



ITWissen

Das große Online-Lexikon
für Informationstechnologie

VIDEOKOMPRESSION

KLAUS LIPINSKI (Hrsg.)



Index

- 3ivx**
- Artefakt**
- AVC**, *advanced video coding*
- B-frame**, *bidirectional frame*
- BT.601**
- CABAC**, *context based adaptive binary arithmetic coding*
- Chrominanz**
- Cinepak-Codec**
- CODEC**, *compression/decompression*
- DCT**, *discrete cosine transformation*
- DivX**
- ENDEC**, *encoder/decoder*
- Farb-Subsampling**
- Farbdifferenzsignal**
- GOP**, *group of picture*
- H.261**
- H.263**
- H.264-Kompression**
- H.264/AVC**, *H.264 advanced video coding*
- H.320**
- H.323**
- H.324**
- HDX4**
- Indeo**, *Intel video*
- I-frame**, *Intra-Frame*
- JND**, *just noticeable difference*
- JVT**, *joint video team*
- Kompression**
- Kompressionsrate**
- Luminanz**
- MJPEG**,
- MP4-Dateiformat**
- MPEG**, *moving picture experts group*
- MPEG-Kompression**
- MPEG-1**
- MPEG-2**
- MPEG-4**
- Nero Digital**
- P-frame**, *predicted frame*
- PES**, *packetized elementary stream*
- SMP**, *software motion picture*
- Sorenson-Codec**
- SSIM**, *structural similarity index*
- Theora-Codec**
- TS**, *transport stream*
- VC-1**, *video codec 1*
- Verlustbehaftete Kompression**
- Video-Codec**
- Videodateiformat**
- Videokompression**
- Videosignal**
- WMF**, *Windows media format*
- WMV**, *Windows media video*
- XviD**

Impressum:

Herausgeber: Klaus Lipinski

Darstellformate V1

Copyright 2007

DATAKOM-Buchverlag GmbH

84378 Dietersburg

Alle Rechte vorbehalten.

Keine Haftung für die angegebenen
Informationen.

Produziert von Media-Schmid

www.media-schmid.de

3ivx

Codierung und Abspielmöglichkeiten der Video-Codecs DivX, XviD und 3ivx

Wiedergabe	Wiedergabe-Codecs
DivX	DivX, FFDSHOW, 3ivx Decoder
XviD	XviD, FFDSHOW, 3ivx Decoder
3ivx	3ivx Decoder
Codierung	Codec
DivX	DivX Pro Vers., DivX freeware
XviD	Free GPL XviD Codec
3ivx	Free Trial 3ivx Filer Suite

3ivx ist eine kommerzielle Software die die Wiedergabe von *MPEG-4-Video*, *MPEG-4-Audio* und das *MP4-Dateiformat* unterstützt. Sie kann für die Wiedergabe der verschiedensten Implementierungen von *MPEG-4* eingesetzt werden: für *DivX* und *XviD* sowie für *MPEG-4*-Implementierungen von Apple, Philips und andere, für die die entsprechenden *Codecs* bereitgestellt werden müssen. 3ivx besteht aus den drei Technologien

Video, Audio und Dateiformat, die im Core verankert sind und mit verschiedenen Front-Ends, die sich als Shell um den Core legen, ausgestattet werden können. Der Core kommuniziert dabei über den Front-End mit der Anwendung, die vom Front-End durch eine Tool- und Filtersammlung unterstützt wird. Beispiele dafür sind das DirectShow-Filter, Komponenten von Quicktime und Codecs für Video for Windows. 3ivx umfasst sechs verschiedene Filter für die Codierung und Wiedergabe der verschiedenen Audio- und Video-Files. Die Filtersammlung kann mit verschiedenen Windows-Programmen gebündelt werden und bietet dadurch Vorteile bei der DirectShow und bei Video for Windows (VfW).

Bei der Codierung erzeugt der 3ivx-Codec *MPEG-4-Videos*, die allerdings in *MP4-Containern* gespeichert werden. Als Audioformat benutzt 3ivx die *AAC-Kompression*.

Artefakt artefact

Bei der Codierung und Decodierung von *Grafik-, Audio- und Videosignalen* treten ebenso wie bei der *Kompression* und Dekompression Fehler auf, die als *Artefakte* bezeichnet werden.

Diese Artefakte machen sich in Sprach-, Audio- und Video-Übertragungen als Störungen oder Geräusche bemerkbar. Typische Artefakte sind die kleinen Quadrate bei der *MPEG-Kompression*.

AVC
advanced video coding

AVC ist eine *Videokompression*, die in den ITI-Standard *H.264* eingebracht wurde und in Kombination *H.264/AVC* eingesetzt wird. Die Bezeichnung "Advanced" bezieht sich dabei auf die Verbesserung des Kompressionsverfahrens gegenüber *MPEG-4*, die sich auf viele Detailverbesserungen bezieht. So benutzt AVC eine Blockgröße von 4x4 Pixel, mit der Bilddetails besser dargestellt werden und nutzt die Bildinformation von vorhergehenden Bildern besser für die Vorhersage aus.

B-frame
Bidirectional-Frame
bidirectional frame

Die Bidirectional-Frames sind in der *MPEG-Kompression* die am höchsten komprimierten Bilder. Diese Videobilder können mit Kompressionsfaktoren von bis zu 50:1 abgelegt werden. Wie der Name sagt, werden Bidirectional-Frames aus den Informationen vorausgegangener und nachfolgender Bilder gewonnen. In diesen Bildern werden Bildteile aus vorangehenden Bilder übernommen.

BT.601
(ITU-R BT.601)

Standard	CCIR 601 NTSC	CCIR 601 PAL/SECAM	CIF	QCIF
Luminanz- auflösung (Y)	720 x 480	720 x 576	352 x 288	176 x 144
Chrominanz- Auflösung (C)	360 x 480	384 x 576	176 x 144	88 x 72
Chroma- Subsampling	4 : 2 : 2	4 : 2 : 2	4 : 2 : 0	4 : 2 : 0
Halbbilder	60	50	30	30
Zeilensprung	ja	ja	nein	nein
Datenrate (Mbit/s)	167,6 216 270	165,9 216 270	36,5	9,1

Das Dateiformat für Digitalvideo mit Zeilensprungverfahren, so wie es von NTSC und dem PAL-Fernsehstandard her bekannt ist, wurde zuerst von der CCIR unter CCIR-601, später dann von der ITU unter dem Standard ITU-R BT.601 spezifiziert. BT.601 beschreibt die Konvertierung der analogen Fernsehsignale mit Bildwiederholfrquenzen von 50 Hz und 60 Hz und Zeilenfrequenzen von

