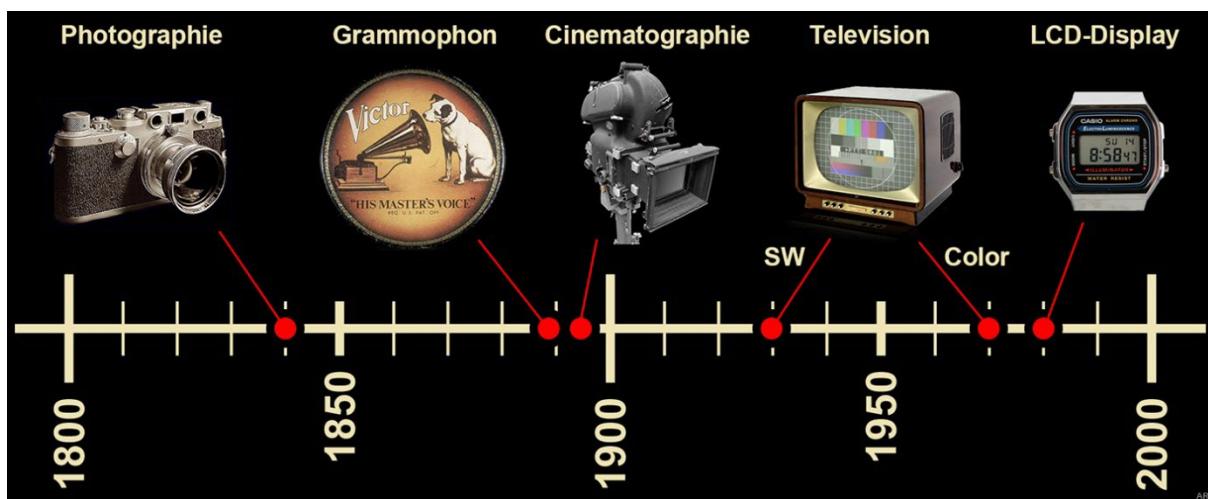




## BILDER CODIEREN & KOMPRIMIEREN THEORIE

Was möchten wir hier lernen?

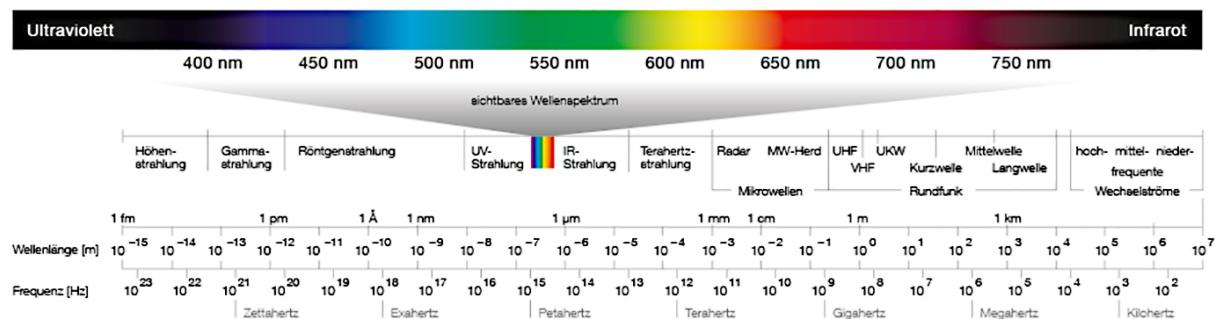
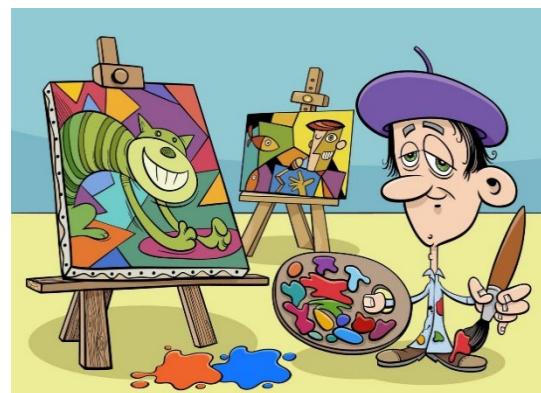
- Farbmodelle Additiv/RGB, Subtraktiv/CMYK
- Alphakanal bei PNG / Transparenzfarbe bei GIF / Animated GIF
- dpi und ppi
- Rastergrafik (Bitmap / Vektorgrafik)
- Speicherbelegung eines unkomprimierten Bildes (Berechnung)
- Bewegtbild (Video)
- Halbbildverfahren Interlaced / Vollbild Progressiv
- HD-Videonormen (HD720, HD1080)
- Codecs / Containerformate
- Speicherbelegung eines unkomprimierten Videos (Berechnung)
- Analog/Digitalwandlung
- Audioqualitätsvergleiche wav, mp3
- Bildkomprimierung (Einfache Verfahren)
- Farbmodell YCrCb (als Vorbereitung auf Subsampling)
- Bildkomprimierung mit Subsampling
- Bildkomprimierung mit DCT bei JPG
- Interframekomprimierung, GOP-Sequenz





## Wie kommt Farbe in unsere Welt?

- Elektromagnetische Wellen werden mit der Wellenlänge oder der Frequenz angegeben.
- Licht besteht aus elektromagnetischen Wellen.
- Nur ein Teil der Frequenzen des Lichts sind für das menschliche Auge sichtbar.

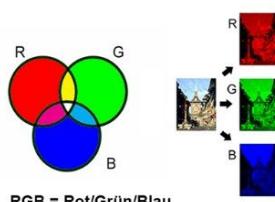


Trifft Licht auf einen Gegenstand, so werden gewisse Frequenzen absorbiert (aufgenommen) und andere reflektiert (zurückgeworfen): Wir sehen eine bestimmte Farbe.

Schwarz absorbiert alle Frequenzen. Weiss reflektiert alle Frequenzen. Rot zum Beispiel entsteht, indem die roten Anteile reflektiert und der Rest absorbiert werden.

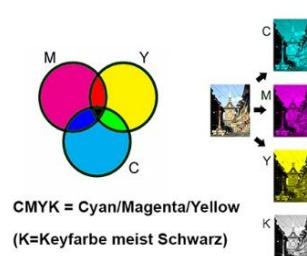
## Die Farbmodelle:

Additives Farbsystem



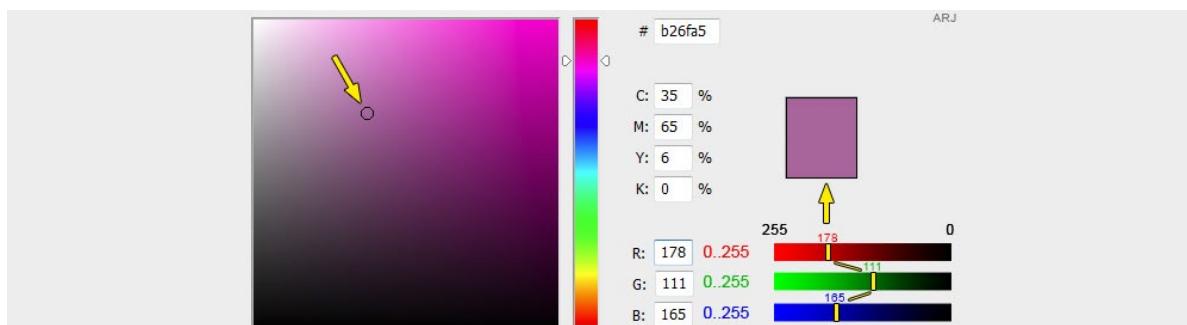
RGB = Rot/Grün/Blau

Subtraktives Farbsystem



CMYK = Cyan/Magenta/Yellow  
(K=Keyfarbe meist Schwarz)

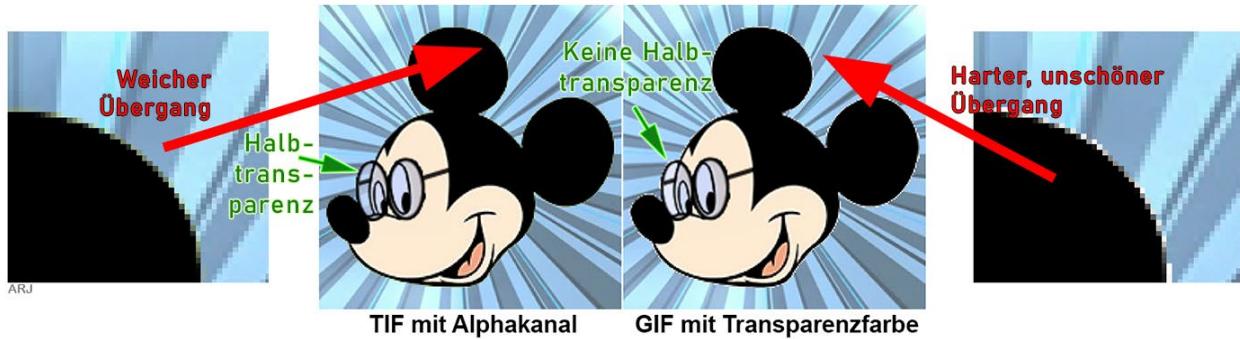
Screenshot einer Bildbearbeitungs-SW: Jedes Pixel enthält drei Subpixel R, G und B. Jedes Subpixel wird mit z.B. 8 Bit aufgelöst. Der hier abgebildete Farbwähler zeigt die Farbe des gewählten Bildpunktes in RGB 3x8=24b und auch CMYK-Werten an.





