

8

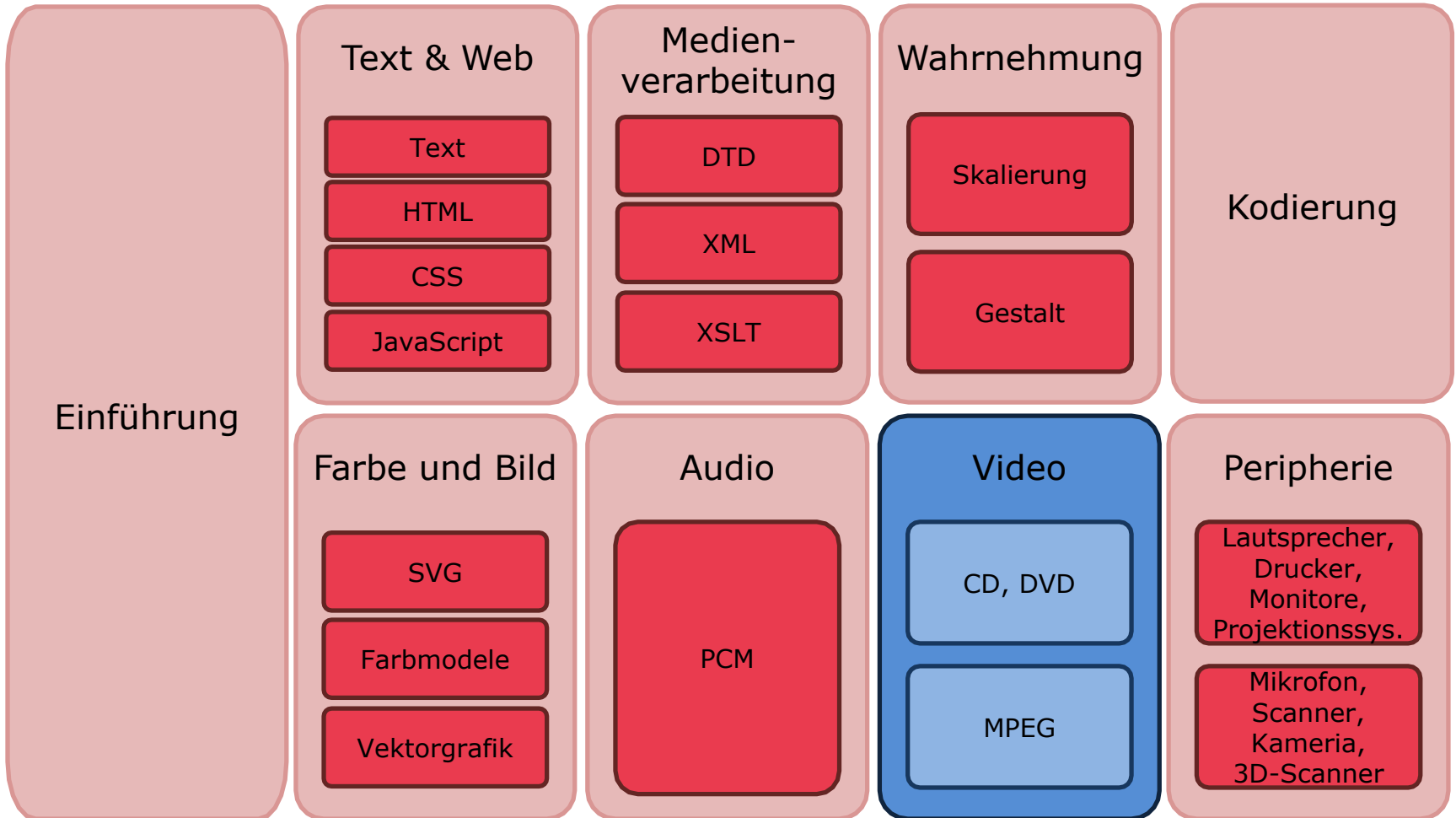
mit Material von G. Weber (TU Dresden):

Einführung in die Medieninformatik

Video

Prof. Dr.-Ing. Tenshi Hara
fhd-emi@lern.es

VORLESUNGSÜBERSICHT



Historie

VORLÄUFER



Thaumatrope
(Wunderscheiben)
(1825, Trentsensky & Vieweg)



Phenakistiskop
(Lebensrad)
(1833, Stampfer, Plateau)



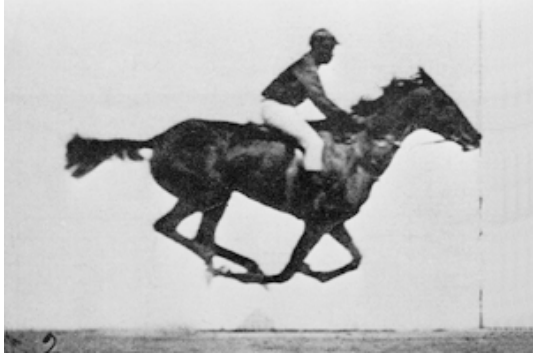
Kineograph
(Daumenkino)
(1868, Linnet)



Praxinoskop
(Tätigkeitsseher)
(1877, Reynaud)

SCHWARZ/WEIß-FILM

- 1879: Zoopraxiskop – Projektion durch ein rotierendes Stroboskop



Zoopraxiskop-Reihe von Muybridge (1879)



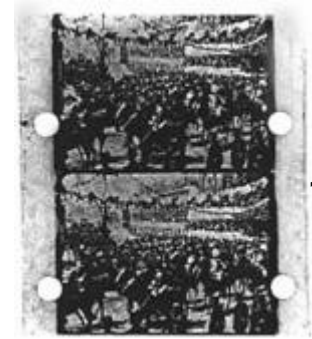
Le Prince film, 1888

54mm



Skladanowsky, 1895

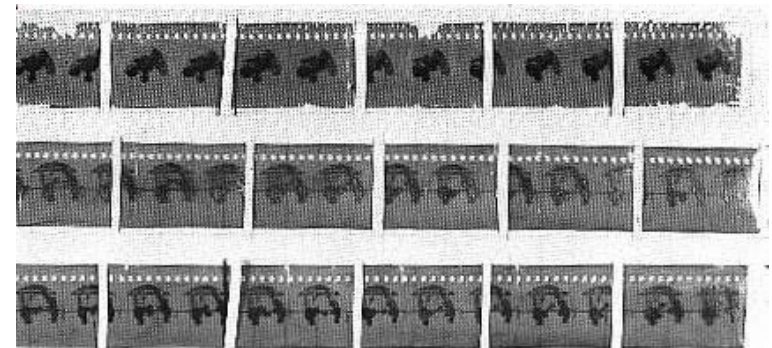
54mm



Lumièr 1900

70mm

- 1888: linearer Zelluloidfilm
Photoemulsion auf Zelluloid aufgebracht
- langer Kampf um Standardisierung;
inzwischen dominiert 35mm-Film
- meist 24 Bilder pro Sekunde
(1000m 35mm-Film \approx 36½ min)



35mm, Edison (1890)

FARBFILM

- bis ca. 1905 handcolorierte Einzelbilder
- 1906: Kinemacolor
 - 32 Bilder pro Sekunde
 - abwechselnd rot und grün belichtet
 - rotierender Farbfilter vor Objektiv
- 1917: Technicolor (Process I)
 - Zwei-Farben-Verfahren (Purpur und Grün)
 - Doppelprisma hinter Objektiv
- 1932: Technicolor (Process IV)
 - Aufspaltung mit Prisma auf drei Einzelbilder
 - Drei-Farben-Dye-Transfer-Verfahren
- 1932: Agfacolor
 - als Farbfilter wirkende Körner vor dem Film
 - erstes chromogenes Negativ/Positiv-Verfahren



