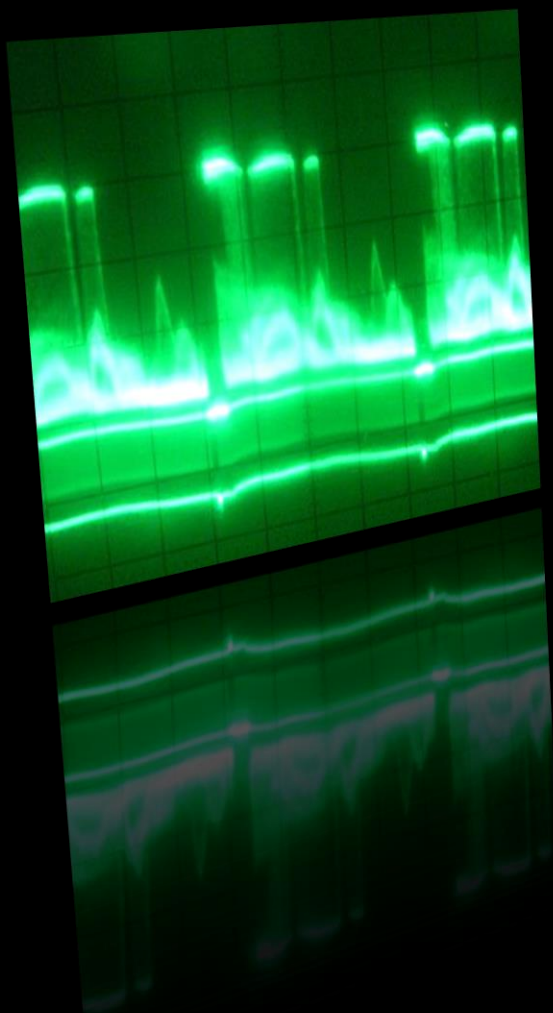


Videotechnik v2023.1 MEDT Pi



INHALT

Bewegte Bilder	6
Videoformate	7
HDTV-Format.....	7
Common Image Format	8
Seitenverhältnis	9
Safe-Areas bei 16:9 Produktionen	10
Bildverschmelzungsgrenze und Bildwiederholfrequenz.....	11
Zeilensprung und progressive Abtastung	12
Pulldown	14
Bildraum und Objektraum	14
Polarisation.....	15
Ortsfrequenz	15
Bildpunkte	17
Einflussfaktoren auf Bildqualität im Video/TV	18
Abbildungsfehler	18
Gammakorrektur der RGB-Signale	19
Digitale Videosignale	20
Farbkonvertierung	20
Komprimierung	21
Chroma-Subsampling.....	21
Das digitale Komponentensignal YCbCr	23
SDI Serial Digital Interface	24
Signalqualität bei SDI.....	25
HDMI.....	27
Stecker und Kabel.....	28
BNC	28
Cinch	28
DIN-Stecker	29
HDMI.....	29

Koaxialkabel	30
Messtechnik.....	31
Der Wellenformmonitor	32
EBU-Farbbalken als Testsignal im Wellenformmonitor	33
EBU-Farbbalken als Testsignal im Vektorskop.....	34
Hauttonkorrektur mit dem Vektorskop.....	35
A/V-Datenstrukturen.....	36
CODECS	36
Container	37
Container und CODECs Übersicht	37
CODECS ohne Daten-Komprimierung	38
CODECS mit verlustfreier Daten-Komprimierung	38
CODECS mit verlustbehafteter Daten-Komprimierung	38
Verlustbehaftete strategien der Datenkomprimierung.....	39
Intraframe-Komprimierung.....	39
Interframe-Komprimierung.....	39
Group of Pictures GOP.....	39
Exkurs: Multimediaframework FFMPEG.....	42
Bezugsquelle	42
Installation.....	42
Syntax	43
Ausgabevorschau	43
Erste Schritte.....	43
Überblick weiterer Optionen	45

Abbildungsverzeichnis

Abbildung 1: BT.709 und BT.2020 Farbraumabdeckung.....	8
Abbildung 2: Nicht-Anamorph und Anamorph. Links ist die entzerzte Version.....	9
Abbildung 3: Safe-Area bei Premiere	10
Abbildung 4: Safe Areas für vertical video in UHD.....	10
Abbildung 5: Interlaced und Progressive	12
Abbildung 6: Oberes und unteres Halbbild eines Frames	13
Abbildung 7: Links: Zwei Halbbilder vereint (50i), Rechts: Progressive Abtastung (25p)	13
Abbildung 8: Bildraum und Objektraum.....	14
Abbildung 9: Schema Transversalwelle	15
Abbildung 10: Kontrast abhängig von Leuchtdichte	15
Abbildung 11: Kontrastempfindlichkeit Y vs. Ortsfrequenz in Linienpaare pro Grad X	16
Abbildung 12: Siemensstern.....	16
Abbildung 13: Ermittlung der notwendigen Zeilenauflösung und Betrachtungsabstand	17
Abbildung 14: Betrachtungsabstand, Videoformat und Panelgröße in Relation.....	18
Abbildung 15: Links linear und rechts mit Gamma 2,2	20
Abbildung 16: Farbunterabtastung	21
Abbildung 17: 4:4:4, 4:2:2 und 4:2:0 im Vergleich	22
Abbildung 18: NRZI-Kodierung	24
Abbildung 19: TV-Studio Schema	25
Abbildung 20: Zu langes Kabel, falsch terminiert, Rise- und Fall-Time nicht symmetrisch.....	26
Abbildung 21: Mangelhaftes SDI-Signal	26
Abbildung 22: Veränderung der Signalqualität bei unterschiedlichen Leitungslängen.....	26
Abbildung 23: BNC-Stecker	28
Abbildung 24: Cinch-Stecker.....	28
Abbildung 25: Mini-DIN-Stecker 4polig.....	29
Abbildung 26: HDMI Typ A, Typ C (Mini), Typ D (Micro)	29
Abbildung 27: Max. Kabellängen eines Herstellers für professionelle Koaxkabel.....	30
Abbildung 28: Leader LV5490.....	31
Abbildung 29: IRE.....	32
Abbildung 30: Wellenformmonitor-Ansicht (Lumetri)	32
Abbildung 31: EBU 100/75 Farbbalken in Premiere	33
Abbildung 32: EBU 100/75 Farbbalken Y/C im Wellenformmonitor in Premiere.....	33
Abbildung 33: EBU 100/75 Farbbalken am Tektronix Oszilloskop.....	34
Abbildung 34: Vektorskop mit Premiere-internem HD-Farbbalken (ARIB)	34
Abbildung 35: ARIB Farbbalken. Die markierten Bereiche sind im Vektorskop als überpegelt dargestellt.....	35
Abbildung 36: Hauttöne weisen Farbstich auf, befinden sich neben der Linie	35
Abbildung 37: Schnellkorrektur des Hauttöns	35

Abbildung 38: Anforderungen an Streaming-Codecs Übersicht	38
Abbildung 40: GOP mit I-, B- und P-Frames	40
Abbildung 41: 6-Frame-GOP mit $m = 3$, $n = 6$	40
Abbildung 42: Medien-Exporteinstellung mit GOP-Konfiguration	41
Abbildung 43: Beispiel für einen 15 Frame-Long-GOP mit $m = 3$ und $n = 15$	41

Einleitung

Ein Video (vom lateinischen *video = sehen*) besteht aus einer **Abfolge von Einzelbildern in einer bestimmten Frequenz (Bildwiederholfrequenz)**. Diese werden grundsätzlich mit einer Optik und einem Bildwandler (Sensor oder Filmmaterial) erzeugt.

Es gelten **ähnliche fototechnische Prinzipien** (Blende, Belichtungszeit, Dynamik...) wie bei einer (DSLR/M)-Kamera. Zusätzlich wird jedoch der Datenstrom in ein bestimmtes Format **kodiert bzw. dekodiert (Codec)**, zB. MPEG oder h.264 und gespeichert. Als **Speichermedien** dienen Film, Magnetband, Solid-State-Speicher, Festplatten oder ähnliche Speicher mit entsprechenden Lese- und Schreibraten.

BEWEGTE BILDER

Beim Menschen entsteht eine **Illusion der Kontinuität ab ungefähr 12 Bildern pro Sekunde** (siehe zB. Daumenkino), falls die dargestellten Objekte langsam bewegt sind. Ab 16 Bildern pro Sekunde entsteht auch für schnellere Bewegungen ein solcher Eindruck. Der Psychologe Max Wertheimer und der Arzt/Forscher Peter Max Roget erkannten auf Basis der Gestaltpsychologie zwei dafür verantwortliche Phänomene:

- Die **Nachbildwirkung** (Trägheit der Retina)
- Der **stroboskopische Effekt**, nach dem unser **Gehirn Einzelbilder anders verarbeitet** als eine **Abfolge von Einzelbildern**. Diese **verschmelzen** zu einem **Bewegungseindruck**, der von den Einzelbildern entkoppelt betrachtet werden kann.

Es entsteht somit eine zusätzliche **Informationsdimension** (*für weitere Infos recherchiere den **Kuleshov-Effekt***). Das Gesamte ist somit auch in diesem Fall mehr als die Summe seiner Teile. Der Fotograf **Eadweard Muybridge (19. Jhdt.)** legte mit seinen Serienaufnahmen den Grundstein für die Entwicklung der Laufbildmedien. Sein **Zoopraxiskop** war der Vorläufer der Projektionstechnik in Kinos.

*Recherchiere und beschreibe den **Kuleshov-Effekt**. Zeige weitere Dimensionen der Informationsvermittlung von Bewegtbildern (Schnitttechniken, Montage, ...)*

*Recherchiere den **Reaction Shot**, der von Steven Spielberg etabliert wurde und der auf dem Kuleshov Effekt basiert.*



VIDEOFORMATE

Wir unterscheiden grundsätzlich folgende **Parameter der Videoformate**:

- **Seitenverhältnis**: 4:3, 16:9, 21:9...
- **Bildpunkte** bzw. **Anzahl an Pixel gesamt und x/y**: 1920 px x 1080 px = 2 073 600 px
- **Bildwiederholfrequenz**: 25 Hz, 50 Hz...
- **Farbtiefe**: 8 bit, 10 bit...
- **Abtastung**: interlace i (Zeilensprung) oder progressive p (Progressiv)
- **Audio**: Stereo, Mehrkanal

In diesem Skript beschäftigen wir uns hauptsächlich mit dem **HDTV-Standard** und dem **SDTV-Standard**. UHD TV, 8K usw. können in den jeweiligen Spezifikationen der ITU¹, EBU² oder SMPTE³ nachgelesen werden.

HDTV-FORMAT

In **Europa** wird **HDTV** laut folgender Spezifikation definiert:

<https://tech.ebu.ch/docs/tech/tech3299.pdf>

- **System 1** (S1) with **1280** horizontal samples and **720** active lines in **progressive** scan with a frame rate of **50 Hz**, **16:9** aspect ratio. Abbreviated **720p/50**.
- **System 2** (S2) with **1920** horizontal samples and **1080** active lines in **interlaced** scan with a frame rate of **25 Hz**, **16:9** aspect ratio. Abbreviated **1080i/25**.
- **System 3** (S3) with **1920** horizontal samples and **1080** active lines in **progressive** scan and a frame rate of **25 Hz**, **16:9** aspect ratio. Abbreviated **1080p/25**.
- **System 4** (S4) with **1920** horizontal samples and **1080** active lines in **progressive** scan at a frame rate of **50 Hz**, **16:9** aspect ratio. Abbreviated **1080p/50**.

Grundlage für diese Spezifikationen sind unter anderem die Empfehlungen der ITU, nämlich **ITU-R BT.709** und **ITU-R BT.2020**, welche die Regeln zur A/D-Wandlung eines analogen Komponentensignals beschreiben.

	BT.709	BT.2020
Bildpunkte	HD	UHD, 8K
Bildwiederholfrequenz	25, 24, 30, 50, 60	24, 25, 30, 50, 60, 100, 120
Zeilensprung/Progressiv	Beide	Nur progressiv
Farbraum	BT.709	BT.2020
Bittiefen	8, 10	10, 12
Farbmodell	RGB, YCbCr	RGB, YCbCr

¹ International Television Union mit Sitz in Genf

² European Broadcast Union

³ Society of Motion Picture and Television Engineers

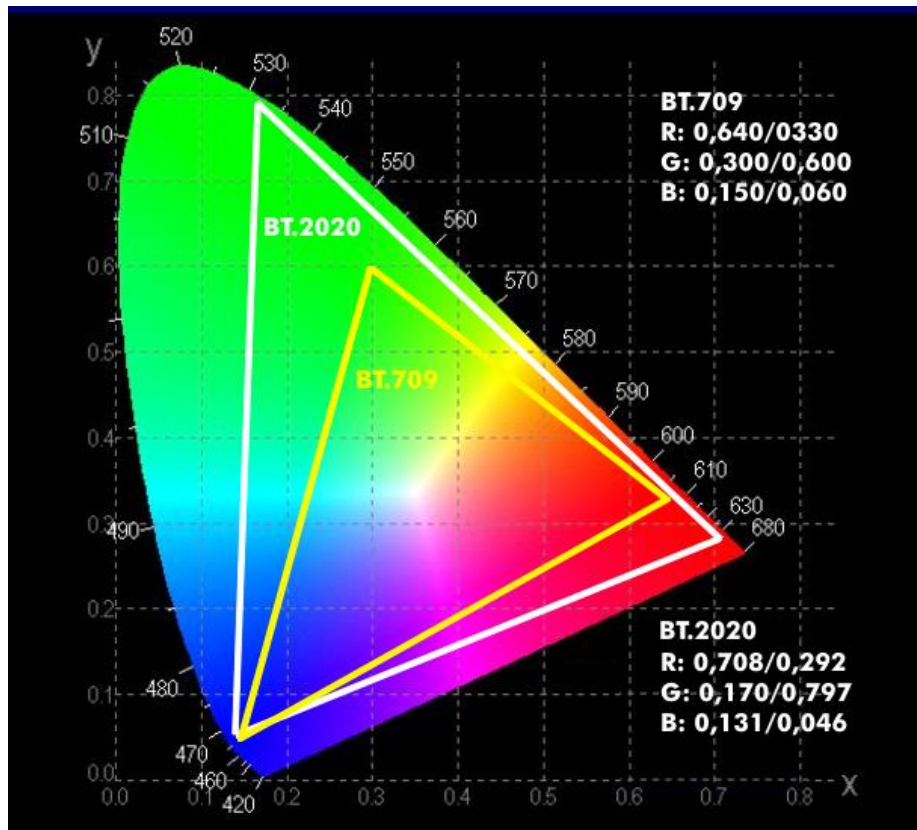


Abbildung 1: BT.709 und BT.2020 Farbraumabdeckung

Für die **Übertragung der Signale im Produktionsumfeld** wird **HD-SDI** verwendet, auf welches später im Skript eingegangen wird.