*CCIR-Standards für
Digitalvideo*

BT.601-Spezifikationen für den PAL-Fernsehstandard

BT.601-Format für PAL	Spezifikationen
Signal-komponenten	Luminanz, Y Chrominanz, Cb, Cr
Bildwiederholfr.	25 Hz
Halbilder	50
Vollbilder	25
Zeilen	625
Abtastwerte sichtbares Luminanz	864 pro Zeile 720 pro Zeile
Chrominanz	360 pro Zeile für Cb und Cr
Abtastfrequenz	13,5 MHz
Codierung	PCM ind 8 Bit oder 10 Bit
Quantisierungs-pegel	0 und 255 für Sync. 1 bis 254 für Video 16 bis 235 für Luminanz 16 entspricht Schwarz 235 entspricht Weiß

15,625 kHz und 15,750 kHz, wie sie im PAL-Fernsehstandard und in NTSC verwendet werden. Das BT.601-Dateiformat spezifiziert eine Abtastfrequenz von 13,5 MHz. Daraus ergeben sich für den PAL-Standard mit einer Zeilenlänge von 64 µs insgesamt 864 Pixel, von denen 720 Pixel sichtbar sind. NTSC hat ebenfalls 720 sichtbare Pixel, obwohl sie 858 Abtastwerte pro Zeile aufweist. In der Vertikalen erfolgt die Darstellung mit 576 sichtbaren Pixeln (PAL) und 486 bei NTSC. Aus den horizontalen und vertikalen

Pixelwerten errechnet sich das Seitenverhältnis der Pixel, die nicht quadratisch sind, sondern beim PAL-Standard mit einer Bilddarstellung von 4:3 ein Seitenverhältnis von 1,0667 von Pixelbreite zu -höhe haben. Bei anamorphical Widescreen mit 16:9-Darstellung, werden die Pixel rechteckiger mit einem Verhältnis von 1,422 (PAL). Weiterhin wird in BT.601 festgelegt, dass die *Chrominanzsignale* Cb und Cr von jedem zweiten Pixel abgetastet werden, also mit einem *Farb-Subsampling* von 4:2:2. Die Farbtiefe kann 8 Bit oder 10 Bit sein, wie sie in der Studioteknik verwendet wird, die Bildwiederholffrequenz beträgt 25 Vollbilder pro Sekunde (PAL). Die resultierende

Videozeile eines BT.601-Signals nach BT.656



Datenrate beträgt bei einer Farbtiefe von 8 Bit bei 216 Mbit/s, bei 10 Bit 270 Mbit/s. Der Datenstrom, dessen Format in BT.656 spezifiziert ist, ergibt sich aus der Abtastrate von 13,5 MHz multipliziert mit der Farbtiefe von 8 Bit für das *Luminanzsignal* sowie der Abtastrate von 13,5/2 MHz einer Farbtiefe von 8 Bit sowohl für das Cb-Signal, als auch für das Cr-Signal. Die einzelnen Komponenten des BT.601-Datenstroms können parallel über mehrere Leitungen als "Y", Cb und Cr, sie können aber auch mit 27 MHz im Zeitmultiplex in der Folge Cb,Y,Cr,Y,Cb,Y,Cr...usw. übertragen werden.

Der eigentliche Bildbereich ist in BT.601 insofern eingeschränkt, als dass der Schwarzwert erst beim Digitalwert 16 und der Weißwert beim Digitalwert 235 beginnt. Rechnet man mit diesem spezifizierten Helligkeitsbereich, kommt man auf eine Datenrate von 165,9 Mbit/s.

Die sichtbaren Zeilen beginnen beim PAL-Standard 132 Abtastwerte nach Beginn des entsprechenden analogen Synchronimpulses und enden 12 Abtastwerte vor dem nächsten Synchronimpuls.

CABAC
context based adaptive binary arithmetic coding

Cabac ist ein sehr effizientes Codierverfahren, das anstelle der Lauflängencodierung (RLE) in *H.263* benutzt wird und mit dem die Datenrate um bis zu 20 % reduziert werden kann. Cabac ist sehr flexibel und reduziert Redundanzen zwischen den Symbolen durch Kontextbildung.

Chrominanz
(C, chrominance)

Der Farbeindruck eines Videobildes wird durch die Chrominanz (C) bestimmt, die die Werte für die Farbsättigung und den Farbton enthält. Obwohl das menschliche Auge Farben wesentlich schlechter auflöst als Helligkeitsinformationen, werden durch den Farbanteil eine verbesserte Detailerkennbarkeit und ein echter Informationsgewinn erreicht. Der Farbanteil wird in allen Fernseh- und

Zusammenhänge zwischen dem Luminanzsignal und den Chrominanzsignalen

Luminanzsignal	$Y = 0,299 R + 0,587 G + 0,224 B$
Farbdifferenzsignale	$U = 0,493 \times (B - Y)$ $V = 0,877 \times (R - Y)$

Videosystemen mit reduzierter Bandbreite gegenüber der *Luminanz* (Y) übertragen. Das Chrominanzsignal besteht im

Video aus zwei Farbkomponenten, den *Farbdifferenzsignalen* (Y-R) und (Y-B), die in Quadraturmodulation mit einem Phasenversatz von 90° moduliert werden. Diese Farbdifferenzsignale werden mit "U" bzw. Cb für B-Y und "V" bzw. Cr für R-Y bezeichnet. Aus den Farbdifferenzsignalen und dem Luminanzsignal (R+G+B) können die Farbkomponenten für Rot (R), Grün (G) und Blau (B) berechnet werden.

Cinepak-Codec

Das Komprimierungsverfahren von Cinepak umfasst das Video- und das Audiosignal. Die Videos werden mit einer Farbtiefe von 24 Bit komprimiert, bei einem Kompressionsfaktor von ca. 8:1. Bei dem Cinepak-Codec werden dem Verfahren nach aus einer Videosequenz Differenzbilder erzeugt und diese komprimiert. Cinepak-Formate erkennt man an den Extensions *.avi und *.mov für Quicktime.

CODEC,
compression/
decompression
(Kompression/
Dekompression)

Das Kunstwort Codec steht für *Compression* und *Decompression*. Bei einem Codec handelt es sich um eine software- oder hardwaremäßige Funktionseinheit, die *Audio- oder Videosignale* nach vorgegebenen Verfahren in Echtzeit verändert. Die Verfahren sind von der ITU standardisiert und u.a. in den ITU-Empfehlungen *H.261, H.263, H.320, H.321* und *H.323* beschrieben. Codecs werden in der Multimediatechnik, der Audio- und Video-Kommunikation eingesetzt. So bei der Bildtelefonie, bei Videokonferenzen, bei der Audiowiedergabe in oder bei der Internet-Telefonie.

Bei den *Video-Codecs* unterscheidet man zwischen den *Intra-Frame-Codecs* und den *Inter-Frame-Codecs*. Erstere versuchen so viele Informationen wie möglich aus einem einzelnen Bild zu tilgen; die Codecs, die mit dem *Inter-Frame-Verfahren* arbeiten, vergleichen zwei aufeinander folgende Vollbilder und ermitteln die Differenz zwischen beiden. Dieses Verfahren wird vorwiegend bei Video eingesetzt.

