

Audio- und Videotechnik Zusammenfassung

1 Signale

- **Hörschwelle:** Geringster bei 1kHz wahrnehmbarer Schalldruck. ($p = p_0 \Rightarrow L_p = 0\text{dB}$)
- **Schwingung Auslöschung:** Frequenz, gleiche Amplitude, 180° Phasenverschiebung.
- **Zeilensprungverfahren:** Bild setzt sich aus zwei aufeinanderfolgenden Halbbildern zusammen. Halbbilder werden zeitlich versetzt aufgenommen. Auftreten von Kamm-Effekten bei falschen Halbbilddominanz. Halbierung der Zeilenauflösung.
- **De-Interlacing:** Aufhebung des Zeilensprungverfahrens \rightarrow algorithmisch. Einsatz: Wenn interlaced (per Zeilensprung) aufgenommene Bilder progressiv wiedergegeben werden sollen. (z.B. Computerdisplay).
- **Schalldruckpegel:** Formel: $20 \cdot \log_{10}(x)$, x ist Faktor der Schalldruckveränderung \rightarrow bspw. bei $+12\text{dB} = 20 \cdot \log_{10}(4)$ also eine Vervierfachung bei einer Veränderung von $+12\text{dB}$.

2 Bildwandler

- **CCD-Bildwandler:**
 - Full-Frame Bildwandler
 - Frame-Transfer Bildwandler
 - Interline-Transfer Bildwandler
 - Frame-Interline-Transfer Bildwandler

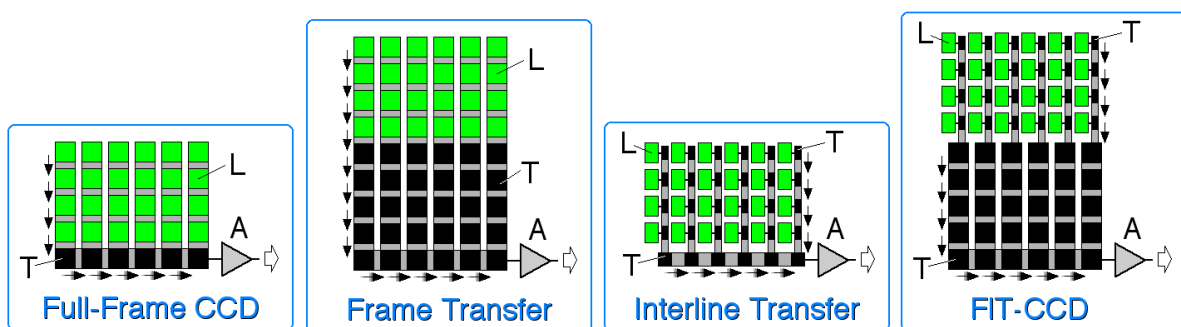


Abbildung 1: CCD-Bauweisen

- **CMOS:** Gegenüber CCD-Sensoren besteht bei CMOS-Sensoren der Vorteil, dass die Elektronik direkt das Spannungssignal jedes einzelnen Pixels auslesen kann, ohne die Ladungen verschieben zu müssen.
- **Halbleiter-Bildwandler:** Die Bildsensoren sind helligkeitsempfindlich und nutzen den inneren photoelektrischen Effekt. Sensoren basieren auf Photodioden.

3 Digitalisierung

- **Digitalisierung** eines analog-kontinuierlichen Signals:
 - **Abtastung** (Sampling): Senkrechtes Raster über das analoge Signal legen. Messung des Signals zu unendlich vielen Zeitpunkten → diskontinuierliches-analoges Signal
 - **Quantisierung**: Waagrechtes Raster über das analoge Signal legen. Wertezuweisung der abgetasteten Werte aus einer endlichen Wertemenge → diskontinuierliches-diskretes Signal