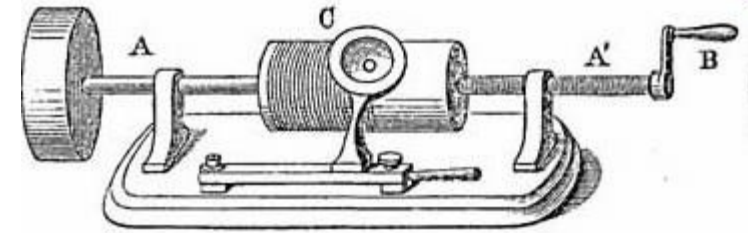
A Visit to the Seaside
(1908, Kinemacolor)

The Wizard of Oz
(1939, Technicolor IV)

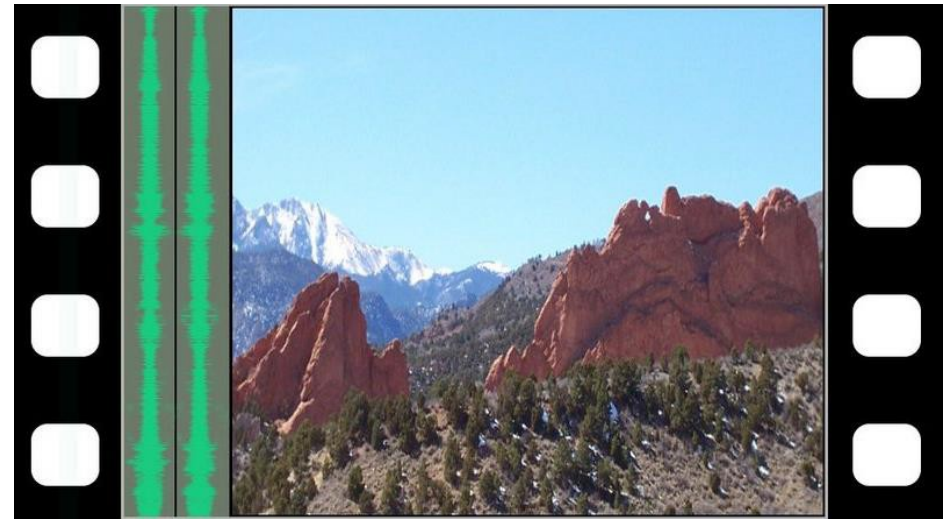


TONFILM

- 1877: Phonograph
 - Nadelton-Verfahren
 - Nadel schreibt und liest Schwingungen auf Walze
- 1922: Lichttonverfahren
 - Tonspur neben Bild
 - Lichtdurchlässigkeit proportional zur Amplitude des Tonsignals
- Später: Magnetton



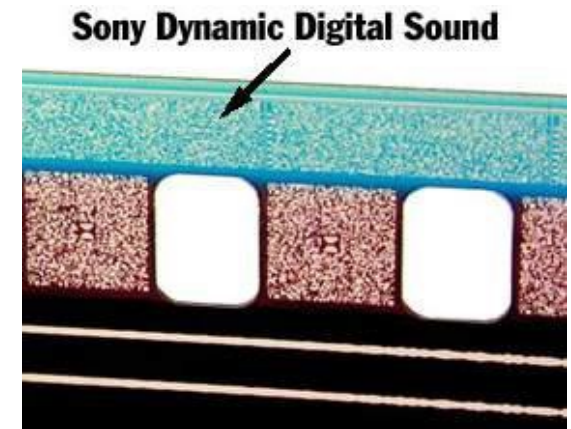
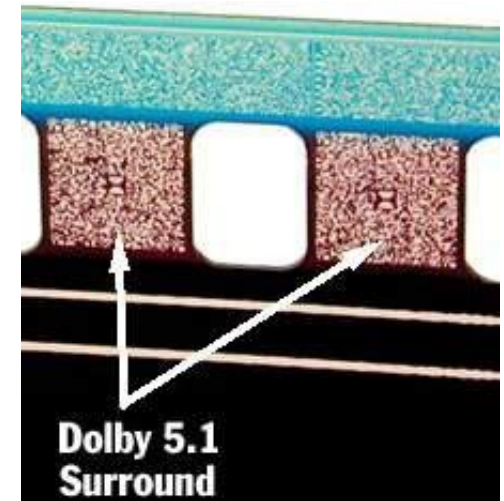
Edison Phonograph



Film mit Stereotonspur

TONFILM – HEUTE

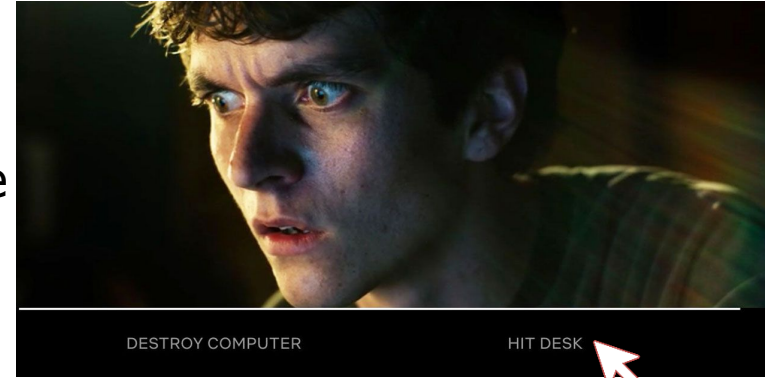
- 4-Kanalton wird mit der Matrixmethode in die Lichtkanäle moduliert
- Dolby 5.1 Surround wird entweder auf CDs mitgeliefert oder digital in Filmperforation gespeichert
- 8-Kanal SDDS wird komprimiert digital auf Filmaußenrand gespeichert



INTERAKTIVE FILME UND RENDERING

Interaktive Filme

- Aufnahme und interaktives Wechseln mehrerer Perspektiven oder Erzählstränge
- Beispiele:
 - Aspen Movie
 - Black Mirror: Bandersnatch



Rendering

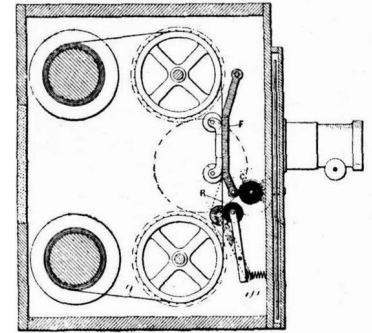
- Einzelne Bilder werden berechnet und zum Film zusammengesetzt
- Beispiele:
 - Pixar-Filme (Geri's Game, ...)
 - Big Buck Bunny



KAMERAS

Funktionsprinzip folgt bis heute dem Kinematograph (Edison, 1891)

- unbelichteter Film wird von einer Spule in Filmkanal geleitet
- Belichtung eines Bildes
- Blende deckt Bildfenster ab
- Film wird zu zweiter Spule geführt
- Entwicklung des Films
- wenn Film durch CMOS ersetzt wird:
digitale Kamera
(zu Kamerazubehör siehe <https://youtu.be/RkTaMyatsTo>)



Aufriss von Acres' Cinetic Camera (1895)



moderne
(analoge)
Kamera



moderne
(digitale)
Kamera mit
Zubehör

PROJEKTOREN

Laterna Magica – Zauberlaterne
(Athanasius Kircher, 1671)

- Projektion auf Rauch



Projektoren für 35mm

- Kinetoscope
(Thomas Edison, 1892)
- nur ein Betrachter
- Cinématographe
(Lumièr Brüder, 1895)
- erste Projektion auf Leinwand



ANALOGE KINOTECHNIK

Links sind drei „Teller“ auf denen ein kompletter Film liegen kann.

Über die kleinen Rollen läuft der Film in den Projektor und zurück auf einen freien Teller, wo er aufgewickelt wird.

Rechts neben dem Projektor, relativ unten, sieht man eine Schaltmatrix, mit der Licht, Vorhang, Bild und Tonformat, etc. gesteuert werden.



DIGITALE KINOTECHNIK

Links: Tonschrank mit Verstärkern, etc.

Rechts daneben: Kühlturm und Analogprojektor

Mittig: kleiner Tisch mit Player-Anschluss (DVD, BD, ...)