## Transparenzfarbe für die Bildmaskierung bei GIF:

Beim GIF-Format kann eine Farbe bestimmt werden, die nicht angezeigt wird sondern transparent zu einem Hintergrund wirkt. Damit lassen sich z.B. auf Webseiten Grafiken mit geschwungenen oder runden Formen auf Hintergründen platzieren. Da aber nur zwei Zustände existieren, nämlich durchlässig oder blockierend, werden schräge, nicht orthogonal angelegte Kanten treppenartig ausgeschnitten, was zu unschönen Rändern führt.

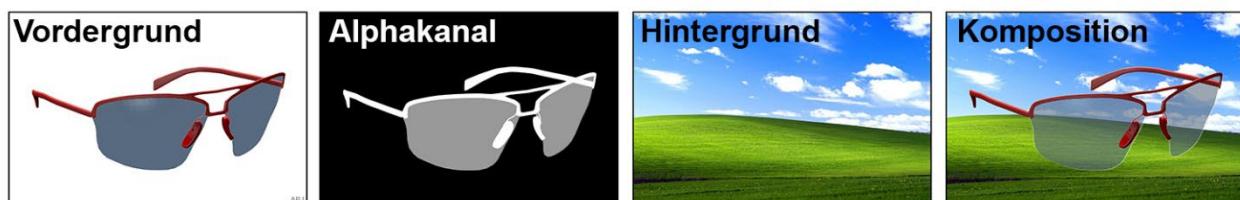


**Bemerkung zu GIF:** GIF und JPG waren anno 1987 die Pioniere in der Webgrafikwelt, weil sie ab der ersten Stunde von allen Webbrowsern unterstützt wurden. Die Transparenzfarbe beim GIF bot die Möglichkeit, Grafiken jenseits der Rechtwinkligkeit in Webseiten-hintergründe zu integrieren. Dieses Feature machte ihm im Jahr 1996 aber PNG mit seinem echten Alphakanal streitig. Und auch das letzte Alleinstellungsmerkmal, die Animation (Animated GIF), ging ab dem Jahr 2008 mit der Einführung von APNG (Animated Portable Network Grafic) verloren. APNG wird aber nur von aktuellen Webbrowsern unterstützt.

## Alphakanal - Ein zusätzlicher Kanal zur Bildmaskierung:

Eine Weiterentwicklung der Transparenzfarbe ist ein eigener Graustufenkanal, der die Lichtdurchlässigkeit bestimmt. Will man Bildkompositionen erstellen, braucht es eine Transparenzinformation zwischen den sich überlagernden Bildern. Unter einer Bildkomposition versteht man zum Beispiel ein rundes Firmenlogo auf einem tapetenartigen Hintergrund. Die Transparenzinformation liefert dabei z.B. ein sogenannter Alphakanal.

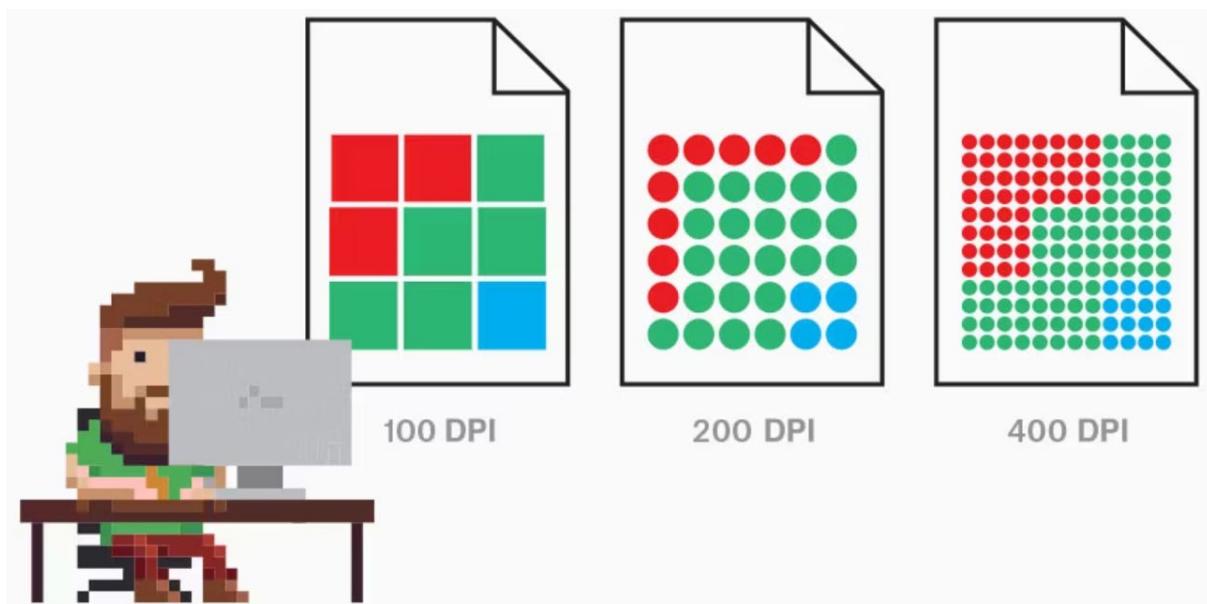
Falls in der Bilddatei ein solcher Alphakanal existiert, finden sich dort wie bereits erwähnt, eine zusätzliche Information, nämlich die der Transparenz. Der Alphakanal ist neben den drei Farbkanälen Rot/Grün/Blau also ein weiterer Kanal in Form eines Graustufenbildes, der in Rastergrafiken die Durchsichtigkeit der einzelnen Pixel (Bildpunkte) bestimmt.





## DPI und PPI:

- Mit **dpi (Dots per Inch)** meint man die Punktedichte bzw. Auflösung bei Druckern: Anzahl Farbpunkte pro Inch. (1 inch entspricht 2.54 cm). Zoll ist übrigens die deutsche Übersetzung für Inch. Gängige Druckerauflösungen starten bei 300dpi und gehen bis zu ca. 4000dpi. Die richtige Wahl der dpi's ist abhängig von der gewählten Papiergrösse und dem Verwendungszweck.
- Mit **ppi (Pixel per Inch)** meint man die Pixeldichte bei Displays. Ein modernes Apple-iPad besitzt bis zu ca. 260 ppi. Es stellt sich die Frage, ob das menschliche Auge in der Lage ist, mit seiner max. Winkelauflösung von einer Bogenminute eine solch hohe Auflösung bei kleinem Abstand zum Display auch aufzulösen. PC-Monitore werden üblicherweise nicht in ppi angegeben, sondern in Anzahl Pixel horizontal bzw. vertikal zusammen mit der Displaydiagonalen in Zoll oder Zentimeter



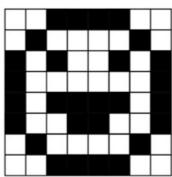


## Rastergrafik / Vektorgrafik: Unterschiede und Einsatzgebiete:

### Rastergrafik oder Bitmap



Originalbild



Bitmapbild

0	0	1	1	1	1	0	0
0	1	0	0	0	0	1	0
1	0	1	0	0	1	0	1
1	0	0	0	0	0	0	1
1	0	1	1	1	1	0	1
1	0	0	1	1	0	0	1
0	1	0	0	0	0	1	0
0	0	1	1	1	1	0	0

Bitmap Binär

3	C
4	2
A	5
8	1
B	C
9	9
4	2
3	C

Bitmap Hex

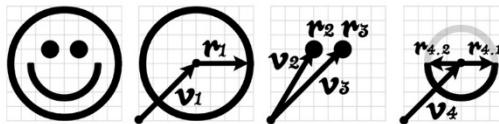


250x250Pixel

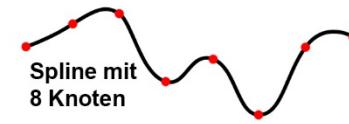


25x25Pixel

### Vektorgrafik



$v_1=(4/4) \quad r_1=4$   
 $v_2=(3/5) \quad r_2=0.5$   
 $v_3=(5/5) \quad r_3=0.5$   
 $v_4=(4/4) \quad r_4=2.5$   
 $r_{4.1}=(6.5/4) \quad r_{4.2}=(1.5/4)$



Spline mit  
8 Knoten

### Typische Anwendung von Vektorgrafiken → Schriftzeichen (Font)



Geometrische  
Beschreibung  
des Buchstabens  
(Vektorgrafik)



So sieht es aus,  
wenn der Buchstabe als  
Rastergrafik abgerufen  
und anschliessend stark  
vergrössert würde:  
Es entstehen  
Treppenartefakte.