COMMON IMAGE FORMAT

Das **CIF** definiert **Bildpunkte, Bildwiederholfrequenz und Seitenverhältnis** des Bildes und der Pixel. Die Bruttodatenrate ist mit **1,485 Gbit/s** bei allen Übertragungen gleich.

	Nettozeilen	Bruttozeilen	Nettopixel/Zeile	Bruttopixel/Zeile
1080p24	1080	1125	1920	2750
1080i/p25	1080	1125	1920	2640
720p50	720	750	1280	1980

$$2750 \times 1125 \times 24 \times 20 =$$

$$1.485.000.000$$

Bruttozeilen * Bruttozeilen * Frames * Bittiefe pro pixel

20 bits da 4:2:2 --> 30 Bits falls 4:4:4

SEITENVERHÄLTNIS

Aspect Ratio = Seitenverhältnis. Unterscheide SAR, DAR und PAR. Wenn die Seitenverhältnisse für SAR und DAR übereinstimmen spricht man von **quadratischen Pixeln** (square pixel), falls nicht, von **nicht-quadratischen Pixeln** (nonsquare Pixel).

Storage-, Display- und Pixel-Aspect Ratio spielen dann eine Rolle, wenn zB. **anamorph** aufgezeichnet wird wie bei DV-HDTV (1440 x 1080). Dabei wird in **4:3 anamorph aufgezeichnet (SAR)** und bei der Wiedergabe (DAR) auf 16:9 entzerrt. Dabei ist das Pixel-Seitenverhältnis (PAR) 4:3, also **nicht-quadratisch**. Ein SAR von 1:1 entspricht dem DAR und somit quadratischen Pixeln.

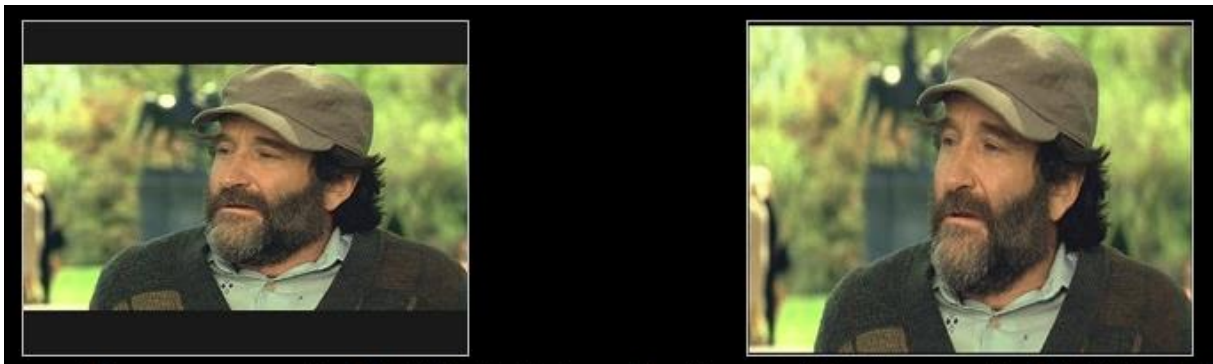


Abbildung 2: Nicht-Anamorph und Anamorph. Links ist die entzerrte Version.

Seitenverhältnis (DAR)	Beschreibung
4:3	35 mm und 16 mm Film, SD-TV
1,43:1	IMAX 70 mm Film
3:2	Fotografie
14:9	Werbung, Kompromiss zwischen 4:3 und 16:9
16:9	Video und HDTV-Standard
2,2:1	70 mm
2:1	50er Jahre Standard Amerika, RED One
2,39:1	Panavision
2,76:1	Ultra Panavision

SAFE-AREAS BEI 16:9 PRODUKTIONEN

In der Videosoftware bzw. an der Kamera können sogenannte **sichere Bereiche** eingeblendet werden. Diese sind vor allem bei der **Aufnahme wichtig** und definieren Pufferzonen bei der Produktion.



Abbildung 3: Safe-Area bei Premiere

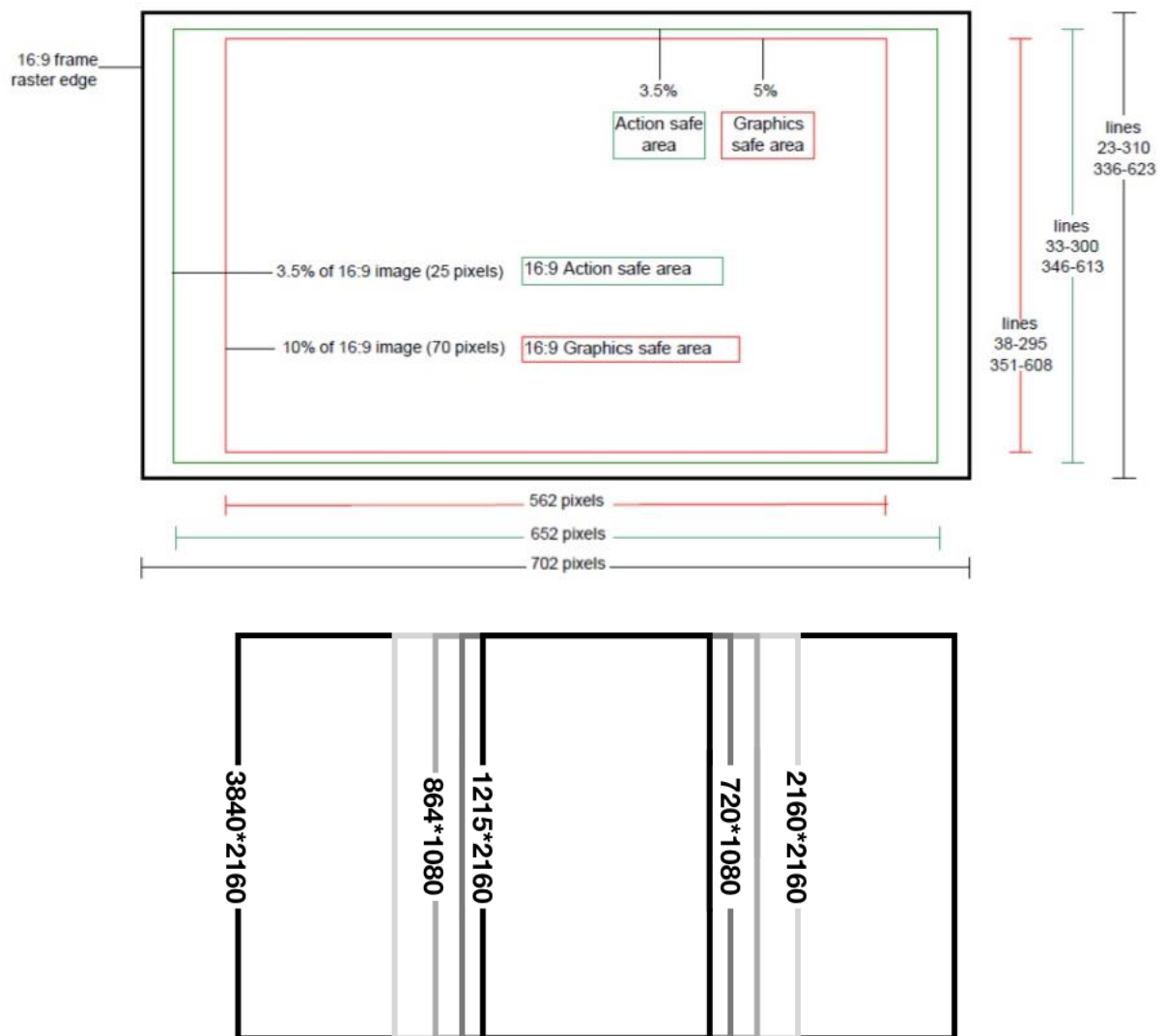


Abbildung 4: Safe Areas für vertical video in UHD

BILDVERSCHMELZUNGSGRENZE UND BILDWIEDERHOLFREQUENZ

Die **Bildverschmelzungsfrequenz** unseres Auges (Sehsinn) liegt bei ungefähr 16 Bildern pro Sekunde (16 Hz). Für Laufbildmedien wurden folgende Werte gewählt:

- 24 Bilder/Sek weltweit für **Film**
- 25 Bilder/Sek in Europa; 30 Bilder/Sek in USA/Japan für **Fernsehen**

Der Unterschied zwischen Europa und USA/Japan liegt an der historischen Verkoppelung mit der jeweiligen Frequenz des Wechselstromnetzes (Europa 50 Hz, USA/Japan 60 Hz).

Die **Bildverschmelzungsfrequenz** ist jene Frequenz, ab welcher eine abgebildete Bewegung als fortlaufend und zusammengehörig empfunden wird.

Die **Bildwiederholfrequenz** ist die tatsächliche Anzahl an Bildern pro Zeiteinheit.

Bildwiederholrate **Empfehlungen:**

- TV 50i
- Film look 24p, 25PsF, evtl. 48/50p
- Kino 24p
- Sport/Action 50p

In Deutschland wurde der erste Kinofilm 2012 mit 48p veröffentlicht (Der Hobbit). Der erste **HFR**⁴-3D-Projektor mit 60 Bildern pro Sekunde wurde im Jahre 2013 im österreichischen Cinepoint in Tirol installiert.

DCI Spezifikationen

Video:

- 2048×1080 2K bei 24 oder 48 Hz und
- 4096×2160 4K bei 24 Hz; 12 bit Farbtiefe im XYZ-Farbraum.
- ISO 15444 (Motion JPEG 2000) mit 250 Mbit/s Datenrate

Audio:

- Bis zu 16 Kanäle
- 24 bit linear PCM, 48 oder 96 kHz Abtastrate, unkomprimiert

Die gesamte Datenrate (Bild und Ton) beträgt damit 113 bis 129 GByte/h (112,5 GByte/h für Bild, 0,52 bis 16,59 GByte/h für Ton).

⁴ High Frame Rate

ZEILENSPRUNG UND PROGRESSIVE ABTASTUNG

Jedes Bild erzeugt ein Vollbild im Fall von **progressive scan**. Die Abtastung beginnt an der oberen linken Ecke des Bildes und endet an der rechten unteren Ecke.

Alternativ zum progressive scan können **bei der Aufnahme zwei Bilder zu zwei verschiedenen Zeitpunkten** abgetastet werden, wobei die Zeilen derart verschachtelt sind, dass **geradzahlige und ungeradzahlige Zeilen** jeweils ein eigenes **sogenanntes Feld (engl. field) bilden**. Dies wird als **unteres bzw. oberes Halbbild** (upper field, lower field) bezeichnet. Diese Abtastung nennt man **interlaced (= Zeilensprungverfahren)**. Dabei entstehen ein zeitlicher Verschluss und eine höhere zeitliche Auflösung.

Bei der Wiedergabe wird das interlace-Material mittels **de-interlacing** in Vollbilder umgewandelt.

Achtung: Bei **50i** werden **50 Halbbilder** abgetastet, die Datenmenge und Bildwiederholffrequenz entsprechen technisch jedoch **25p**, weshalb auch **50i-Footage** eine Bildwiederholffrequenz von **25 Bildern** anzeigt (siehe zB. Premiere).



Abbildung 5: Interlaced und Progressive

Bei **PsF⁵** speichert man **Vollbildmaterial als Halbbilder**. Jedes Vollbild wird dabei in **zwei Halbbildern** ohne zeitlichen Verschluss der einzelnen Zeilen **innerhalb eines Halbbildes** geteilt.

Das HDTV-Format **1080i50** wird dann im **PsF-Verfahren** übertragen, wenn progressives Bildmaterial mit 25 Vollbildern pro Sekunde vorliegt und gesendet wird. Die Vorteile durch die höhere zeitliche Auflösung bei interlace-Abtastung **entfallen jedoch**.

⁵ Progressive Segmented Frame

Progressive Abtastung:

- Film-Look
- Für Nachbearbeitung (Post) geeignet
- Geringe Bewegungsunschärfe
- Ortsauflösung höher (spatial resolution)

Interlace Abtastung

- Zeitliche Auflösung höher (temporal resolution)
- Video-Look
- Für Broadcast/TV-Workflow geeignet



Abbildung 6: Oberes und unteres Halbbild eines Frames



Abbildung 7: Links: Zwei Halbbilder vereint (50i), Rechts: Progressive Abtastung (25p)

PULLDOWN

Da LCD-Panels nur Vollbilder darstellen können, ist Inverse Telematik und 2:2, bzw. 3:2 Pull-down nötig.

- **Inverse Telematik (IVTC):** Bei Film-Material notwendig, das für den amerikanischen Fernsehstandard NTSC aufbereitet wurde, um das mit Hilfe von Halbbildern auf 29,97 Hz gestreckte Material wieder auf 23,976⁶ Hz zurückzuführen.
- **3:2-Pull-down** ist ein Verfahren zur Umwandlung eines Filmsignals mit 24p in ein NTSC-Fernsehsignal mit 29,97 Hz. Dabei wird das Material um 0,1 % verlangsamt.
- **2:2-Pull-down** bezeichnet die Verschiebung mehrerer Halbbilder eines Videomaterials um ein Halbbild nach vorn oder nach hinten. Normalerweise besteht ein Einzelbild aus einem ungeraden und einem geraden Halbbild, die bei progressivem Bildmaterial wieder ein Vollbild ergeben. Es kommt jedoch häufig vor, dass diese Reihenfolge um ein Halbbild verschoben wird. Mit dem 2:2-Pull-down ist es möglich, diese Verschiebung wieder zu korrigieren und somit wieder ein progressives Bildmaterial zu erhalten.

Es lassen sich je zwei Halbbilder durch Zwischenspeichern des Signals und Umsortieren der Zeilen nahtlos zu einem Vollbild zusammenfügen. Ein Kinofilm wird somit bei PAL mit 25 Vollbildern dargestellt. Dabei ist anzumerken, dass dadurch der Film etwa um 4,166 Prozent beschleunigt wird, im Gegensatz zum 3:2-Pull-down bei NTSC (PAL Speedup).

Obige Maßnahmen erfolgen in der Schnittsoftware, viel öfters aber im Endgerät/Player.

BILDRAUM UND OBJEKTRAUM

Zur Wiederholung seien an dieser Stelle Bildraum und Objektraum in der Übersicht dargestellt.

