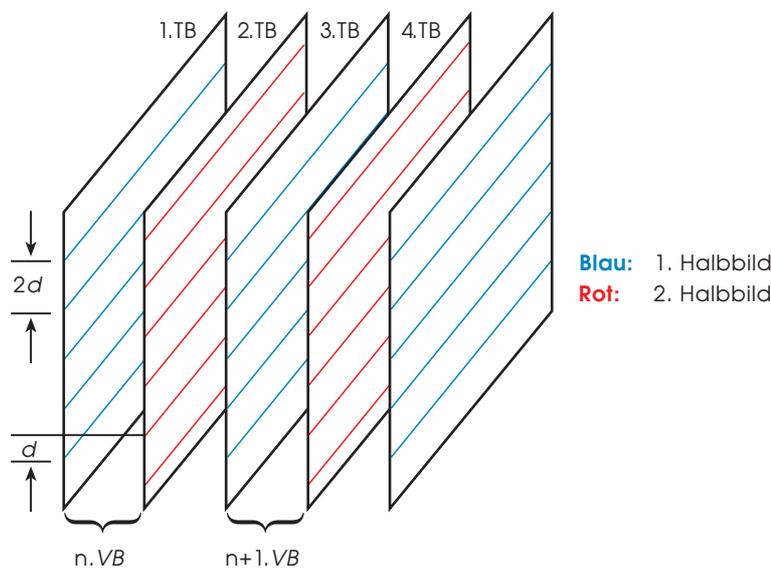


Abb. 1.4-3: Bildfolge bei der Zeilensprungabtastung

Die so genannte Zeilenfrequenz f_h (auch: Zeilenwechselfrequenz, Horizontalfrequenz) errechnet sich direkt gemäß der Randbedingungen von Abb. 1.4-2 aus der Halbbildfrequenz $f_v = 50$ Hz, der Anzahl der nominellen Zeilen $N = 625$:

$$f_h = \frac{N}{2} \cdot f_v = 15625 \text{ Hz.} \quad 1.4-5$$

Für das amerikanische TV-System mit $f_v = 59.94 \text{ Hz}$ und $N = 525$ ergibt sich:

$$f_h = \frac{N}{2} \cdot f_v = 15734.264 \text{ Hz.} \quad 1.4-6$$

3 Beim Übergang vom Schwarzweiß- zum Farbfernsehsystem wurde eine geringfügige Änderung der Zeilen- und Halbbildfrequenz vorgenommen, um Übersprechstörungen zwischen Farb-, Luminanz- und Tonsignal weitestgehend unterdrücken zu können.

1.4.2 Das FBAS-Signal

Das Farbustastsynchron-Signal (FBAS-Signal) enthält alle Bild- und Synchroninformationen so, wie sie zeitsequenziell übertragen werden. Abb. 1.4-4 zeigt beispielhaft den Verlauf dieses Signals für eine Zeile (Europäische Fernsehnorm).

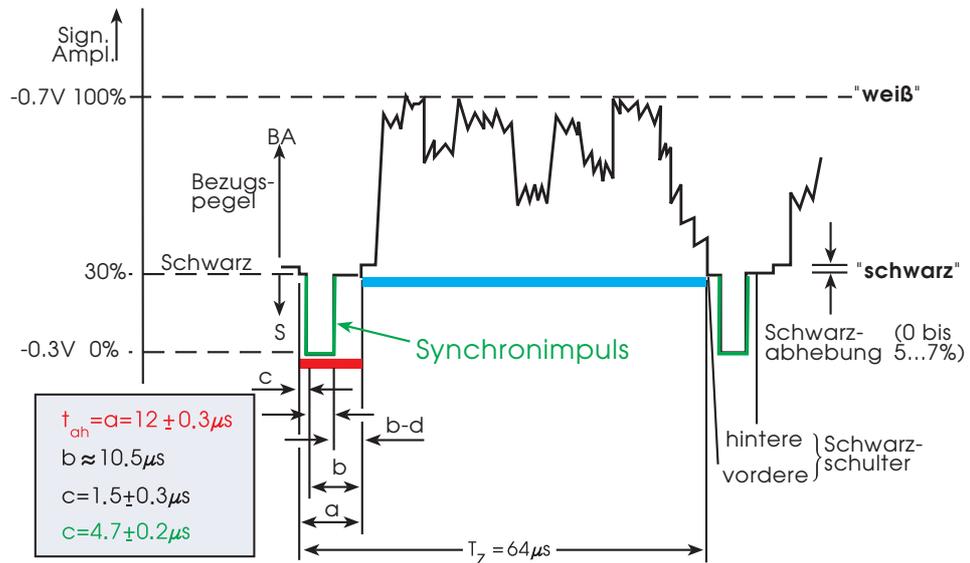


Abb. 1.4-4: FBAS-Signal

Man kann zwischen dem aktiven Bildbereich (*blau*) und dem nicht sichtbaren Bereich (*horizontale Austastlücke t_{ah}*) einer Zeile (*rot*) unterscheiden. Innerhalb des aktiven Zeilenbereiches verändert sich das Signal abhängig von der Helligkeit der korrespondierenden Bildvorlage: Für ein *weiß* wird eine Spannung von +0.7 V erzeugt, ein *schwarz* liegt nahe bei 0 V. Dazwischen liegen alle übrigen Grauwerte. Im nicht sichtbaren Bereich ist ein Synchronimpuls mit -0.3 V Amplitude enthalten. Er ist etwa 4.7 μs lang und trägt dafür Sorge, dass die einzelnen Zeilen am Empfänger wieder exakt synchron übereinander geschrieben werden können.

Die nominelle Zeilendauer T_z ergibt sich aus der Zeilenfrequenz f_h nach

$$T_z = \frac{1}{f_h} = 64 \mu\text{s}. \quad 1.4-7$$

Die horizontale Austastlücke t_{ah} beträgt etwa 12 μs . Somit verbleibt für die aktive Zeilendauer ein Wert von 52 μs .

Die *Farbinformation* der Bildvorlage wird beim europäischen Farbfernsehen in der Form einer hochfrequenten Schwingung (*Farbhilfsträger*, $f_{sc} = 4.43 \text{ MHz}$ auf der Basis der *quadraturmodulierten* und

Farbhilfsträger

*gammakorrigierten*⁴ Farbdifferenzsignale

Farbdifferenzsignale

$$\begin{aligned} U &= E_U = 0.493(E'_B - Y) \text{ und} \\ V &= E_V = 0.877(E'_R - Y) \end{aligned} \quad 1.4-8$$

in das Schwarzweißsignal eingebettet. In Matrixschreibweise lässt sich der Übergang der Farbkomponenten E'_R , E'_G , und E'_B

nach Y , U und V dann unter Verwendung der Gl. 1.3-6 folgendermaßen ausdrücken:

$$\begin{pmatrix} U \\ V \\ Y \end{pmatrix} = \begin{pmatrix} -0.147 & -0.289 & +0.436 \\ +0.615 & -0.515 & -0.100 \\ +0.299 & +0.587 & +0.114 \end{pmatrix} \cdot \begin{pmatrix} E'_R \\ E'_G \\ E'_B \end{pmatrix} \quad 1.4-9$$

In der digitalen Bildtechnik werden Komponentensignale ebenfalls oft verwendet. Die Farbdifferenzsignale verwenden dann jedoch andere Konstanten in der Tabelle 1.4-1. Zur Unterscheidung ist dann oft auch die Schreibweise C_B und C_R üblich (vgl. Abschnitt 2.2.2).

Für die Synchronisation des Farbsignals wird auf der *hinteren Schwarzschulter* (vgl. Tabelle 1.4-1) ein weiteres Signal, der so genannte *Farbburst*, übertragen, welcher aus etwa 10 Schwingungen des Farbhilfsträgers besteht.