- **Rastergrafik, Bitmap oder Pixelmap:** Durch z.B. ein Foto-Objektiv auf einen CCD-Sensor (CCD = Charge Coupled Device) oder CMOS-Sensor (z.B. als APS = Active Pixel Sensor) eingefangene Bildpunkte. (CMOS ist übrigens die Bezeichnung für spezielle und nicht nur lichterfassende Halbleiterbauelemente = Complementary Metal Oxide Semiconductor)
- **Vektorgrafik:** Eine Vektorgrafik wird aus grafischen Primitiven wie Linien, Kreisen, Polygonen oder allgemeinen Kurven (Splines) beschrieben. Da die Vektorgrafik verlustlos skalierbar ist, kann sie im Gegensatz zur Rastergrafik ohne entstehende Artefakte beliebig vergrössert werden. Bei Schrift-Fonts (z.B. TTF = TrueType Font) handelt es sich übrigens um Vektorgrafiken. Die Vektorgrafik beschreibt ein Bild mathematisch und muss bei jeder Ausgabe auf einen Bildschirm oder einen Matrixdrucker wie Inkjet- oder Laserprinter Pixel für Pixel in eine Rastergrafik der gewünschten Auflösung (Breite x Höhe) umgerechnet (gerendert) werden. Anders sieht das beim Ausgabegerät Stiftplotter aus, wo ein in einer Ebene aufgehängter und in x-Achse und y-Achse verschiebbarer Zeichenstift die Grafik zu Papier bringt: Dieser versteht prinzipbedingt direkt Vektordaten.
  - **TTF:** TrueType-Fontformat (.ttf)
  - **PostScript-Fontformate:** Type-1-Schriften (.pfm)
  - **OpenType:** TrueType- oder PostScript-flavoured (.ttf oder .otf)
  - **Fontssätze:** ARIAL, COURIER NEW, COMIC SANS MS, FUTURA, FRUTIGER, HELVETICA etc. (*Nicht alle Fontsätze beinhalten den kompletten ASCII- und schon gar nicht den vollständigen UTF8-Zeichensatz. Gewisse Fancy-Schriftarten beschränken sich auf die Zeichen A-Z, 0-9, inklusive dem Bindestrich -*)



## Wieviel Speicher belegt mein Bild?

Um sich eine Vorstellung machen zu können, welches Bildformat wieviel Speicher einfordert, wurde das folgende Regenbogenfarbenbild (1000 Pixel x 1000 Pixel) etwas genauer untersucht:



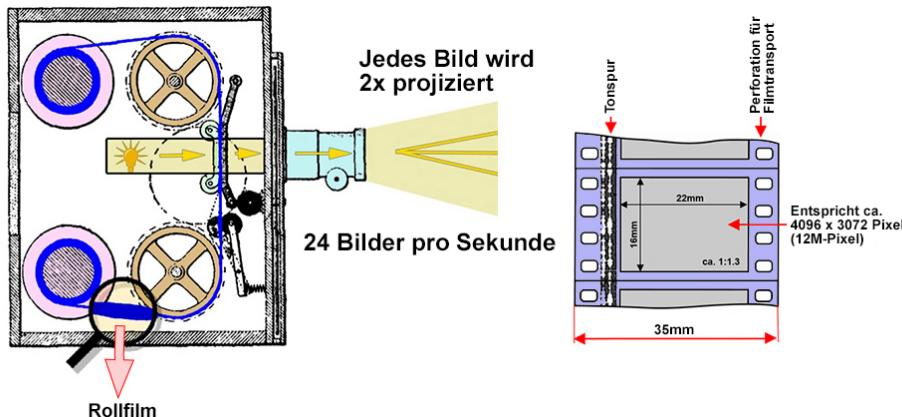
- **Theoretische Speichergrösse:**  $1000 \times 1000 = 1'000'000$  Bildpunkte. Das bedeutet pro Pixel (3 Byte):  $1'000'000 \times 3 = 3'000'000$  Byte.
- **GIF-Datei:** 71'903 Byte (Artefakte deutlich sichtbar, wegen Beschränkung auf 256 unterschiedliche Farben. Das Alleinstellungsmerkmal von GIF ist übrigens: Animated GIF)
- **PNG:** 356'476 Byte (keine sichtbaren Artefakte)
- **TIF:** 3'021'496 Byte (kein Informationsverlust. Das Archivformat schlechthin)
- **TIF mit LZW-Komprimierung:** 507'660 Byte (kein Informationsverlust)
- **JPG mit tiefster Kompressionsrate:** 271'201 Byte (keine sichtbaren Artefakte)
- **JPG mit hoher Kompressionsrate:** 32'810 Byte (Blockartefakte deutlich sichtbar)

Wer will, kann die sechs Bilder selber untersuchen:

<https://www.juergarnold.ch/Videotechnik/Samples.zip>

Bis jetzt haben wir immer nur Einzelbilder betrachtet. Reiht man mehrere Einzelbilder aneinander und projiziert sie in regelmässigen Abständen, nennt man dies Film bzw. Video.

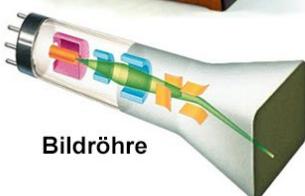
**Kino:** Den Kinofilm gibt es seit ca. 1894. Man hat sich beim klassischen Kino weltweit auf eine Bildwiederholfrequenz von 24 Vollbilder/sec. geeinigt. Auch die Digitalisierung hat beim Kinofilm Einzug gehalten. Heutige Filme werden komplett digital produziert und in den Kinos auch digital auf Beamern vorgeführt.





**Fernsehen:** Television oder auf Deutsch Fernsehen gibt es in der Schwarzweissversion seit ca. 1929 und in der Farbversion ab ca. 1967. Fernsehen wurde erst durch die Massentauglichkeit von Elektronenstrahlröhren bzw. Bildröhren möglich. Um das Bildflimmern auf ein erträgliches Mass zu reduzieren, müssen wie beim Kinofilm 50 Bilder pro Sekunde angezeigt werden. Dies war mit der damaligen Technik allerdings nicht möglich, weil die erforderlichen Bandbreiten fehlten und man mit vernünftigem Aufwand auch keine Bilder zwischenspeichern konnte. Darum bediente man sich eines Tricks: Man lieferte 50 halbe Bilder pro Sekunde und zwar einmal alle ungeraden Zeilen, gefolgt von den geraden Zeilen.

#### Interlaced / Halbbildverfahren:



50 Halbbilder pro Sekunde



Kammbildung beim Halbbildverfahren und schneller Bewegung.  
Milderung dieses Effekts bietet das sogenannte Deinterlacing.

- Interlaced Mode:** Es Halbbilder nacheinander (ungerade/gerade Zeile) gesendet und im Fernseher infolge der Trägheit der Bildröhren als Komplettbild wahrgenommen. Der Vorteil besteht darin, dass weniger Bandbreite erforderlich ist. Dem gegenüber steht der Nachteil des sogenannten Kammeffekts bei schnell bewegten Bildern, weil diese zeilenweise nacheinander aufgezeichnet werden.
- Progressive Mode:** Es werden komplett Bilder nacheinander empfangen. Vorteil dabei, man erhält immer ein komplettes Bild. Nachteil: Mehr Bandbreite erforderlich.

#### Eine Auswahl an aktueller Videonormen:

Norm	Auflösung	Progressiv	Interlaced	Format	Bitrate
<b>NTSC-SD</b>	640 x 480		60 fps	4:3	
<b>PAL-SD</b>	768 x 576		50 fps	4:3	
<b>ATSC-HD-720p60</b>	1280 x 720	60 fps		16:9	19 Mb/s
<b>ATSC-HD-720i60</b>	1280 x 720		60 fps	16:9	
<b>EBU-HD-720p50</b>	1280 x 720	50 fps			
<b>ATSC-HD-1080p60</b>	1920 x 1080	60 fps		16:9	25Mb/s
<b>ATSC-HD-1080i60</b>	1920 x 1080		60 fps	16:9	
<b>EBU-HD-1080p50</b>	1920 x 1080	50 fps		16:9	
<b>EBU-HD-1080i50</b>	1920 x 1080		50 fps	16:9	25Mb/s
<b>UHD-1 (4k)</b>	3840 x 2160	≤ 120 fps			10.2Gb/s
<b>UHD-2 (8k)</b>	7680 x 4320	≤ 120 fps			24 Gb/s

SD=Standard Definition / HD=High Definition / NTSC=National Television System Committee / PAL Phase Alternating Line / ATSC=Advanced Television System Committee / EBU=European Broadcasting Union / UHD=Ultra High Definition



### Fassungsvermögen und Datenrate von optischen Datenträgern:

- CD 1.4Mb/s (650 ... 900 MB)
- DVD 10Mb/s (4.7GB pro Layer bei max. 2 Layer pro Seite)
- BluRay 432Mb/s (25GB pro Layer bei max. 2 Layer pro Seite)

### Datenrate von Schnittstellen: (Auswahl)

- Mobile 5G ≤10Gb/s
- USB-C 20Gb/s
- HDMI 2.1 42.7Gb/s
- SAS-5 45Gb/s
- DisplayPort 2.0 77.37Gb/s

**Codecs und Containerformate:** Zwei Begriffe, die sie nicht verwechseln sollten.

- **CODEC:** Das Kofferwort "CoDec" bestehend aus Coder und Decoder bezeichnet ein Verfahren, das Daten oder Signale digital kodiert und dekodiert. Meist werden beim Kodievorgang die analogen Signale nicht verlustlos digitalisiert, sondern es wird eine Dynamikreduktion des analogen Signals sowie eine Datenkompression des digitalen Signals vorgenommen, die je nach Ausmass und Verfahren zu Qualitätsverlusten bei der Rückwandlung des digitalen Datenstroms in die analogen Signale führt. Bild und Ton können betroffen sein aber auch die Kontinuität der Wiedergabe.

Beispiele:

- **Video-Codecs:** MPEG-2, MPEG-4, AVCHD, H.261, H.263, H.264, H.265 Windows Media Video wmv und weitere.
- **Audio-Codecs:** MPEG1 Layer 2, MPEG1 Layer3 → MP3, Windows Media Audio wma, aac Advanced Audio Coding und weitere.

- **CONTAINER:** Container sind Behälter, der unterschiedliche Dateitypen enthalten kann. Mit dem Format wird die Art und Weise beschrieben, wie die Datenformate innerhalb des Containers angeordnet sind (innere Struktur). Audio/Videocontainerformate können zumindest einen Audio- und einen Videostream enthalten. Einige Formate ermöglichen zusätzlich die Einbettung von Untertiteln und Menüstrukturen oder anderen Inhalten.

Beispiele:

- **Container:** Microsoft AVI → Audio-Video Interleave, Apple Quicktime mov, DVD, MPEG-2 Stream, Blu-ray Disc und weitere.