Da Codecs mit unterschiedlichen Kompressionsalgorithmen arbeiten, kann es zu starken Qualitätsunterschieden kommen und zwar in Bezug auf die Bild- oder Tonqualität, die sich in der Kontinuität in der Darstellung und Übertragung ausdrückt.

Die wichtigsten Hardware-Codec für Video sind Motion-JPEG, Px64 nach H.261 und

H.320, und *MPEG* in den verschiedenen Varianten. An Software-Codecs sind zu nennen Cinepak, *Indeo*, *DivX*, Quicktime, Video for Windows und Windows Media Video (*WMV*).

Wichtige Audio-Codecs sind aacPlus, mp3pro, Ogg Vorbis, RealAudio, Windows Media Audio, QDesign u.a.

Da Codecs im Allgemeinen durch viele Patente und firmeninternes Wissen belegt sind, werden häufig nur wenige Details über deren Algorithmen und Verfahren publiziert.

Die Abkürzung Codec wird häufig auch im Zusammenhang mit einem Codierer/ Decodierer verwendet.

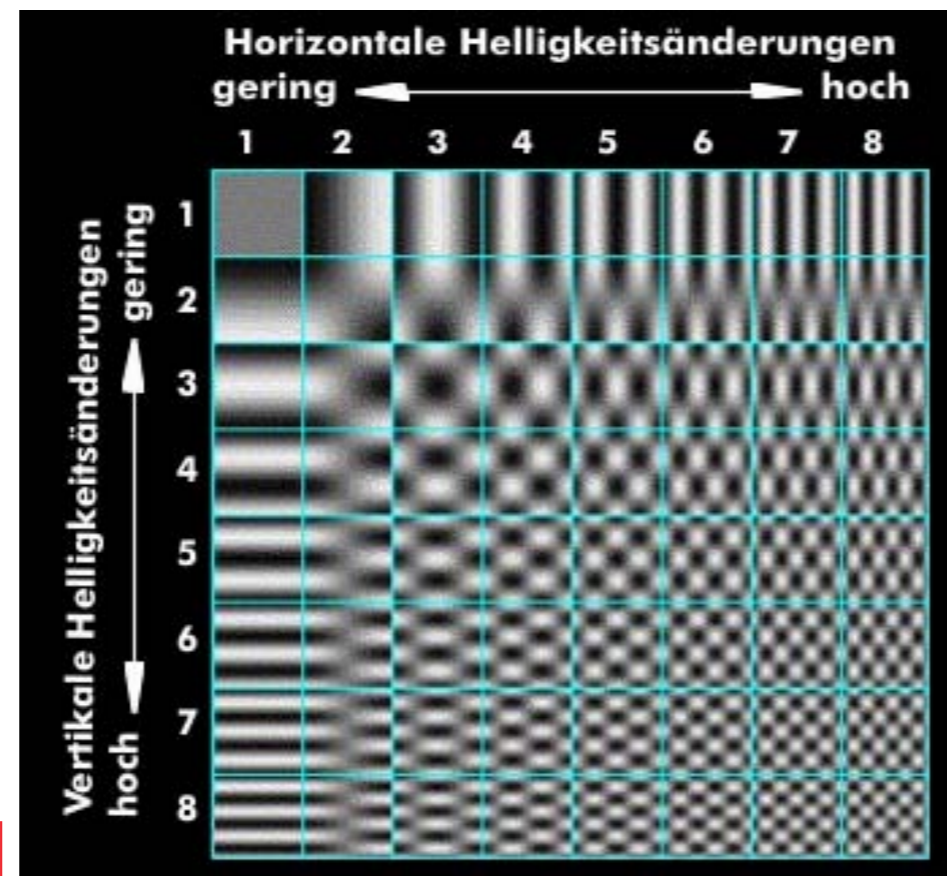
DCT-Transformation
DCT, discrete cosine transformation

Die diskrete Cosinus-Transformation (DCT) ist ein Verfahren zur Umsetzung von zeitbezogenen Signalen in frequenzbezogene. Der Unterschied zur Fourier-Transformation besteht darin, dass die DCT-Transformation, die u.a. die Basis für JPEG bildet, die Umsetzung erfolgt in zwei Dimensionen, wobei nicht alle

Frequenzen gleichmäßig behandelt werden.

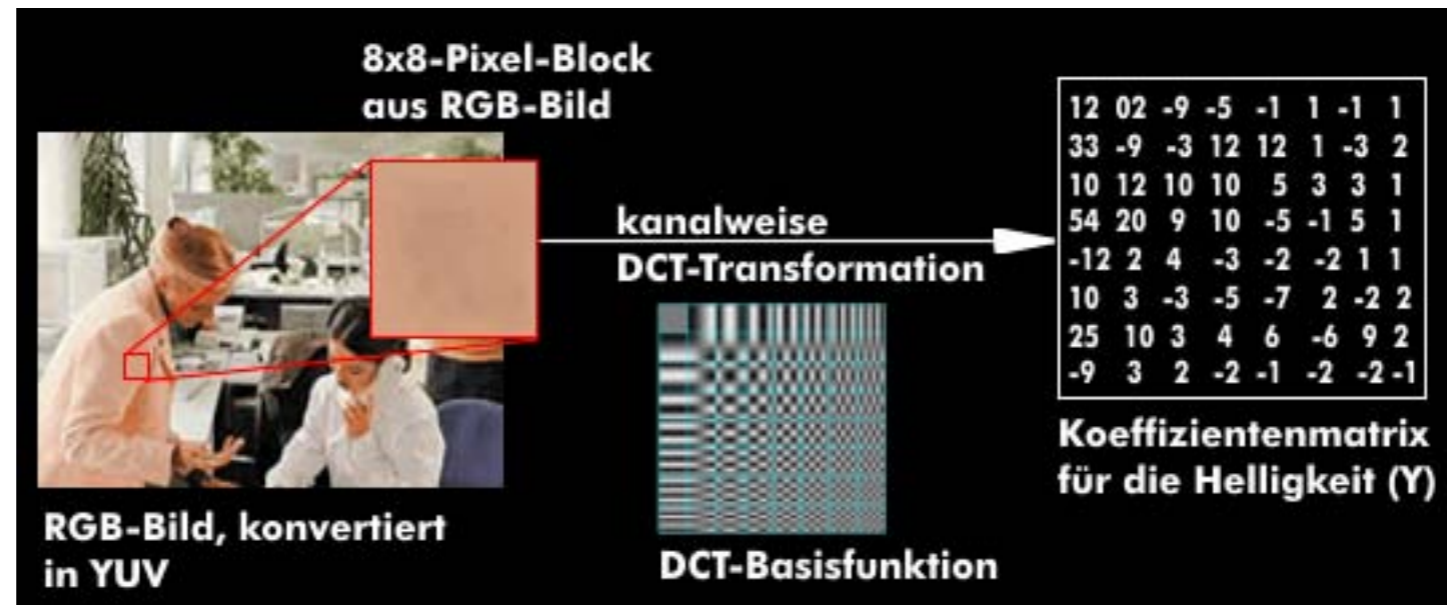
Tiefere Frequenzen, die einen langsamen Anstieg oder Abfall der Helligkeit repräsentieren, werden anders behandelt als hohe Frequenzen, die für einen schnellen Helligkeitswechsel und damit für die Auflösung und Kontur stehen. Die Unterscheidung zwischen tieferen und höheren Frequenzen ist wichtig, da dadurch die Bildschärfe beeinflusst werden kann.