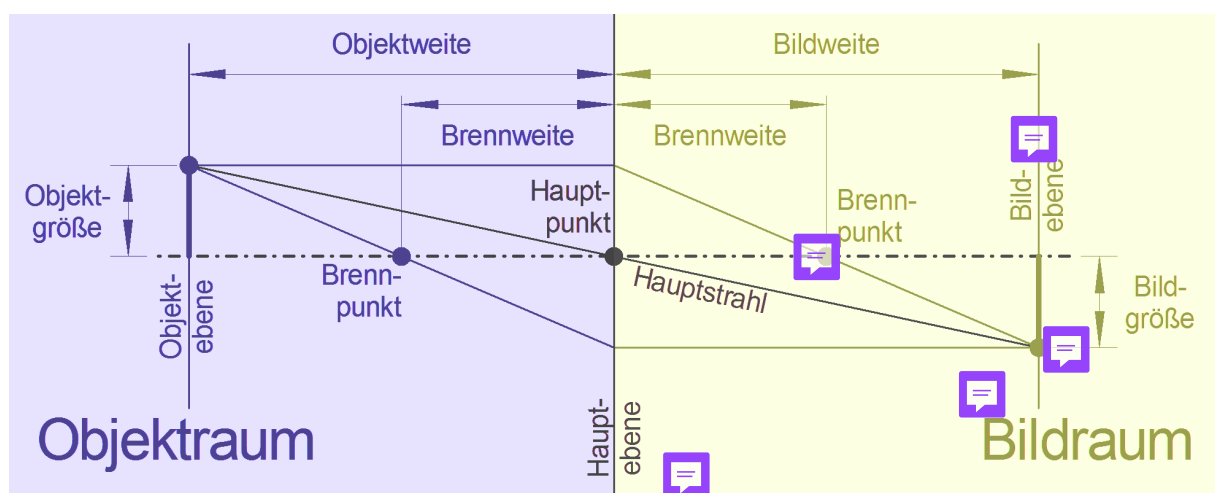


Abbildung 8: Bildraum und Objektraum

⁶ Das NTSC-Äquivalent zu 24p

POLARISATION

Das Licht als **elektromagnetische Welle** ist gleichzeitig eine sogenannte **Transversalwelle**, deren **Amplitude senkrecht zur Ausbreitungsrichtung** steht.

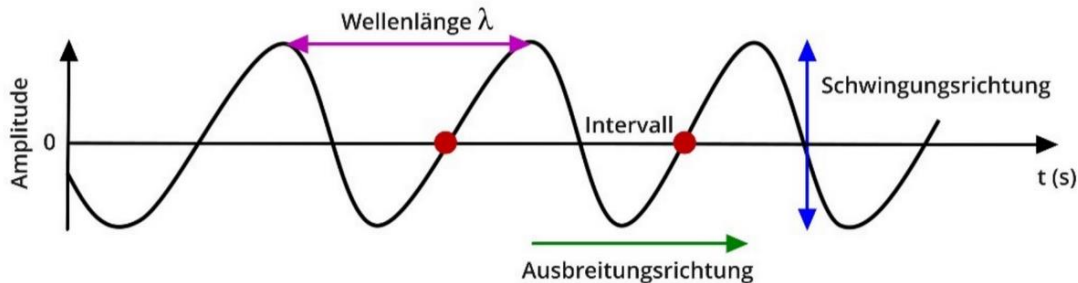


Abbildung 9: Schema Transversalwelle

Dabei beschreibt die **Polarisation** die Richtung der Schwingung. Wenn aus einem Mix verschiedener Polarisationsrichtungen eine bestimmte Richtung herausgefiltert wird, entsteht polarisiertes Licht. Polarisation spielt beim **Polfilter** und **3D-Projektion** (Strahlteilung für **rechte/linke** Augen) eine Rolle in der Videotechnik.

ORTSFREQUENZ

Die **Ortsfrequenz f_r** bezeichnet in der **Wahrnehmungspsychologie** die Anzahl der **Linienpaare bzw. Perioden pro Grad Sehwinkel**.

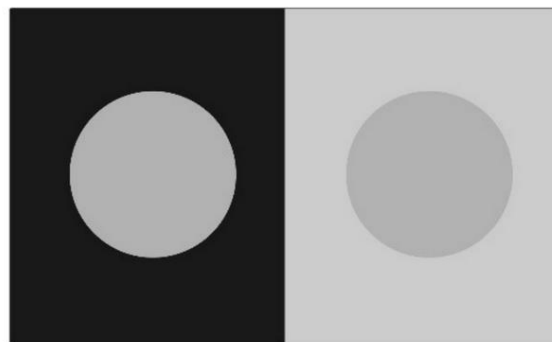


Abbildung 10: Kontrast **abhängig von Leuchtdichte**

Bei einer **optimalen** Funktionalität des Auges kann eine Rasterung von bis zu **neun Linienpaaren pro Grad Sehwinkel** erkannt werden. Steigt die Frequenz sinkt der empfundene Kontrast proportional.

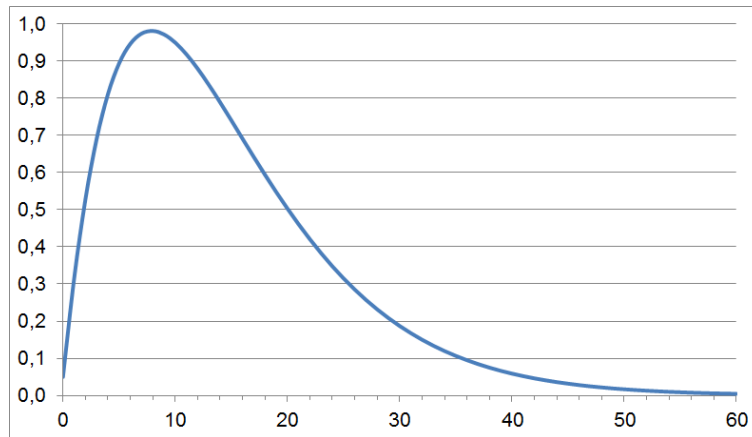


Abbildung 11: Kontrastempfindlichkeit Y vs. Ortsfrequenz in Linienpaare pro Grad X

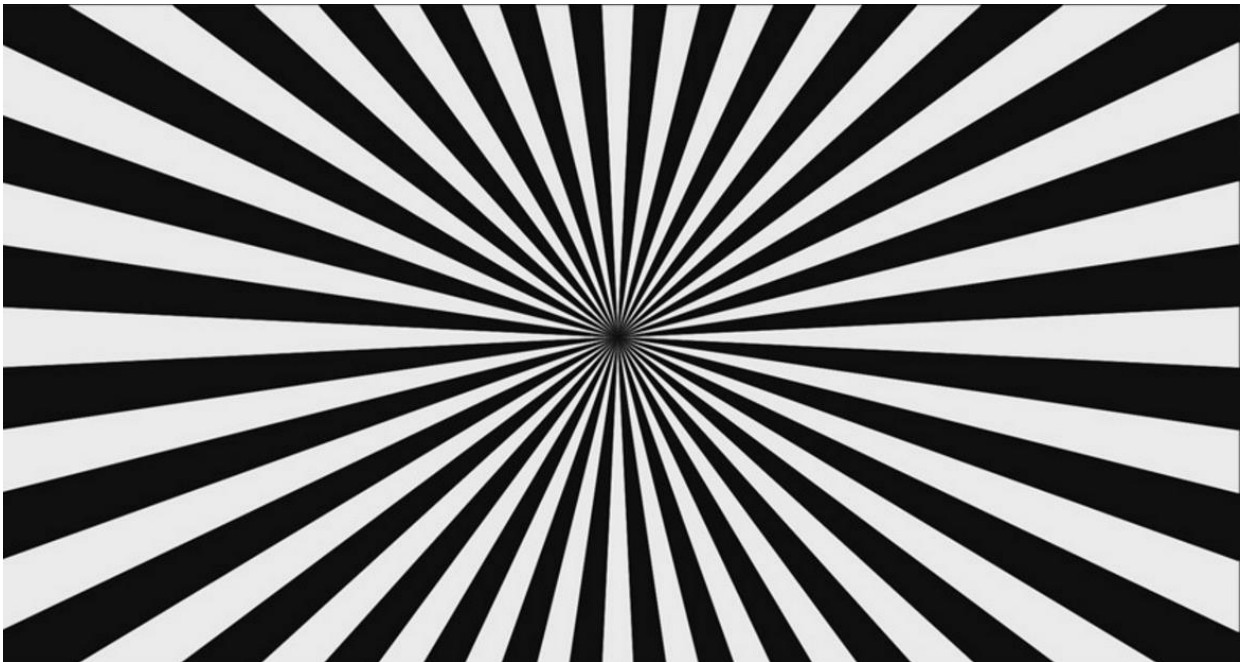


Abbildung 12: Siemensstern

Der dargestellte Siemensstern verfügt über eine maximale Ortsfrequenz von zirka 330 Linienpaaren pro Bildhöhe (FHD). Die Details in der Mitte des Bildes gehen dabei komplett verloren, während nach außen hin die Kanten schärfer wirken. Der Siemensstern setzt das menschliche Wahrnehmungsvermögen mit der optischen Abbildungsleistung in Relation (zB. Objektive).

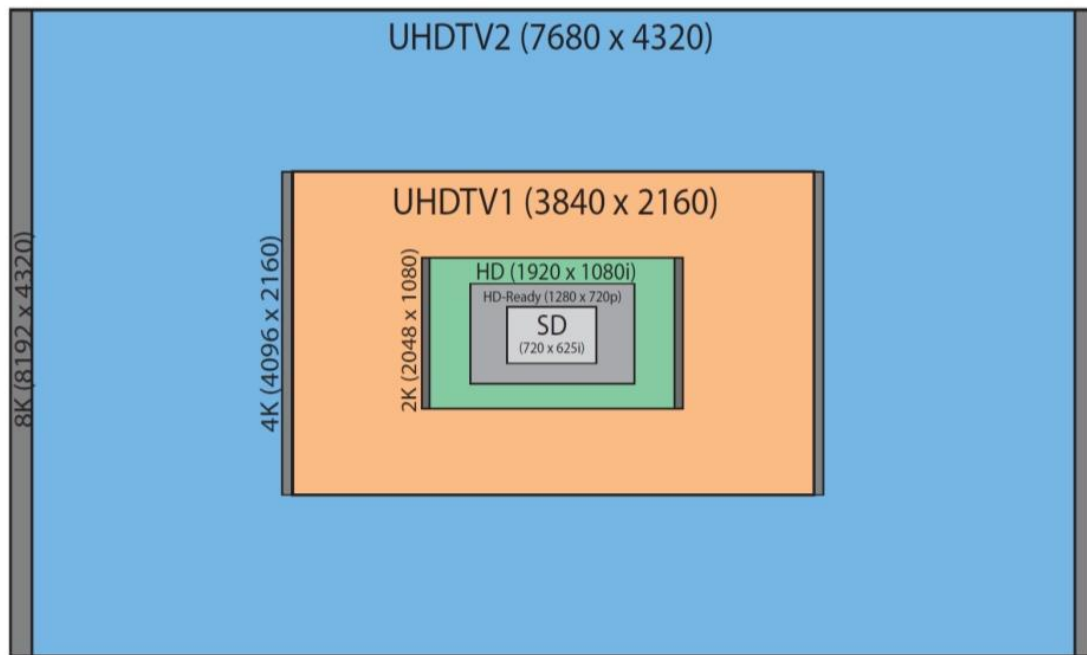
Eine Modulation mit geringer Ortsfrequenz entspricht also groben Strukturen im Bild, und eine Modulation mit großer Ortsfrequenz entspricht feinen Strukturen im Bild.

Das Verhältnis der Modulation im Bildraum zur Modulation im Objektraum in Abhängigkeit von der Ortsfrequenz wird **Modulationsübertragungsfunktion** $T(f_r)$ (englisch: modulation transfer function = MTF) oder auch **Kontrastübertragungsfunktion** genannt.

$$T(f_r) = \frac{\text{Bildraum } f_r}{\text{Objektraum } f_r}$$

BILDPUNKTE

Die Anzahl der Pixel wurde seit dem SD-Standard sukzessive erhöht. Meist wurde dabei die **Anzahl der Zeilen verdoppelt**. Die Anzahl der Spalten ergibt sich dann aus dem Seitenverhältnis.



Aufgrund der **angularen Sehschärfe**⁷ des menschlichen Sehsinns, welche **etwa 1 Bogenminute im Durchschnitt** ($1' = 1/60^\circ = 0,01666\dots^\circ$) beträgt, kann für entsprechende Bilddiagonalen ein optimaler Betrachtungsabstand ermittelt werden.

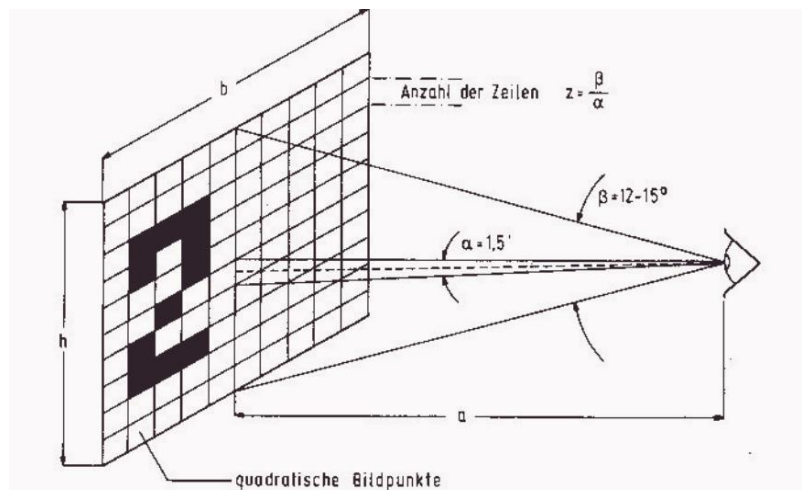


Abbildung 13: Ermittlung der notwendigen Zeilenauflösung und Betrachtungsabstand

⁷ Auflösungsvermögen, bei dem zwei Sehobjekte noch als getrennt wahrgenommen werden