Mögliche **Video-Codecs** für den **BluRay-Container**: H.264/MPEG4, AVC, VC-1, MPEG2

Mögliche **Audio-Codecs** für den **BluRay-Container**: Dolby-Digital, DTS (Digital Theater System High Definition), PCM (Pulse Code Modulation)



## Wie berechne ich den theoretisch benötigten Speicherplatz für einen Spielfilm?

Die Aufgabe besteht darin, den Speicherbedarf eines **unkomprimierten HD720i50**-Videos mit einer normalen Spielfilmlänge von **90 Minuten** zu berechnen. Der Audioanteil soll hier nicht berücksichtigt werden.

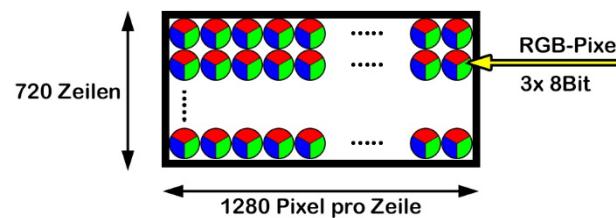
Zur Info: **HD720i50** bedeutet...

- **HD** bedeutet High Definition im Format 16:9
- **720** gibt die Anzahl Bildzeilen an. Daraus ergibt sich gemäss dem Format 16:9 die Anzahl Bildpunkte (Pixel) pro Zeile:  $720/9 \cdot 16 = 1280$
- **i50** bedeutet 50 Halbbilder pro Sekunde. Dies entspricht 25 vollen Bildern pro Sekunde.
- Jeder Pixel hat **drei Subpixel** Rot/Grün/Blau, wobei jedes Subpixel eine Farbtiefe von 8Bit aufweist. Der einfacheitshalber rechnen wir mit 3Byte pro Pixel anstatt den  $3 \cdot 8\text{Bit} = 24\text{Bit}$ . Dies hat vorerst mal einen praktischen Hintergrund: Die Zahlen werden damit etwas kleiner.

Der Speicherbedarf **eines einzelnen Bildes** berechnet sich wie folgt:

$$S_1 = 720 * 1280 * 3\text{Byte}$$

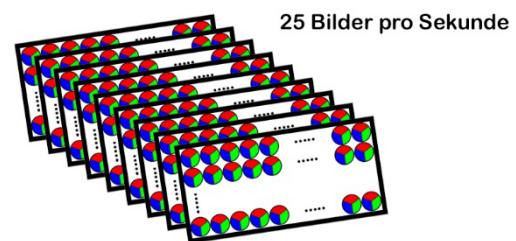
**S<sub>1</sub> = 2'764'800 Byte** (Dies entspricht 22'118'400 Bit)



Der Film besteht aber aus mehreren Bildern. Darum geht die Rechnung so weiter: Eine Sekunde Film ergibt den folgenden Speicherbedarf:

$$S_2 = 2'764'800 \text{ Byte} * 25$$

**S<sub>2</sub> = 69'120'000 Byte** (Dies entspricht 552'960'000 Bit)



Da der Film aber nicht nur eine Sekunde, sondern ganze 90 Minuten dauert, braucht es noch etwas mehr:

$$S_3 = 90 \text{ Minuten} * 60 * 69'120'000 \text{ Byte}$$

**S<sub>3</sub> = 373'248'000'000 Byte** (Dies entspricht 2'985'984'000'000 Bit)

Exakt für solche grossen zahlen haben wir die Massvorsätze kennengelernt.  
Aus 373'248'000'000 Byte werden also:

**S<sub>3</sub> = 347.6 GiB** (Achtung: IEC-Präfix weil Größenangabe eines Speichers)

**Welche Medien (Speicher, Download) wären dazu geeignet?** «Physisch» erwirbt man HD-Filme als optischen Speicher: Eine solche BluRay ist in der Lage, pro Layer 25GB, bzw. bei **2 Seiten zu je 2 Layer** maximal **100GB** zu speichern. Wir benötigten also **vier ganze BluRays**. Sollte der Spielfilm auf nur einer BluRay-Disc Platz haben, muss ganz kräftig komprimiert werden.

**Und als Videostream?** Möchte man den Film als Videostream über eine Netzwerkverbindung beziehen, würde dies einen Netzwerkanschluss erfordern, der 552'960'000 Bit pro Sekunde bzw. rund **553 Mbit/Sekunde** liefern kann.



**Und als kompletten Download von einem Medienserver im Internet?** Geht man davon aus, dass ein heutiger (Stand 2023), durchschnittlicher Internetzugang eine maximale Download-Rate (Best-Effort, Brutto-Datenrate) von **1Gbit/Sekunde** bietet und man somit mit einer **Nettodatenrate** von **900Mb/s** rechnen kann, dann würde der Download von 2'985'984'000'000 Bit exakt 3'317.76 Sekunden oder rund **55.3 Minuten** dauern.

**Analogen Audiosignal und A/D-Wandlung:** Obwohl mit Bildern nicht wirklich verwandt, möchten wir uns kurz dem Thema «Audio» widmen. So sieht z.B. Musik aus, wenn man die Schallwellen graphisch sichtbar macht:

### Kontinuierliches Audiosignal

y-Richtung: Amplitude, Spannung in Volt

X-Richtung: Zeit in (Milli-)Sekunden



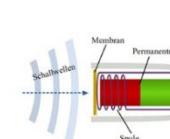
Früher wurde dieser «Verlauf/Kurve» z.B. in eine Vinyl-Scheibe eingraviert und man konnte das so konservierte auf einem Grammophon abspielen.



Die Schallwellen werden vom Mikrofon erfasst und vom Lautsprecher wiedergegeben.

## Mikrofon

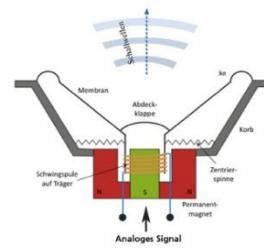
Für die digitale Verarbeitung ist eine A/D-Wandlung erforderlich.



## Lautsprecher

Für die analoge Ausgabe ist eine D/A-Wandlung erforderlich.

(Audioverstärker gibt es auch in „digitaler“ Form (Class-D). Ein digitales Audiosignal wird mittels eines geeigneten Verfahrens, beispielsweise durch Pulsweitenmodulation (PWM), in eine Folge von Pulsen gebracht. Dadurch kann der Verstärker im Schaltbetrieb gefahren werden, wodurch die Schaltelemente (Transistoren) entweder maximal leitend oder maximal isolierend sind und somit nur zwei Zustände kennen. Somit kann auf eine D/A-Wandlung verzichtet werden.)

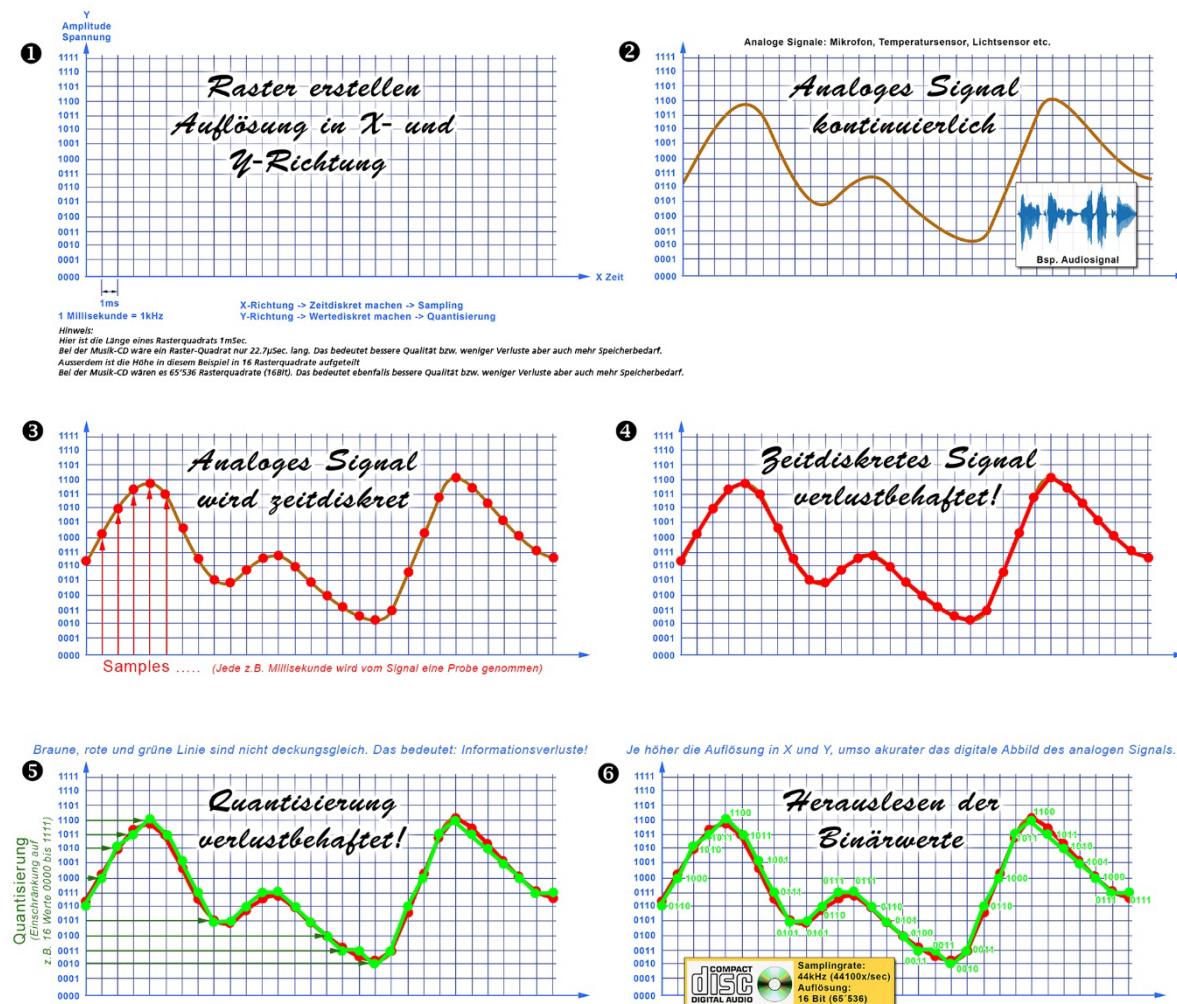




Heute ist aber alles Digital. Der Computer, der unter vielem anderen auch an die Stelle des Grammophons getreten ist, versteht direkt kein «Analog».