Rechts: eigentlicher Digitalprojektor



Formate

ANFORDERUNG: ZEITLICHE AUFLÖSUNG

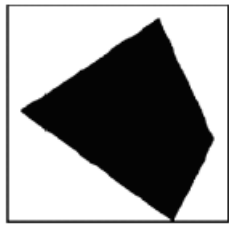
- flüssige Bewegung
 - ab ca. 20 Bilder/Sek.
 - Filmbildfrequenz: 24Hz
- Flimmern
 - durch Dunkelpausen zwischen Bildern wiederholter Erregung und Hemmung der Neuronen
→ unangenehmes Flackern
 - Flimmerverschmelzungsfrequenz bei ca. 50Hz; steigt mit Bildhelligkeit
- Film
 - Bilder 2x projiziert
 - Verdopplung der Dunkelpausen
 - Flimmerfrequenz 48Hz
- Computermonitore
 - hohe Helligkeit, hohe Aufmerksamkeit, großer horizontaler Sehwinkel
→ Bildwiederholffrequenz >72Hz
- TFT/LED/OLED-Displays
 - keine Dunkelphasen
→ kein Flimmern

ANFORDERUNG: HELBIGKEITSEMPFINDUNG UNTERSTÜTZEN

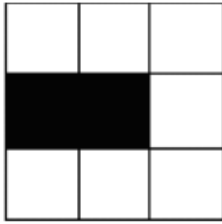
- Abhängig vom Adaptionszustand des Auges und der Umfeld-Beleuchtung
- Auge kann sich an Leuchtdichteunterschiede anpassen, dauert aber bis zu 30 Minuten (mondlose Nacht → Sonnen-beschienenes Schneefeld)
- bei adaptierten Augen können ca. 200 Helligkeitswerte unterschieden werden
- Fernsehen
 - Streulichter schränken Kontrastempfindung ein
 - Umfeld-Leuchtdichte sollte ca. 10% der Spitzenleuchtdichte des Bildschirms betragen

ANFORDERUNG: ALIASING VERMEIDEN

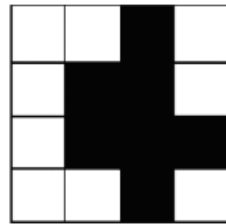
Treppentufen und Geisterfrequenzen



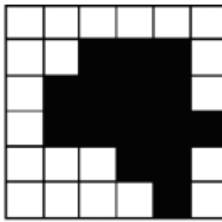
Original quadrilateral



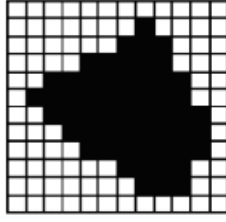
3 x 3



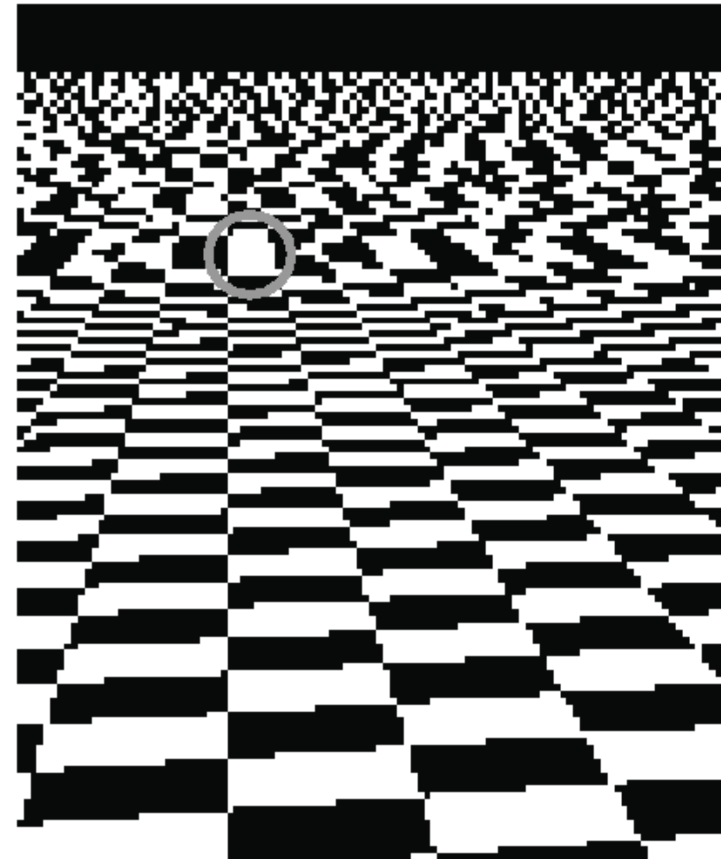
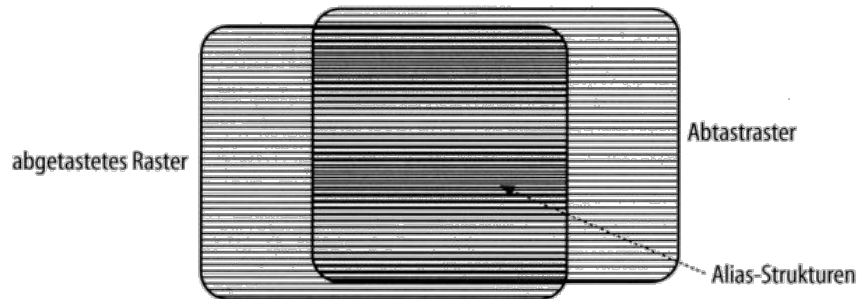
4 x 4



6 x 6

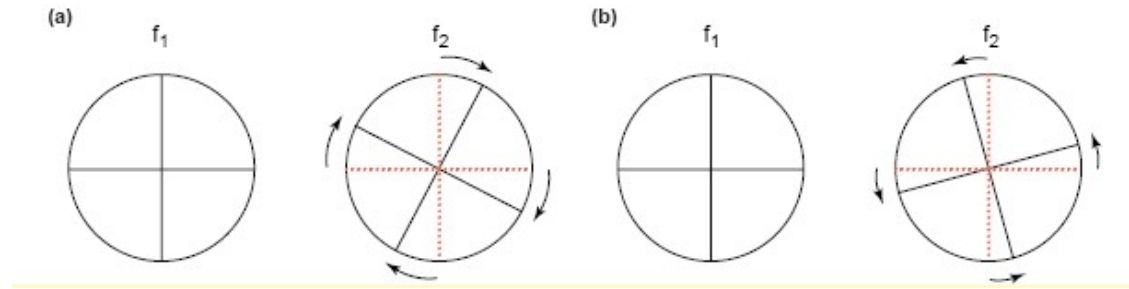
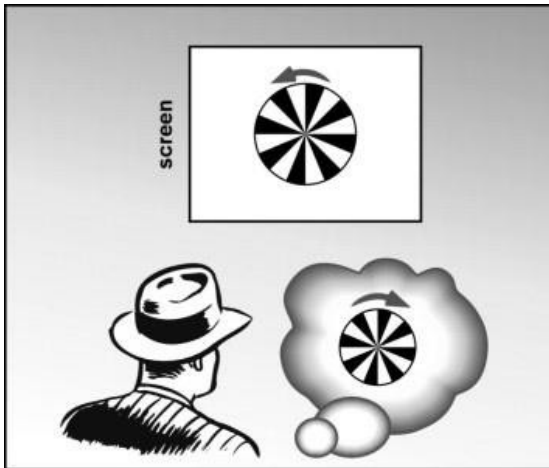