Parameter analoger
Fernsehsysteme

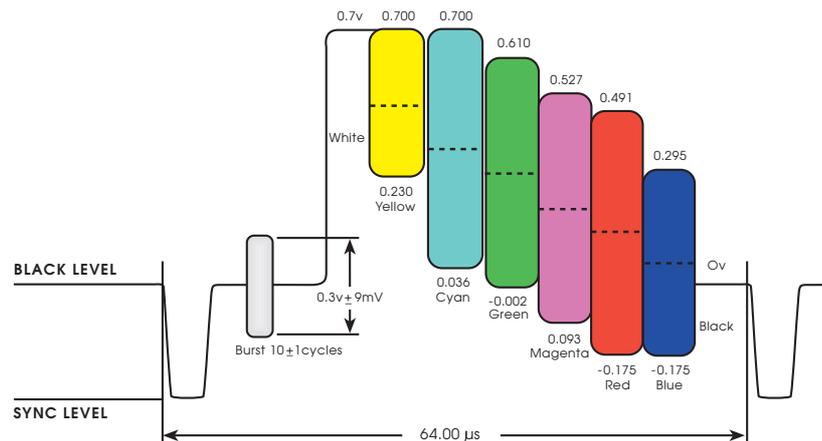
4 Der Zusammenhang zwischen elektrischem Signalwert und korrespondierendem Helligkeitswert muss kameraseitig bewusst nichtlinear sein, damit bei der Wiedergabe über einen Kathodenstrahlmonitor ohne weitere Maßnahmen die Darstellung insgesamt wieder linear ist. Die Nichtlinearität des Monitors wird also bereits auf der Kameraseite berücksichtigt und entsprechend ausgeglichen.

Tab. 1.4-1: Parameter analoger Fernsehstandards nach ITU-R.BT 470

	Europa (PAL, Norm G/B)	USA (NTSC, Norm M)
Zeilenzahl, nominell N :	625	525
Zeilenzahl, aktiv N_{aktiv} , ($N_{aktiv,digital}$):	575 (576)	485 (480)
Zeilendauer T_z :	64 μ s	63.55 μ s
Zeilenfrequenz f_z :	15625 Hz	15734 Hz
Horizontale Austastlücke t_{ah} :	12 μ s	10.9 μ s
Vollbildfrequenz f_B :	25 Hz	29.97 Hz
Halbbildfrequenz f_{HB} :	50 Hz	59.94 Hz
Halbbilddauer T_{HB} :	20 ms	16.6 ms
Vertikale Austastlücke t_{av} :	25 Zeilen	20 Zeilen
Bild-Seitenverhältnis:	4:3	4:3
Farbhilfsträgerfrequenz f_{sc}	4.43 MHz	3.58 MHz
Grenzfrequenz Bildsignal $f_{B_{max}}$:	5 MHz	4.2 MHz

Tabelle 1.4-1 fasst noch einmal die wichtigsten Zahlen der beiden analogen Fernsehstandards *PAL* (Europa) und *NTSC* (USA) zusammen. Weitere Details, insbesondere auch über die Unterschiede in der Toncodierung, lassen sich in [1.7] und [1.8] nachlesen.

Die nachfolgende Abb. 1.4-5 zeigt den Signalverlauf einer Zeile eines 75%-FBAS-Farbbalkentestsignals (*PAL*-Standard). Es veranschaulicht die Signale des Luminanzsignals (Graustufentreppe) und das überlagerte, hochfrequente Farbartsignal der Farbdifferenzkomponenten U und V .

**Abb. 1.4-5:** 100/75%-Farbbalkentestsignal nach dem *PAL*-Standard

Selbsttestaufgabe 1.4-1:

Beschreiben Sie kurz die Bildfeldzerlegung bei der analogen Bildtechnik!

Selbsttestaufgabe 1.4-2:

Warum wird zwischen nomineller und aktiver Zeilenzahl unterschieden?

Selbsttestaufgabe 1.4-3:

Welche Aufgabe hat die horizontale Austastlücke?

Selbsttestaufgabe 1.4-4:

Geben Sie eine Transformationsgleichung an, mit der man aus dem Farbvektor (E'_R, E'_G, E'_B) den Vektor (Y, U, V) für ein EBU-Farbsystem ermitteln kann!

2 Digitale Bilddarstellung und -übertragung

2.1 Abtastung und Quantisierung

Das Prinzip der Fernsehübertragung, das in Abb. 1.4-1 gilt auch bei der digitalen Erzeugung und Übertragung von Bildern. Auch hier gibt es eine Bildfeldzerlegung in Zeilen und sequenziell aufeinander folgender Bilder. Zwei Prozesse kommen hier jedoch hinzu: Die örtliche und zeitliche Abtastung des analogen Bildsignals sowie die Diskretisierung (*Quantisierung*) der Werte.

Abb. 2.1-1 zeigt den Prinzip der Signalaufbereitung. Ausgehend von den analogen Komponentensignalen Y und $B-Y$, $R-Y$ wird eine Analog-Digital-Wandlung in den drei Signalkanälen vorgenommen (Abtastung: Diskretisierung und Quantisierung). Die digitalen Komponentensignale werden anschließend zunächst zu einem N -Bit breiten Zeitmultiplexsignal zusammengefasst. Für die Übertragung wird oft zudem noch ein Parallel-Seriell-Wandler nachgeschaltet, um die Übertragung in einem seriellen Komponentensignal (*Digital Serial Components, DSC*) zu erreichen.

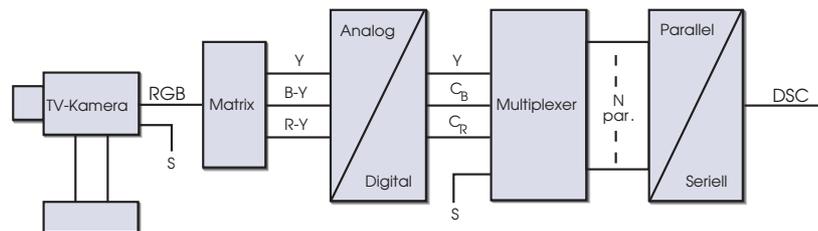


Abb. 2.1-1: Aufbau einer Bildaufnahmeeinheit mit digitalem Signalausgang (nach [2.2])

Die für die Analog-Digital-Wandlung erforderliche zeitliche Abtastung bzw. die örtliche Diskretisierung geschieht dabei in folgender Weise: Zunächst wird das analoge Eingangssignal auf eine maximale Signalfrequenz f_{max} mit einem Tiefpass bandbegrenzt. Dieses Signal wird dann periodisch mit kurzen Impulsen abgetastet. Dieses ist der Vorgang der Diskretisierung. Es entsteht eine Folge von analogen Spannungswerten, die auf einen beschränkten Zahlenraum „gerundet“ werden. Diesen Vorgang nennt man Quantisierung. Als Ergebnis entsteht am Ausgang eine unendliche Folge von N -bit breiten Codeworten, welche das ursprüngliche analoge Eingangssignal repräsentieren. 2^N gibt hierbei den Zahlenraum an, welcher für die Quantisierung benutzt wird. In der digitalen Bildtechnik ist eine Wortbreite von $N = 8$ und $N = 10$ üblich. Daraus ergeben sich für eine Graustufendarstellung 256 bzw. 1024 verschiedene Werte. Diese Anzahl reicht in den meisten Anwendungsfällen aus, ohne dass der Betrachter Graustufensprünge an ursprünglich kontinuierlich verlaufenden Helligkeitsverläufen wahrnimmt.

Entscheidender für die Bildqualität eines digitalen Bildsignals ist vielmehr die richtige Wahl der Abtastfrequenz f_T . Aus der Signaltheorie ist bekannt, dass für eine fehlerfreie Rekonstruktion des analogen Eingangssignals aus der o.g. Zahlenfolge gelingen kann, wenn die Abtastfrequenz f_T mindestens doppelt so hoch ist, wie die höchste vorkommende Frequenz des Eingangssignals.

Diesen Zusammenhang nennt man das *Abtasttheorem nach Shannon*, welches formal lautet:

Abtasttheorem nach
Shannon

$$f_T \geq 2 \cdot f_{\max}. \quad 2.1-1$$

Da im allgemeinen nicht von einer idealen Bandbegrenzung des eingangsseitig verwendeten Tiefpasses ausgegangen werden kann, muss die Abtastfrequenz in der Praxis oft sehr viel größer sein, als die doppelte Nutzfrequenz.

Die nachfolgenden Beispielbilder veranschaulichen noch einmal den Einfluss der Diskretisierung und der Quantisierung. Abb. 2.1-2 zeigt ein Beispielbild, bei dem die Grauwerte im Rahmen der Darstellungsgenauigkeit als orts- und wertekontinuierlich angesehen werden sollen.