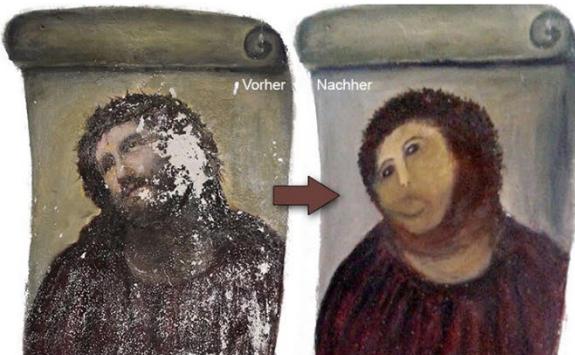
Somit muss das analoge Audiosignal zuerst digitalisiert werden. Dies erledigen sogenannte Analog-Digital-Wandler (DAC). Und das geht so:





## Wie bring ich einen ganzen Spielfilm auf eine BluRay?

Wie die vorangegangene Berechnung eines Spielfilms ergeben hat, müssen Bilder meist kräftig komprimiert werden. In einem früheren Kapitel haben wir die verlustlose Komprimierung behandelt. Nun folgt die bei Bild und Ton unverzichtbare verlustbehaftete Variante. Diese bietet zwar wesentlich mehr Potential zur Datenreduktion, hat aber auch den Nachteil, dass bei der Komprimierung und Dekomprimierung Daten verloren gehen und das Resultat nicht mehr ganz dem Original entspricht.



**Nicht ganz so wie es vorher war, aber so ähnlich:**

*In Spanien hat eine Amateur-Restauratorin versucht,  
ein Jesus-Fresko aufzurüsten und die alte  
Kirchenmalerei bis zur Unkenntlichkeit überpinselt.*

Man versucht mit geeigneten Verfahren in einer Datei den Informationsgehalt derart zu reduzieren, dass das subjektive Empfinden des Datei-Konsumenten die Datenreduktion möglichst nicht wahrnimmt. Dabei gibt es verschiedene Ansätze und oft auch Kombinationen davon.

### Die naheliegenden Techniken zur Datenreduzierung:

Die mathematisch weniger aufwändigen Verfahren reduzieren z.B. die Bildauflösung, die Farbtiefe oder die Anzahl der gezeigten Bilder:

- **Bildgrösse reduzieren:** Reduziert man bei einem Original-Bild von 1000 Pixel x 1000 Pixel Grösse sowohl Länge als auch Breite um die Hälfte, erhält man eine viermal kleinere Datei.
- **Farbauflösung reduzieren:** Reduziert man zum Beispiel bei einem TrueColor-Bild die Farbauflösung von ursprünglich pro Grundfarben (Rot,Grün,Blau) 8 Bit bzw. 256 Abstufungen auf 4 Bit bzw. 16 Abstufungen), ist man zwar mit nur noch 4'096 unterschiedlichen Farbe weit entfernt von TrueColor mit 16'777'216 unterschiedlichen Farbe, hat die Dateigrösse damit aber halbiert.
- **Bildwiederholrate reduzieren:** Beim Kino wird seit den Anfängen mit 24 Bilder pro Sekunde gerechnet. Aber schon ab 12 Bilder pro Sekunde würde eine Bewegung einigermassen flüssig zu erkennen sein. Ersparnis: 50%
- **Samplingrate reduzieren:** Wird ein analoges Signal wie z.B. das elektrische Signal eines Mikrofons digitalisiert, läuft das wie folgt ab: In regelmässigen Abständen wird die Amplitude des elektrischen Signals gemessen und in seinen digitale Wert umgewandelt. Bei Audioaufnahmen sind übliche Werte 44.1kHz. Gemäss Abtasttheorem kann damit ein analoges Audiosignal bis 22kHz, was der theoretischen oberen Grenzfrequenz des menschlichen Gehörs entspricht, aufgelöst werden. Reduziert man nun die Samplingrate auf z.B. 8kHz, scherbelts zwar gehörig aus dem Lautsprecher, einem Gespräch kann man aber immer noch folgen, wie die historische Telefongrenzfrequenz von 4kHz beweist. Theoretische Ersparnis bei einer Samplingrate-Reduktion von 44.1kHz auf 8kHz: Datei wir ca. 5x kleiner. Man könne nun auch noch an der Auflösung der Amplitude herumschrauben und anstatt z.B. 16 Bit nur 8 Bit auflösen, was die Dateigrösse nochmals halbieren würde.



- **Eine Farbtabelle benutzen:** Anstatt jedes einzelne Pixel mit den drei Grundfarben RGB in 3x8Bit zu beschreiben, kann man die Farbinformation eines Pixels auch in einer Farbtabelle referenzieren. Das lohnt sich allerdings nur, wenn das reduzierte Farbangebot keine Rolle spielt. Mit Dithering kann man zwar etwas tricken, indem man benachbarte Pixel mit unterschiedlichen Farbinformationen versieht, die dann bei der Betrachtung aus einer gewissen Distanz zu einer imaginären Drittfarbe verschmelzen.

Bsp.: GIF mit 8 Bit Farbauflösung (ergibt 256 Farben).



Bereits etwas trickreicher ist das folgende Verfahren, dass sich Subsampling oder Farbunterabtastung nennt.

### Subsampling (Farbunterabtastung):

Im Folgenden werden wir über Farbbilder und Graustufenbilder sprechen. Daher zuerst ein kleiner Exkurs in die Biologie.

**Die Farbwahrnehmung des Menschen:** Unser Auge ist nicht für alle Lichteindrücke gleichermassen empfänglich. Für Hell/Dunkel (Graustufen) verfügt unser Auge über ca. 100 Millionen Helligkeitsempfindliche Stäbchen, wohingegen nur ca. 7 Millionen Zäpfchen der Farbempfindlichkeit dienen und diese dazu noch bei unterschiedlichen Wellenlängen (Farben) unterschiedlich stark reagieren.

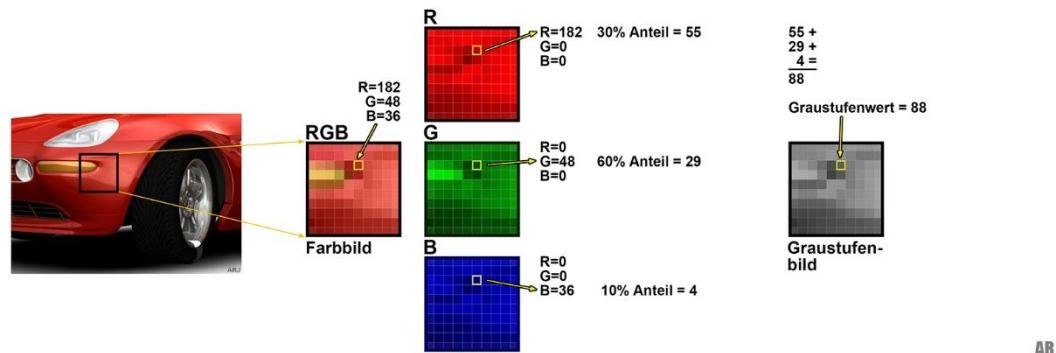


Mit diesem Wissen soll nun aufgezeigt werden, wie ein Farbbild in ein Graustufenbild umgewandelt wird. In bestimmten Fällen ist es nämlich nötig, das Farbbild in ein Graustufenbild umzuwandeln. Zum Beispiel bei dem im nächsten Abschnitt behandelten Farbmodell YC<sub>b</sub>C<sub>r</sub>.

Bei der Umwandlung werden die drei Farbanteile RGB verschieden gewichtet. Dies könnte man mit folgendem historischem Hintergrund erklären: Der frühzeitliche Mensch als Jäger und Sammler war vor allem auf eine hohe Auflösung im Grünbereich (Wälder, Wiesen etc.) angewiesen, um Beute oder herannahende Gefahr besser erkennen zu können. Die Farbe Blau war da eher seltener und darum weniger wichtig. Da diese Aspekte des Farbensehens des menschlichen Auges berücksichtigt werden müssen. So wird beispielsweise Grün heller wahrgenommen als Rot, dieses wiederum heller als Blau.



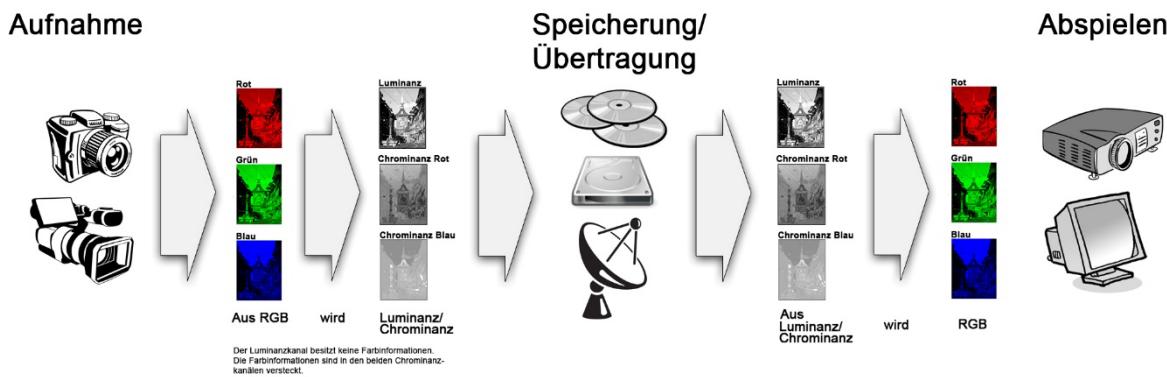
Diese unterschiedliche Gewichtung wird in folgender Umrechnungsformel berücksichtigt:  
Luminanz (Y) = 0.3 x Rot + 0.6 x Grün + 0.1 x Blau



- Unser Auge ist für Grün am empfindlichsten, darum hat Grün im Graustufenbild den grössten Anteil von 60%.
- Weniger empfindlich ist man für Rot. Darum hat Rot nur 30% Anteil.
- Am wenigsten empfindlich ist man für Blau. Sein Anteil ist 10%.