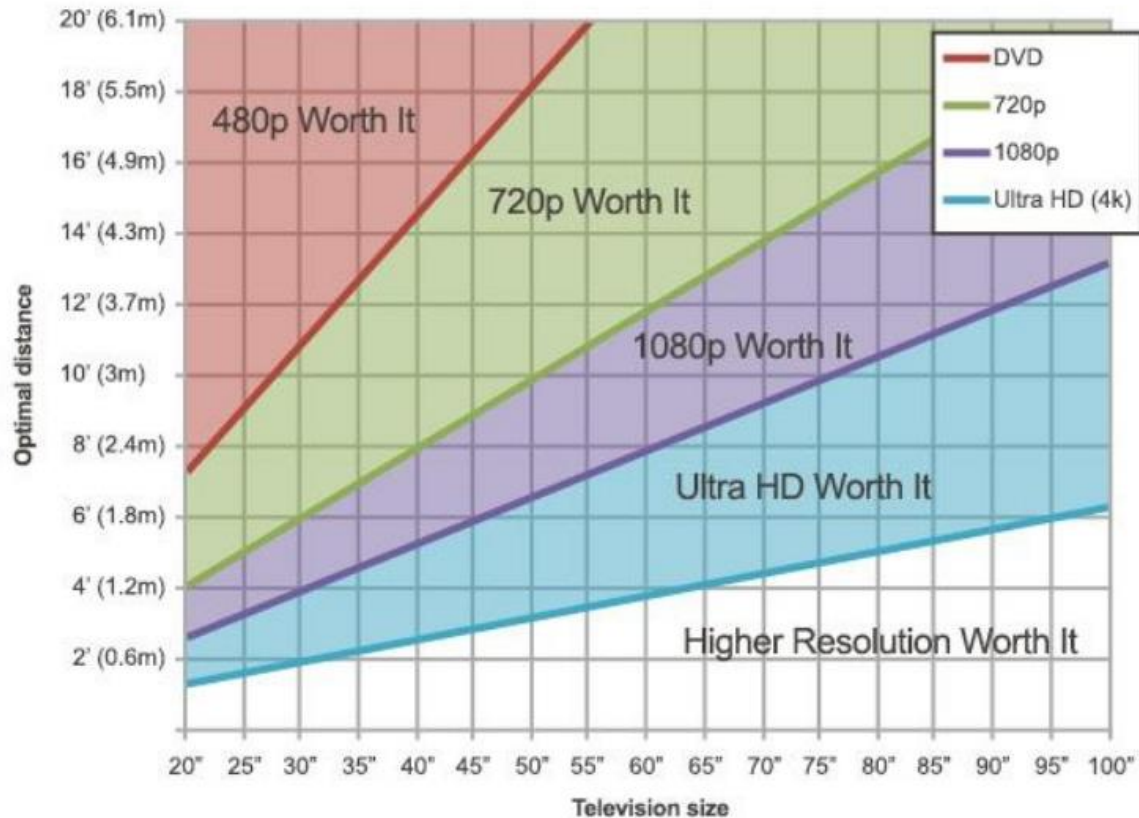


Abbildung 14: Betrachtungsabstand, Videoformat und Panelgröße in Relation

EINFLUSSFAKTOREN AUF BILDQUALITÄT IM VIDEO/TV

Für hochqualitative Video- und TV-Produktionen sind neben der Kamera also **technisch folgende Parameter relevant**:

- **Anzahl Bildpunkte** (HD, FHD, UHD...)
- **Bildwiederholfrequenz** (24, 25, 50, ...)
- **Dynamikumfang** (SDR, HDR)
- **Farbraum** (BT.709, BT.2020...)
- **Abtastung progressiv oder Zeilensprung**
- **Quantisierungstiefe** (8 bit, 10 bit, 12 bit)

ABBILDUNGSFEHLER

Diese können am Aufnahmegerät, der Verarbeitung oder am Wiedergabegerät liegen.

- **Klötzchenbildung/Artefakte**: Codierungs und Übertragungsfehler
- **Banding/Clouding**: Treppeneffekte bei Verläufen.
- **Blooming**: Dunkle Bildbereiche sind zu hell.
- **Crosstalk/Ghosting**: Nur bei der 3D Wiedergabe.
- **Soap Effekt**: Bei der Generierung von Zwischenbildern wird ein Effekt forciert, der an billige Studioaufnahmen erinnert (soap opera).
- **Halo Effekt**: Nachzieheffekt oder helle Säume an Kanten.

GAMMAKORREKTUR DER RGB-SIGNALE

Aus den RGB-Signalen vom Sensor wird mittels Gammakorrektur ein R'G'B'-Signal erzeugt. Die **Gammakorrektur** ist nötig, damit die linear anwachsenden Größen die der Sensor liefert in ein dem menschlichen Sehsinn entsprechendes, nicht-lineares Signal konvertiert wird.

$$\text{Ausgangsgröße} = \text{Eingangsgröße}^{\frac{1}{\gamma}}$$

Die menschliche Wahrnehmung ist anatomisch so, dass dunklere Tonwerte heller und helle Tonwerte weniger hell wahrgenommen werden. Die Empfindlichkeit von Bildwandlern ist jedoch in den Lichtern höher als in den Tiefen, weshalb nach der Gammakorrektur mehr Tonwerte für die Tiefen zur Verfügung stehen. Bei linearer Kodierung müssten mindestens 11 bit zur Verfügung stehen, damit keine Treppeneffekte (Posterisierung) sichtbar sind, durch gammakorrigierte Kodierung reichen 8 bit.

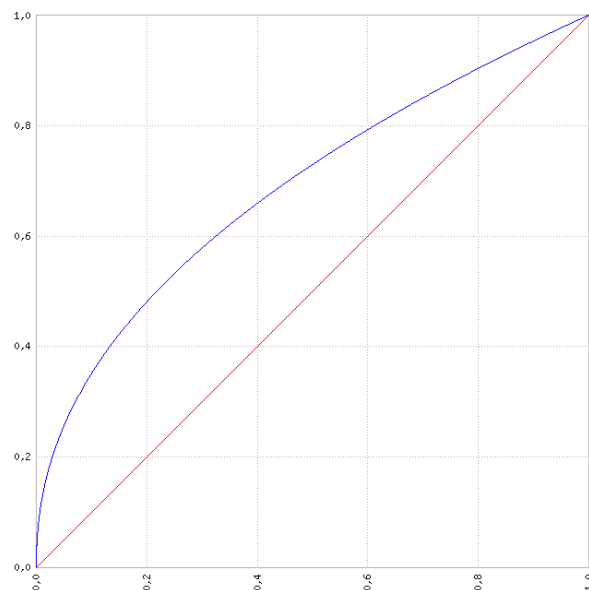
Originalszene



Linear kodiert



Gammakorrigiert



Gammakurve (blau) für ein Gamma von 2,2; Lineare Werte des Bildwandlers (rot)

$$f(x) = x^{\frac{1}{2.2}}$$



Abbildung 15: Links linear und rechts mit Gamma 2,2

DIGITALE VIDEOSIGNALE

Das analoge **Y'P'bP'r-Signal** wird mit der entsprechenden **Abtastfrequenz digitalisiert** und **serialisiert**.

FARBKONVERTIERUNG

Aus den analogen R'G'B'-Informationen des Sensors wird ein analoges Y'P'bP'r- und ein **digitales Y'CbCr-Signal** erzeugt. Bei der **Umwandlung von R'G'B' nach Y'CbCr** wird folgendes Schema verwendet, abgeleitet von der Helligkeitsempfindlichkeit des menschlichen Sehsinns:

$$Y' = a * R' + b * G' + c * B'$$

$$Cb = \frac{B' - Y'}{d}$$

$$Cr = \frac{R' - Y'}{e}$$

	BT.601	BT.709	BT.2020
a	0.299	0.2126	0.2627
b	0.587	0.7152	0.6780
c	0.114	0.0722	0.0593
d	1.772	1.8556	1.8814
e	1.402	1.5748	1.4747

Dabei zu berücksichtigen sind bei der Digitalisierung **15 Tonwerte Foot- und Headroom für Cb und Cr** und ein **Tonwertbereich von 16 – 235 für Y'**.

Die **Signale Y' bzw. C'r und C'b** nehmen einen Pegelbereich von jeweils **700 mV** ein.

KOMPRIMIERUNG

Bei der digitalen Kodierung wird meist eine Komprimierung zur **Reduktion der Datenmenge** angewandt. Dabei unterscheidet man zwischen **verlustfreier** (lossless) und **verlustbehafteter** (lossy) Komprimierung.

Verlustfreie Kompression bzw. Redundanzreduktion

- **Tokenbasierte Kompression (Wörterbuchmethode):** Redundante Informationen werden durch Tokens ersetzt. Die Tokens müssen in einem Wörterbuch existieren.
- **Entropiekodierung:** zb. Huffman-Code; Häufig verwendete Informationen werden mit kurzen, seltene mit langen Codes abgespeichert.

Verlustbehaftete Komprimierung: Aufgrund der Anatomie des Auges ist die Farbauflösung unserer Wahrnehmung schlechter als die Helligkeitsauflösung. Somit kann die Farbinformation in gewissen Grenzen reduziert werden. Solch eine Methode wird **Chroma-Subsampling** genannt.

Neben der Redundanzreduktion wird bei der verlustbehafteten Komprimierung die **Irrelevanzreduktion** eingesetzt. Man eliminiert dabei Informationen, die nur zu nicht oder kaum wahrnehmbaren Fehlern führen. Die Irrelevanzreduktion wird später im Skript noch behandelt.

CHROMA-SUBSAMPLING

Ein Videosignal kann unterabgetastet werden. Da Helligkeitsunterschiede stärker wahrgenommen werden als Farbunterschiede, kann das **Chrominanzsignal mit einer geringeren Abtastfrequenz digitalisiert werden als das Luminanzsignal**. Als Abtastverhältnis bezeichnet man das Verhältnis der drei Abtastraten untereinander.

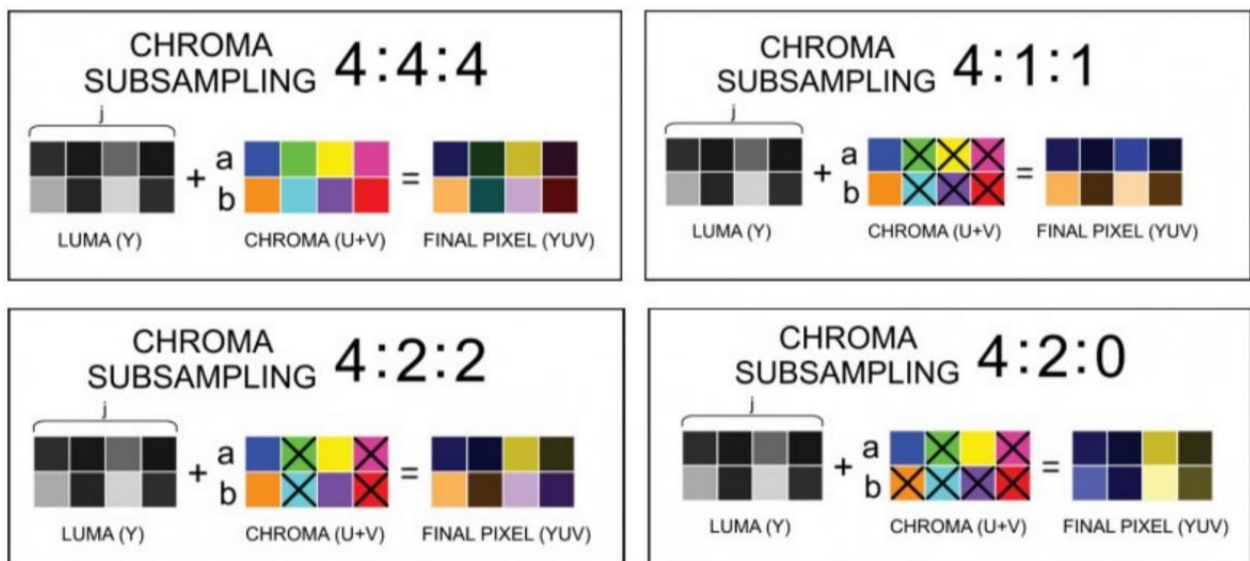


Abbildung 16: Farbunterabtastung

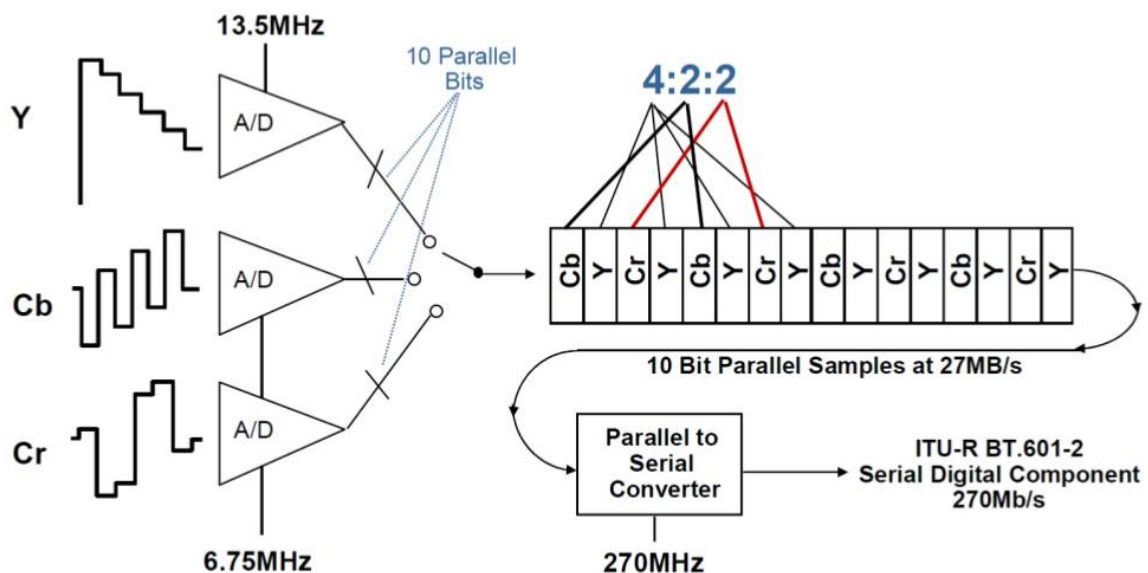


Abbildung 17: 4:4:4, 4:2:2 und 4:2:0 im Vergleich

DAS DIGITALE KOMPONENTENSIGNAL YCBCR

Laut **ITU-Standard ITU-R BT.709** ist definiert:

- **Y**: 74,25 MHz Abtastfrequenz
- **Cr, Cb**: 37,125 MHz
- **Seitenverhältnis**: 16:9
- **Interlaced** und **progressive** und progressive PsF i/p/psf
- **1920 Spalten, 1080 Zeilen bei 60 Hz, 50 Hz, 30 Hz, 25 Hz and 24 Hz** Bildwechselfrequenz
- **1280 Spalten, 720 Zeilen bei 60 Hz, 50 Hz, 30 Hz, 25 Hz and 24 Hz** Bildwechselfrequenz
- **8 oder 10 bit** Quantisierung
- **4:2:2** Subsampling
- Wertebereich: **16 – 235 bei 8 bit**
- Sync-Signal im Headroom



In der obigen Abbildung ist zu erkennen wie die Samples 4:2:2 unterabgetastet und serialisiert werden. Auf 2 Y-Samples kommt nur jeweils 1 Cb- oder 1 Cr-Sample.

SDI SERIAL DIGITAL INTERFACE

SDI wird von der SMPTE spezifiziert und ist eine Weiterentwicklung der analogen PAL- und NTSC-Standards. Über alle **HD-SDI-Verbindungen** können **Bild, Ton und Metadaten** übertragen werden.