Abb. 2.1-2: Orts- und wertekontinuierliches Originalbild (analoges Eingangssignal)

Abb. 2.1-3 zeigt, wie sich eine Diskretisierung (örtliche Abtastung) auswirkt, wenn das Abtasttheorem nicht eingehalten wird, weil entweder der Eingangstiefpass keine ausreichende Bandbegrenzung gewährleistet bzw. die Abtastfrequenz für das Eingangssignal zu niedrig gewählt wurde.

Beispiel Diskretisierung



Abb. 2.1-3: Ortsdiskretisierung (Werte weiterhin kontinuierlich)

Beispiel Quantisierung

Abb. 2.1-4 zeigt die Störungen, wenn zwar das Abtasttheorem eingehalten wird, die Grauwerte jedoch nicht mehr auf einen ausreichend großen Zahlenraum quantisiert werden.



Abb. 2.1-4: Werte zu stark quantisiert, (Ortsauflösung kontinuierlich bzw. Abtasttheorem eingehalten)

Abb. 2.1-5 visualisiert die Störungen, welche durch eine zu grobe Diskretisierung und einer zu geringen Werte-Quantisierung entstehen.



Abb. 2.1-5: Werte zu stark quantisiert, örtliche Diskretisierung zu grob

Das durch die Quantisierung verursachte Störsignal wird allgemein auch als *Quantisierungsrauschen* bezeichnet. Dabei geht man davon aus, dass bei klein gewählten Quantisierungsstufen eine Zufallsfolge entsteht, welche mit einem gleichverteilten vom ursprünglichen Signal weitgehend unabhängigem Spektrum korreliert. Damit lässt sich näherungsweise ein *Signal-Störabstand* ermitteln:

Signal-Störabstand

$$S_Q \approx (6 \cdot N + 10.8) \text{ dB}. \quad 2.1-2$$

Ein Herleitung findet sich in [2.1] und [2.3].

2.2 Digitale Bildformate, Farbdarstellungen und Standards

2.2.1 Farbtastraster

Die bei der Diskretisierung entstandenen Abtastwerte beziehen sich üblicherweise auf das aktive Bild⁵ mit einer Anzahl von n Pixel je Zeile und N_{aktiv} Zeilen je Bild. Die Werte können sowohl *progressives* Bildmaterial repräsentieren oder aus einer *Zeilensprungabtastung* hervorgehen. Obwohl dem Zeilensprungverfahren viele Nachteile anhaften, basieren viele digitale Bildnormen weiterhin auf diesem Standard. Insbesondere in den klassischen Fernseh-Umgebungen wird aus Kompatibilitätsgründen noch das Zeilensprungverfahren auch in der digitalen Bildtechnik verwendet (Beispiel: *Digital Video Broadcast - DVB*, vgl. Kapitel 6).

Auf der Suche nach Einsparpotentialen der digitalen Repräsentation eines analogen Bildes werden bereits bei der Abtastung Eigenschaften des menschlichen visuellen Systems in Bezug auf die *Kontrastempfindlichkeit* genutzt. In Abb. 1.2-6 (Abschnitt 1.2.3) wurde die Kontrastempfindlichkeit für farbart- und helligkeitsmodulierte räumliche Gitter betrachtet. Als Ergebnis konnte festgehalten werden, dass das räumliche Auflösungsvermögen für reine Farbartübergänge geringer ist, als bei Helligkeitsübergängen. Signaltheoretisch betrachtet dürfen demnach die Farbart beinhaltenden Signalanteile (= *Farbdifferenzsignale*) $B-Y$ und $R-Y$ vor der Abtastung stärker bandbegrenzt werden als das Luminanzsignal Y . Damit ist auch eine niedrigere Abtastfrequenz für die Farbdifferenzsignale möglich. Um einfache örtliche Bezüge der digitalisierten Codeworte für Luminanz und Chrominanz zu haben, stehen die Abtastfrequenzen im Verhältnis 2:1 oder 4:1 zueinander. Zur Unterscheidung werden diese digitalisierten und örtlich unterabgetasteten Farbdifferenzsignale mit

$$C_B = (E'_B - E'_Y)_{\perp, \downarrow s} \quad 2.2-1$$

Digitale
Farbdifferenzsignale C_B
und C_R

bzw.

$$C_R = (E'_R - E'_Y)_{\perp, \downarrow s} \quad 2.2-2$$

bezeichnet, wobei E'_Y , E'_B und E'_R die gammakorrigierten Signale von Y , B und R bezeichnen, das Symbol \perp den Abtastvorgang beschreibt und $\downarrow s$ die Reduktion um den Faktor s im Verhältnis zur Abtastung des Luminanzsignals Y angibt.

Mit diesen Vorüberlegungen werden die so genannten *4:4:4-* und *4:2:2-Abtastformate* definiert. Sie bezeichnen das Verhältnis der Abtastfrequenzen zwischen Y , C_B und C_R , bezogen auf eine Basisfrequenz von

$$f_{\perp, Basis} = 3.375 \text{ MHz}. \quad 2.2-3$$

5 d. h. Bildbereiche der horizontalen oder vertikalen Austastlücke bleiben unberücksichtigt.

So ergibt sich für ein Abtastformat von 4:4:4 eine Abtastfrequenz von

$$f_{\perp} = 13.5 \text{ MHz} \quad 2.2-4$$

sowohl für das Luminanzsignal Y als auch für beide Farbdifferenzsignale C_B und C_R . Es geschieht keine örtliche Auflösungsreduktion in der Farbe. Dem gegenüber ergeben sich beim 4:2:2-Abtastformat Frequenzen für die Farbdifferenzsignale von

$$f_{\perp, \downarrow 2} = 13.5 \text{ MHz} / 2 = 6.75 \text{ MHz}. \quad 2.2-5$$

Die Wahl der Basisfrequenz $f_{\perp, Basis}$ mit dem Wert von 3.375 MHz ist nicht willkürlich, sondern sorgt für eine starre Verkopplung mit der Zeilenfrequenz der jeweiligen Farbfernsehnorm:

$$625\text{-Zeilen-System (PAL-Norm Europa): } 3.375 \text{ MHz} = 216.0 \cdot 15625 \text{ Hz}$$

$$525\text{-Zeilen-System (NTSC-Norm USA): } 3.375 \text{ MHz} = 214.5 \cdot 15734 \text{ Hz}.$$

Die Verkopplung sorgt für ein orthogonales Abtastraster, so dass die digital abgetasteten Bildpunkte übereinander und in jedem Teilbild immer an der gleichen örtlichen Position liegen.

Abb. 2.2-1 veranschaulicht die Luminanz- und Chrominanzpositionen für verschiedene Abtastformate. Angegeben sind die Luminanzwerte (weiß) sowie die Chrominanzwerte C_B (blau) und C_R (rot) und deren örtlicher Zugehörigkeitsbereich (rot umrandet). Darüber hinaus ist der Datenreduktionsfaktor angegeben, der sich mit dem jeweiligen Abtastformat gegenüber einer 4:4:4-Abtastung erzielen lässt.