Vom Verfahren her werden bei der DCT-Transformation die Einzelbilder in



DCT-Teilspektren

DCT-Transformation eines 8x8-Pixelblocks in eine Koeffizientenmatrix bei JPEG



8x8-Blöcke zerlegt. Jeder Block wird als zweidimensionaler Vektor mit 64 Pixelwerten betrachtet, der die Helligkeits- und Farbwerte enthält, wobei

jeder Block auch einen bestimmten Frequenzbereich repräsentiert. Die Helligkeits- und Farbblöcke werden getrennt behandelt.

Wie aus den DCT-Basisfunktionen ersichtlich enthält der erste Wert (links oben) den Mittelwert für die Helligkeit von allen 64 Pixel des 8x8-Blocks. Die weiteren Werte in der Horizontalen repräsentieren die Helligkeitsänderungen, beginnend bei geringen Änderungen und ansteigend bis hin zu den kürzesten Änderungen. So dass der zweite Wert (2) in der Horizontalen ein Maß für den Helligkeitsverlauf von links nach rechts ist. Der dritte (3) zeigt an, ob sich die Helligkeit bis zu Mitte und danach wieder bis zum Rand ändert; der Wert spiegelt somit eine zweigeteilte Achse wider. In der Vertikalen werden die gleichen Kriterien angesetzt, allerdings für die vertikale Bildachse.

Aus diesen Daten wird eine Koeffizientenmatrix erstellt, deren Daten bei JPEG mittels Quantisierung komprimiert werden. Die DCT-Transformation selbst bewirkt keine *Datenkompression*.

DivX DivX ist eine *Videokompression*, entwickelt von DivXNetworks und basierend auf "MPEG-4 Part 2 Video Coding". Die patentierte *Kompression* arbeitet sehr effizient und reduziert Videodateien auf etwa ein Zehntel des unkomprimierten Dateiformats. DivX komprimiert ausschließlich Video, der Ton muss mit einer Audiokompression,

*Codierung und
Abspielmöglichkeiten der
Video-Codecs DivX, XviD
und 3ivx*

Wiedergabe	Wiedergabe-Codecs
DivX	DivX, FFDSHow, 3ivx Decoder
XviD	XviD, FFDSHow, 3ivx Decoder
3ivx	3ivx Decoder
Codierung	Codec
DivX	DivX Pro Vers., DivX freeware
XviD	Free GPL XviD Codec
3ivx	Free Trial 3ivx Filer Suite

wie beispielsweise MP3, komprimiert werden. Da DivX mit Übertragungsraten von ca. 600 kbit/s übertragen werden kann, eignet sich dieses Verfahren auch für das Abspielen von Video aus dem Internet. DivX basiert auf einer Optimierung der psychovisuellen Wahrnehmung des Menschen. Wie bei der Audiokompression von MP3 wurden bei DivX bestimmte Algorithmen in die

Videokompression aufgenommen, die eine gesteigerte Datenreduktion ermöglichen. Dazu gehören beispielsweise der Rate Distortion Algorithmus oder die Peak Signal to Noise Rejection (PSNR). Die Verbesserung der Kompression ist auch auf eine interne Architekturänderung des Decoders zurückzuführen, der gleichzeitig mehrere Datenblöcke verarbeitet.

DivX-Dateien müssen immer mit der Version abgespielt werden, mit der sie auch komprimiert wurden, da neuere DivX-Versionen nicht kompatibel sind zu älteren Versionen.

<http://www.divx.com>

ENDEC,
encoder/decoder

Eine elektronische Funktionsgruppe zur Umwandlung (Codierung) analoger Signale (z.B. Sprache) in digitale und Decodierung digitaler Signale in analoge. Der Endec unterstützt in einer Einheit beide Funktionen und ist häufig mit anwendungsspezifischen Eigenschaften ausgestattet.

Farb-Subsampling
color subsampling

Das *Videosignal* besteht aus dem *Luminanzsignal* für die Helligkeit und dem *Chrominanzsignal* für die Farbe. Da das Auflösungsvermögen des menschlichen Auges für Farbe wesentlich geringer ist als für Helligkeitsinformationen, werden bei allen videorelevanten Farbmodellen die Chrominanzsignale mit einer wesentlich

geringeren Bandbreite als das Luminanzsignal "Y" übertragen. Bei einer Video-Bandbreite von 5 MHz für das Luminanzsignal stehen den *Farbdifferenzsignalen*, "U" bzw. Cb und "V" bzw. Cr, nur 1 MHz bis 2 MHz zur Verfügung. Das bedeutet auch, dass die Farbdifferenzsignale ohne sichtbare Qualitätseinbußen mit einer geringeren Sampletiefe quantisiert werden können.

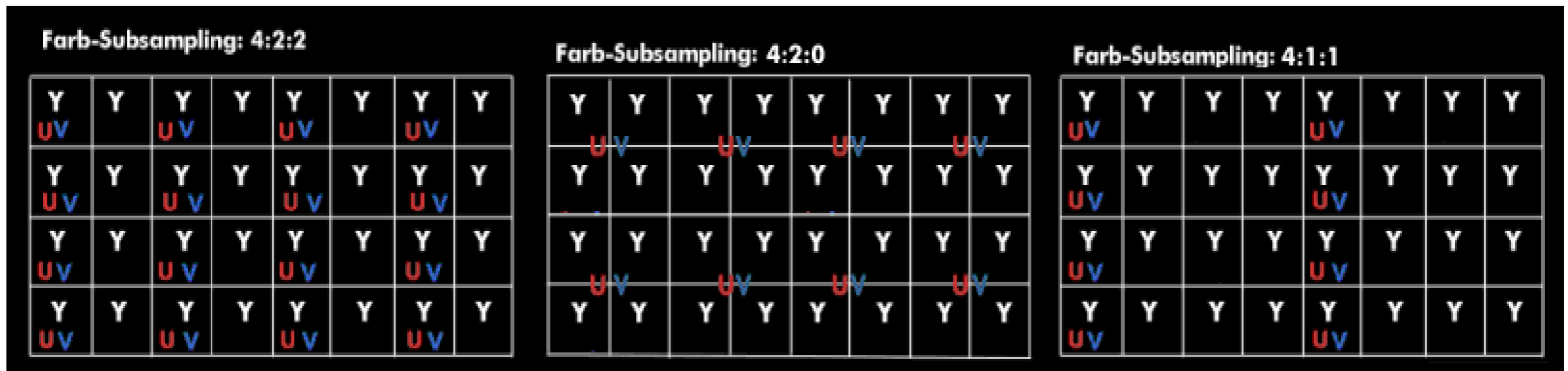
Das Farb-Subsampling sagt aus, in welchem Verhältnis die Helligkeitsabtastung zur Farbabtastung steht. Die Angabe 4:x:x bedeutet, dass die Luminanz mit dem Faktor 4, Cb mit dem Abtastverhältnis von x und Cr ebenfalls mit dem Abtastverhältnis von x quantisiert werden. Die Basisabtastrate beträgt 3,375 MHz, d.h. die vierfache beträgt 13,5 MHz. Mit dieser Abtastrate wird die Luminanz abgetastet.