- Kodierung: NRZI⁸
- Amplitude: 800 mV

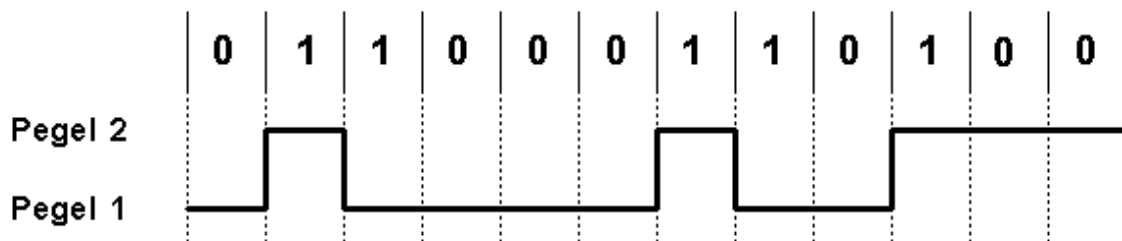


Abbildung 18: NRZI-Kodierung

Standard	Name	Bitraten in Gbit/s	Videoformate
SMPTE 259M	SD-SDI	0,27, 0,36, 0,143 und 0,177	480i (NTSC), 576i (PAL)
SMPTE 292M	HD-SDI	1,485 und 1,485/1,001	720p, 1080i (HDTV)
SMPTE 372M	Dual Link HD-SDI	2,970 und 2,970/1,001	1080p
SMPTE 424M	3G-SDI	2,970 und 2,970/1,001	1080p, 4K (Dual Link)
SMPTE ST 2081	6G UHD-SDI	6	4K
SMPTE ST 2082	12G UHD-SDI	12	4K

Bei HD-SDI werden die **Y bzw. Cb/Cr Daten parallel zu je 10 bit** in einem 20 bit breiten Wort abgebildet. Die Werte 0 bis 3 und 1020 bis 1023 für Signalisierungen und Steuerworte reserviert. Unter anderem wird mit diesen Steuerworten der Beginn bzw. das Ende einer Videozeile signalisiert.

⁸ Non Return to Zero Inverted: Eine 1 bedeutet eine Pegeländerung



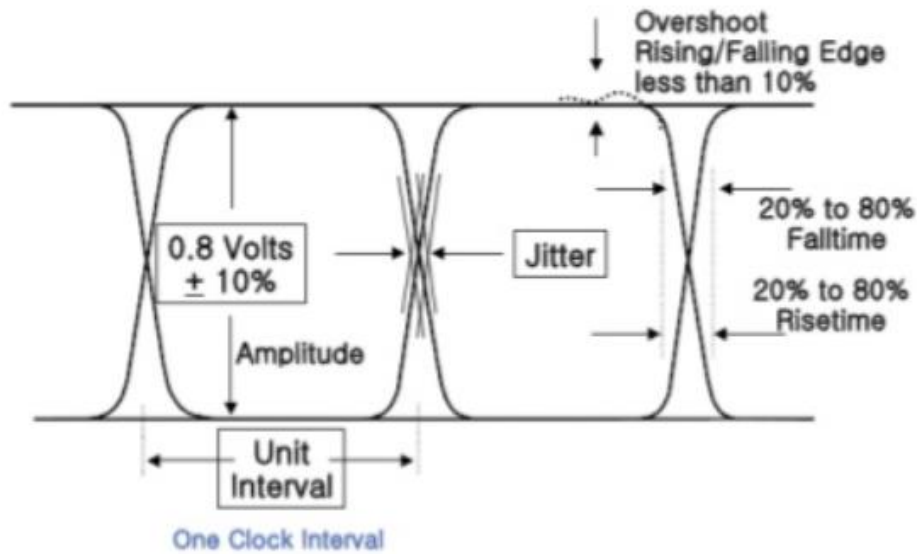
Abbildung 19: TV-Studio Schema

SIGNALQUALITÄT BEI SDI

Bei SDI-Verkabelungen müssen vor allem folgende Parameter des Signals kontrolliert werden:

- Amplitude 800 mV
- Jitter⁹
- Störungen/Noise, Rise- und Falltime an den Flanken
- Frequenzbereich und Intervalldauer
- Overshoot unter 10 %

9 Fluktuation



Die Signalform wird oft als Auge/Eye bezeichnet. Abweichungen davon lassen Rückschlüsse auf Störfaktoren zu. Durch eine **3R-Regeneration** kann die Amplitude, Wellenform und Phasenlage/Timing korrigiert werden.

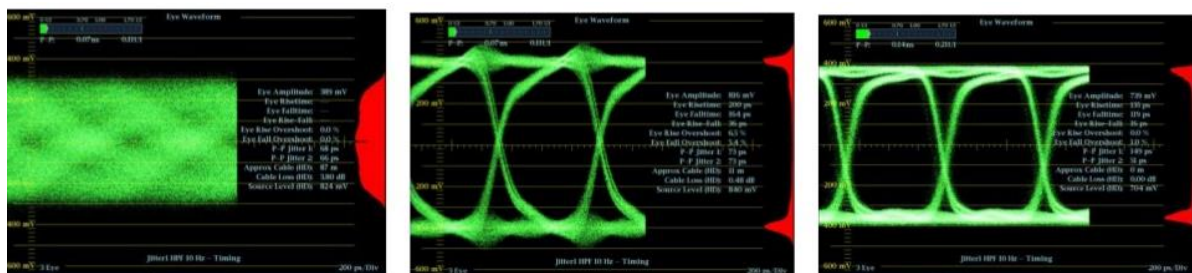


Abbildung 20: Zu langes Kabel, falsch terminiert, Rise- und Fall-Time nicht symmetrisch

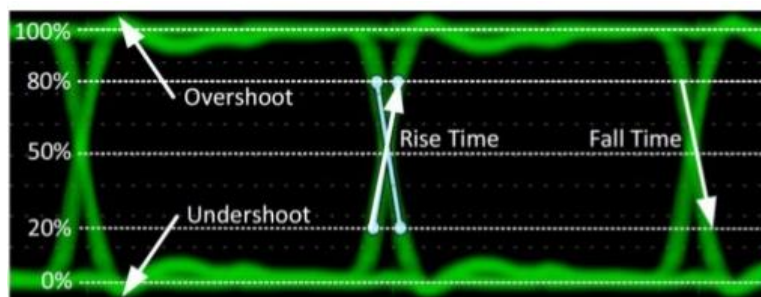


Abbildung 21: Mangelhaftes SDI-Signal

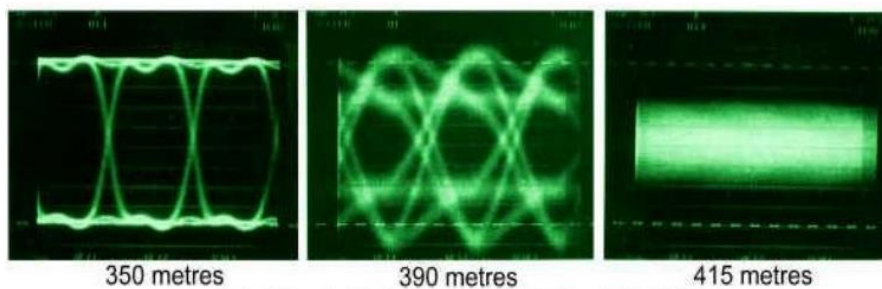


Abbildung 22: Veränderung der Signalqualität bei unterschiedlichen Leitungslängen

HDMI

HDMI ist in verschiedenen **Spezifikationen** im Laufe der Zeit erweitert worden und eine semi-professionelle AV-Schnittstelle. Folgende Übersicht zeigt die wichtigsten Features.

HDMI-Version	1.0	1.1	1.2	1.3	1.4	2.0	2.0a	2.1
sRGB	✓	✓	✓	✓	✓	✓	✓	✓
YCbCr	✓	✓	✓	✓	✓	✓	✓	✓
8-Kanal-PCM, 192 kHz, 24-bit	✓	✓	✓	✓	✓	✓	✓	✓
CEC ¹⁰ (Steuerfunktionen wie Play, Standby...)	✓	✓	✓	✓	✓	✓	✓	✓
Deep Color (30-bit-Farbtiefe)	✗	✗	✗	✓	✓	✓	✓	✓
Auto Audio to video synchronization lip-sync	✗	✗	✗	✓	✓	✓	✓	✓
Dolby-TrueHD-Bitstream-Fähigkeit	✗	✗	✗	✓	✓	✓	✓	✓
DTS-HD-Master-Audio-Bitstream-Fähigkeit	✗	✗	✗	✓	✓	✓	✓	✓
3D	✗	✗	✗	✓	✓	✓	✓	✓
1920 × 1080 (1080p), 120 Hz	✗	✗	✗	✗	✓	✓	✓	✓
Ethernet channel	✗	✗	✗	✗	✓	✓	✓	✓
Audio return channel (ARC)	✗	✗	✗	✗	✓	✓	✓	✓
4K	✗	✗	✗	✗	✓	✓	✓	✓
4K 50/60p	✗	✗	✗	✗	✗	✓	✓	✓
4K 3D	✗	✗	✗	✗	✗	✓	✓	✓
32-Kanal-Audio-Support	✗	✗	✗	✗	✗	✓	✓	✓
1536-kHz-Audio-Support	✗	✗	✗	✗	✗	✓	✓	✓
1080p 3D HFR	✗	✗	✗	✗	✗	✓	✓	✓
21:9 Cinemascope support	✗	✗	✗	✗	✗	✓	✓	✓
Multi Stream Audio und Video	✗	✗	✗	✗	✗	✓	✓	✓
High Dynamic Range	✗	✗	✗	✗	✗	✗	✓	✓
4K 3D 50/60p support	✗	✗	✗	✗	✗	✗	✗	✓
Unterstützung von 14 und 16 Bit Farbtiefe	✗	✗	✗	✗	✗	✗	✗	✓

¹⁰ CEC: Consumer Electronics Control

STECKER UND KABEL

An dieser Stelle ist es sinnvoll, eine **Unterteilung in Stecker und Kabel** zu treffen. Da dieses Thema bereits im 2. Jahrgang besprochen wurde, sind die Kabel und Stecker hier nur auf Basis ihrer Relevanz für die Videotechnik dargestellt.

BNC

Der Stecker verfügt über einen **Bajonettverschluss**, der verhindert, dass sich Stecker und Buchse versehentlich voneinander lösen und die Störanfälligkeit erheblich reduziert. In der Videotechnik werden hauptsächlich zum **Kabelstandard RG 59** passende UG 88 U-C59 Stecker verbaut. Die Stecker werden gelötet oder gecrimpt.



Abbildung 23: BNC-Stecker

CINCH

In der Videotechnik werden häufig folgende farbige **Kodierungen für die Stecker** verwendet:

- **Rot:** Rechter Kanal Audio
- **Weiß:** Linker Kanal Audio
- **Gelb:** Composite Video
- **Grün:** Y Component Video
- **Blau:** Pb Component Video
- **Rot:** Pr Component Video



Abbildung 24: Cinch-Stecker

DIN-STECKER

4polige Mini-DIN-Stecker werden zur Übertragung von **Y/C-Videosignalen** genutzt.



Abbildung 25: Mini-DIN-Stecker 4polig

HDMI

HDMI-Stecker werden bei semiprofessionellen Verkabelungen genutzt.



Abbildung 26: HDMI Typ A, Typ C (Mini), Typ D (Micro)

KOAXIALKABEL

Im Bereich der digitalen Videoanwendungen ist die SDI oder HD-SDI Technik weit verbreitet. Beide Varianten verwenden **75 Ω Koaxialkabel** zur Signalübertragung.

Koaxialkabel sind zweipolige Kabel mit konzentrischem Aufbau. Sie bestehen aus einem Innenleiter, der in konstantem Abstand von einem hohlzylindrischen Außenleiter umgeben ist. Der Außenleiter schirmt den Innenleiter vor Störstrahlung ab.

In der Studio- und Videotechnik kommen Kabel mit 75 Ω Wellenwiderstand/Impedanz zum Einsatz. Die Kabel für die Videotechnik werden als **RG¹¹ 59 Kabel** bezeichnet.



Die maximale Kabellänge hängt von der Qualität des Koaxkabels ab. Bei RG 59 Kabeln liegt die **maximale Länge bei einem HD-SDI-Signal bei ca. 100 m**.

Für **3G-SDI** betragen die erreichbaren Längen ca. 70 % der HD-SDI Werte, für 6G-SDI ca. 50 % der HD-SDI Werte. Mit speziell konstruierten Kabeln sind jedoch auch Strecken von 350 m bei HD-SDI erreichbar.

Ein nicht zu unterschätzender Faktor ist die **temperaturabhängige Dämpfung des Signals**. Ein RG 59 Kabel zeigt zB. bei 800 MHz eine Dämpfung von 34 dB bei 20 °C. Bei höheren Temperaturen/Frequenzen nimmt auch die Dämpfung zu. Bei **Signalabzweigern** muss auf die unterschiedliche Dämpfung der Tap- und Signal-Ausgänge geachtet werden.

		HD				2K		
Signal		HD-SDI (HD)				3G-SDI (2K)		
based on		-20dB @ 1/2 clock				-20dB @ 1/2 clock		
Clock		1.5 GHz				3 GHz		
Nyquist Limit		750 MHz				1.5 GHz		
Data Rate		1.5 Gb/s				3 Gb/s		
Specification		SMPTE 292M (SMPTE 372M)				SMPTE 424M		
Application		720p-1080i (Dual Link 1080p)				1080p		
Part No.	Type	ft.*	m*	m	m	ft.*	m*	m
179DT	0.3/1.4	115	35	40		81	25	
1855A	0.6/2.6	215	66	80		149	45	78
1855x/70080x	0.6/2.8	284	87	100		170	52	
1505x/70081x	0.8/3.7	304	93	120		210	64	
1694x/70082x	1.0/4.6	371	113	140	230	256	78	140
1794x/70083x	1.3/5.7	469	143	200*		320	98	
7731x/70084x	1.6/7.2	580	177	230		393	120	188

*calculated

Abbildung 27: Max. Kabellängen eines Herstellers für professionelle Koaxkabel

MESSTECHNIK

Grundsätzlich stehen der **Wellenformmonitor** (Software, zB. Premiere oder Hardware-Oszilloskop) und das **Vektorskop** zur Verfügung. Das Histogramm wird eher in der Fototechnik genutzt, da es sich weniger für Bewegtbilder eignet.



Abbildung 28: Leader LV5490

DER WELLENFORMMONITOR

Der **Wellenform-Monitor** ist ein Kontrollmittel für das Videosignal. Er zeigt die Amplituden der aufgenommenen Bildpunkte und ermöglicht eine **Kontrolle der Belichtung bzw. Überbelichtung/Unterbelichtung**.

Die **Y-Achse** ist entweder in mV oder IRE¹² angezeigt. IRE ist eine Pegelangabe für analoge Fernsehsignale. 0 - 100 IRE ist der Signalhub zwischen schwarz und weiß. Pegel unter 0 IRE markieren den Synchronimpuls.