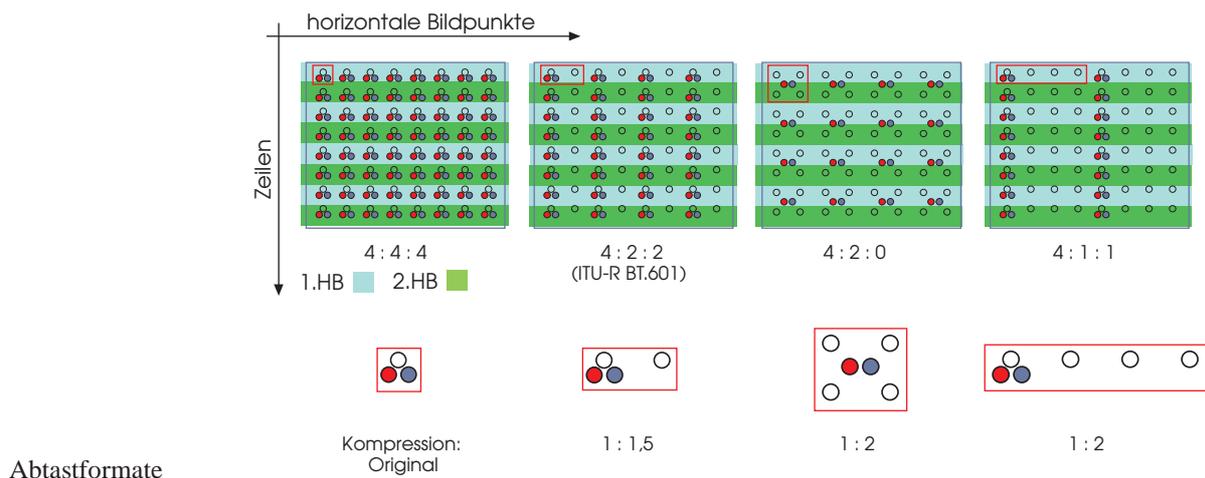


Abb. 2.2-1: Luminanz- und Chrominanzpositionen für verschiedene Abtastformate

Es ist zu erkennen, dass die Abtastformate 4:2:2 bzw. 4:1:1 zu einer örtlichen Auflösungsreduktion rein horizontal um den Faktor 2 bzw. 4 führen. Tatsächlich spielt das 4:4:4-Abtastformat in der digitalen Bildübertragung kaum eine Rolle. Spezialanwendungen im virtuellen Studio (*Blue-Box-Verfahren*) oder in der Medizin- und Radartechnik sind die Domänen für dieses Format. Das klassische digitale Studio-signal hingegen kommt mit einem Abtastformat von 4:2:2 aus (vgl. Abb. 2.2-1b).

Für die Übertragung zum Endnutzer wird eine weitere Reduzierung der örtlichen Farbauflösung vorgenommen. Dabei gibt es ausgehend vom 4:2:2-Format zwei verschiedene Varianten, die in unterschiedliche Länderstandards integriert wurden.

4:1:1-Abtast- und Übertragungsformat

4:1:1 Abtastung

Leicht zu realisieren ist eine weitere horizontale Unterabtastung um den Faktor 2 (vgl. Abb. 2.2-1d). Dann ergibt sich insgesamt eine horizontale Reduzierung der Farbauflösung auf 25 %, vertikal bleibt die volle Auflösung erhalten. Bei der Magnetbandaufzeichnung (DV-Bildcodierung, *DVCPro 25*, vgl. Kapitel 4) und der datenreduzierten Übertragung von Videosignalen basierend auf 525-Zeilensystemen (z. B. USA) wird dieses Format verwendet.

4:2:0-Abtast- und Übertragungsformat

4:2:0 Abtastung

Um horizontal und vertikal in etwa eine gleiche örtliche Farbauflösung zu haben, wird in vielen Anwendungen oft statt der zusätzlichen horizontalen Auflösungsreduktion die vertikale Reduktion bevorzugt (vgl. Abb. 2.2-1c). Diese zunächst recht plausible Vorgehensweise entpuppt sich aber gerade bei Zeilensprungsignalen als etwas tückisch. Verwendet man nämlich 4 im Quadrat örtlich benachbarte Luminanzpixel, auf die sich ein C_B-C_R -Paar beziehen soll, wird man feststellen, dass die übereinander liegenden Luminanzwerte aus verschiedenen Halbbildern und damit zeitlich aus unterschiedlichen Bewegungsphasen stammen (vgl. Abb. 1.4-3). Die Signalverarbeitung zur Berechnung der Chrominanzwerte des 4:2:0-Abtastrasters geschieht also in dem man durch eine örtlich-zeitliche Mittelwertbildung virtuelle „Zwischenzeilen“ errechnet, auf die sich die Chrominanzwerte beziehen. Dieser Vorgang wird oft auch als Formatkonversion bezeichnet.

Man muss sich daher im Klaren darüber sein, dass neben der vertikalen Auflösungsreduktion auch eine zeitliche Bewegungsverschleifung in den Farbkomponenten unvermeidlich ist. Diese Problematik ist aber in den Fällen weniger gravierend, in denen das analoge Eingangsbild auf dem PAL-Standard basiert, da bereits hier für die Chrominanz-Luminanz-Trennung ein örtlich-zeitliches Filter im Chrominanzpfad verwendet wird, das eine ähnliche Filterung vornimmt, wie die, die bei der 4:2:0-Abtastung erforderlich ist. Das ist auch der Grund, warum das 4:2:0-Abtastformat oft auf 625-Zeilensystemen (z. B. PAL-Europa) aufgesetzt wird (*DV-Codierung* –Kapitel 4, *DVB-Verfahren*–Kapitel 6).

Bei der Schreibweise „4:2:0-Abtastformat“ muss man zudem beachten, dass die Zahlen nicht mehr die Systematik mit den Vielfachen des Basiswertes der Abtastfrequenz anwenden; es handelt sich hierbei vielmehr um eine symbolische Schreibweise zur Unterscheidung der verschiedenen Abtastraster. Tatsächlich kann man schließlich noch das Abtastformat 4:1:0 definieren, das eine 4-fache horizontale und 2-fache vertikale Unterabtastung der Farbkomponenten bezeichnet.

2.2.2 Digitale Studionorm ITU-R BT.601

Die erste digitale Studionorm wurde vom *Comité Consultatif International des Radiocommunications (CCIR)* 1982 als Empfehlung (*Recommendation*) 601

„*Encoding Parameters of Digital Television for Studios*“ verabschiedet. In der Folgezeit gab es neben der Umorganisation des *CCIR* zum heutigen *ITU (International Telecommunication Union)* auch einige Ergänzungen zur ursprünglichen Norm 601. Für die nachfolgenden Betrachtungen soll die Version *ITU-R BT.601-4* aus dem Jahr 1994 herangezogen werden (vgl. [2.5]).

Diese Norm gibt die digitale Darstellung von 525-Zeilen- und 625-Zeilen-Systeme an, welche nach dem 4:2:2-Abtastformat codiert werden.

Die Pegel werden vor der Abtastung und Quantisierung wie folgt normiert: Das Y-Signal deckt analog einen Spannungsbereich zwischen 0 bis 700 mV ab. Dieser Bereich wird digital auf einen Bereich zwischen 0 und 1.0 V normiert. Damit auch die Farbdifferenzsignal C_B und C_R des Farbbalkensignals (vgl. Abb. 1.4-4) auf einen Bereich zwischen -0.5 V und +0.5 V abgebildet werden können, dürfen die korrespondierenden analogen Signale nur zwischen -350 mV und +350 mV liegen. Dies wird möglich, wenn die Farbdifferenzsignale C_B und C_R abweichend von Gl. 2.2-1 und Gl. 2.2-2 unter Berücksichtigung der Gleichung für die EBU-Primärvalenzen (vgl. Gl. 1.3-5) wie folgt gebildet werden:

Farbdifferenzsignale bei
ITU-R BT.601

$$C_B = 0.564 \cdot (E'_B - E'_Y) \quad C_B = -0.169E'_R - 0.331E'_G + 0.500E'_B \quad 2.2-6$$

$$C_R = 0.713 \cdot (E'_R - E'_Y) \quad C_R = 0.500E'_R - 0.419E'_G - 0.081E'_B. \quad 2.2-7$$

In diesem Zusammenhang beachte man den Unterschied bei Normierung zwischen der analogen Bildübertragung nach Gl.1.4-8 und den o. g. Gl. 2.2-6 und Gl. 2.2-7.

Die nachfolgende Tabelle (Tabelle 2.2-1) fasst die wesentlichen Parameter der Norm *ITU-R BT.601* für das 525-Zeilen- und 625-Zeilen-System zusammen. Einige Besonderheiten sind zu beachten. So sind beispielsweise für beide Fernsehsysteme die gleiche Abtastfrequenz (Y : 13.5 MHz, C_R , C_B : 6.75 MHz) und die gleiche Anzahl von aktiven Pixeln (Y : 720, C_R , C_B : 360) festgelegt worden. Das hat zur Konsequenz, dass sich die Anzahl der in die Austastlücke fallenden Abtastwerte zwischen beiden Normen geringfügig unterscheidet.