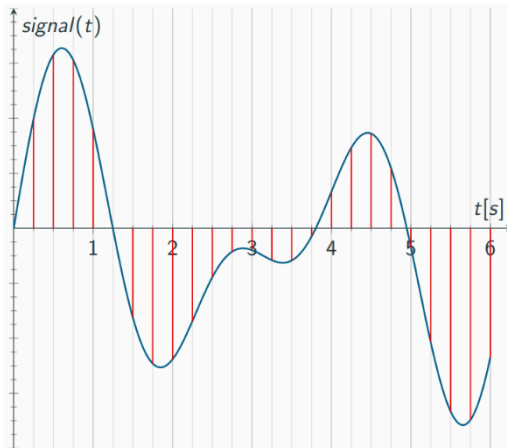


Abbildung 2: Abtastung

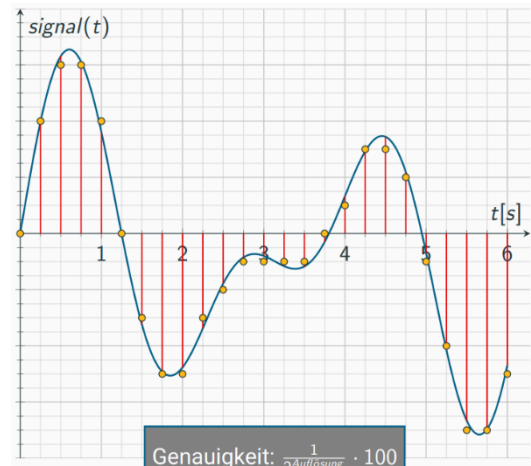


Abbildung 3: Quantisierung

- **Nyquist-Shannon Abtasttheorem**: Abtastfrequenz $> 2 \cdot$ maximale Frequenz des Signals, um Informationsverlust zu vermeiden.
- **Color Subsampling**
Luminanz Y: Lediglich andere Darstellung für RGB Werte.
 - 4:4:4 → $3 \cdot n$ Bit je Pixel (**Jeder** Pixel mit voller Farbinformation)
 - 4:2:2 → $2 \cdot n$ Bit je Pixel (Jeder **zweite** Pixel mit voller Farbinformation)
 - 4:2:0 → $1,5 \cdot n$ Bit je Pixel
 - 4:1:1 → $1,5 \cdot n$ Bit je Pixel (Jeder **vierte** Pixel mit voller Farbinformation)
- **Bitrate**: in bit / Byte pro Sekunde bspw. kb/s oder MB/s
- **Datenmenge**:
 - BZ = Bildzeilen; Asp = Bildseiten-Verhältnis; Ft = Farbtiefe
 - 30 min, 3 Kanal, 4:2:0, 50fps, progressiv → $BZ \cdot BZ \cdot Asp \cdot Ft \cdot 3 \cdot 50 \cdot 60 \cdot \frac{1}{2}$
 - 1 min, 3 Kanal, 4:2:2, 50fps, interlaced → $BZ \cdot BZ \cdot Asp \cdot Ft \cdot 2 \cdot 25 \cdot 60$
- **Quantisierungsfehler**: Je **geringer** der Signalpegel ist, desto **stärker** wirken sich Quantisierungsfehler aus. Quantisierungsfehler sind kleiner, je feiner die Auflösung. Bspw. Beeinflussung der Anzahl der Farb- / Helligkeitswerte.

4 Datenkompression

Entropiekodierung	RLE (Lauflängenkodierung), Huffman-Kodierung, Arithmetische Kodierung...	Entropiekodierung (Entropy Coding) <ul style="list-style-type: none">• Keine Datenanalyse• Verlustfrei (lossless compression)								
Quellenkodierung	<table><tr><td>Prädiktion</td><td>DPCM, Delta Modulation</td></tr><tr><td>Transformation</td><td>FFT, DCT</td></tr><tr><td>Layered Coding</td><td>Subsampling, Subband Coding</td></tr><tr><td colspan="2">Vektor Quantisierung</td></tr></table>	Prädiktion	DPCM, Delta Modulation	Transformation	FFT, DCT	Layered Coding	Subsampling, Subband Coding	Vektor Quantisierung		Quellenkodierung (Source Coding) <ul style="list-style-type: none">• Datenanalyse und Semantik• Verlustbehaftet (lossy compression)
Prädiktion	DPCM, Delta Modulation									
Transformation	FFT, DCT									
Layered Coding	Subsampling, Subband Coding									
Vektor Quantisierung										
Hybridkodierung	JPEG, HEIF, MPEG, H.261, H.264, H.265 Proprietäre Verfahren: Google, Sorensen, Real, Microsoft	Hybride Verfahren (Hybrid Coding) <ul style="list-style-type: none">• Entropie- und Quellenkodierung								