12 x 12



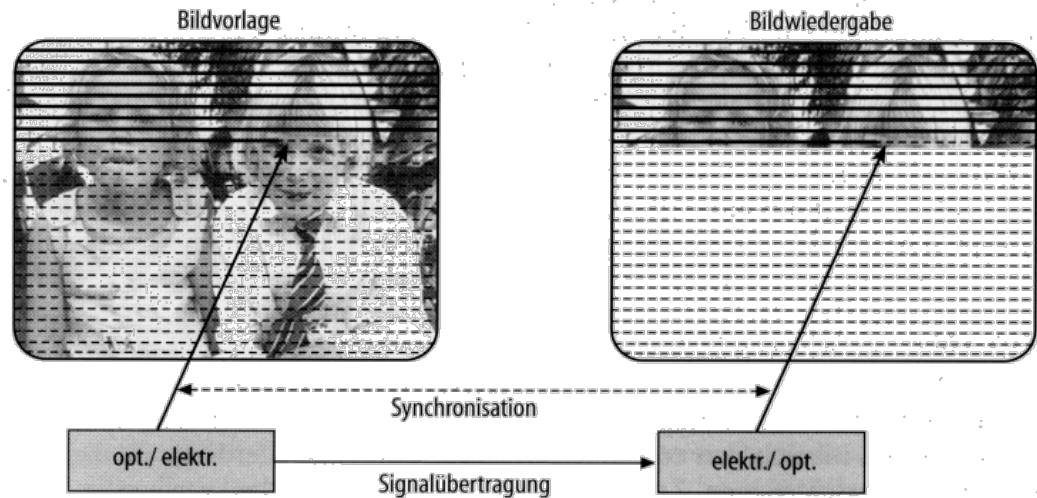
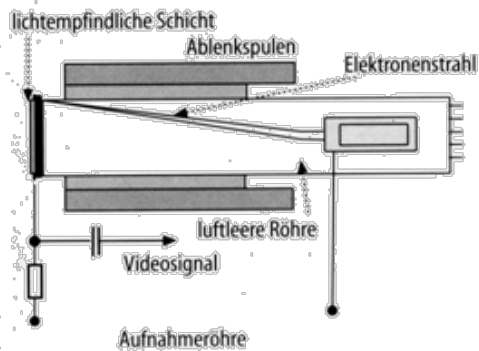
ZEITLICHES ALIASING

- Wagenrad scheint sich rückwärts zu drehen
→ die Abtastfrequenz zu gering
- Stillstand falls genau um 90° verdreht
- wird vermutlich durch neuronale Adaption verursacht

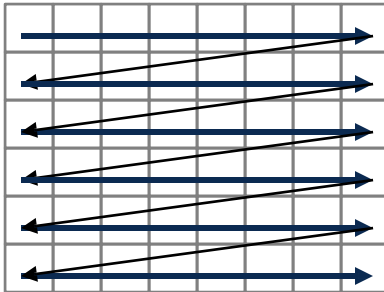


VIDEOSIGNAL (SCHWARZ/WEIß ANALOG):

- zeilenweise Abtastung der Helligkeitswerte
- serielle Übertragung der Bildpunkte
- synchrone Darstellung der Punkte auf Monitor



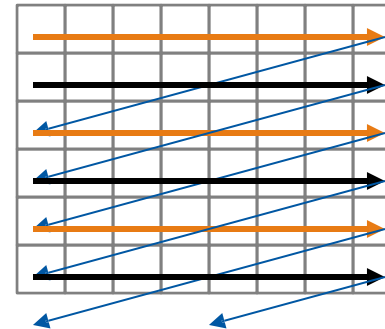
PROGRESSIV/INTERLACED



Progressiv

Vollbilder – Frames

typische Unschärfe:



1. Scan Field 1
1. Scan Field 2
2. Scan Field 1
2. Scan Field 2

Interlaced

2 Halbbilder

(Fields in doppelter Frequenz)

typische Artifacts:



DE-INTERLACING

- beim Abspielen von TV-Signalen auf 100Hz Röhren bzw. allen LCD/TFT-Panelen und Plasmabildschirmen muss von interlaced auf progressiv umgerechnet werden
- **Weave**
 - gleichzeitige Darstellung der Halbbilder
 - i.O. falls zum gleichen Zeitpunkt aufgenommen
 - bei TV jedoch verschiedene Zeitpunkte → führt zu Kammeffekt
- **Unschärfe**: Weave mit anschließendem Weichzeichnen
- **Bobbing**: jedes Halbbild wird durch Interpolation in Vollbild umgewandelt um so 50Hz zu erreichen
 - jede Zeile mit Nachbarinformationen auffüllen
 - Bild „springt“ auf und ab (bobbing), da erste und letzte Zeile jeweils ohne Nachbar
- **Blending**: wie Bobbing aber Mittelung der Halbbilder zu einem Vollbild, das 2mal dargestellt wird
- **Motion Compensation**: wie Weave, ungerade Zeilen werden jedoch an gerade Zeilen angepasst

ANALOGUE FORMATE (1/2)

BAS

- Bild: Helligkeit auf 75% Signalamplitude skaliert
- Austastung: Zeilenrücklauf: 100% Ampl. für $5\mu s$
- Synchronisation: 100% Ampl. für $10\mu s$

FBAS

- Farbsignale werden überlagert

Komponentenvideo

- getrennte Übertragung von drei Farbsignalen nach RGB, YUV, oder YIQ

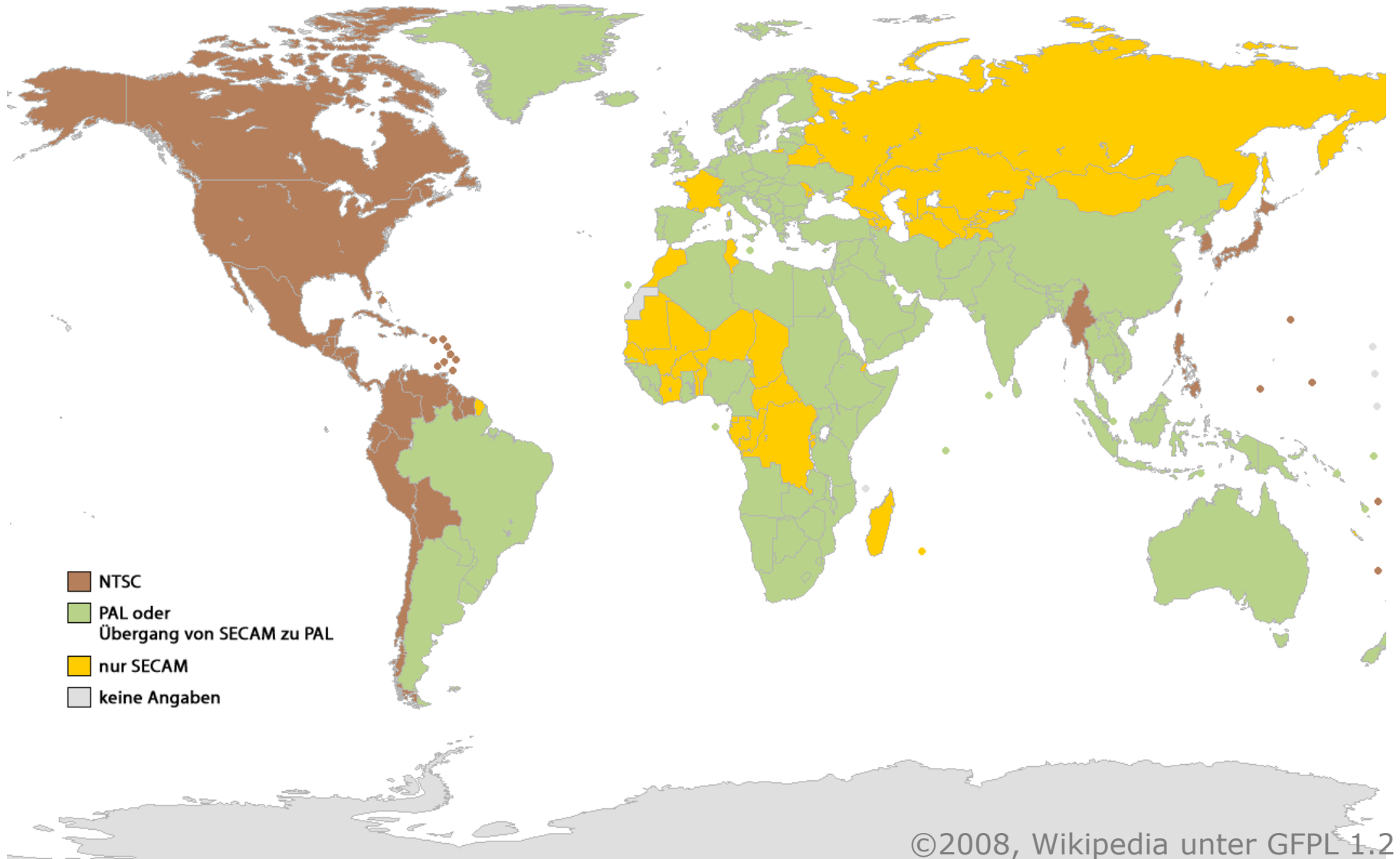
Y/C-Video
(separiertes Video)