Bei der Farbunterabtastung möchte man nun Daten derart reduzieren, dass die Luminanz (Helligkeit) möglichst unangetastet bleibt und man bei den für unsere Augen weniger relevanten Farben Abstriche macht. Das gelingt allerdings bei RGB nicht. Wir benötigen einen anders definierten Farbraum.

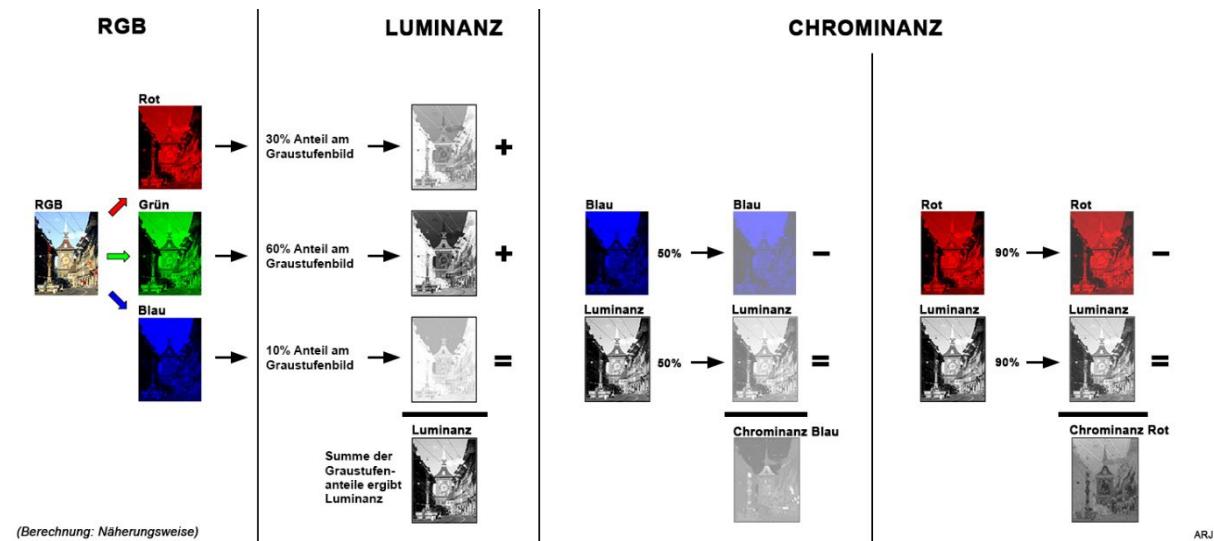
**Das Helligkeits-Farbigkeits-Modell oder Luminanz/Chrominanz-Modell:** Die Kamera zeichnet das Bild in RGB auf. Abgespeichert und datenreduziert wird es im Luminanz/Chrominanz-Farbmodell. Damit das Bild wieder angezeigt werden kann, muss es dekomprimiert und vom Luminanz/Chrominanz-Farbmodell in RGB zurückgerechnet werden.



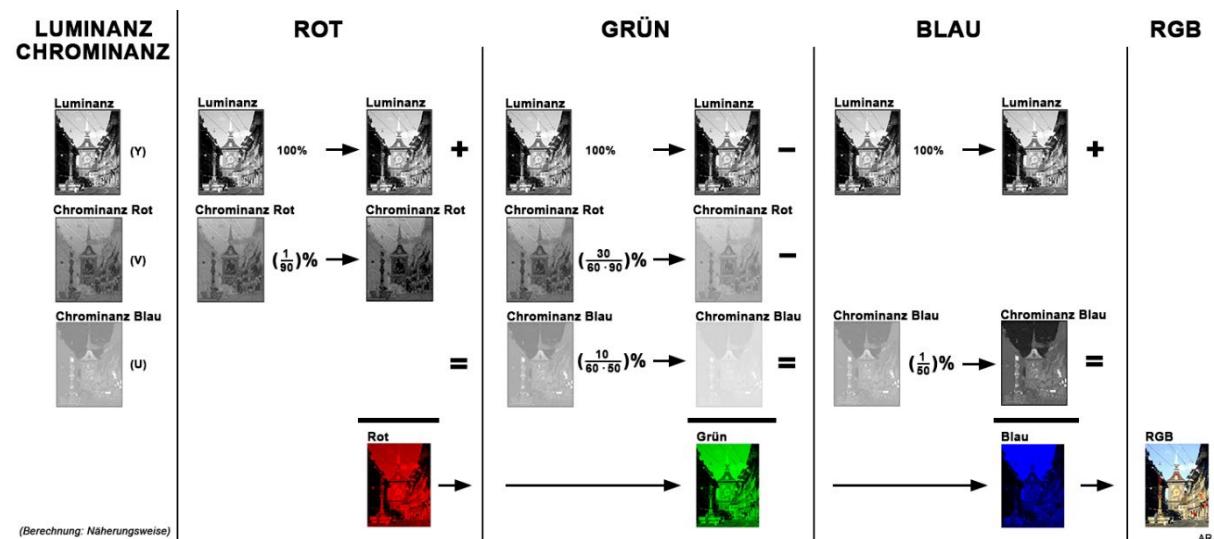
**Die Umwandlung RGB zu Luminanz/Chrominanz ( $YC_bC_r$ ):** Die drei Farbkanäle Rot, Grün und Blau werden in die drei Kanäle Luminanz (Y), Chrominanz Blau ( $C_b$ ) und Chrominanz Rot ( $C_r$ ) umgewandelt.

Bei der Umwandlung in den Luminanzkanal (Graustufenbild), hat jeder Farbkanal den Anteil, dem ihm gemäss Physiologie des Auges zusteht: Grün 60%, Rot 30%, Blau 10%).

Bei der Berechnung der Chrominanzkanälen kommt reine Algebra zur Anwendung:



**Die Zurückwandlung Luminanz/Chrominanz ( $YC_bC_r$ ) zu RGB: Der Vollständigkeitshalber sei hier auch gezeigt, wie der Weg zurück erfolgt.**





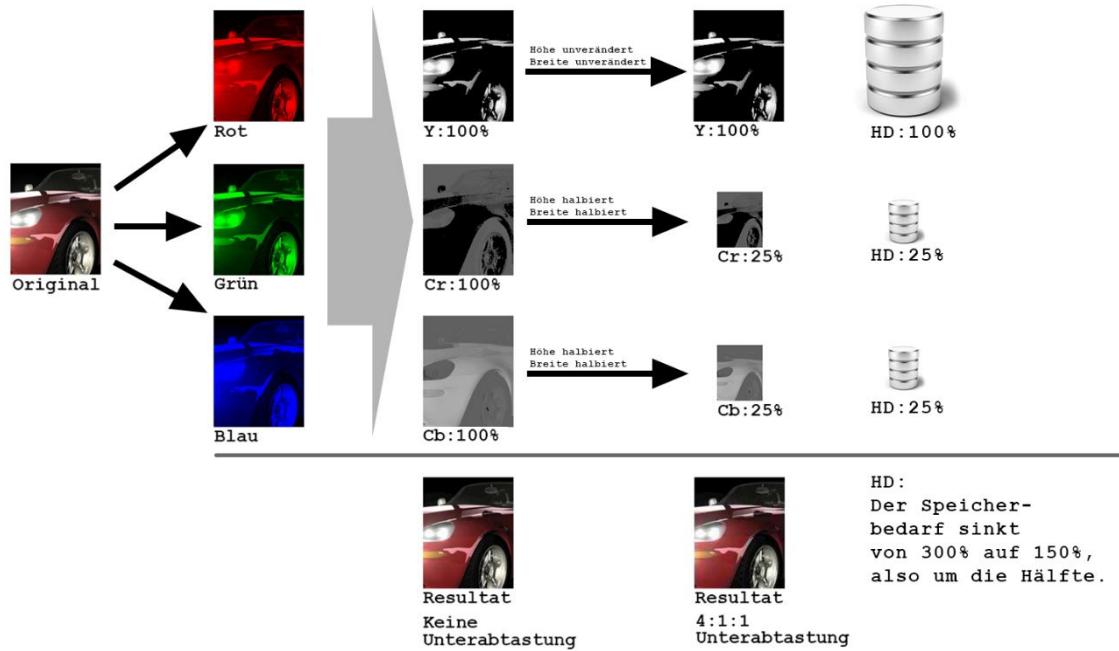
## Zusammenfassend kann gesagt werden:

- Das **YC<sub>b</sub>C<sub>r</sub>-Helligkeits-Farbigkeits-Modell** (gemäss CCIR-601 bzw. IEC 601-Standard) wurde für das Digitalfernsehen entwickelt. Ausserdem wird es für digitale Bild- und Videoaufzeichnung, bei JPG-Bildern, MPEG-Videos und damit auch bei DVDs, sowie den meisten anderen digitalen Videoformaten verwendet. Es teilt die Farbinformation in die (Grund-)Helligkeit Y und die Farbigkeit, bestehend aus den zwei Farbkomponenten C<sub>b</sub> (Blue-Yellow Chrominance) und C<sub>r</sub> (Red-Green Chrominance) auf.
- Die **Helligkeit** entspricht der Hellempfindlichkeit des Auges, die im grünen Spektralbereich am grössten ist. Chrominance oder kurz Chroma bedeutet Buntheit.
- Die **unterschiedliche Wahrnehmung** von Y gegenüber den C<sub>b</sub>- und C<sub>r</sub>-Kanälen entspricht der Entwicklung der Farb- und Helligkeitsverteilung in der Natur: Im Laufe der Evolution hat sich der menschliche Sehsinn daran angepasst. Das Auge kann geringe Helligkeitsunterschiede besser erkennen als kleine Farbtonunterschiede, und diese wiederum besser als kleine Farbsättigungsunterschiede. So ist ein Text grau auf schwarz geschrieben gut zu lesen, blau auf rot geschrieben bei gleicher Grundhelligkeit jedoch nur sehr schlecht.
- Die **Analogie zum menschlichen Sehsinn** wird für einen grossen Vorteil von YC<sub>b</sub>C<sub>r</sub> genutzt: die Farbunterabtastung (engl. chroma subsampling). Dabei wird die Abtastrate und damit die Datenmenge der Chrominanz-Kanäle C<sub>b</sub> und C<sub>r</sub> gegenüber der Abtastrate des Luminanz-Kanals Y reduziert, ohne dass es zu einer spürbaren Qualitätsverringerung kommt. So kann man z. B. mit der JPEG-Komprimierung eine nicht unerhebliche Datenmenge einsparen.
- Das **YUV-Farbmodell** der analogen Fernsehtechnik wird manchmal fälschlicherweise mit YC<sub>b</sub>C<sub>r</sub> für digitale Darstellung von Farbvideosignalen gleichgesetzt. Einfachheitshalber bedienen wir uns für die folgenden Erklärungen bei dem YC<sub>b</sub>C<sub>r</sub>-ähnlichen YUV-Farbmodell. Wer es ganz genau wissen will, konsultiert die einschlägige Fachliteratur oder informiert sich bei den entsprechenden Beiträgen auf Wikipedia.

**Und so funktioniert die Farbunterabtastung/Subsampling:** Beim Subsampling oder Farbunterabtastung wird die Bildauflösung reduziert. Allerdings nicht in RGB. Dazu ein Beispiel: Reduziert man ein HD1080-TV-Bild von 1920x1080 auf 960x540 hat man zwar die Dateigrösse geviertelt, möchte man das Bild aber wieder auf einem HD1080-Projektor anschauen, wirkt es unscharf, weil es von 960x540 wieder auf 1920x1080 gedehnt werden muss. Bei der Unterabtastung wählt man einen anderen Weg: Man wandelt das RGB-Bild in YC<sub>b</sub>C<sub>r</sub> (Luminanz/ChrominanzBlau/ChrominanzRot) um. Anstatt wie beim vorangegangenen Beispiel alle drei Kanäle (RGB) in der Pixelauflösung zu reduzieren, wird dies bei Chroma-Subsampling nur in den beiden Chrominanz-Kanälen getan. Damit erhält man zwar eine etwas ungenauere Einfärbung, die Bildschärfe (Luminanz) bleibt aber vollständig erhalten. Es sind folgende Chroma-Subsamplings vorgesehen: 4:4:4, 4:2:2, 4:1:1, 4:2:0. Und was wird damit eingespart? Bei 4:1:1 bedeutet dies Y=100%, C<sub>b</sub>=25%, C<sub>r</sub>=25% und somit gegenüber dem Original (300%) Dateigrösse halbiert.



Beispiel zu Subsampling:





## Anspruchsvollere Verfahren wie DCT, das z.B. bei JPG verwendet wird:

Wirklich schwierig zu verstehen sind erst die Komprimierungsverfahren, die auf höherer Mathematik beruhen. Hier braucht schon fast einen Hochschulabschluss, um die Technik dahinter zu verstehen wie z.B. bei der DCT (Discrete Cosine Transformation) für JPG-Komprimierung, wo in Kombination zwar einige bereits besprochene Verfahren wie Huffman, RLC, Chroma-Subsampling zur Anwendung kommen, aber eben auch die diskrete Kosinustransformation, wo uns die mathematischen Kenntnisse dazu fehlen. Wer's trotzdem wissen will, hier ein vereinfachter Erklärungsversuch:

### Die sechs DCT-Verarbeitungsschritte:

**Erster Schritt:** Als erster Schritt wird das **RGB-Bild** in ein **YC<sub>b</sub>C<sub>r</sub>-Bild** umgewandelt. Danach wird die Auflösung in den beiden Chrominanzkanälen reduziert. (**Chroma-Subsampling**) Die drei Kanäle (1xLuminanz, 2xCrominanz) werden ab hier separat verarbeitet. Jeder Kanal wird nun in 8x8 Pixelblöcke aufgeteilt.

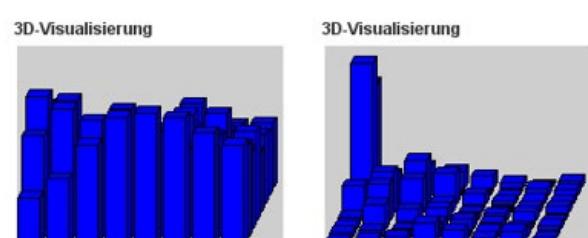
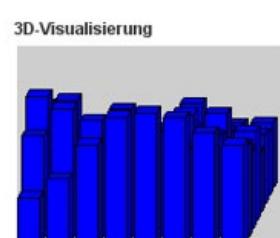


**Zweiter Schritt:** Auf jeden dieser 8x8 Pixelblöcke wird die diskrete **Kosinustransformation** angewandt. Vereinfacht ausgedrückt bedeutet dies, dass die 8x8 Bildwertematrix vom Bildbereich in den sogenannten Frequenzbereich transformiert wird. Im Bildbereich prägen Unterschiede in den Helligkeitswerten die Wertematrix, wohingegen im Frequenzbereich die Schnelligkeit der Helligkeitsänderungen entscheidend sind. (Etwas vereinfacht ausgedrückt: Scharfe Bilder ergeben schnelle Helligkeitsänderungen und damit viele hohe und unterschiedliche DCT-Werte, unscharfe langsamere und somit tiefere und weniger unterschiedliche DCT-Werte.)



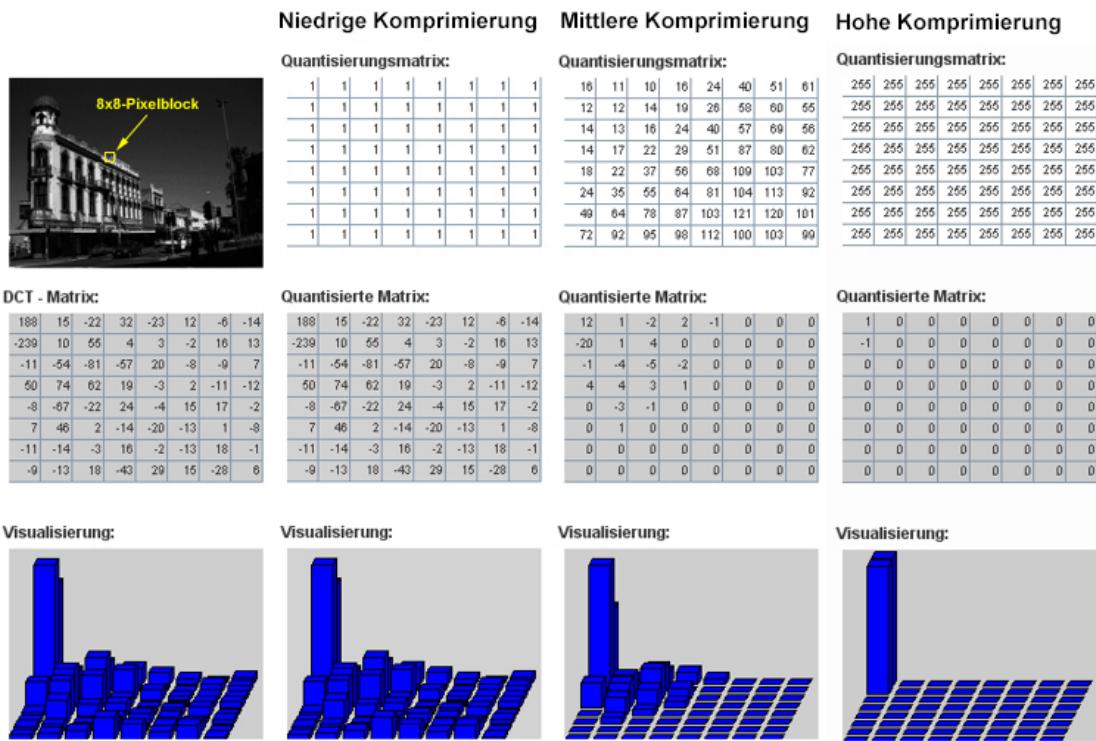
Originalwerte im 8x8 Block															
114	115	117	114	109	117	113	115	120	114	117	117	121	113	120	110
118	119	121	114	117	112	119	125	163	149	141	115	130	193	143	105
204	163	133	155	141	163	191	159	236	228	178	163	168	204	163	138
180	228	208	227	186	157	179	163	80	119	179	228	236	227	200	179

DCT-transformierte Werte															
188	15	-22	32	-23	12	-6	-14	-239	10	55	4	3	-2	16	13
-11	-54	-81	-57	20	-8	-9	7	50	74	62	19	-3	2	-11	-12
-8	-67	-22	24	-4	15	17	-2	7	46	2	-14	-20	-13	1	-8
-11	-14	-3	16	-2	-13	18	-1	-9	-13	18	-43	29	15	-28	6





**Dritter Schritt:** Die DCT-transformierten Werte werden nun **quantisiert**. Unter Quantisierung ist etwa dasselbe zu verstehen, wie das Notenrunden bei Schulprüfungen: Wenn die Schulnote mit einer Genauigkeit von 1/10 festgehalten wird, hat man mehr mögliche Notenwerte, als wenn die Note auf 0.5 gerundet wird. Allerdings verliert man mit der größeren, weil ungenauerer 0.5-er Notenskala auch an Aussagekraft. Dasselbe gilt für die quantisierten DCT-Werte. Bis jetzt hat man allerdings noch keine eigentliche Datenreduktion erreicht. Abhängig von der Stärke der Quantisierung der DCT-Werte (im Extremfall werden die meisten Werte zu 0) erhält die 8x8-Matrix eine geeignete Form, für RLC und VLC.