Bei einem Abtastverhältnis von 4:4:4 ist die Auflösung der beiden Farbdifferenzsignale identisch der des Luminanzsignals. Alle Signale werden mit 13,5 MHz abgetastet. Dieses Subsampling liefert zwar die beste Bildqualität, es wird aber nur selten angewendet, und wenn, dann in der Produktion um die Farbverluste klein zu halten.

Beim Farb-Subsampling von 4:2:2 wird bei jedem Pixel der Luminanzwert abgetastet, bei jedem zweiten Pixel die beiden Chrominanzsignale. Das Helligkeitssignal wird mit 13,5 MHz und die Farbdifferenzsignale "U" und "V" mit 6,75 MHz abgetastet. Dieses Farb-Subsampling entspricht *ITU-R 601*.

Bei einem Subsampling von 4:1:1 werden die Farbdifferenzsignale bei jeder vierten

Farb-Subsampling mit Luminanz- und Chrominanzsignalen



Abtastung des Luminanzsignals abgetastet. Das bedeutet, dass vier Farbpixel einen U- und V-Wert haben. Die Abtastrate für die beiden Chrominanzsignale ist damit nur ein Viertel von der des Luminanzsignals.

Beim Subsampling von 4:2:0, erfolgt die Abtastung der Chrominanzsignale für jeweils vier quadratisch angeordnete nebeneinander liegende Pixel. Wie bei den anderen Subsampling-Verfahren auch, wird das Luminanzsignal bei jedem Pixel abgetastet. Im Prinzip handelt es sich um eine Kombination zweier Verfahren, die auf die geraden und ungeraden Zeilen angewendet werden: Die ungeraden Zeilen werden im Verhältnis von 4:2:2 abgetastet, die geraden im Verhältnis von 4:0:0. Die Datenreduktion liegt wie bei der 4:1:1-Abtastung bei 50 %. Das 4:2:0-Subsampling wird bei den meisten DV-Formaten beim PAL-Fernsehstandard eingesetzt.

Farbdifferenzsignal
(color-difference signal)

Farbdifferenzsignale sind Farbsignale, die aus den Primärfarben Rot (R), Blau (B) und dem *Luminanzsignal* (Y) gebildet werden. Im PAL-Fernsehstandard werden die Farbdifferenzsignale mit "U" und "V" bezeichnet, bei NTSC mit "I" und "Q". Sie enthalten die *Chrominanz* und werden im Gegensatz zum Luminanzsignal mit einer geringeren Bandbreite übertragen. Die Farbdifferenzsignale werden in Verbindung mit dem *Farb-Subsampling* zur Datenreduktion verwendet, so beispielsweise bei JPEG.

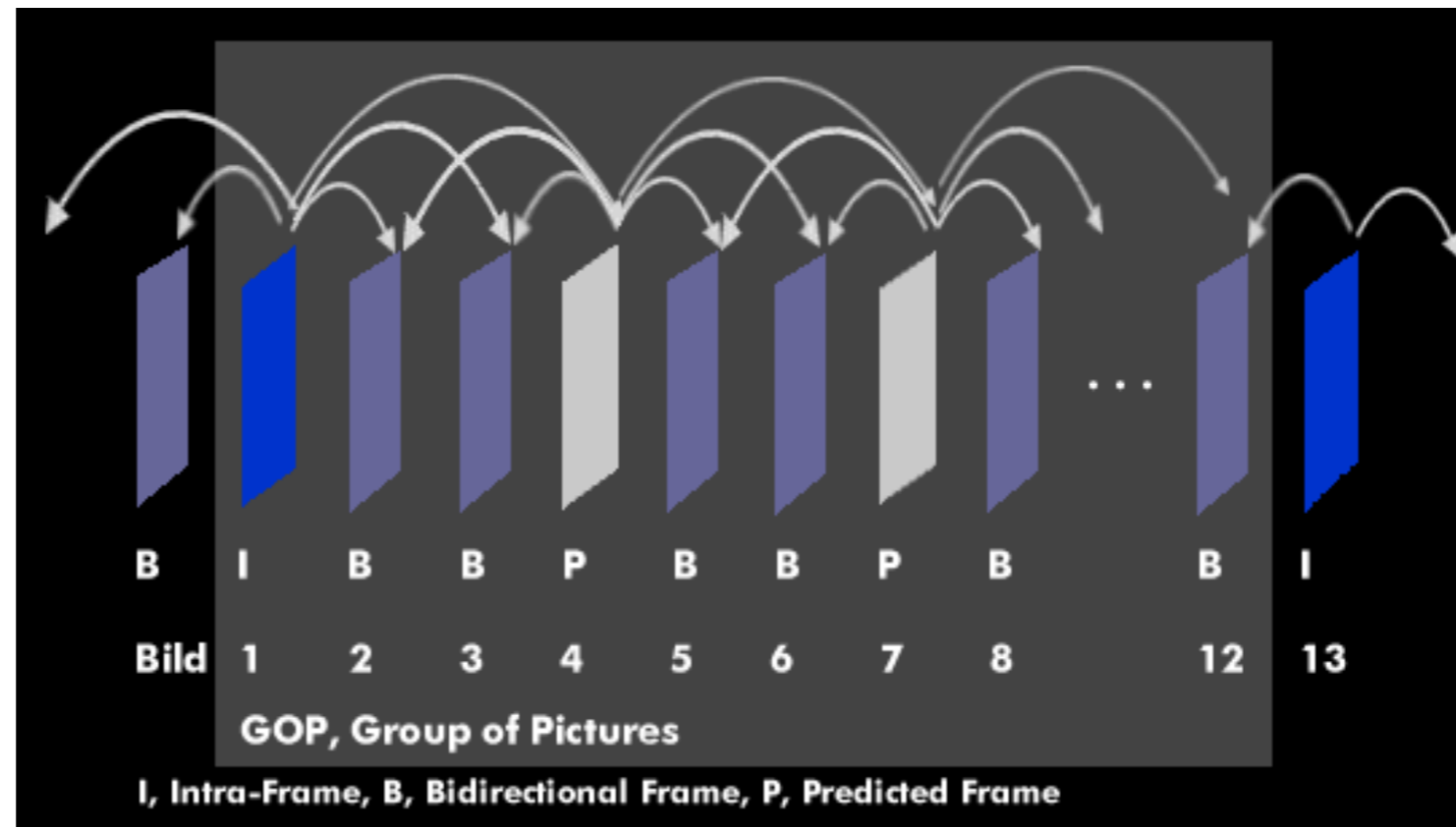
Zwischen den Farbdifferenzsignale, dem Luminanzsignal und den Primärfarben (R), (G) und (B) besteht ein direkter mathematischer Zusammenhang: $U = B - Y$, $V = R - Y$. Das U-Signal wird auch mit Cb und das V-Signal mit Cr bezeichnet.