- Trennung von Helligkeit & Chrominanzsignalen

ANALOGUE FORMATE (2/2)

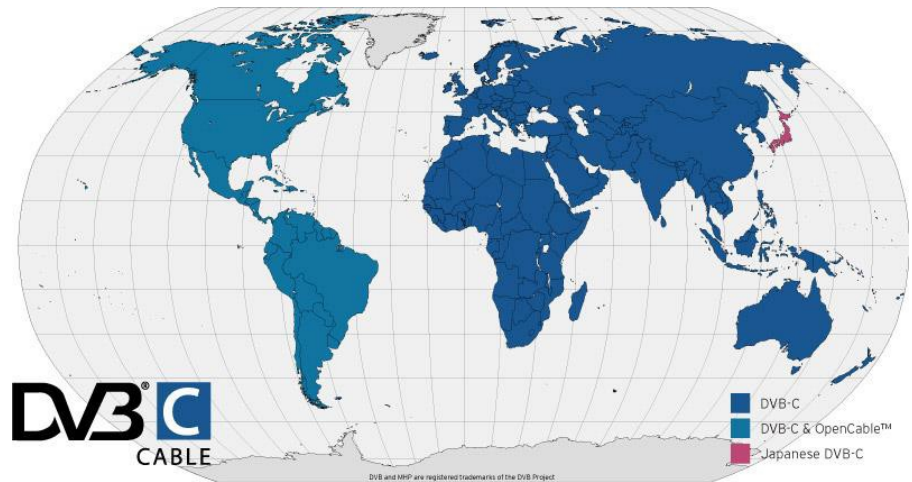
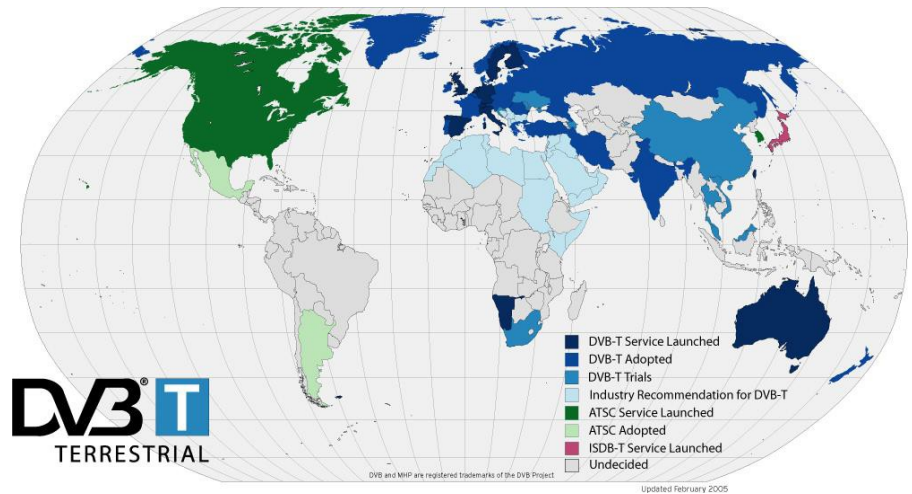
- PAL (Phase Alternating Line System)
 - Standard in Europa, Südostamerika, Vorderasien, Australien und Teilen Afrikas
 - interlaced FBAS in YUV
 - 833x625 quadratische Pixel, 720x576 sichtbar, 50 Fields = 25 Frames pro Sekunde
 - Bandbreite: Y 5,5 MHz, U/V 1,8 MHz
- NTSC (National Television Systems Committee)
 - Standard in Nord- und Südwestamerika sowie Ostasien
 - interlaced FBAS in YIQ
 - 700x525 quadratische Pixel, 29,97 Frames pro Sekunde
 - ohne Kontrolldaten 485 Zeilen, 480 sichtbar
 - Bandbreiten: Y/I/Q 4,2/1,0/0,6 MHz
- Secam (Sequential Couleur avec Memoire)
 - Standard in den meisten ehemaligen Sowjetrepubliken und Teilen Afrikas
 - wie PAL mit Bandbreiten: Y/U/V 6/2/2 MHz

TV-STANDARDS DER ERDE



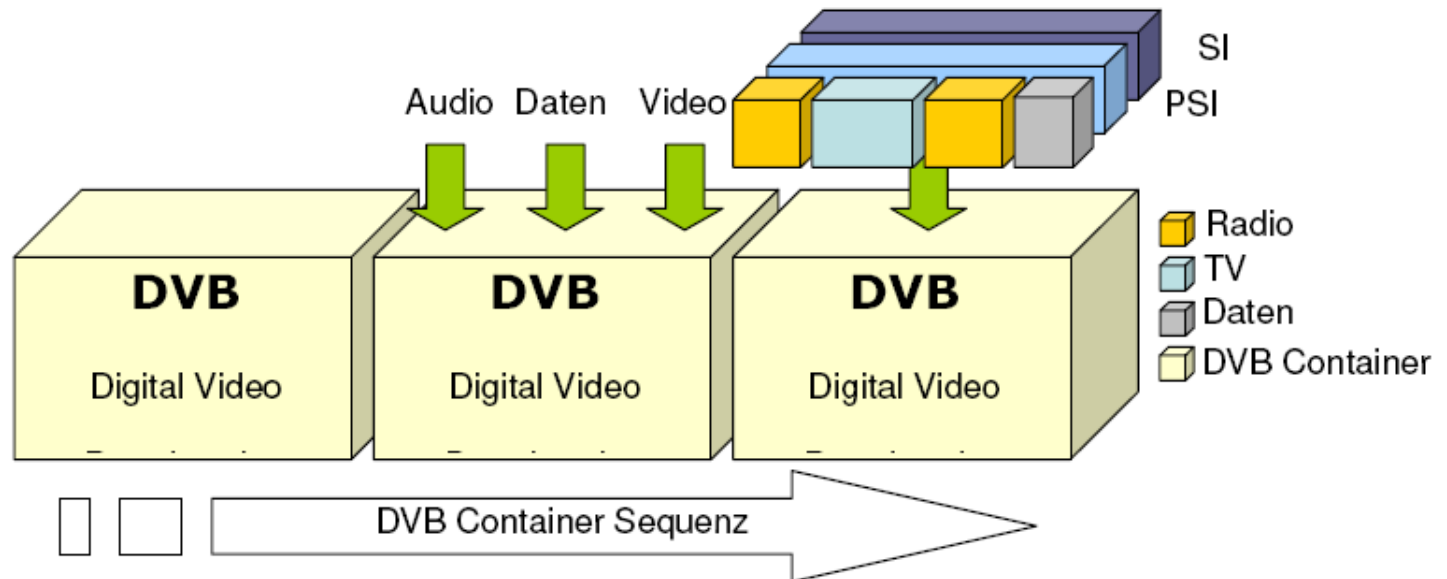
DIGITALFERNSEHEN: DVB

- DVB-Standard seit 1993 (siehe www.dvb.org)
- MPEG2 basiertes Fernsehen
- interaktive Dienste basieren auf MPEG Digital Storage Command Control (DSM-CC)
- Daten werden als Stream gesendet



DIGITALES TV

- DVB (Digital Video Broadcast)
- DVB-T, DVB-C, DVB-S
- Digital-TV auf Mobiltelefon via DVB-H, UMTS und GSM
- Multimedia Home Plattform (MHP)
- Hybrid Broadcast Broadband TV (HbbTV)
- Container transportiert Video, Audio und Daten
- Verteilung mittels Objektkarussell



DTV mit seinen Verbreitungskanälen (inkl. Objektkarussell) sowie mit CE-HTML, Javascript und Medienformaten CE-HTML besteht aus