- **Redundanzelimination:** Informationen, die aus räumlichen oder statistischen Informationen / Abhängigkeiten abgeleitet werden. Verlustlos.
- **Irrelevanzelimination:** Nicht wahrnehmbare oder unwichtige Informationen. Verlustbehaftet.
- **Quellenkodierung:** Orientierung an der menschlichen Wahrnehmung. Elimination von irrelevanten Daten. Verlustbehaftete Kodierung. Trennung zwischen relevanten und irrelevanten Daten.
- **Entropiekodierung:** Quelldaten und dekodierte Daten sind gleich (**verlustlose** Kodierung). Ausnutzung der statistischen Eigenschaften der Daten.
- **Entropiekodierung** mit Basis auf **statistischen** Eigenschaften:
 - Arithmetische Kodierung
 - Huffman-Kodierung
- **Huffman-Kodierung:**
 - Wahrscheinlichkeiten der Größe nach sortieren
 - Code-Baum aufstellen, beginnend mit der kleinsten Wahrscheinlichkeit
 - 1 und 0 an die Kanten schreiben
 - Kanten verfolgen und Kodierung erstellen, bsp. 101011
 - **Mittlere Codewortlänge:** Summe von allen Wahrscheinlichkeiten * Bit-Länge
- **Irrelevanzelimination:** Quantisierung, Color-Subsampling, Motion Compensation
- **DCT-Transformation**
 - Intervallschift
 - Anwendung der DCT
 - Rundung (Quantisierung)
 - RLE
- **RLE:** sehr gut komprimierbar Bsp.: 999999999999999911111111

- **LZW**: sehr gut komprimierbar Bsp.: 12121212121212121212
- **Zick-Zack-Reihenfolge**: Der Koeffizienten-Block wird in einer Zick-Zack-Reihenfolge durchlaufen. Hierdurch werden besonders lange Folgen von Nullen erreicht. Experimente haben gezeigt, dass die Wahrscheinlichkeit eines nicht-Nullwertes (AC-Koeffizient) mit höheren Indizes in Zick-Zack-Reihenfolge abnimmt.
- **JPEG**: Hybride Kodierung

5 Videokodierung

- **I-Frame**: Intraframe → Kodierung ohne Referenz auf andere Bilder, Intrakodierung
- **P-Frame**: Predictive Frame → Kodierung mittels vorhergehendem Referenzbilds
- **B-Frame**: Bidirectional Frame → Kodierung mittels vorhergehendem und nachfolgendem Referenzbilds
- **GOP**: (Group of Pictures) Gruppe von Bildern, bestehend aus I-, B-, P-Frames. Beginnend immer mit I-Frame. Reihenfolge wird beim Encoding festgelegt.