**Vierter Schritt:** Nun folgt eine Zick-Zack-Scan.



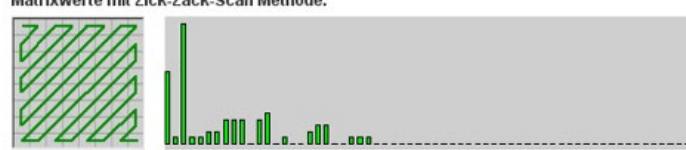
Quantisierte Matrix:

12	1	-2	2	-1	0	0	0
-20	1	4	0	0	0	0	0
-1	-4	-5	-2	0	0	0	0
4	4	3	1	0	0	0	0
0	-3	-1	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

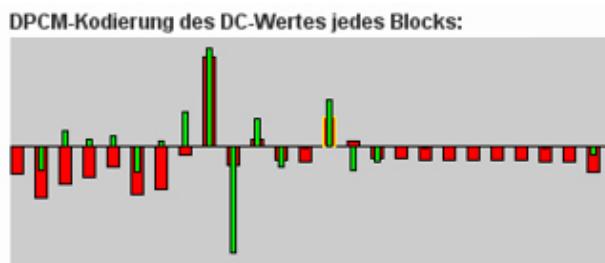
Matrixwerte von links nach rechts und von oben nach unten:



Matrixwerte mit Zick-Zack-Scan Methode:



**Fünfter Schritt:** Abgesehen vom Subsampling werden erst jetzt mit RLC die Daten reduziert.



## BLF- und Variable Length Integer Kodierung der AC-Werte:

### **Var Run Length-Encoding:**

run Length-Encoding:  
1 -20 -1 1 -2 2 4 -4 4 0 4 -5 0 -1 0 0 -2 3 -3 0 0 1 -1 0  
0 0

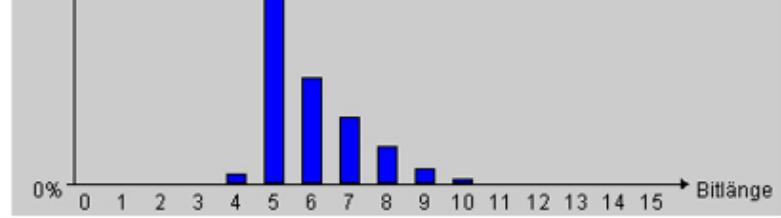
#### Nach Run Length-Encoding:

**Final Length Encoding:**  
**1 -20 -1 1 -2 2 4 -4 4 1 4 -5 1 -1 2 -2 3 -3 2 1 -1 1 39**

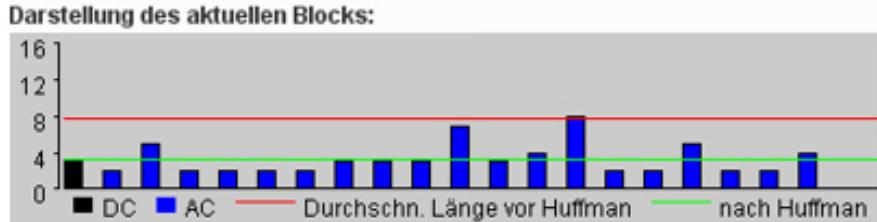
## Verteilung der Langen nach

Erstellung der Lungenentzündung VZ-Neuerung:

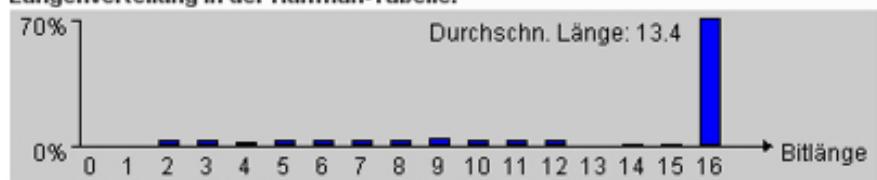
47/20



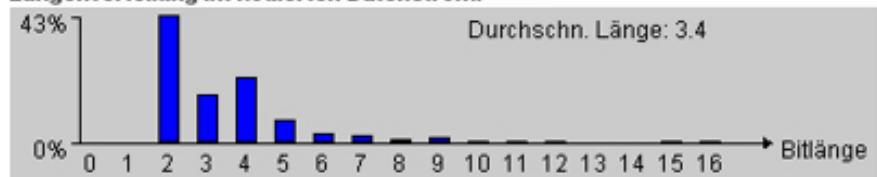
**Sechster Schritt:** Darauf folgt ein VLC in Form von Huffman



#### Längenverteilung in der Huffman-Tabelle:

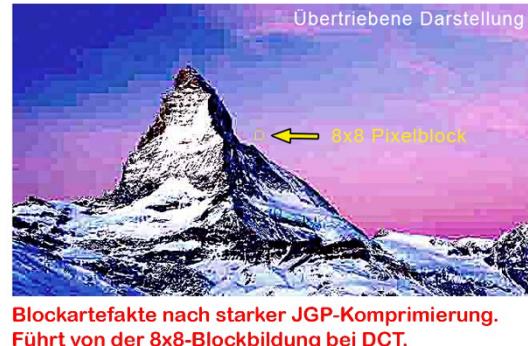


#### Längenverteilung im kodierten Datenstrom:





**Schlussfolgerung zu DCT:** Wenn bei der DCT zu stark quantisiert wird, gleichen sich die Farben der 8x8 Pixelblöcke einander an und es entsteht eine Blockstruktur. Man spricht dann von **Blocking-Artefakten**:

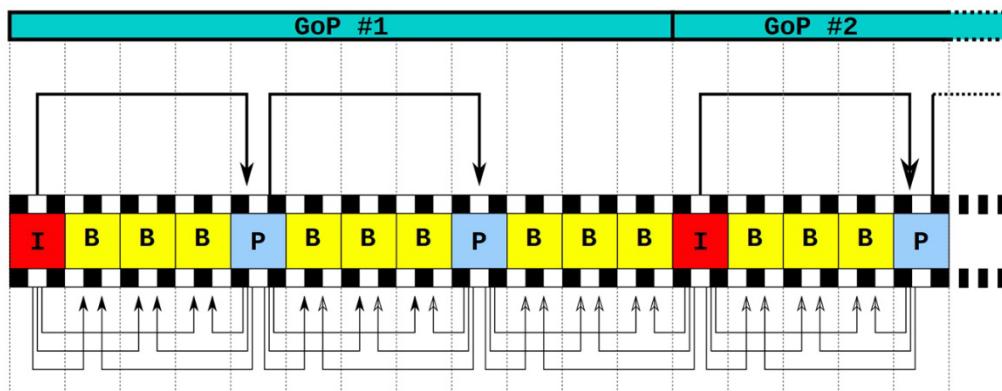


### Intraframe- und Interframe-Komprimierung:

Bisher haben wir ausschliesslich Verfahren betrachtet, die sich auf ein **einzelnes Bild** beziehen. Man bezeichnet dies auch **Intraframe-Komprimierung**. Es ist aber auch möglich innerhalb einer **Bildsequenz** Bandbreite oder Speicher zu sparen. Dies nennt man dann **Interframe-Komprimierung**. Und das geht so: Anstatt in einem Videostream pro Sekunde 24 komplette Bilder zu verschicken (oder speichern), könnte man ja nach einem Vollbild (i-Frame) nur noch die Änderung des Bildinhalts zum vorangegangenen Bild übermitteln.



Allerdings darf man das nicht übertreiben. In regelmässigen Abständen sollte ein Vollbild verfügbar sein, damit die Filmwiedergabe bei einem Übermittlungsfehler neu synchronisieren kann. Man spricht dabei von einer **GOP** (Group-of-Pictures). Je nach Dynamik der Filmszene gibt das eine immense Einsparung an Daten. Filmt man 10 Minuten lang seine geschlossene Haustüre ohne weitere Action, würde dies bei GOP25 einer Datenreduktion von so etwa 96% bedeuten. BTW: Ein ähnliches Verfahren kennen wir beim Backup, wo zwischen Fullbackup und inkrementellem Backup unterschieden wird.





### **Ein paar Worte zur Audiokomprimierung (MP3):**

MP3 ist eigentlich MPEG-1/2 Audio Layer 3 und damit ein Verfahren zur verlustbehafteten Kompression digital gespeicherter Audiodaten. Zur Komprimierung, die gegenüber einer Original-WAV-Datei durchaus bis zu über 85% betragen kann, werden psychoakustische Effekte der menschlichen Wahrnehmung von Tönen und Geräuschen ausgenutzt. Von MP3-Playern unterstützt werden Datenraten zwischen 8kb/s und 320kb/s. Wer mehr dazu erfahren will, dem sei z.B. der Wikipedia-Eintrag zu MP3 empfohlen.