- XHTML
- CSS-TV Profil
- DOM
- XMLHttpRequest
- Javascript
 - API abhängig vom Hersteller
 - Unterstützung von Eingaben per Fernbedienung
- OIPF-Medienformate (mehrere Audio-/Videoformate)

mit Fernsehsignal wird mitgeliefert

- linearer Broadcast (TV-Programm)
- Application Information Table (AIT) mit URL für CE-HTML-Anwendung die beim Drücken auf die Trigger-Taste aktiviert wird
- DSCMCC-Karussell enthält Daten und löst Ereignisse aus

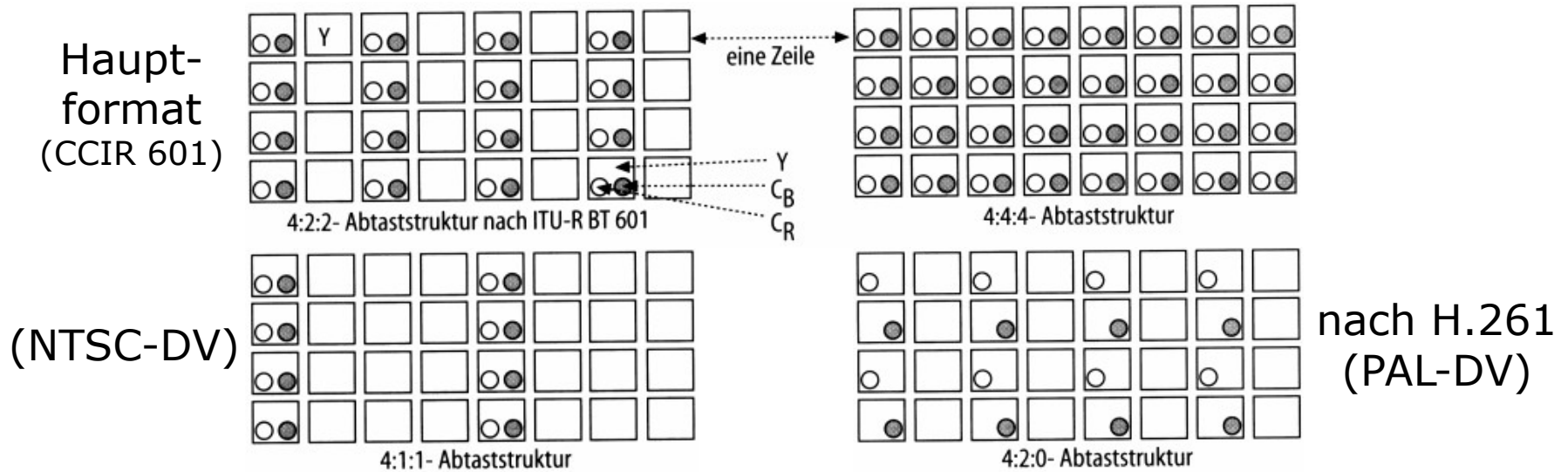
Darstellung erfolgt auf mehreren Bildebenen

- (schwarzer) Hintergrund
- Video-Ebene
- Untertitelebene/Gebärdendolmetscher
- HbbTV-Anwendungsebene
- gerätespezifische Informationen (z.B. Lautstärkeinstellung, Uhrzeit)

separater Kanal über HTTP für anwendungsspezifische Inhalte des TV-Anbieters

DIGITALE FORMATE (1/2)

- farbwahrnehmungsbezogenes Chroma-Subsampling $Y:C_R:C_B$
- Y-Komponente wird mit 3,375 MHz abgetastet
- erste Zahl ist ein Vielfaches der Y-Abtastrate
- zweite Zahl ist auch ein Vielfaches der Y-Abtastrate aber beschreibt die Abtastraten für C_R und C_B
- dritte Zahl deutet die räumliche Position an



DIGITALE FORMATE (2/2)

digitaler Bildstandard CCIR 601/ITU-R BT 601

- 50/60Hz, 13,5MHz Pixel
- 720x625x8Bit, 4:2:2,

Camcorder Standards

- D1/2/3/5/6/9, DCT, Digital Betacam
- DV (Digital Video), DVCpro, DVCAM, DVCAM50, ...

Digital Video Broadcasting

- DVB-S, DVB-C, DVB-T, DVB-H (www.dvb.org)
- Kompression mit MPEG2

Audio Video Interleave (AVI):
Containerformat für beliebige Codecs

Motion by JPEG (MJPEG)

- JPEG basierte Kompression der Halb-/Vollbilder
- kein Standard → Probleme beim Austausch

Motion Pictures Expert Group (MPEG)

- weit verbreiteter Standard in verschiedenen Versionen (1, 2, 4, 7, 21)

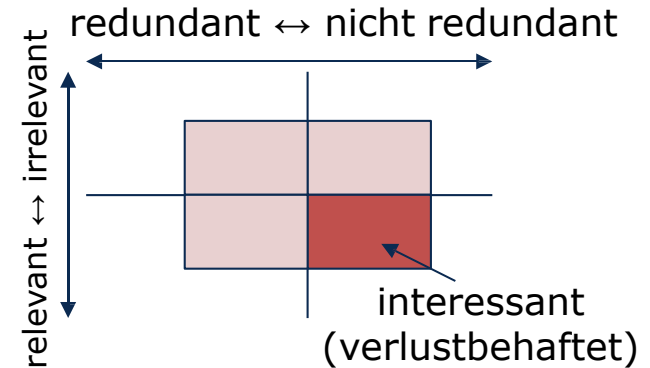
VERGLEICH DER VIDEOFORMATE

Abkürzung	Name	Auflösung	Seitenverhältnis	phisches Seitenverhältnis	Bildfrequenz	interlaced
SQCIF	Semi QCIF	128 × 96	4:3		29,97	
QCIF	Quarter CIF	176 × 144	4:3	11:9	29,97	
SIF	Source Input Format	352 × 240	4:3, 16:9	11:7,5		
SIF (NTSC)		360 × 240	4:3	3:2	29,97	
SIF (PAL)		360 × 288	4:3	5:4	25	
CIF	Common Intermediate Format	352 × 288	4:3, 16:9	11:9	29,97	
4CIF	2×2 CIF	704 × 576	4:3	11:9	29,97	
9CIF	3×3 CIF	1056 × 864	4:3	11:9	29,97	
16CIF	4×4 CIF	1408 × 1152	4:3	11:9	29,97	
TV		720 × 576	4:3, 16:9	5:4		
PAL	Phase Alternating Line					
D-1 PAL	D-1 PAL	768 × 576	4:3	1:1	25	ja
(PAL)		352 × 576	4:3, 16:9	11:18	25	ja
(PAL)		480 × 576	4:3, 16:9	5:6	25	ja
(PAL)		544 × 576	4:3, 16:9	17:18	25	ja
PAL-Wide	PAL Widescreen	1024 × 576	16:9		25	ja
NTSC	National Television Standards Committee	720 × 480	4:3, 16:9	3:2	29,97	ja
D-1 NTSC	D-1 NTSC	720 × 540	4:3		29,97	ja
(NTSC)		352 × 480	4:3, 16:9	11:15	29,97	ja
(NTSC)		480 × 480	4:3, 16:9	1:1	29,97	ja
(NTSC)		640 × 480	4:3		29,97	ja
(NTSC)		544 × 480	4:3, 16:9	17:15	29,97	ja
HDTV 720	High Definition Television 720 lines	1280 × 720	16:9		50/60	nein
HDTV 1035	HDTV 1035 lines	1920 × 1035	16:9	> 1,855	50/60	nein
HDTV 1080	HDTV 1080 lines	1920 × 1080	16:9		50/60	nein
HDTV 1152	HDTV 1152 lines	1440 × 1152		5:4	50/60	nein
2K		2048 × 1080		> 17:9		
		2048 × 1536		4:3		
DCI 2K	Digital Cinema Initiatives 2000	2048 × 1556		< 4:3		
		3840 × 2160	16:9			
4K		4096 × 2160		> 17:9		
DCI 4K	DCI 4000	4096 × 3112		< 4:3		
UHDV	Ultra High Definition Video	7680 × 4320	16:9		60	nein