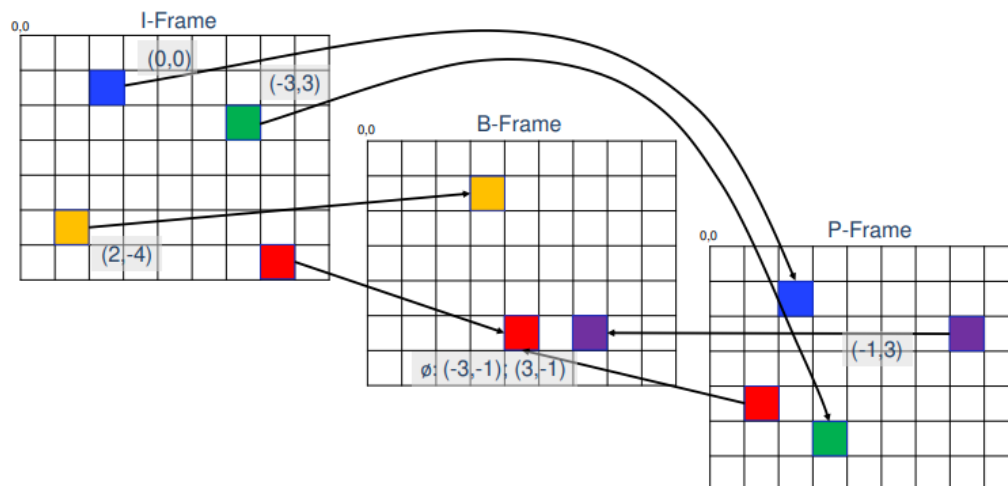


Abbildung 4: Vorwärts- und Rückwärtsreferenzierung

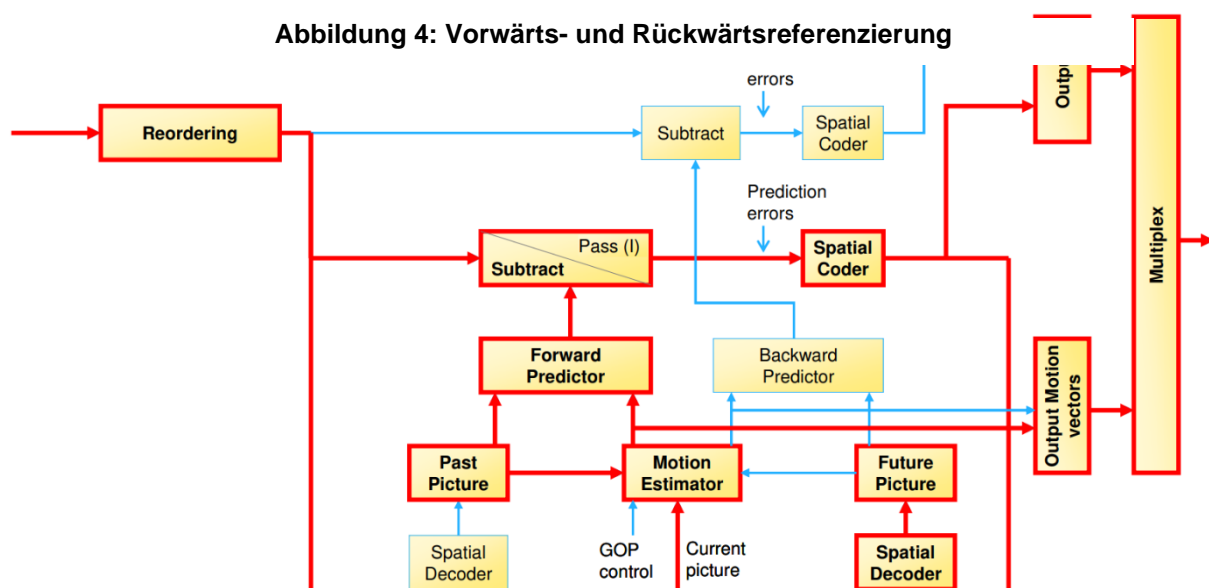


Abbildung 5: Bidirektionaler Kodierer - P-Frame

- **Kompression eines P-Frames:** Die Kodierung des P-Frames erfolgt unter Berücksichtigung des vorhandenen I-Frames. Dieses steht in dem **Past Picture** – Speicher zur Verfügung und im **Motion Estimator** erfolgt eine Bewegungsschätzung. Als Ergebnis liegt dabei anschließend pro Makroblock ein Bewegungsvektor (sofern Schätzung bzgl. Bewertungsfunktion ein befriedigendes Ergebnis erzielt). Der **Forward Predictor** nutzt diesen Vektor um das I-Frame zu verschieben. Das geschätzte Bild wird dann vom aktuellen Bild **subtrahiert** (Schätzfehler). Der Schätzfehler wird räumlich komprimiert (**Spatial Coder**). Das P-Frame wird als Schätzfehler und Menge von Bewegungsvektoren in den Datenstrom geschrieben (**Output Motion Vektors**). Analog wie im vorhergehenden Schritt wird auch das P-Frame lokal dekodiert und in einen Bildspeicher geschrieben (**Future Picture**), welcher zum Zweck der B-Frame Kodierung genutzt wird.
 - **Recording:** Bilder Werden für bidirektionale Kodierung ggf. umgeordnet.
 - **Subtract/Pass:** Ermittlung des Schätzfehlers. Subtraktion der Schätzung des aktuellen Bilds.
 - **Spatial Coder:** Intraframe-Kodierung des Schätzfehlers. DCT und Quantisierung Subtract.
 - **Spatial Decoder:** Interne Decodierung des Schätzfehlers. Dadurch erhält der Bildspeicher exakt die Daten, die später auch im Decoder vorliegen.
 - **Past Picture:** Bild-Speicher für vorhergehendes Referenzbild
 - **Future Picture:** Bild-Speicher für nachgehendes Referenzbild
 - **Motion Estimator:** Bewegungsabschätzung aus aktuellem Bild und vorhergehendem Referenzbild. Ergebnis: je ein Bewegungsvektor pro Makroblock
 - **Forward Predictor:** Erzeugt aus dem vorhergehenden Referenzbild und dem Bewegungsvektor aus dem Motion Estimator eine vorwärtsgerichtete Bewegungsabschätzung (für ein P-Frame oder B-Frame) Backward Predictor
 - **Output Motion Vectors:** Schreibt die Bewegungsvektoren in den Multiplexer
 - **Output:** Schreibt die Schätzfehler in den Multiplexer
 - **Multiplex:** Erzeugt den Datenstrom, Multiplex von Bewegungsvektoren und Schätzfehlern.