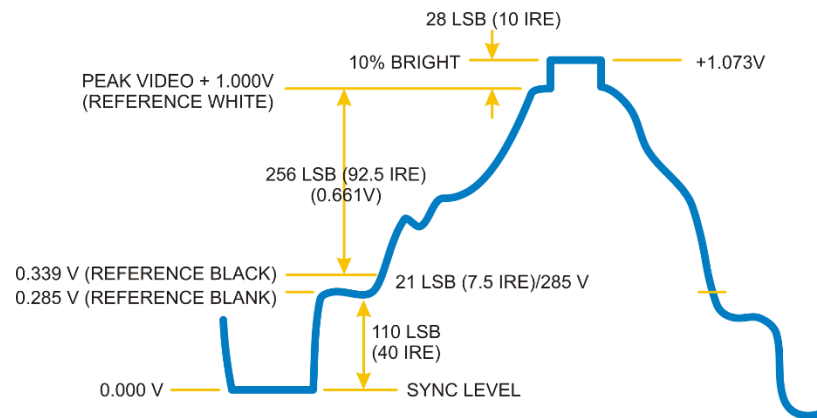


Abbildung 29: IRE

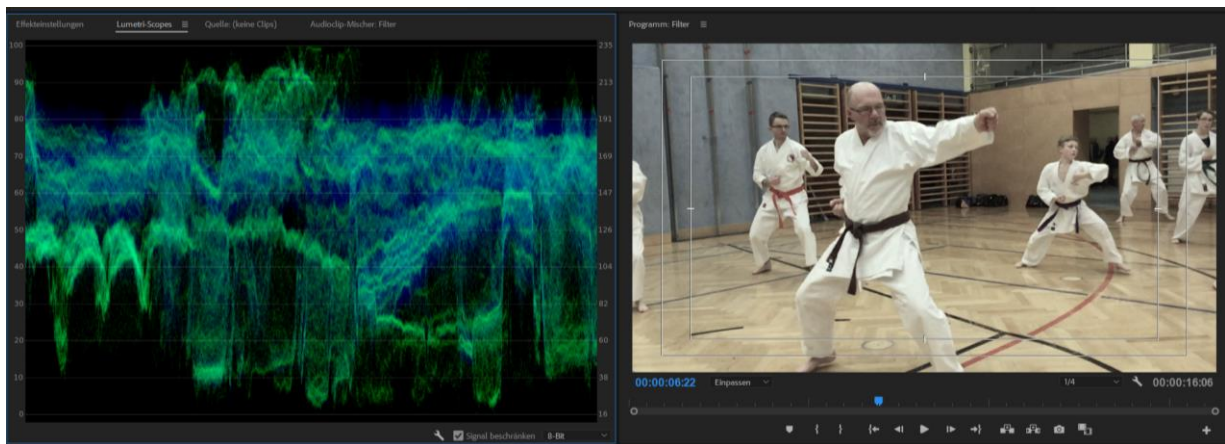


Abbildung 30: Wellenformmonitor-Ansicht (Lumetri)

¹² Institute of Radio Engineers

EBU-FARBALKEN ALS TESTSIGNAL IM WELLENFORMMONITOR

Wie bereits beim RGB-Signal erwähnt, werden **Farbbalken als gängige professionelle Testsignale** eingesetzt. Üblicherweise werden diese durch Testgeneratoren erzeugt.

- 8 Balken (Weiß, Schwarz und 6 Primärfarben YCGMRB)
- Weiß links, Schwarz rechts
- Nützlich um Verbindungen, Farbqualität, Farbkonsistenz, Pegel usw. zu testen
- 100/100 oder 100/75 Balken; 100/75 Farbbalken besteht aus 100 % Weiß und 75 % ausgesteuerten Farbbalken

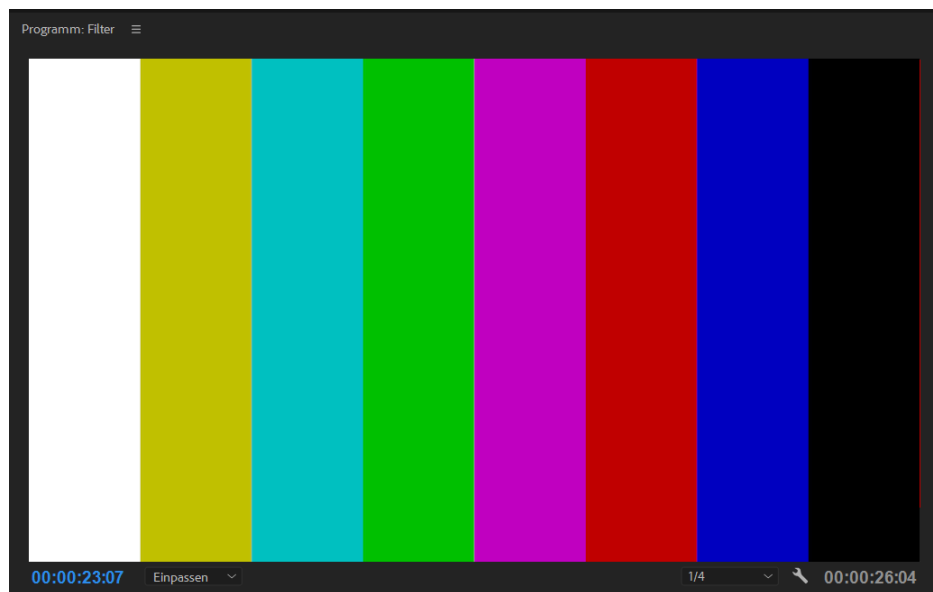


Abbildung 31: EBU 100/75 Farbbalken in Premiere

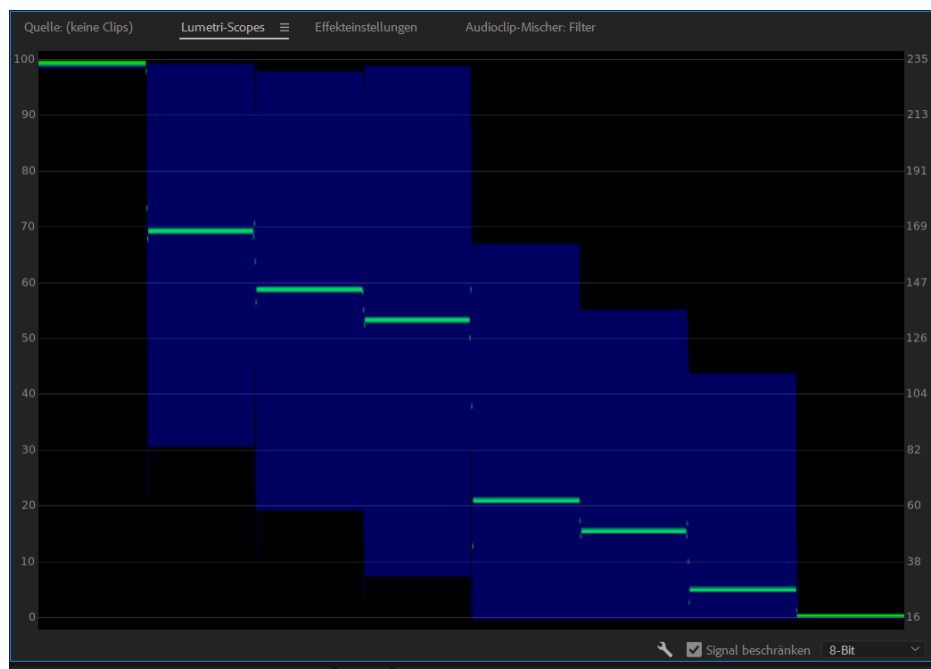


Abbildung 32: EBU 100/75 Farbbalken Y/C im Wellenformmonitor in Premiere

Beachte die Angaben die Y-Achsen in obiger Abbildung.

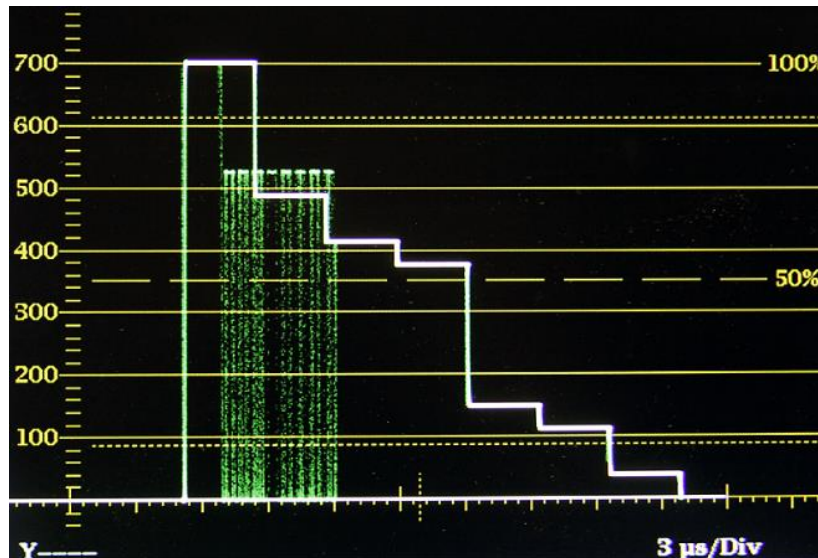


Abbildung 33: EBU 100/75 Farbbalken am Tektronix Oszilloskop

EBU-FARBBALKEN ALS TESTSIGNAL IM VEKTORSKOP

Das Vektorskop stellt die **Sättigung und den Farbton** der Pixel dar.

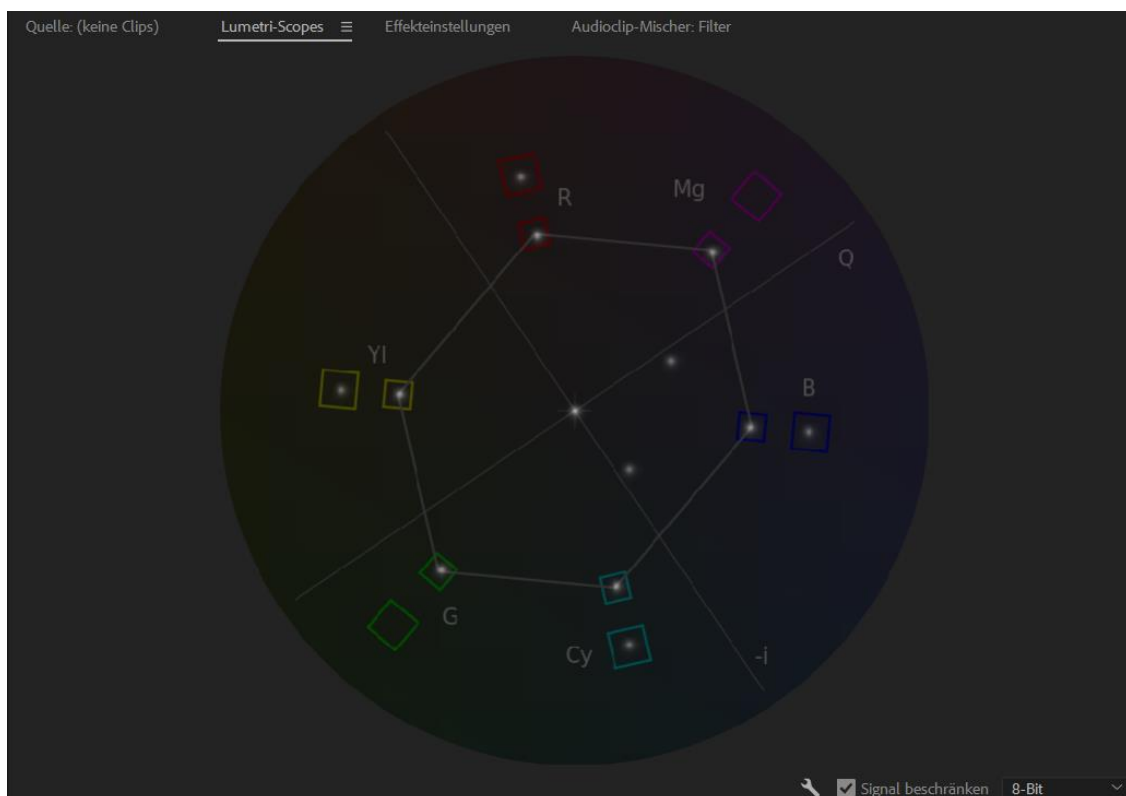


Abbildung 34: Vektorskop mit Premiere-internem HD-Farbbalken (ARIB)

Beachte in der obigen Darstellung, dass nur jene Pegelbereiche in den **durch Linien verbundenen Primärfarbkästchen broadcasttauglich** sind.



Abbildung 35: ARIB Farbbalken. Die markierten Bereiche sind im Vektorskop als überpegelt dargestellt.

HAUTTONKORREKTUR MIT DEM VEKTORSKOP

Anhand der **Hauttonlinie** kann eine Anpassung der Hauttöne durchgeführt werden. Idealerweise crop man das Gesicht und korrigiert dann entsprechend.

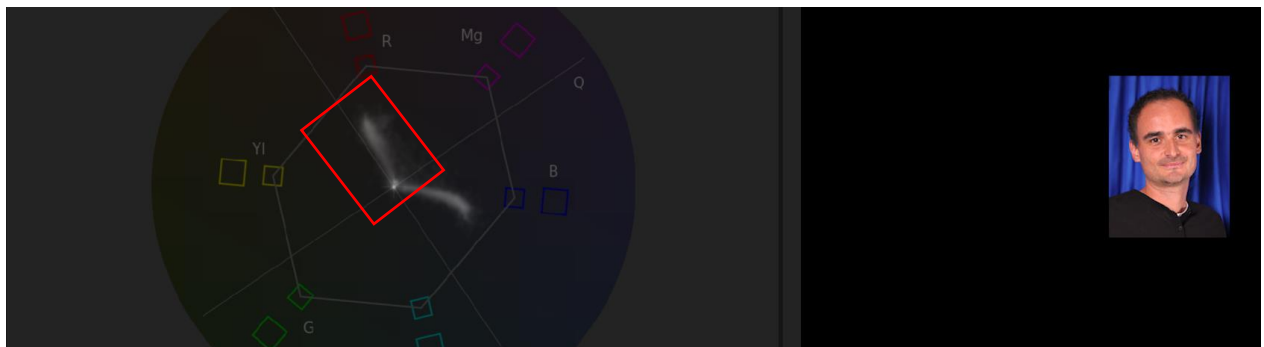


Abbildung 36: Hauttöne weisen Farbstich auf, befinden sich neben der Linie

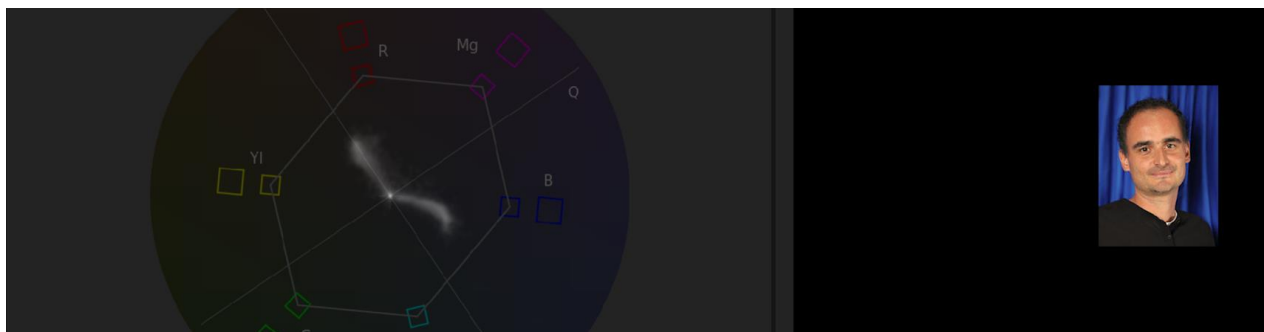


Abbildung 37: Schnellkorrektur des Hauttons

A/V-DATENSTRUKTUREN

Video- und Audiodaten werden in speziellen **Datenformaten** gespeichert. Dabei werden **Elementarstreams** in Containern organisiert. Die Elementarstreams werden mit CODECs kodiert und dekodiert. **Wichtig:** *Unterschiede die Begriffe Video-/Audiodatenformat von Video-/Audioformat.*

Video-/Audio-Datenformate spezifizieren Container und CODECs zur Speicherung. AVI ist zB. ein Video-Datenformat, MP3 ein Audio-Datenformat.