Cb und Cr werden im ITU-Standard *BT.601* für Digitalvideo benutzt. Cb entspricht der Farbabweichung Blau-Gelb und Cr der Farbabweichung Rot-Türkis.

GOP
group of picture

Bei der *MPEG-Kompression* werden auf Grund der Ähnlichkeit von aufeinander folgenden Videobildern Makroblöcke mit unterschiedlichen Informationsinhalten gebildet, die so genannten Group of Pictures (GOP). Innerhalb dieser Bildgruppen werden die Bilder nacheinander komprimiert, ausgehend von dem datenintensiven Anfangsbild der Gruppe, dem *Intra-Frame* (I-Frame). Dieses Bild bildet die

GOP-Struktur bei der MPEG-Kompression



Informationsbasis für die nachfolgenden Bilder und definiert den Bezug innerhalb der Videosequenz. Das I-Frame wird ähnlich JPEG komprimiert. Die nachfolgenden Bilder sind

die so genannten *Bidirectional-Frames* (B-Frames), weil sie gleichermaßen von den vorhergehenden Bildern als auch den nachfolgenden abhängen und diese als Referenz nutzen. Es handelt sich um besonders hoch komprimierte Bilder, da lediglich die Unterschiede zu den benachbarten Bildern als Informationen erhalten. Nach einer B-Frame-Gruppe folgt ein *Predicted Frame* (P-Frame), das auf der Prädiktion vom Intra-Frame basiert. Das P-Frame orientiert sich an dem Referenz-Frame, dem I-Frame, bildet die Referenzpunkte für die I-Frames und ist für die Dekompression besonders wichtig.

Die Reihenfolge der verschiedenen Frames wird bei der *MPEG*-Codierung festgelegt. Es ist allerdings darauf zu achten, dass nur das Intra-Bild zur Bearbeitung benutzt werden kann. Eine typische GOP-Struktur beginnt mit dem I-Frame, endet vor dem nächsten I-Frame und sieht folgendermaßen aus: IBBPBBPBBPBB.

Neben der Video-Codierung umfasst die MPEG-Kompression auch die mehrkanalige Audio-Codierung, die in *MPEG-1*-Audio und *MPEG-2* umgesetzt wird.

H.261 Der H.261-Standard beschreibt die *Kompression* von Bildsequenzen für Videokonferenzen und Bildtelefonie und legt die Grundprinzipien der digitalen Bewegtbildübertragung und Datenkompression fest. Das Verfahren basiert auf JPEG und ist für ISDN-Anwendungen mit $n \cdot 64$ -kbit/s-Kanälen ausgelegt. Die *Kompressionsraten* liegen bei 100:1 bis 2.000: 1. H.261 ist Teil von *H.320*. In H.261 wird eine maximale Verzögerungszeit für die Kompression und Dekompression von zusammen 150 ms festgelegt. Als Dateiformate für das Komponentenvideo kommen das CIF-Darstellformat und QCIF zum Einsatz. H.261 arbeitet nach dem *YUV-Farbmodell*, mit *DCT-Transformation* und Huffman-Codierung und hat, ebenso wie *MPEG*, einen Vorhersage-Mechanismus für die nachfolgende Bildsequenz mit *Intra-Frames* und *Predicted-Frames*.

H.263 Video-Codierung ist für die Kommunikation audiovisueller Dienste mit Bitraten unter 64 kbit/s. Der Codec H.263 ist eine Erweiterung von *H.261* und ersetzt diesen teilweise.

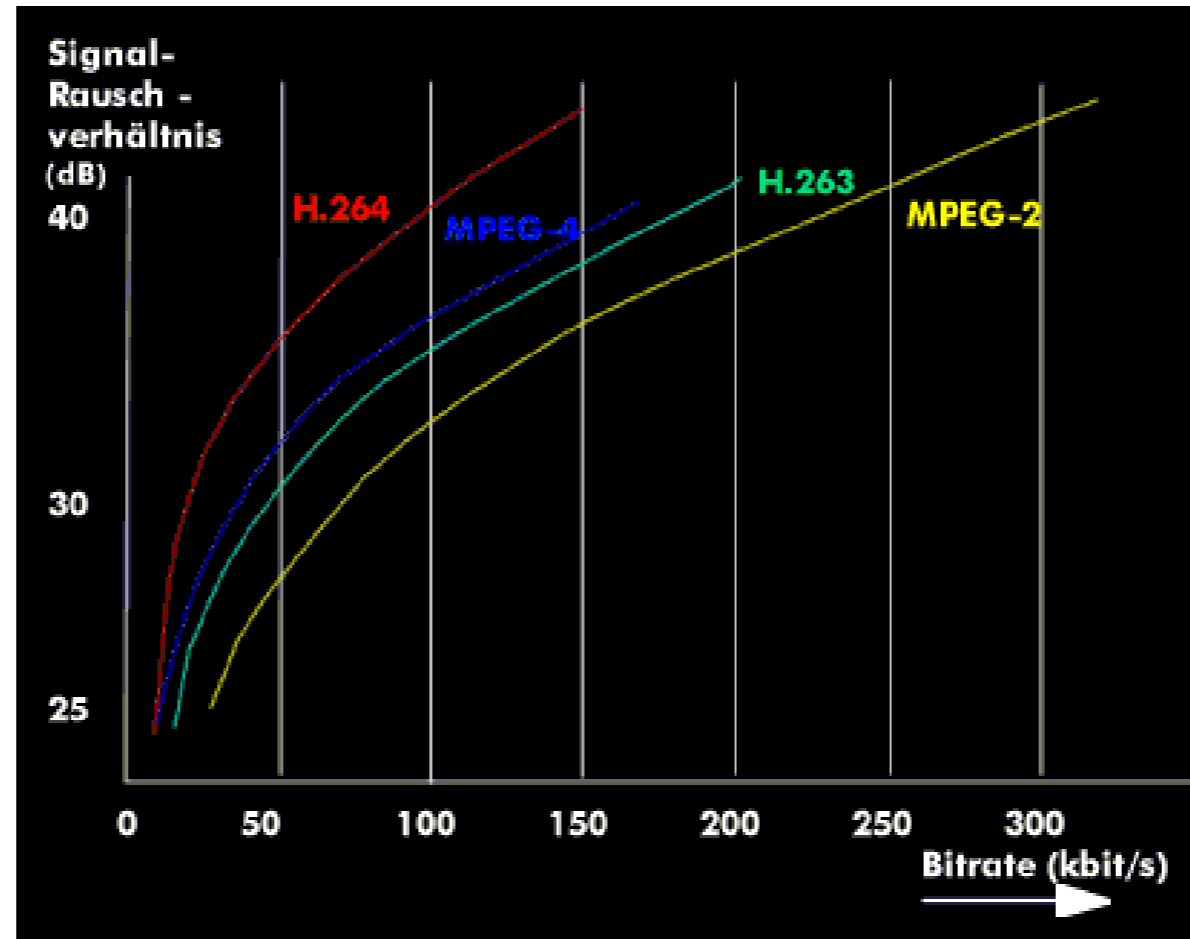
Mit H.263 wird die Qualität von niedrig auflösenden Bildern in der Größe einer Kreditkarte auf außergewöhnlich detaillierte und qualitativ hochwertige Images erweitert. Dies kann z.B. bei Anwendungen in der Telemedizin eingesetzt werden. H.263 ist rückwärtskompatibel zum Quarter Common Intermediate Format (QCIF) und optional auch zum CIF-Darstellformat und zu SQCIF (Sub-QCIF) des *H.320*-Standards.