- **CBR und VBR:** (Konstante Bitrate, Variable Rate) VBR liefert eine bessere Bildqualität bei dynamischen Inhalten. VBR optimiert die Bandbreite hinsichtlich einer konstanten Kodierqualität.
- **Datenvolumen für eine Videoübertragung:** in bspw. kBit/s oder Byte/s

- **Radrotation:**
 - Rad: 6 Hz, Framerate: 24 Hz → Stillstand (Framerate / Radfrequenz glatt teilbar)
 - Rad: 11 Hz, Framerate: 24 Hz → Schnelle Vorwärtsdrehung (Doppelte Radfrequenz in Nähe von Framerate, aber kleiner)
 - Rad: 13 Hz, Framerate: 24 Hz → Eindruck einer Rückwärtsdrehung (Doppelte Radfrequenz in Nähe von Framerate, aber größer)
 - Rad: 25 Hz, Framerate: 24 Hz → Langsame Vorwärtsdrehung (Radfrequenz minimal größer als Framerate)

6 Video-Standards / MPEG

- **MPEG-Videostandards:** Beschreiben die Datensyntax und Semantik des Bitstroms und sorgen für Kompatibilität beim Dekodieren.
- **MPEG-2:** Digital versatile Disk, Digitales TV – Standardauflösung (720x576)
- **Hierarchische Kodierung:** Wird bei variierendem Quality-of-Service eingesetzt. Base-Layer Daten werden geschützt als Enhancement-Layer Daten.
- **H.264/AVC:** Prädikation im Ortsbereich. Deblocking-Filter für Verringerung von Block-Artefakten. Einsatz: Digitales TV – HD-Auflösung; Blu-Ray Disk.
- **H.264/AVC Veränderungen zu Vorgängern**
 - Verbesserte Motion Compensation
 - Prädikation im Ortsbereich / Räumliche Intraframe-Prädikation
 - kleinere Blöcke für die Transformationskodierung
 - 4x4 Integer Block Transform Deblocking
 - Verbesserte Entropiekodierung
- **H.265/HEVC:** (High Efficiency Video Coding) Color-Subsampling: 4:2:0, Farbtiefe 8 Bit; Typische Auflösungen von Filmmaterial: 2k - DVB T2, 4k - UHD, 8k

7 Audiokompression / Psychoakustisches Modell

- **Zeitliche Maskierung**
 - **Vorverdeckung:** Liegen Maskierungsimpuls und Druckimpuls sehr dicht beieinander (<20ms), so ist eine Vorverdeckung feststellbar. → Laute Impulse werden vom Gehirn schneller verarbeitet als leises Rauschen. Der laute Impuls überholt also schwächere Signale.
 - **Nachverdeckung:** Spielt man einen lauten Ton, der dann plötzlich stoppt, so dauert es eine Weile, bis man leise Klänge wieder wahrnehmen kann. In Abhängigkeit der Amplitude des Maskierungstons ist die Amplitude des Druckimpulses der Maskierungsschwelle bis ca. 10 ms nach Abschalten des Maskierungstons

unverändert. Erst beim Abfallen der Hörschwelle. Nach ca. 200ms wieder Ruheshwelle erreicht.