KODIERUNG UND KOMPRIMIERUNG

Redundanz-Reduktion

- Verzicht auf mehrfach vorhandene Information(en)
- auf Signalstatistik beruhend
- i.d.R. verlustfrei
- Beschreibung von Gruppeneigenschaften (Muster)

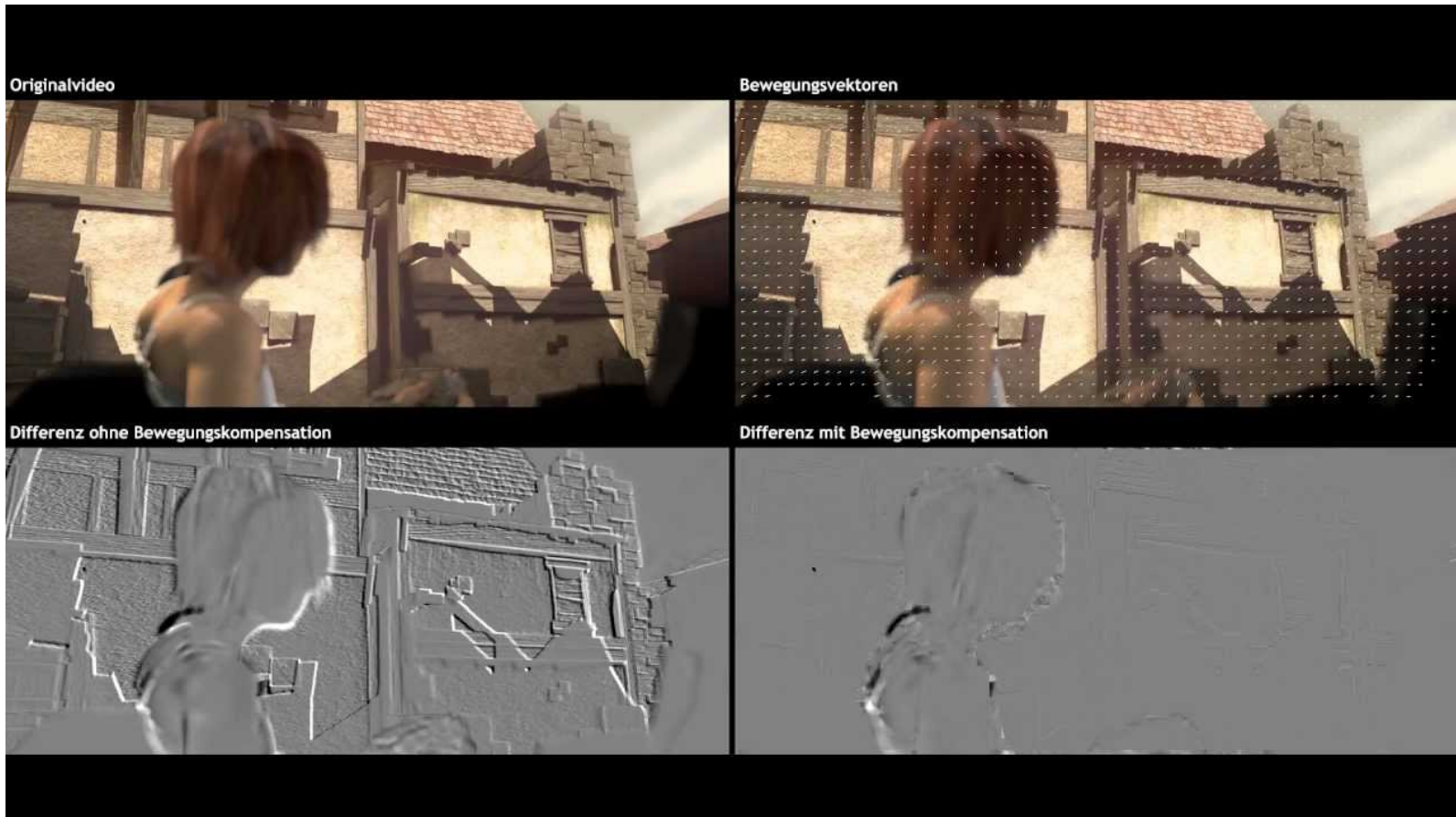


Irrelevanz-Reduktion

- Verzicht auf kaum erkennbare Signalanteile
- auf Wahrnehmungsphysiologie beruhend
- verlustbehaftet

KODIERUNG

Wie bei Audiodaten, ist es auch bei Videodaten sinnvoll, das nächste Bild aus den vorigen Bildern vorherzusagen und nur die Differenz zur Vorhersage zu kodieren.



©2016 CC-BY-SA 3.0 DE, Bernd Schöler
<https://encodingwissen.de/hintergrund/grundlagen/video-kompression/bewegte-differenzen/>

H.261 UND H.263 (1/2)

H.261 (1984-1990)

- Videokompression für Video-Telefonie über ISDN-Kanalbündelung (je Kanal 64 kBit/s)
- weniger als 150ms Verzögerung durch Kompression und Dekompression
- Videoformat: CIF/QCIF, $Y C_R C_B$ 4:2:0
- Kompression auf 1,5 MBit/s \approx 24 gebündelte ISDN-Kanäle

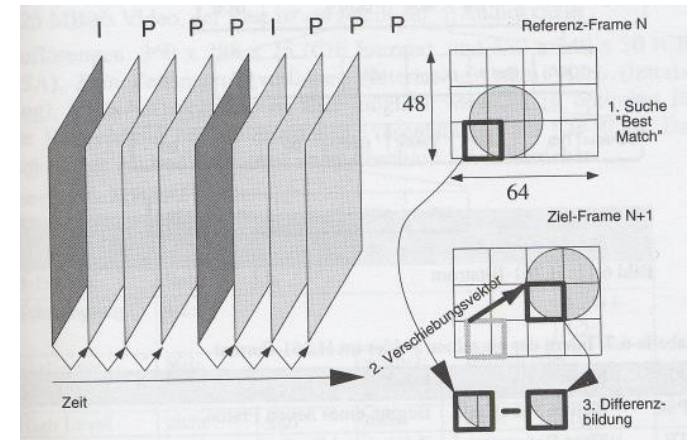
H.263

- Effizienzsteigerung durch
 - Weglassen unwichtiger Teile und
 - Verringerung der Bittiefe von Zusatzinformation
- ergibt ungefähr halbe Datenrate

H.261 UND H.263 (2/2)

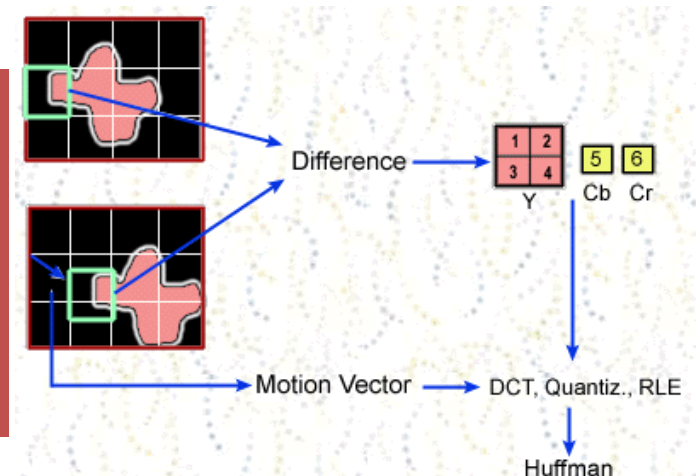
I Intra- Frames (JPEG)

- Unterteilung in 16x16 px große Blöcke
- darin vier Subblöcke für Helligkeit
- zwei Blöcke für Chroma
- DCT, anschließend Huffman



P Predicted Frames

- Unterteilung in 16x16 px Blöcke
- für jeden Block Suche nach „Best Match“ in vorherigem dekodierten Frame
- dann Differenzkodierung