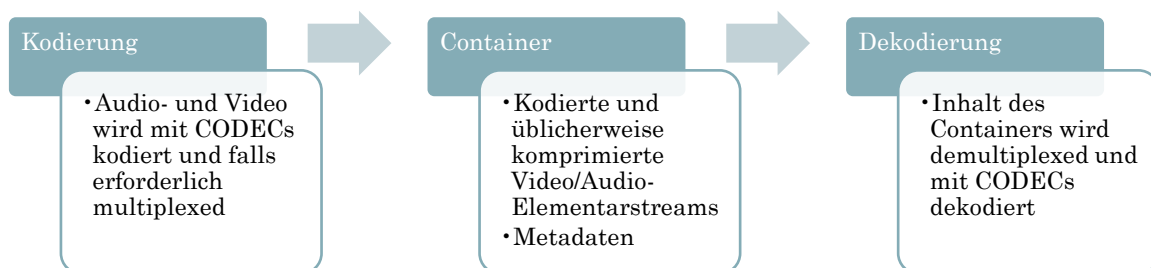
Videoformate spezifizieren folgende Eigenschaften: Seitenverhältnis, Anzahl Pixel, Bildwiederholrate, Quantisierungstief. Dies wurde in den vorigen Kapiteln bereits besprochen.

Audioformate spezifizieren folgende Eigenschaften: Anzahl Kanäle, Quantisierungstiefe, Samplefrequenz.

CODECS

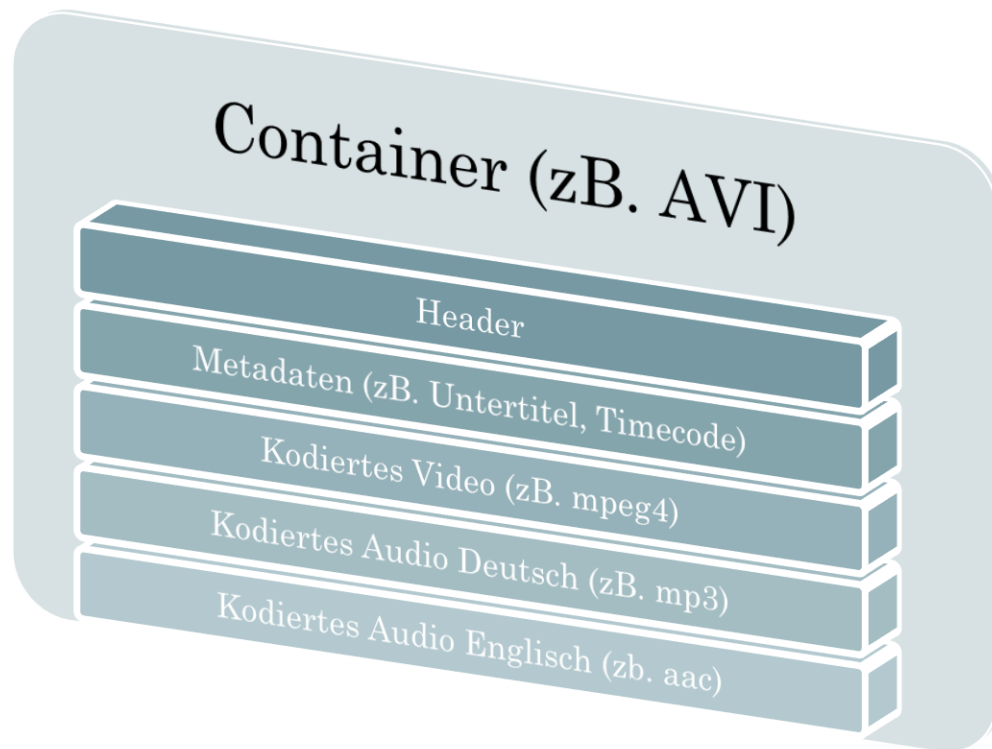
CODECs (Abkürzung für **Coder/Decoder**) sind Programme, welche Strategien zur Kodierung und Dekodierung von Audio- und Videomaterial umsetzen. Oft wird bei der Kodierung ein **Kompri-mierungsalgorithmus** zur Datenreduktion eingesetzt. Ein CODEC ist ua. auch notwendig, um Daten nach der A/D-Wandlung in eine Datei zu schreiben um diese für Streaming, Weiterverarbeitung oder Download bereitzustellen. Die **Elementarströme** werden meist in Container verpackt.

Betrachte zur Veranschaulichung folgendes Prozessdiagramm bei dem die Kodierung, das Verpacken in einen Container und die Dekodierung von Videodaten exemplarisch dargestellt sind.



CONTAINER

Multimediadaten werden in **Container** verpackt, welche **Elementarstreams** und **Metadaten** beinhalten. So kann eine AVI-Videodatei aus mehreren MPEG-Elementarstreams für mehrsprachigen Ton, Video sowie Metadaten wie Untertiteln bestehen.



Das **Kombinieren von Audio- und Videostreams** in einen Container wird mit einem **Multiplexer** durchgeführt. Bei der Wiedergabe werden die Datenströme wieder mit einem **Demultiplexer** getrennt um danach vom jeweiligen Dekodierer dekodiert zu werden. Daher kommen auch die Bezeichnungen **Muxen** bzw. **Demuxen**.

CONTAINER UND CODECS ÜBERSICHT

Über die **Optionen und Parameter von A/V-Containern** kann auf folgender Wikipedia-Seite eine Übersicht aufgerufen werden: https://en.wikipedia.org/wiki/Comparison_of_container_formats

Eine **Liste der CODECs** kann hier abgerufen werden: https://en.wikipedia.org/wiki/List_of_codecs

Welcher CODEC/Container zum Einsatz kommt hängt von der Aufgabe bzw. der Vorgabe ab.

Eines der wichtigsten Containerformate im Broadcast ist **MXF** mit XDCAM bzw. AVC-Videodatenströmen und AES3-Audio.

CODECS OHNE DATEN-KOMPRIMIERUNG

Es folgt eine Übersicht der wichtigsten **non-compression**-CODECs.

Audio: LPCM, WAV, AIFF

Video: RGB 4:4:4, YUV 4:4:4

CODECS MIT VERLUSTFREIER DATEN-KOMPRIMIERUNG

Es folgt eine Übersicht der wichtigsten **lossless**-CODECs.

Audio: FLAC, ALAC, WavPack, MPEG-4 ALS

Video: h.264 lossless, h.265 lossless, motionJPEG 2000, Dirac Pro lossless, Apple Animation RLE, AV1, Grass Valley lossless, magicYUV

CODECS MIT VERLUSTBEHAFTETER DATEN-KOMPRIMIERUNG

Es folgt eine Übersicht der wichtigsten **lossy-compression**-CODECs.

Audio: AC3, DTS, MP3, AAC, Opus, OGG Vorbis, WMA, AES3

Video interframe: AV1, MPEG-2, h.264, h.265, MPEG-4, OGV Theora, WMV, XDCAM, AVC

Video intraframe: MotionJPEG, MotionJPEG 2000, Apple ProRes 422/444, AVC-Intra

RAW (Bayer-Mosaik): CinemaDNG, ArriRAW, Apple ProRes RAW

	H.264	HEVC	VP9	AV1
Revenue Side				
Reach				
Computers	100%	Only Safari/Edge/ No Chrome or Firefox	95%	80% in 6 months/CPU not yet known
Mobile with hardware	100%	50%	50%	2 years
Smart TV/STB/OTT	100%	100%	under 100%	2 years
Features				
Live	100%	Many options	Little support	Challenging
Live transcode	100%	Many options	Wowza & Nimble	Challenging
Low latency	100%	Some options	unknown	Challenging
HDR	Not optimal (reach of 10-bit AVC unknown)	100%	HLG; no Dolby Vision or HDR10	2 years
Quality	1 - lowest of the bunch	H.264 @ 60%	H.264 @ 60%	VP9 @ 70%
Cost Side				
Encoding time	1x	~ 4x H.264	~2X H.264	1000+ x VP9
Content royalty cost	PPV/Subscripion	Velos?	None	No current std. support
FUD Factor	Nokia/Motorola	Others not in pool	Feels low risk	Some risk

Abbildung 38: Anforderungen an Streaming-Codecs Übersicht

VERLUSTBEHAFTETE STRATEGIEN DER DATENKOMPRIMIERUNG

Bei Videodaten werden jedoch **interframe** und **intraframe** Komprimierungsstrategien unterschieden.

INTRAFRAME-KOMPRIMIERUNG

Sogenannte **I-Frames** sind **intraframe-komprimiert**. Diese können **ohne Referenz** zu anderen Frames der Bildfolge dekodiert werden.

INTERFRAME-KOMPRIMIERUNG

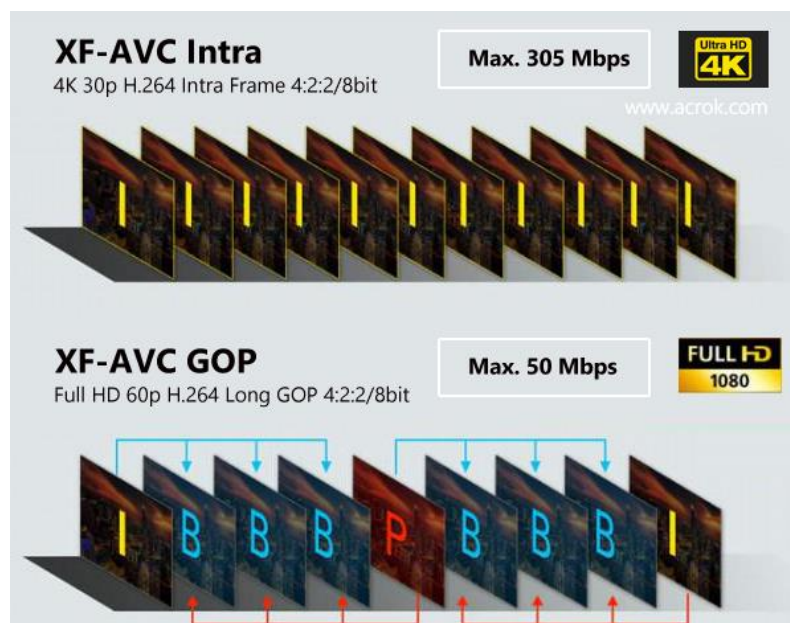
Eine **Gruppe von Bildern wird in Relation zueinander kodiert** und redundante Informationen reduziert. Vereinfacht formuliert wird in sog. **B- und P-Frames** nur gespeichert, was sich in Relation zu anderen Frames verändert.

- **P-Frames** werden in **Relation zu zum naheliegendsten, vergangenem I- oder P-Frame** kodiert.
- **B-Frames** werden in **Relation zu zwei Frames, einem vergangenem und einem zukünftigem I- oder P-Frame** kodiert.

B- und P-Frames **können nicht ohne Referenz zu anderen Frames** dekodiert werden. Dies ist ein wichtiger Faktor in der Schnittsoftware.

GROUP OF PICTURES GOP

Anhand der **Zielbitrate** wird eine **Anzahl an I-, B- und P-Frames definiert**. Eine GOP beinhaltet ein I-Frame und eine spezifizierbare Anzahl an B- und P-Frames. Um auch die Bewegung einzelner Objekte im Bild zu erfassen, wird das Bild in sogenannte **Makroblöcke** der Größe 16 x 16 Pixel zerlegt.



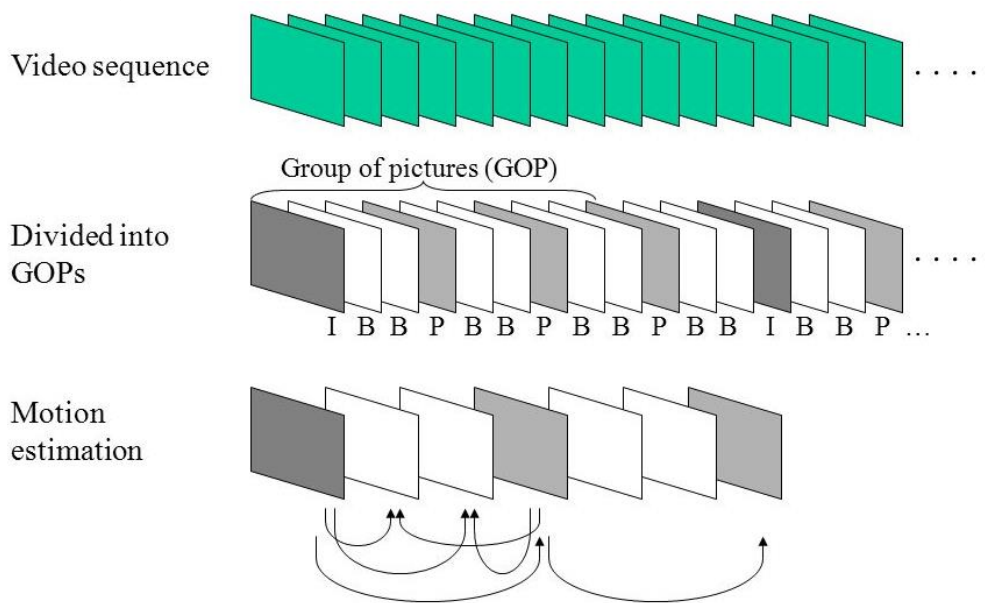


Abbildung 39: GOP mit I-, B- und P-Frames

Die Länge der **GOPs** ist **variabel einstellbar** und **beginnt immer mit einem I- oder B-Frame** und **endet immer mit einem I- oder P-Frame**.

GOP-Muster werden dabei durch die beiden Werte n und m beschrieben. Das n gibt dabei den **Abstand zwischen zwei I-Frames** (und somit auch die Anzahl der Frames pro GOP) und das m gibt den **Abstand zwischen einem I- und einem P-Frame oder zwei P-Frames** an.