- **Frequenzmaskierung**

- **Verdeckung:** laute Töne einer bestimmten Frequenz können leisere Töne einer anderen Frequenz verdecken, so dass diese nicht mehr gehört werden können.
- **Maskierungstöne:** geben an, bei welchem Hörpegel Sinustöne gerade noch wahrnehmbar sind, wenn man sie mit einem gleichbleibenden Ton fester Frequenz und Lautstärke hört.
- Horizontale Achse: Frequenzvariation des maskierten Signals
- Vertikale Achse: Zeigt, um wieviel dB der Intensitätspegel des Signals über der Hörschwelle liegen muss, damit man es auch mit Maskierungston noch hört.
- Beispiel, Maskierungston 1200 Hz: Ein Signal, dessen Hörpegel z.B. unterhalb von der Kurve für 80 dB bleibt, wird von einem 1200 Hz Maskierungston mit 80 dB Hörpegel völlig verdeckt.

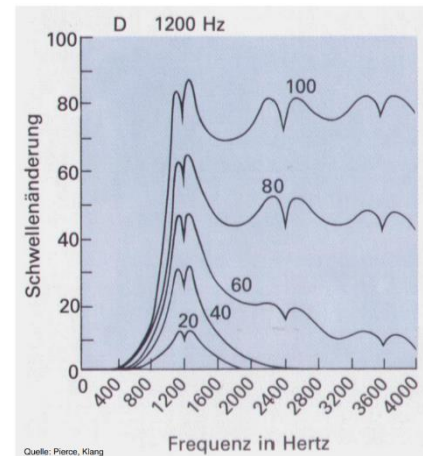


Abbildung 6: Maskierungston 1200 Hz

- **Nutzung des Psychoakustischen Modells:** Über ein psychoakustisches Modell wird die Quantisierung gesteuert. Es wird gerade so quantisiert, dass das dann entstehende Quantisierungsrauschen gerade unterhalb der Hörschwelle bleibt! Anwendung jeweils innerhalb der kritischen Frequenzbänder.
- **Sprachkodierung Vocoder-Verfahren:** Schlechte Sprachqualität bei sehr kleinen Bitraten. Sprachqualität ist auch bei großen Bitraten nur mittelmäßig.
- **G.722:** Zerlegung des Eingangssignals in zwei Frequenzbereiche. Auflösung Bit/Sample ist in allen Frequenzbändern gleich.
- **Kritische Frequenzbänder:** Schallreize werden zusammengefasst wahrgenommen. Im Bereich 0-16 kHz lassen sich 24 kritische Frequenzbänder (Frequenzgruppen) identifizieren. Ein Maskierungston aus einem kritischen Band wirkt vorrangig innerhalb seines Bands maskierend.
- **Wichtige Eigenschaften für Audiokompression:** Zeitliche Maskierung, Adaptive Quantisierung, Frequenz Maskierung