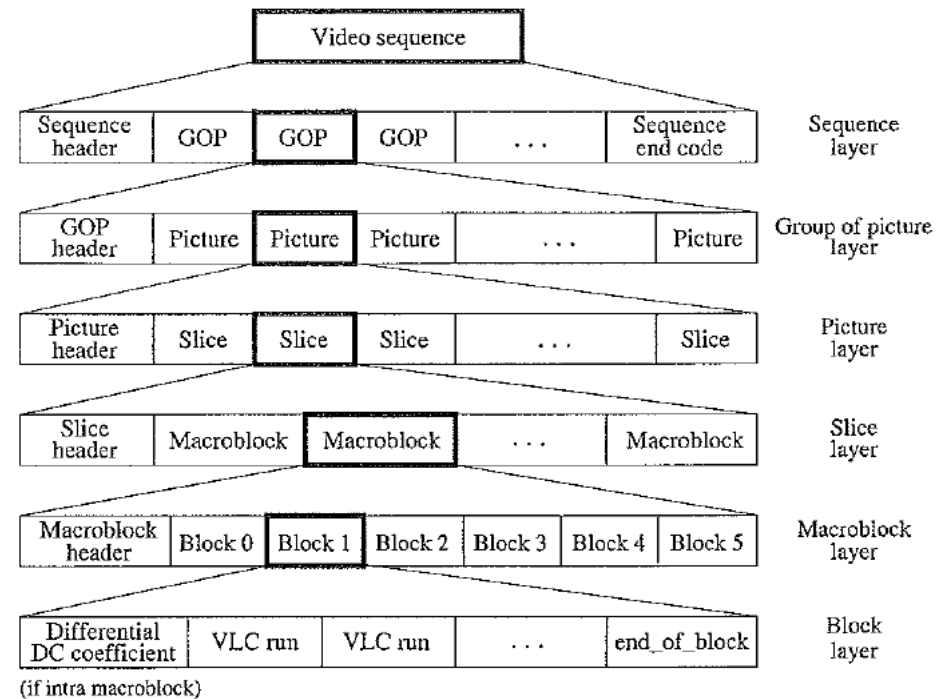
MOVING PICTURES EXPERTS GROUP (MPEG)

weit verbreitete Standards für die komprimierte Speicherung von digitalem Video mit Audio

- MPEG-1 (1992): Video auf Speichermedien (nur progressiv)
- MPEG-2 (1993): MPEG-1 und TV sowie HDTV
- MPEG-3 (1993): HDTV (in MPEG-2 aufgenommen)
- MPEG-4 (1998): Interaktives Multimedia, Audiovisuelle Objekte (seit 1998 in 33 Parts laufend erweitert, zuletzt 2015; enthält auch MPEG-4 AVC in Part 10 (H.264))
- MPEG-7 (2001): MM Description Interface
- MPEG-21 (2003): Digital Multimedia Framework
- MPEG-H (2013): High Efficiency Coding and Media Delivery in Heterogeneous Environments (enthält in Part 2 MPEG-4 HEVC (H.265); zuletzt 2019 aktualisiert)
- MPEG-I (vrs. 2020): Coded Representation of Immersive Media (enthält in Part 2 MPEG-4 VVC (H.266))

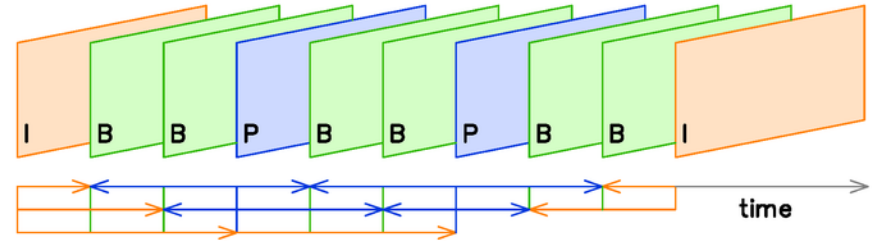
VOKABULAR

- **Sequence:**
alle Pictures vom *Sequence Header* bis zum *Sequence End Code*
- **Group of Pictures (GOP):**
alle Pictures zwischen 2 GOP-Headern (12/25 bzw. 15/30 fps)
- **Picture:**
Field (Halbbild) oder Frame (Vollbild)
- **Macroblock:**
16x16 Pixel
- **Slice:**
16 Zeilen x (n x 16 Pixel)
- **Block:**
8x8 Elemente (Luminanz, Chrominanz oder DCT-Koeffizient)
- **Sample:**
einzelner Helligkeits- oder Farbwert eines Pixels



MPEG-1 (H.262)

- Zusätzlich zu I- und P-Frames werden B-Frames eingeführt, die bidirektional vorhergesagt werden
- I-Frames wieder nach JPEG-Prinzip
- P-Frames werden aus Vorgänger I/P-Frame wie bei H.261 kodiert
- B-Frames werden aus Vorgänger I/P-Frame und übernächstem Nachfolger B/P-Frame oder Nachfolger I-Frame kodiert (Zweirichtungsdelatkodierung)
- dadurch Umsortierung bei Streaming notwendig



Darstellungsreihenfolge													
Bild	1	2	3	4	5	6	7	8	9	10	11	12	13
Typ	I	B	B	P	B	B	P	B	B	P	B	B	I
Übertragungsreihenfolge													
Bild	1	4	2	3	7	5	6	10	8	9	13	11	12
Typ	I	P	B	B	P	B	B	P	B	B	I	B	B
	GOP			GOP			GOP			GOP			

MPEG-2

Unterschiede zu MPEG-1

- Best Match-Suche nicht nur nach Frames, sondern auch nach Fields
- Makroblöcke mit Farb-Subsampling von 4:2:2 oder 4:4:4
- Frame-Größe bis zu 16.383x16.383 Pixel
- nichtlineare Quantisierungstabelle
- verschiedene Profile
 - räumliche Skalierbarkeit bedeutet, dass derselbe Datenstrom in verschiedenen Auflösungen angezeigt werden kann
 - zeitliche Skalierbarkeit bedeutet, dass der Datenstrom adaptiv auf Veränderungen der Bandbreite reagiert

ÜBERBLICK DER STANDARDS

	MPEG-1	MPEG-2	MPEG-4
Haupt-anwendung	CD-ROM, Digital Storage Media	Broadcast, DVD, HDTV	Webauthoring, Multimedia, Compression, Wireless Mediaphone
Video-Bitraten	1,5 Mbps	4,6 Mbps	20 Kbps - 6 Mbps
Bildauflösung	352x240 720x480	176x144 352x288 1280x720	176x144 352x288 720x480 1280x720
Audio-Qualität	Stereo CD	Surround Sound	Sprache, Musik, CD, HD-Surround

H.264 UND H.265

H.264 (MPEG-4 AVC)

- statt DCT auf 8×8-Blöcken findet Integertransformation auf 4×4-Blöcke statt
- neben Huffman-Kodierung auch leistungsfähigere arithmetische Kodierung
- Makroblöcke können in unterschiedliche große Subblöcke unterteilt werden
 - eigener Bewegungsvektor für jeden Subblock
 - räumlich schärfer abgegrenzte, komplexe Bewegungen besser kodierbar
- Bildvorhersage auch innerhalb von I-Frames möglich
- P- und B-Frames können auf bis zu 16 vorhergehende (beliebig weit zurückliegende) Frames referenzieren (Long-Term Prediction)
 - effizientere Kodierung von periodischen Bewegungen

H.265 (HEVC, MPEG-4 VVC)

- variable Block- und Macroblock-Größe
- bessere Kodierung von Transformationen

VERGLEICHSRECHNUNG

Folge einer Fernsehsendung im US-Werbeblock-Format (42 Minuten) in Full-HD bei 29,97 fps mit Dolby Digital 5.1 bei 640 kb/s:

- unkomprimiert: 437,56 GB
- MPEG-1: 180,22 GB
- MPEG-2 (CBR): 60,72 GB
- H.264 (VBR): 15,18 GB
- H.265 (VBR): 7,79 GB
- H.266 (Draft): 3,98 GB