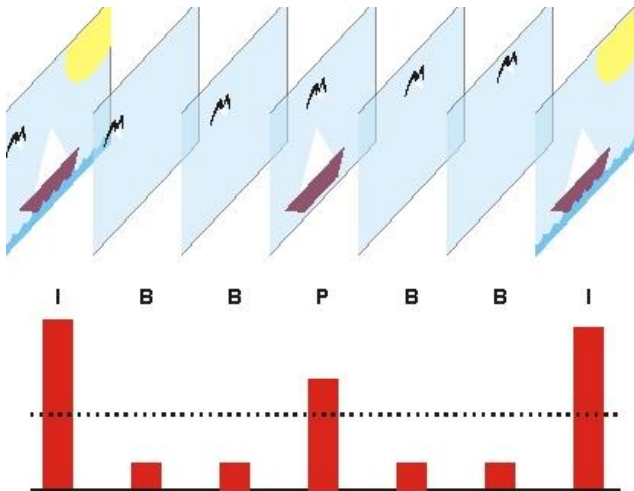


Abbildung 40: 6-Frame-GOP mit $m=3$, $n=6$.

Obige Abbildung: Das I-Frame beinhaltet die volle Bildinformation. Die B-Frames beinhalten nur die Information, welche von den I- oder P-Frames abweicht. Das P-Frame beinhaltet die Änderungen in Relation zum letzten I- oder P-Frame.

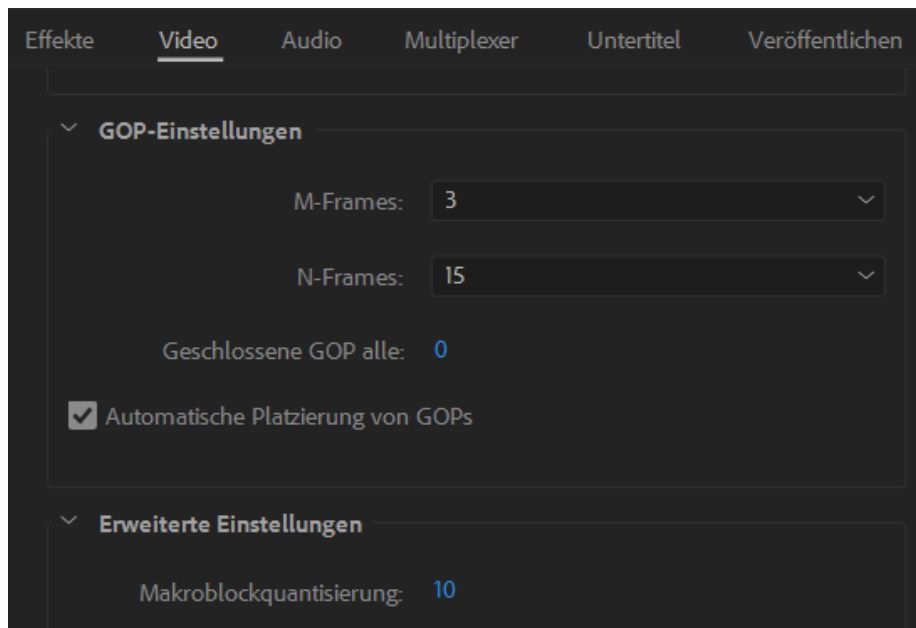


Abbildung 41: Medien-Exporteinstellung mit GOP-Konfiguration

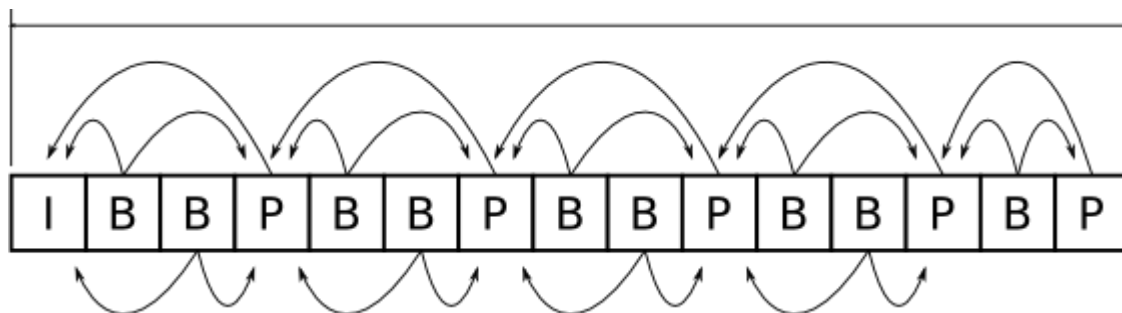


Abbildung 42: Beispiel für einen 15 Frame-Long-GOP mit $m = 3$ und $n = 15$

EXKURS: MULTIMEDIAFRAMEWORK FFMPEG



FFmpeg ist das führende **Multimediaframework zur Aufnahme, Konvertierung/Transkodierung, Enkodierung, Dekodierung, Mux/Demux, Manipulation und Streaming von Audio-, Bild- und Videomaterial**. Neben der Anwendung in der Kommandozeile werden häufig die FFmpeg-Bibliotheken und API in der Softwareentwicklung eingesetzt. So basieren zB. die Audio- und Videofunktionen in Audacity, Handbrake Blender, Chrome, VLC media player usw. auf den ffmpeg Bibliotheken. Weiters sind viele grafische Benutzeroberflächen (GUI) erhältlich. Wir beschäftigen uns mit dem Einsatz von FFmpeg als Kommandozeilentool. In den folgenden Abschnitten sollen die Installation und Arbeitsvorbereitung erläutert werden.

BEZUGSQUELLE

FFmpeg ist quelloffen und sowohl die Exe als auch der Quelltext können unter <https://ffmpeg.org> heruntergeladen werden. Die **Dokumentation** ist unter <https://ffmpeg.org/documentation.html> abrufbar. Ein **Wiki mit Infos zu CODECs, Streaming, Filtern** usw... kann hier aufgerufen werden: <https://trac.ffmpeg.org/wiki/WikiStart>

Paketbestandteile:

- **FFmpeg**: A/V Encoder und Decoder
- **FFplay**: Medienplayer
- **FFprobe**: Medieneigenschaften anzeigen
- **FFserver**: Broadcast-Server für Multimedia-Streaming via HTTP und RTSP
- **Libcodecs, lib...**: Bibliotheken mit Filtern, CODECs, Tools und Processing-Algorithmen

INSTALLATION

Lade das entsprechende Paket herunter und **entpacke das Archiv** an einen Ort deiner Wahl. Starte eine Kommandozeile und wechsle in das Verzeichnis bin (mit dem DOS-Befehl *cd*). Lass dir mit *dir* anzeigen, ob du im richtigen Verzeichnis bist. Alternativ kann auch eine Umgebungsvariable mit dem Ordnerpfad angelegt werden.

Es kann auch sinnvoll sein, die ffmpeg.exe mit *ren* in f.exe umzubenennen, man spart dadurch 5 Zeichen.

SYNTAX

```
ffmpeg [global options] [input file options] -i Eingabedatei
[output file options] [Ausgabedatei]
```

Beispiel: `ffmpeg -y -i video.mp4 -s vga video.mkv`

`-y`: global option

`-i video.mp4`: Eingabedatei

`-s vga`: output option

`video.mkv`: Ausgabedatei

AUSGABEVORSCHAU

Um nicht immer eine Datei zu erzeugen, kann der Output auch an ein Display umgeleitet werden. Dafür nutzen wir *ffplay* mit einem SDL¹³.

Beispiel: `FFplay -i video.mp4`

ERSTE SCHRITTE

Teste ob ffmpeg aufgerufen werden kann:

```
ffmpeg
```

Es wird die Version und die Version der Bibliotheken angezeigt.

Unterstützte CODECs und Filter ausgeben:

```
ffmpeg -codecs
```

```
ffmpeg -pix_fmts
```

```
ffmpeg -protocols
```

```
ffmpeg -filters > info.txt
```

```
ffmpeg -formats >> info.txt
```

Teste den SDL mit der bereitgestellten Datei:

```
ffplay -i video.mp4
```

In der Kommandozeile werden Infos zur Datei angezeigt und ein directMedia-Fenster öffnet sich.

Sehen wir nach, worum es sich bei der bereitgestellten Datei handelt:

```
ffprobe -i aa.ts
```

¹³ Simple DirectMedia Layer

Wir sehen nun, dass es sich um einen MPEG-TransportStream handelt, seine **Dauer**, die **Bitrate** und dass im Container **2 Datenströme** sind. Ein AC3 Audio- und ein MPEG2 Videostrom. *Analysiere auch die weiteren Angaben.*

Die Datei ist für die Übung etwas zu lang und der Schwarzvorspann könnte auch entfernt werden. Dazu geben wir folgendes ein:

```
ffmpeg -ss 00:00:10 -i aa.ts -t 00:01:00 -c copy bb.ts
```

-c copy sorgt dafür, dass Video, Audio und Untertitel in den neuen Container ohne Rekodierung kopiert werden.

Transkodieren wir nun die gekürzte Datei nach MKV.

```
ffmpeg -i bb.ts cc.mkv
```

Der MPEG2 Videodatenstrom wird nun default in h.264 transkodiert, der Audiodatenstrom wird default in vorbis transkodiert. *Analysiere die weiteren Angaben. Beachte, dass sich auch die Bitrate geändert hat. Das liegt daran, dass wir default-Werte nutzen, da wir keine eigenen Optionen angegeben haben.*

Wir transkodieren nochmal in einen Matroska-Container, diesmal geben wir jedoch Optionen zu Bitrate und Encoder an:

```
ffmpeg -i bb.ts -c:v libx264 -c:a aac -b:v 5000k -b:a 64k dd.mkv
```

Kontrolliere mit *ffprobe* das Ergebnis.

Stellen wir das Video nun auf den Kopf:

```
ffmpeg -i dd.mkv -vf "transpose=2, transpose=2" -c:v libx264 -c:a aac -b:v 5000k -b:a 64k ee.mkv
```

Fügen wir nun Audio ein. Da wir schon eine Audiospur haben, muss mit *-map* eine Zuordnung erstellt werden. Kontrolliere die Audiodatei zuvor mit *ffprobe*:

```
ffmpeg -i ee.mkv -i sound.mp3 -c:v copy -c:a aac -b:a 64k -map 0:v:0 -map 1:a:0 ff.mkv
```

Beachte, dass der Ton nun länger als die Videospur ist. Beschneide den clip nun wieder auf 1 Minute. *Beachte die Videobitrate vorher und nachher.*

Reduzieren wir nun auf 720p, drehen das Video und kodieren in h.265:

```
ffmpeg -i gg.mkv -vf "transpose=2, transpose=2",scale=1280:720 -c:v libx265 -b:v 2500k -c:a copy hh.mkv
```

Sieh dir nun die bereitgestellte Untertitel-Datei *untertiteltest.srt* in notepad an. Diese fügen wir nun ins Video ein.

```
ffmpeg -i hh.mkv -i untertiteltest.srt -c copy ii.mkv
```

Kontrolliere mit *ffprobe* ob die Untertitel nun im Datenstrom 0:2 enthalten sind und diese im VLC-Player angezeigt werden. Die Untertitel müssen noch im Player aktiviert werden.

Die Untertitel können auch mit dem Videodatenstrom verrechnet werden:

```
ffmpeg -i ii.mkv -vf subtitles=ii.mkv -c:a copy -c:v libx265 -b:v 2500k  
jj.mkv
```

Wellenformen können in Echtzeit ausgegeben werden:

```
ffplay ii.mkv -vf "split[a][b];[a]wave-  
form=e=3,split=3[c][d][e];[e]crop=in_w:20:0:235,lutyuv=v=180[low];[c]crop=i  
n_w:16:0:0,lutyuv=y=val:v=180[high];[d]crop=in_w:220:0:16,lutyuv=v=110[mid]  
; [b][high][mid][low]vstack=4"
```

Berechnungen können auch hardware-beschleunigt angestellt werden:

```
ffmpeg -i bbb.mp4 -vf scale=1280:720 -r 50 -c:v h264_nvenc -b:v 2000k  
bbb.avi
```

Erstelle einen UDP-Stream

```
ffmpeg -i bbb.avi -vcodec libx264 -f mpegts udp://127.0.0.1:23000  
ffplay udp://127.0.0.1:23000
```

Screenshot aus dem Video grabben:

```
ffmpeg -i jj.mkv -ss 00:00:15 screenshot.png
```

Animiertes GIF erstellen:

```
ffmpeg -ss 00:03:21 -i bbb.mp4 -t 00:00:03 -r 20 -pix_fmt rgb8 -vf  
scale=320:200 anim.gif
```

ÜBERBLICK WEITERER OPTIONEN

-r Bildwiederholfrequenz

-vf fps=fps=xx Filter für Bildwiederholfrequenz

-b Bitrate

-minrate -maxrate Bei variabler Bitrate Grenzen angeben

-bufsize

```
ffmpeg -i in.avi -b 0.5M -minrate 0.5M -maxrate 0.5M -bufsize 1M out.mkv
```

-fs maximale Dateigröße

-f forciert Format

-vf scale Skalierung. Mit iw und ih kann proportional skaliert werden. *-vf scale=iw/2:ih/2*

-vf crop=ow:oh:x:y

```
ffmpeg -i input -vf crop=iw/3:ih:0:0 output
```

-vf pad=w:h:x:y:color Bereich hinzufügen. Beispiel: 16:9 auf 4:3 mit Balken

```
ffmpeg -i hd_video.avi -vf pad=iw:iw*3/4:0:(oh-ih)/2 video.avi
```

-vf unsharp

-vf mp=denoise3d oder -vf hqdn3d oder -nr 0-600

-vf overlay=x:y

```
ffmpeg -i input1 -i input2 -filter_complex overlay=x:y output
```

-vf drawtext

```
ffmpeg -i input -vf drawtext=fontfile=arial.ttf:text=Welcome output
```

-target

-t Dauer

-ss Startdelay

-vf setpts=xxx. Beispiel beschleunigt Video 3x. PTS ist Presentation Timestamp 1/fps. Bei 25 fps ist PTS 0,04

```
ffplay -i input.mpg -vf setpts=PTS/3
```

-af atempo=xxx Faktor für Audiotempo (wichtig bei US/EU-Bildraten)

-metadata

```
ffmpeg -i video.avi -metadata location="London, United Kingdom" -metadata "camera type"="SONY DSC" -metadata "camera mode"="movie" -metadata weather="sunny" video.wmv
```

Weitere Optionen können online nachgeschlagen werden. FFmpeg ist sehr gut dokumentiert.

Quellennachweis:

Professionelle Videotechnik, Ulrich Schmidt, Springer-Verlag Berlin Heidelberg

Mohidien Morade, Präsentationen

EBU, ITU, SMPTE Dokumentverzeichnis

Sony, Fujifilm

Videotechnik-Vorlesung Rauchberger, Wassermann

Wikipedia

Klaus, Steffen und Andreas Burosch, Medientechnik