H.263 hat gegenüber H.261 eine verbesserte Bildnummerierung, fünf unterschiedliche Auflösungen und verhandelbare Leistungsparameter. Die Genauigkeit ist gegenüber H.261 verringert.

Der Codec H.263 wird auch von Media-Playern wie dem RealPlayer oder Quicktime unterstützt.

H.264-Kompression Im Gegensatz zu anderen *Videokompressionen* arbeitet *H.264* mit kleineren Pixel-Blöcken, die nur 4 x 4 Pixel groß sind. Dadurch verringert sich der Transformationsaufwand und die Berechnung kommt mit wenigen Grundoperationen

Die verschiedenen Video-Codecs im Vergleich



aus: der Addition, Subtraktion und binären Verschiebeoperationen. Bei einer Hardware-Implementierung führt dies zu einer wesentlichen Geschwindigkeitserhöhung. Darüber hinaus reduzieren die kleinen Blockgrößen *Artefakte*, die durch die *Kompression* entstehen; das sind Störsignale, die durch die *verlustbehaftete Kompression* verursacht werden. Bei der Decodierung erfolgt eine

Farbkonversion vom *YUV-Farbmodell* in RGB. Außerdem kann ein Grafik-Overlay erfolgen bei der die Videodarstellung mit Grafiken und Texten überlagert wird.

H.264/AVC,
H.264 advanced video coding

Der ITU-Standard H.264, auch bekannt als *MPEG-4 Part 10* und *MPEG-4 AVC*, ist eine hochentwickelte *Videokompression*, die sich durch qualitativ hochwertige Darstellqualität und einen hohen Kompressionsfaktor auszeichnet. Die Bitratenreduktion gegenüber *MPEG-2* kann bei vergleichbarer Qualität durchaus über 50 % betragen. H.264/AVC, eine Weiterentwicklung von MPEG-4, basiert auf einer neuen Codiermethode, sie ist netzwerkfähig und hat mehrere anwendungsspezifische Profile, die vom Videotelefon über Videokonferenzen, hochauflösendes Fernsehen (HDTV) bis hin zur Archivierung reichen. Die *H.264-Kompression* unterstützt konstante (CBR) und variable Bitraten (VBR) und hat dabei eine moderate Datenrate von 8 Mbit/s.

Zu den H.264-Profilen gehören das Baseline-Profil, Main-Profil und das Extended-Profil.

Das Baseline-Profil zielt auf die Echtzeit-Kommunikation für Endgeräte, und somit auf Videokonferenzen und Videotelefonaten. Es verwendet die fortlaufenden *I-Frames* und *Predicted-Frames* sowie die CAVLC-Transformation. Aufgrund der robusten Fehlertoleranz und niedrigen Latenzzeiten hat das Baseline-Profil seine besonderen Stärken bei dynamischen Netzen auf denen die Videodaten mit geringer Bandbreite übertragen werden.

Das auf den Broadcastmarkt zielende Main-Profil benutzt hingegen alle Bildframes der Group of Picture (GoP), es unterstützt das Zeilensprungverfahren und benutzt CAVLC- und CABAC-Transformationen.

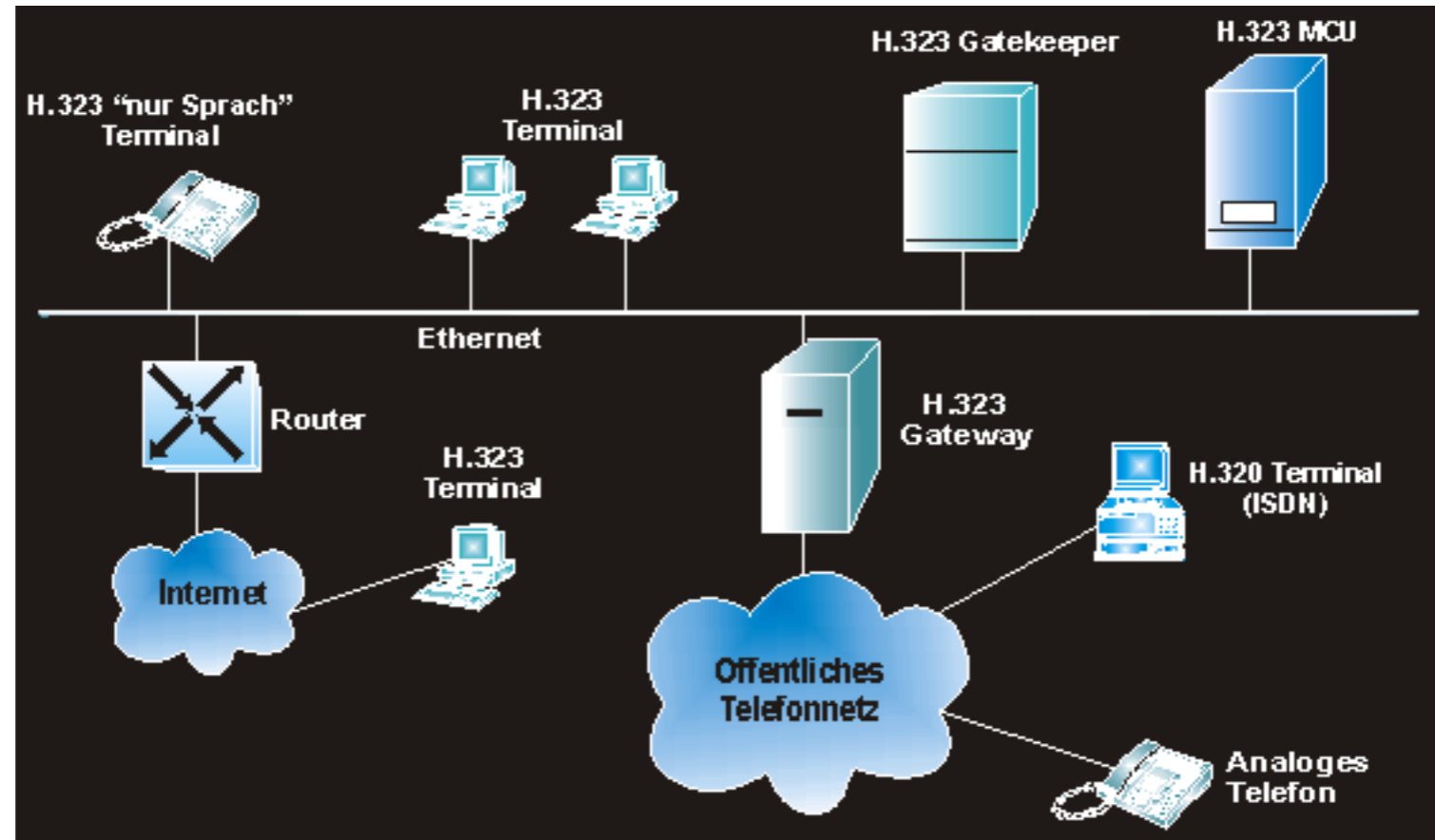
Und im Extended-Profil wird das Zeilensprungverfahren und der Progressive Scan unterstützt. Das Extended-Profil zielt auf die Abwicklung von problembehafteter Kommunikationskanälen ab.