8 Audio-Standards / MPEG

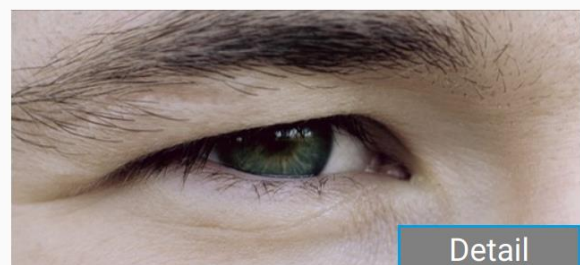
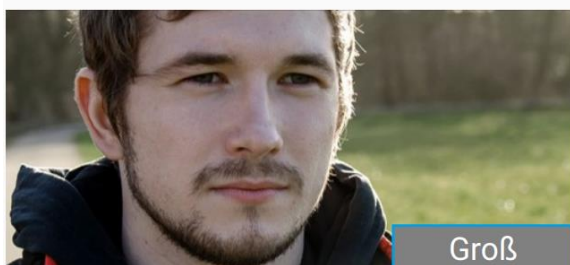
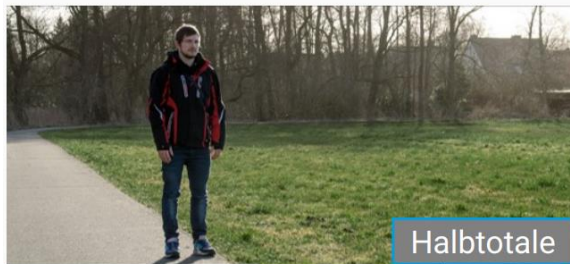
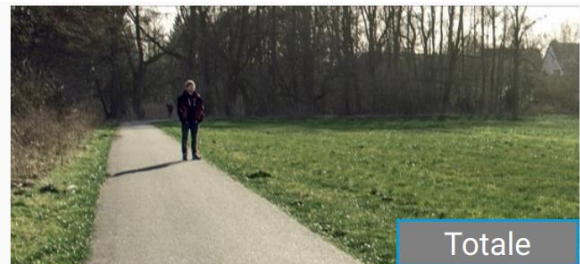
- **AAC**: Standardisiert in MPEG-2 Part 7. Nutzt das psychoakustische Modell für die Komprimierung
- **Räumliche Wahrnehmung**: Interaurale Pegeldifferenz (ILD), Interaurale Zeitdifferenz (ITD)
- **MPEG-1**: PCM mit 32, 44.1 oder 48 kHz; max. Datenrate 448 kbit/s
- **MPEG-2**: PCM mit 16, 22.05, 24, 32, 44.1, 48 kHz; max. 5.1 Kanäle; max. Datenrate 384 kbit/s
- **MPEG-D**: Technologie für Multikanal Audiokompression. Stereo Downmix des Signals vor der Kodierung.

9 Produktion / Beleuchtung

- **Ausarbeitungen**: Exposé → Treatment → Drehbuch → Storyboard
- **Primärschatten**: Sichtbarkeit der Objektstruktur. Hervorhebung der Form des Objektes
- **Sekundärschatten**: Abstand des Objektes vom Boden und Hintergrund. Schaffung einer Beziehung zwischen Objekt und seiner Umgebung.
- **Weiches Licht**: Schattenlose Ausleuchtung. Helligkeitsabnahme in feinen Abstufungen.
- **3-Punkt Ausleuchtung**
 - **Key-Light** (Führungslicht): Grundlage der Beleuchtung, Hartes Licht, 4-5H / 1-2 V
 - **Fill-Light** (Aufhelllicht): Schwächung der Konturen, Kontrastreduktion, Aufhellen der Schattenbereiche, Weiches Licht, 7-8 H / 1-2 V
 - **Back-Light** (Gegenlicht): Trennung Objekt / Hintergrund, 11-1 H / 10-11 V
- **Farbtemperatur**: Beschreibt die Wärme (Rotanteil) bzw. die Kälte (Blauanteil) des Lichtes. Bspw. Hat bedeckter Himmel eine höhere Farbtemperatur als klares Sonnenlicht.
- **Weißabgleich**
 - Kunstlichtaufnahme → Weißabgleich auf Tageslicht = Rotstich
 - Tageslichtaufnahme → Weißabgleich auf Kunstlicht = Blaustich
 - Kunstlichtaufnahme → Weißabgleich auf Kunstlicht = neutral
 - Tageslichtaufnahme → Weißabgleich auf Tageslicht = neutral

		Einstellung im Aufnahmesystem	
		Tageslicht	Kunstlicht
Lichtverhältnisse	Tageslicht	neutral	blau
	Kunstlicht	rot / orange	neutral

- **Einstellgrößen**



- **Aufnahmewinkel / Kamera Perspektive**