Parameter ITU-R
BT.601

Tab. 2.2-1: Wesentliche Parameter der ITU-R BT.601

Fernsehstandard (korrespondiert mit)	525 Zeilen-System (60 Hz) (NTSC Norm M)	625 Zeilen-System (50 Hz) (PAL Norm B/G)
Anzahl aktive Zeilen	480	576
Zeilendauer	63.55 μ s	64 μ s
Anzahl Abtastwerte über gesamte Zeilendauer (nominell)	Y: 858 C_B, C_R : 429	Y: 864 C_B, C_R : 432
Anzahl Abtastwerte je aktiver Zeile (53.33 μ s)	Y: 720 C_R, C_B : 360	
Abtastfrequenz	Y: 13.5 MHz C_B, C_R : 6.75 MHz	
Bandbegrenzung der zu codierenden Signale ($\pm 0,05$ dB)	Y: 5.75 MHz C_R, C_B : 2.75 MHz	
Quantisierung	Linear : 8 Bit oder 10 Bit (bevorzugt)	
Nutzbare Stufenzahl 8 Bit 10 Bit	Y: 220, C_R, C_B : 225 Y: 877, C_R, C_B : 897	
Abtaststruktur	Orthogonal (übereinander liegende Abtastwerte)	

Darüber hinaus sind die Pixel nicht exakt quadratisch. Um dieses zu erreichen müssten für ein 4:3 Bildseitenverhältnis und den 576 abgetasteten aktiven Zeilen aus beiden Halbbildern insgesamt

$$n_{square} = 576 \text{ Pixel} \cdot \frac{3}{4} = 768 \text{ Pixel} \quad 2.2-8$$

Pixel je Zeile vorhanden sein. In einigen Spezialanwendungen werden abweichend von der Norm tatsächlich 768 statt 720 aktive Pixel pro Zeile verwendet.

Aus den Angaben in Tabelle 2.2-1 lässt sich weiterhin entnehmen, dass die aktive Zeilendauer

$$T_{Z,\perp} = \frac{720}{13.5 \text{ MHz}} \approx 53.33 \mu\text{s} \quad 2.2-9$$

von dem korrespondierenden Zahlenwert

$$T_{Z,analog} = 52 \mu\text{s} \quad 2.2-10$$

des analogen Umfeldes abweicht. Das wiederum hat zur Folge, dass nach der Analog-Digital-Wandlung einer analogen Zeile nur effektiv etwa

$$n_{effektiv} = \frac{720 \text{ Pixel}}{53.33 \mu\text{s}} \cdot 52 \mu\text{s} \approx 702 \text{ Pixel} \quad 2.2-11$$

genutzt werden. Es verwundert daher auch nicht, wenn in vielen Übertragungen nur 704 aktive Pixel verwendet werden⁶.

⁶ Aufgrund der besonderen Eigenschaften der allgemein eingesetzten Quellencodierung (blockbasierte DCT, vgl. Kapitel 3) muss die aktive Pixelzahl pro Zeile durch 8 teilbar sein.

Signalpegelbereich bei
ITU-R BT.601

Bei der Quantisierung der Signalpegel auf 8- bzw. 10-Bit Codeworte werden die Extremwerte 0 und 255 (bei 8-Bit-Codierung) bzw. 0 und 1023 (bei 10-Bit-Codierung) nicht verwendet. Sie sind reserviert für spezielle Meldungen, die Synchronisationszwecken dienen (vgl. Abschnitt 2.3). Weiterhin werden beim abgetasteten Luminanzsignal die Bereiche zwischen 1 und 15 bzw. 1 und 63 als so genannten *Footroom* und die Bereiche zwischen 236 und 254 bzw. 941 und 1022 als so genannten *Headroom* verwendet. Diese Bereiche dienen als Über- bzw. Untersteuerungsreserve für Überschwinger des Analogsignals. Bei den Chrominanzsignalen sind dies die Bereiche zwischen 1 und 15 bzw. 1 und 63 sowie zwischen 241 und 254 bzw. 961 und 1022.

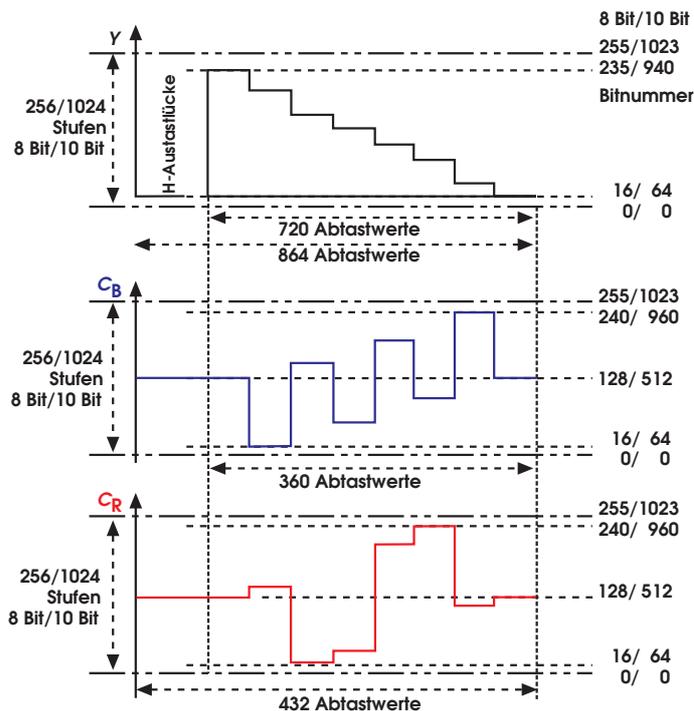


Abb. 2.2-2: Abtast- und Quantisierungswerte nach ITU-R BT.601 (Farbbalken)

In Abb. 2.2-2 sind die Abtast- und Quantisierungswerte nach ITU-R BT.601 eines 100/75%-Farbbalkensignals (vgl. Abb. 1.4-5) für eine 8- und 10-Bit Quantisierung dargestellt.

16:9 und ITU-R BT.601

Beim Übergang von einem Bildseitenverhältnis von 4:3 auf 16:9 (vgl. Gl. 1.4-3 und Gl. 1.4-4) ist es notwendig, die Anzahl der horizontalen Abtastwerte um den Faktor 1.33 auf 960 aktive Pixel zu erhöhen. Damit wird auch eine Abtastfrequenz für die Luminanz von 18 MHz notwendig. Auch dieser Fall wird in der Norm ITU-R BT.601 berücksichtigt.

In der Praxis zeigt sich jedoch, dass eine Erhöhung der horizontalen Abtastwerte bei einem 16:9 Bildseitenverhältnis nicht zwingend notwendig ist. Im Hinblick auf existierende Endgeräte im 16:9 PALplus-Format ist die mit effektiven 702 Abtastwerten erreichbare horizontale Auflösung immer noch höher, als die der analogen Endgeräte. Selbstverständlich sind dann die abgetasteten Pixel noch weniger quadratisch, sondern entsprechen Quadern mit einem Seitenverhältnis von

$$\frac{B}{H_{Pixel}} = \frac{\frac{576}{9}}{\frac{720}{16}} = \frac{64}{45} \approx 1.422. \quad 2.2-12$$

Der Vollständigkeit halber sei hier auch noch darauf hingewiesen, dass auch ein Digitalstandard für das abgetastete FBAS-Compositesignal definiert ist. Um die Störungen bei der Abtastung in Bezug auf den Farbhilfsträger gering zu halten, wird das FBAS-Signal exakt mit der vierfachen Farbhilfsträgerfrequenz f_{sc} abgetastet:

Digitales
Compositesignal

$$f_{\perp, Composite} = 4 \cdot f_{sc} = 4 \cdot (283.75 \cdot 15625 Hz + 25 Hz) \approx 17.73 MHz. \quad 2.2-13$$

Dieser Standard wird heute nur noch bei einigen wenigen digitalen Aufzeichnungsverfahren verwendet. Nachteilig ist die beschränkte 8 Bit-Auflösung, die PAL-typischen Störungen (z. B. Cross-Störungen zwischen Helligkeits- und Farbartsignal) sowie die recht beschränkte Farbauflösung des zugrunde liegenden FBAS-Signals.

Selbsttestaufgabe 2.2-1:

Formulieren Sie das Abtasttheorem nach Shannon!

Selbsttestaufgabe 2.2-2:

Definieren Sie die Begriffe Diskretisierung und Quantisierung!

Selbsttestaufgabe 2.2-3:

Erläutern Sie die Abtastformate 4:4:4, 4:2:2, 4:1:1 und 4:2:0! Welche Besonderheit ist bei 4:2:0 zu beachten?

Selbsttestaufgabe 2.2-4:

Geben Sie die Abtastfrequenz der Luminanz für ein 625-Zeilen-System nach der ITU-R BT.601 Norm an! Begründen Sie die Wahl dieser Abtastfrequenz!

Selbsttestaufgabe 2.2-5:

Wieso wird bei der Quantisierung nicht der komplette Coderaum von 8 bzw. 10 Bit für die Signalpegel genutzt?

Literatur

- [1.1] Poynton, Ch. Digital Video and HDTV - Algorithms and Interfaces Morgan Kaufmann Publishers - 2003
- [1.2] Hauske, G. Systemtheorie der visuellen Wahrnehmung Teubner 1994
- [1.3] Hentschel, H.-J. Licht und Beleuchtung Hüthig 1987
- [1.4] Kubitz, P.P. Der Traum vom Sehen Verlag der Kunst 1997
- [1.5] Lang, H. Farbmeterik und Farbfernsehen Oldenbourg 1978
- [1.6] Smith, W.J. Modern Optical Engineering McGraw Hill 2000
- [1.7] Mäusl, R. Fernsehtechnik - Vom Studiosignal zum DVB-Sendesignal Hüthig-Verlag 2002
- [1.8] Schmidt, U. Professionelle Videotechnik Springer-Verlag 2000
- [1.9] Bonse, Th. Zur Konzeption einer visuell angepassten Beschreibung und Darstellung von Bewegtbildern Shaker Verlag 1996
- [1.10] Kelly, D. H. Motion and Vision. II. Stabilized Spatio-Temporal Threshold Surface Journal of Optical Society of America., Vol 69, No. 10, pp. 1340-1349, 1979
- [1.11] Watanabe, A., H. Sakata, H. Isona Chromatic Spatial Sine-Wave Responses of the Human Visual System NHK Laboratories Notes Ser. No. 198, 1976
- [1.12] EBU Document Tech. 3213 EBU Standard for Chromaticity Tolerances for Studio Monitors, 1975 reissued 1981
- [1.13] Society of Motion Picture and Television Engineers (SMPTE) RP 145 SMPTE C Color Monitor Colorimetry updated 1999
- [2.1] Mäusl, R., siehe [1.7]
- [2.2] Schmidt, U., siehe [1.8]
- [2.3] Reimers, U. Digital Video Broadcasting - The International Standard for Digital Television Springer-Verlag 2001
- [2.4] Poynton, Ch., siehe [1.1]
- [2.5] ITU-R BT.601-4 Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios 1994
- [2.6] ITU-R BT.656-3 Interfaces for Digital Component Video Signals in 525-Line and 625-Line Television Systems Operating at the 4:2:2 Level of Recommendation ITU-R BT.601 (Part A) 1995
- [2.7] Watkinson, J. The Art of Digital Audio Focal Press Oxford, 1994

- [2.8] ITU-R BT.1120 Digital Interfaces for 1125/60/2:1 and 1250/50/2:1 HDTV Studio Signals
- [2.9] Digital Display Working Group Digital Visual Interface - DVI 1999
- [2.10] Ottenbreit, M. SDI & Embedded Audio FKT Jg. 57, Heft 10/03, S. 466ff - Hüthig-Verlag 2003
- [3.1] Mäusl, R., siehe [1.7]
- [3.2] Schmidt, U., siehe [1.8]
- [3.3] Reimers, U., siehe [2.3]
- [3.4] Ohm, J.-R. Digitale Bildcodierung - Repräsentation, Kompression und Übertragung von Bildsignalen Springer-Verlag 1995
- [3.5] Symes, P. Video Compression Demystified - How Compression Really Works; DCT and wavelets; MPEG & DV McGraw-Hill 2001
- [3.6] ITU-T T.81 Information Technology – Coded Representation of Picture and Audio Information – Progressive Bi-level Image Compression 1992
- [3.7] ISO/IEC 10918-1 *identisch mit* [3.6] 1993
- [3.8] Lohscheller, H. Subjectively Adapted Image Communication System IEEE Transactions on Communication, COM-32(12), S. 1316-22 1984
- [3.9] Hilton, M. L., B. D. Jawerth, A. Sengupat Compression of Still and Moving Images with Wavelets Multimedia Systems, Vol. 2, No. 3 1994
- [3.10] Edwards, T Discrete Wavelet Transforms: Theory and Implementation Stanford University, September 1991
- [3.11] Daubechies, I. Orthonormal Bases of Compactly Supported Wavelets Comm. Pure Appl. Math., 41, S. 909–996 1988
- [3.12] Said, A, W. A. Pearlman A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees IEEE Transactions on Circuits and Systems for Video Technology, Vol. 6 1996
- [3.13] Shapiro, J. M. Embedded Image Coding Using Zerotrees of Wavelet Coefficients IEEE Trans. Signal Proc. 41(12), S. 3445–3462 1993
- [3.14] ISO/IEC 15444-1 Information technology - JPEG 2000 image coding system Part 1: Core coding system 2000
- [3.15] Barnsley, M. F., L. P. Hurd Bildkompression mit Fraktalen Vieweg-Verlag 1996
- [3.16] Fisher, Y. (ed.) Fractal Image Compression - Theory and Application Springer-Verlag 1994
- [4.1] Mäusl, R., siehe [1.7]

- [4.2] Ohm, J.-R., siehe [3.4]
- [4.3] Poynton, Ch., siehe [1.1]
- [4.4] Reimers, U., siehe [2.3]
- [4.5] Schmidt, U., siehe [1.8]
- [4.6] Schönfelder, H. Fernsehtechnik im Wandel Springer-Verlag 1996
- [4.7] Symes, P., siehe [3.5]
- [4.8] Wendland, B., H. Schröder Fernsehtechnik, Band II: Systeme und Komponenten zur Farbbildübertragung Hüthig-Verlag 1991
- [4.9] Ziemer, A. Digitales Fernsehen Hüthig-Verlag 2002
- [4.10] Hartwig, S., W. Endemann Digitale Bildcodierung Sonderdruck aus Fernseh- und Kinotechnik 01/92 bis 01/93 Hüthig-Verlag 1993
- [4.11] Fell-Bosenbeck, F. DV-Kompression - Grundlagen und Anwendungen Fernseh- und Kinotechnik, 53. Jahrgang, Nr. 6, 1999, S. 336-345
- [4.12] SMPTE 314M Television - Data Structure for DV-Based Audio, Data and Compressed Video - 25 and 50 Mb/s 1999
- [4.13] IEC 61834 Helical-Scan Digital Video Cassette Recording System Using 6,35 mm Magnetic Tape for Consumer User 1999
- [4.14] Wiswell, D. Panasonic DVCPRO - from DV to HD <http://www.panasonic.com/broadcast>
- [4.15] Cafforio, C., F. Rocca Methods for Measuring Small Displacements of Television Images IEEE Transactions on Information Theory, No.5, 1976
- [4.16] Sikora, T. MPEG Digital Video-Coding Standards in: IEEE Signal Processing Magazine - September 1997, pp. 82-100 1997
- [4.17] Hollmann, Th. Zur Gewinnung von Bewegungsinformation in einer Kamera mit hoher Bildaufnahmefrequenz Shaker Verlag 1994
- [5.1] ITU-T Recommendation H.120: Codec for Videoconferencing Using Primary Digital Group Transmission, 1. Version: 1984, 2. Version: 1988
- [5.2] ITU-T Recommendation H.261: Video Codec for Audiovisual Service at p x 64 kbits, 1. Vorversionen: 1988, Standardisiert: 1991, Letzte Erweiterungen: 1993
- [5.3] ISO/IEC 11172 Information Technology - Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1,5 Mbit/s (MPEG-1) 1992
- [5.4] Heyna, A., M. Briede, U. Schmidt (Hrsg.) Datenformate im Medienbereich Fachbuchverlag Carl Hanser Verlag Leipzig, 2003

- [5.5] Symes, P., siehe [3.5]
- [5.6] Ohm, J.-R., siehe [3.4]
- [5.7] Poynton, Ch., siehe [1.1]
- [5.8] ISO/IEC 13818 Information Technology - Generic Coding of Moving Pictures and Associated Audio Information (*MPEG-2*); Part 1: Systems, Part 2: Video 1994
- [5.9] Reimers, U., siehe [2.3]
- [5.10] Mäusl, R., siehe [1.7]
- [6.1] International Organization for Standardization ISO 9660:1988 Information processing – Volume and file structure of CD-ROM for information interchange, 1988
- [6.2] International Organization for Standardization ISO/IEC 10149:1995 (Yellow-Book) Information technology – Data interchange on read-only 120 mm optical data disks (CD-ROM), 1995
- [6.3] Taylor, J. DVD Demystified - Second Edition McGraw-Hill, 2001
- [6.4] LaBarge, R. DVD Authoring & Production CMP books, 2001
- [6.5] European Telecommunications Standards Institute (ETSI) Norm ETS 300 468 (V1.3.1) Digital Video Broadcasting (DVB) Specification for Service Information (SI) in DVB systems (DVB-SI) 1998
- [6.6] Mäusl, R., siehe [1.7]
- [6.7] Reimers, U., siehe [2.3]
- [6.8] Symes, P., siehe [3.5]
- [6.9] Ziemer, A., siehe [4.9]
- [6.10] European Telecommunications Standards Institute (ETSI) Norm EN 300 421 (V1.1.2) Digital Video Broadcasting (DVB) Framing Structure, Channel Coding and Modulation for 11/12 GHz Satellite Services 1997
- [6.11] Sweeney, P. Codierung zur Fehlererkennung und Fehlerkorrektur Hanser Verlag 1992
- [6.12] Forney, G. D. Burst-Correcting Codes for the Classic Bursty Channel IEEE Transactions on Communication Technology – COM-10 No. 5, 10/1971, S. 772ff
- [6.13] European Telecommunications Standards Institute (ETSI) Norm EN 300 429 (V1.2.1) Digital Video Broadcasting (DVB) Framing Structure, Channel Coding and Modulation for Cable Systems (DVB-C) 1998
- [6.14] European Telecommunications Standards Institute (ETSI) Norm EN 300 744 (V1.1.2) Digital Video Broadcasting (DVB) Framing Structure,

- Channel Coding and Modulation for Digital Terrestrial Television (DVB-T) 1997
- [7.1] Jack, Keith Video Demystified – A Handbook for the Digital Engineer LLH Technology Publishing – 3rd Edition 2001
 - [7.2] International Telecommunication Union Line Transmission of Non-Telephone Signals (ITU-T) Video Coding for Low Bit Rate Communication Recommendation H.263, Version 1 – 12/1995
 - [7.3] International Telecommunication Union Line Transmission of Non-Telephone Signals (ITU-T) Video Coding for Low Bit Rate Communication Recommendation H.263, Version 2 (H.263+) – 01/1998
 - [7.4] International Telecommunication Union Line Transmission of Non-Telephone Signals (ITU-T) Video Coding for Low Bit Rate Communication Recommendation H.263, Version 3 (H.263++) – 11/2000
 - [7.5] International Telecommunication Union Series H: Audiovisual And Multimedia Systems Infrastructure of Audiovisual Services – Coding of Moving Video Video Coding for Low Bit Rate Communication Recommendation H.263, Annex X – 04/2001
 - [7.6] International Organization for Standardization ISO/IEC 14496 Information Technology – Coding of Audio-Visual Objects 2000, Part 1 - 6
 - [7.7] Sikora, T., B. Makai Shape-Adaptive DCT for Generic Coding of Video” IEEE Transactions on Circuits and Systems for Video Technology Bd. 5, Nr. 1, S. 59-62 02/1995
 - [7.8] International Organization for Standardization ISO/IEC 14496-2 Information Technology – Coding of Audio-Visual Objects Part 2: Visual, Amendment 1: Studio Profile 2002
 - [7.9] International Organization for Standardization ISO/IEC 14496-2 Information Technology – Coding of Audio-Visual Objects Part 2: Visual, Amendment 2: Streaming Video Profile 2002
 - [7.10] Waggoner, B. Compression for Great Digital Video CMP Books 2002
 - [7.11] Pereira, F., T. Ebrahimi The MPEG-4 Book IMSC Press Multimedia Series 2002
 - [7.12] Symes, P., siehe [3.5]
 - [7.13] International Telecommunication Union Series H: Audiovisual And Multimedia Systems Infrastructure of Audiovisual Services – Coding of Moving Video Advanced Video Coding for Generic Audiovisual Services Recommendation H.264 – 05/2003 (Prepublish)
 - [7.14] Marpe, D. H.264 / MPEG-4 AVC: The New Video Coding Standard Beitrag zum Wiesbadener *Medien Symposium* “Codierung audiovisueller Daten”, 29.10. – 30.10.2003

- [7.15] Wiegand, Th., G. J. Sullivan, G. Bjoentegaard, A. Luthra u .a. Overview of the H.264/AVC Video Coding Standard IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13, No. 7 pp. 560-576 July 2003
- [7.16] Schäfer, R., Th. Wiegand, H. Schwarz The Emerging H.264/AVC Standard EBU Technical Review Januar 2003
- [7.17] Schäfer, R. Videocodierung für Fernsehen, Bildkommunikation und Multimedia Beitrag zum Wiesbadener *Medien Symposium* “Codierung audiovisueller Daten”, 29.10. – 30.10.2003
- [7.18] Heyna, A., M. Briede, U. Schmidt (Hrsg.), siehe [5.4]

Lösungen zu Selbsttestaufgaben der Kurseinheit

Lösung zu Selbsttestaufgabe 1.2-1:

Der ungefähre Wellenlängenbereich für die sichtbare optische Strahlung liegt zwischen 380 nm und 760 nm.

Lösung zu Selbsttestaufgabe 1.2-2:

Die Sensortypen auf der menschlichen Netzhaut heißen Stäbchen und Zapfen. Die Stäbchen sind für die Helligkeitswahrnehmung, die Zapfen für die Farbwahrnehmung zuständig.

Lösung zu Selbsttestaufgabe 1.2-3:

Der Mach-Effekt äußert sich in der gegenseitigen Beeinflussung örtlich benachbarter Sinneszellen. Es handelt sich hierbei um eine Bandpasscharakteristik. Wird eine Sehzelle durch einen Lichtreiz aktiv, hemmt diese Aktivität gleichzeitig die Aktivität unmittelbar benachbarter Sinneszellen. Der Mach-Effekt ist auch an Objektkanten zu beobachten und führt in der Regel subjektiv zu einer Kontrasterhöhung im Kantenbereich.

Lösung zu Selbsttestaufgabe 1.2-4:

Die Bewegungsverschmelzungsfrequenz ist die minimale Bildwiederholfrequenz, die erforderlich ist, um in der dargestellten Folge von Einzelbildern einen kontinuierlichen Bewegungsverlauf wahrnehmen zu können.

Die Flimmerverschmelzungsfrequenz ist die minimale Bildwiederholfrequenz, die erforderlich ist, um bei vorgegebenen Beleuchtungs- und Betrachtungsbedingungen kein Flimmern oder Flackern mehr wahrzunehmen.

Lösung zu Selbsttestaufgabe 1.2-5:

Werden die einzelnen Kinobilder mehrfach hintereinander projiziert (Realisierung: Shutterblende) kann der Flimmereindruck vermindert werden.

Da diese Maßnahme faktisch einer Bildwiederholung gleichkommt, kann es bei Bewegungen im Bild (Objekt- und/oder globale Bewegungen) zu Ruckelstörungen kommen.

Lösung zu Selbsttestaufgabe 1.2-6:

In Abb. 1.2-6 ist die Kontrastempfindlichkeit des visuellen Systems für farbart- und helligkeitsmodulierte räumliche Gitter dargestellt. Daraus ist, dass die Kontrastempfindlichkeit für Farbartübergänge zu höheren Frequenzen früher abfällt als für Helligkeitsübergänge. So ist das Auflösungsvermögen des menschlichen Gesichtsinns für Farbartübergänge (bei gleicher Helligkeit) wesentlich geringer als für Helligkeitsübergänge.

Lösung zu Selbsttestaufgabe 1.3-1:

Additive Farbmischung: Hier vermischen sich die Farbanteile von selbst leuchtende Farbpunkten. Aus großer Entfernung betrachtet werden die einzelnen Farbpunkte nicht mehr separat aufgelöst und verschmelzen damit additiv zu einer neuen Farbe. Beispiel für dieses Farbmischprinzip ist die Farbdarstellung eines Farb-Kathodenstrahlmonitors.

Subtraktive Farbmischung: Hierbei durchdringt ein spektral weißes Licht mehrere hintereinander liegende Farbstoffe (quasi als Farbfilter). Die verschiedenen Farbstoffe besitzen unterschiedliche spektrale Absorptionseigenschaften und entziehen somit dem Lichtstrahl verschiedene spektrale Anteile. Übrig bleibt der spektrale Anteil des Lichtes, der von keinem Farbstoff absorbiert wurde. Ein Beispiel für die subtraktive Farbmischung ist das Zusammenführen verschiedenfarbiger Farbstoffe, wie es bei der Malerei geschieht. Die subtraktive Farbmischung wird vielfach auch bei Farbdruckern verwendet, die mit drei oder mehr Grundfarben arbeiten. Als Grundfarben werden oft gelb, magenta, cyan und ggf. weitere verwendet.

Lösung zu Selbsttestaufgabe 1.3-2:

Metamere Farben sind Farbvalenzen, die gleich aussehen, nicht aber unbedingt die gleiche spektrale Verteilung besitzen.

Lösung zu Selbsttestaufgabe 1.3-3:

Mit einem EBU- oder SMPTE-Farbmonitor lassen sich nicht alle in der Natur vorkommenden Farbarten darstellen, da nur solche dargestellt werden können, die innerhalb des jeweiligen Farbdreiecks liegen (vgl. Abb. 1.3-5). Es lassen sich Farborte außerhalb des jeweiligen Farbdreiecks zwar in Bezug auf den Farbton korrekt darstellen, nicht aber mit der vollständigen Farbsättigung. Die maximal mögliche Farbsättigung liegt auf der Begrenzung des jeweiligen Farbdreiecks selbst.

Lösung zu Selbsttestaufgabe 1.3-4:

Als praktischen Unbuntpunkt wird der Weißpunkt D65 für beide Farbdreiecke verwendet: $D65(x, y, z) = (0.3127, 0.3290, 0.3582)$.

Lösung zu Selbsttestaufgabe 1.4-1:

Die Bildvorlage wird in horizontale, vertikal aneinander angrenzende Zeilen zerlegt. Dazu wird der Bildinhalt die Zeilen von oben links anfangend zeilenweise nacheinander bis zum Ende der letzten Zeile unten rechts ausgelesen.

Lösung zu Selbsttestaufgabe 1.4-2:

Die nominelle Zeilenzahl ist größer als die aktive Zeilenzahl. Die Differenz ergibt sich aus dem Wunsch, eine gewisse Zeit zwischen der letzten aktiven Zeile des Bildes n bis zum Beginn der ersten Zeile des Bildes $n+1$ zu lassen. Diese Zeit wird empfängerseitig für den vertikalen Rücklauf eines Wiedergabesystems mit Kathodenstrahlröhre benötigt, der nicht beliebig kurz sein kann (vgl. Abb. 1.4-2 roter Sägezahn).

Lösung zu Selbsttestaufgabe 1.4-3:

Ähnlich wie beim vertikalen Rücklauf (vgl. Selbsttestaufgabe 1.4-2) wird auch Zeit für einen horizontalen Zeilenrücklauf vorgesehen, da der für die Ansteuerung notwendige Sägezahn (in Abb. 1.4-2 blau eingezeichnet) in der Rücklaufphase nicht beliebig steil sein kann.

Lösung zu Selbsttestaufgabe 1.4-4:

Die erforderliche Transformationsgleichung ist in Gl. 1.4-9 dargestellt:

$$\begin{pmatrix} U \\ V \\ Y \end{pmatrix} = \begin{pmatrix} -0.147 & -0.289 & +0.436 \\ +0.615 & -0.515 & -0.100 \\ +0.299 & +0.587 & +0.114 \end{pmatrix} \cdot \begin{pmatrix} E'_R \\ E'_G \\ E'_B \end{pmatrix}$$

Lösung zu Selbsttestaufgabe 2.2-1:

Für eine fehlerfreie Rekonstruktion eines abgetasteten Signals muss die Abtastfrequenz f_T mindestens doppelt so hoch sein, wie die höchste vorkommende Frequenz des Eingangssignals. Dann überlappen sich die Spektren nicht und das analoge Ausgangssignal kann theoretisch exakt wieder durch ein geeignetes Rekonstruktionsfilter zurück gewonnen werden.

Formal gilt also:

$$f_T \geq 2 \cdot f_{max}.$$

Lösung zu Selbsttestaufgabe 2.2-2:

Unter der Diskretisierung versteht man eine örtliche und/oder zeitliche Aufteilung des Signals in eine beschränkte Anzahl von (kontinuierlichen) Werten, die das Ausgangssignal repräsentieren.

Unter der Quantisierung versteht man das Runden eines analogen Ausgangssignals in eine beschränkte Anzahl von Wertestufen. In der Regel wird die Quantisierung in $2N$ verschiedene Stufen vorgenommen. In der digitalen Bildtechnik sind Wortbreiten von $N = 8$ und $N = 10$ gebräuchlich.

Lösung zu Selbsttestaufgabe 2.2-3:

Die Abtastformate $M:N:O$, (M , N , O natürliche Zahlen) geben das Abtastverhältnis zwischen Y , C_B und C_R , bezogen auf eine Basisfrequenz von $f_{\perp, Basis} = 3.375$ MHz an. Für 4:4:4 ergibt sich eine Abtastfrequenz von je 13.5 MHz, der Studiostandard arbeitet mit einem Abtastverhältnis von 4:2:2. Hier sind also die Farbdifferenzsignale horizontal um den Faktor 2 niedriger abgetastet (6.75 MHz) als das Helligkeitssignal. Bei 4:1:1 sinkt die Abtastfrequenz für die Farbdifferenzsignale noch einmal um den Faktor 2 auf eine Frequenz von 3.375 MHz ab. Es handelt sich insgesamt also um eine Farbunterabtastung in horizontaler Richtung um den Faktor 4.

Das Abtastformat 4:2:0 gibt nicht mehr das Abtastverhältnis zwischen Y , C_B und C_R an. Vielmehr handelt es sich hierbei um eine horizontale und vertikale Unterabtastung der Farbdifferenzsignale jeweils um den Faktor 2.

Lösung zu Selbsttestaufgabe 2.2-4:

Beim ITU-R Standard BT.601 wird die Luminanz bei einem 625-Zeilen-System mit 13.5 MHz abgetastet. Hier kommt das Abtastformat 4:2:2 zum Einsatz, welches sich aus dem vierfachen der Basisfrequenz $f_{\perp, Basis} = 3.375 \text{ MHz}$ errechnet .

$$f_{\perp} = 4 \cdot 3.375 \text{ MHz} = 13.5 \text{ MHz}.$$

Lösung zu Selbsttestaufgabe 2.2-5:

Beim abgetasteten Luminanzsignal werden die Bereiche zwischen 1 und 15 (8-Bit-Codierung) bzw. 1 und 63 (10-Bit-Codierung) und die Bereiche zwischen 236 und 254 (8-Bit-Codierung) bzw. 941 und 1022 (10-Bit-Codierung) nicht für das eigentliche Videosignal verwendet. Diese Bereiche dienen als Über- bzw. Untersteuerungsreserve für Überschwinger des Analogsignals. Bei den Chrominanzsignalen sind dies die Bereiche zwischen 1 und 15 bzw. 1 und 63 sowie zwischen 241 und 254 bzw. 961 und 1022. Die Werte 0 und 255 (8-Bit-Codierung) bzw. 0 und 1023 (10-Bit-Codierung) sind für spezielle Meldungen, die Synchronisationszwecken dienen reserviert.