Zur Erhöhung der Datenreduktion variiert Advanced Video Coding u.a. die Blockgröße in Abhängigkeit von den Bildänderungen. H.264/AVC eignet sich für die Übertragung breitbandiger Videoströme über das Internet und erreicht Durchsatzraten von etwa 1 Mbit/s.

H.320 H.320 ist einer der dominierenden Standards für Videokonferenzen über ISDN. Festgelegt sind mehrere Normen für die Video- und die Audioübertragung. Bei ISDN beträgt die minimal benötigte Übertragungsrate 64 kbit/s, die maximale 1.920 kbit/s. Der Standard ist ebenso für Mietleitungen mit X.21 gültig, mit einer maximalen Datenrate von 2,048 Mbit/s.

H.320 beschreibt als übergeordneter Standard die technischen Basisfunktionen für die Endgeräte, die die untergeordneten Standards *H.261*, *H.221*, *H.230* und *H.242* erfüllen müssen. Darin enthalten sind die Spezifikationen für den Codec, das Verpacken und Übertragen der multimedialen Daten von Video und Audio sowie die Signalisierung. H.320 unterstützt die wichtigsten Fernsehstandards; den NTSC- ebenso wie den PAL-Fernsehstandard.

H.323



H.323-Strukturmodell

H.323 ist ein internationaler ITU-Standard für die Sprach-, Daten- und Videokommunikation über paketorientierte Netze, der die spezifischen Fähigkeiten von Endgeräten im IP-Umfeld festlegt. H.323, das funktional vergleichbar ist mit dem SIP-Protokoll, wurde für die Übertragung von Multimedia-Applikationen entwickelt und bildet die Grundlage für VoIP und Video-over-IP. Über diesen Standard wird die Echtzeitkommunikation in LANs definiert, die keine garantierte Dienstgüte zur Verfügung stellen können.

Der H.323-Standard besteht aus einer ganzen Reihe von Protokollen für die Signalisierung, zum Austausch von Endgerätefunktionalitäten, zur Verbindungskontrolle, zum Austausch von Statusinformationen und zur Datenflusskontrolle. Der Standard ist mehrfach überarbeitet worden und definiert in der dritten Version die Übertragung von Leistungsmerkmalen.

Der Multimedia-Standard kennt drei verschiedene Signalisierungsfunktionen zwischen dem H.245-Steuerungskanal, dem Q.931-Signalisierungskanal und dem

Videokonferenz-Standards
der ITU

	H.320	H.321	H.322	H.323	H.324
Verabschiedet	1990	1995	1995	1998	1996
Netzwerk	ISDN	ATM/LAN	Paketnetze mit QoS	Paketnetze ohne QoS	PSTN
Videocodierung	H.261 H.263	H.261 H.263	H.263 H.261	H.263 H.261	H.261 H.263
Audiocodierung	G.711 G.722 G.728	G.711 G.722 G.728	G.711 G.722 G.728	G.711 G.722 G.728 G.723 G.729	G.723
Multiplex	H.221	H.221	H.221	H.255.0	H.223
Kontrolle	H.230 H.243	H.242	H.242 H.230	H.245	H.245
Mehrpunkt-Verbindungen	H.231 H.243	H.231 H.243	H.231 H.243	H.323	
Datenübertragung Schnittstelle	T.120 I.400	T.120 AAL I.363 AJM I.361 PHY I.400	T.120 I.400 TCP/IP	T.120 TCP/IP	T.120 V.34-Modem

RAS-Kanal.
Der Standard ist abgeleitet aus dem H.320 Multimedia-Standard für ISDN. Er hat definierte Netzübergänge zwischen ISDN und ATM und garantiert die Interoperabilität der Herstellerprodukte untereinander, wobei vier Hauptgruppen definiert sind: H.323-Gateways, Gatekeeper, Terminals und Multi Control Units (MCU). Für die IP-Telefonie ist der Gatekeeper die wichtigste Komponente, weil er die erforderliche Software für

das Management von so genannten Zonen und Rufdiensten umfasst.

H.324 Die ITU-T-Empfehlung H.324 beschreibt ein Terminal für die Multimedia-Kommunikation mit geringen Bitraten. Darüber hinaus umfasst die Empfehlung H.324 weitere Standards für *Video-Codex*s mit geringen Datenraten, die über V.34 mit analogen Telefonleitungen arbeiten, aber auch über ATM, ISDN und über Mobilfunknetze.

Die Empfehlung dient der Systemspezifikation für ein Videotelefon. Kompatibilität mit H.320, Netzwerk-Codeumsetzung, Multipoint-Unterstützung und Continuous Presence werden zur Verfügung gestellt.

Mit H.324M wird die Einbeziehung der Videotechnologien in die 3. Generation der Mobilfunkgeräte vorangetrieben. Das entsprechende Protokoll heißt 3G-H.324M.

HDX4 HDX4 ist ein lizensierungspflichtiger *Video-Codec*, basierend auf MPG-4. Der HDX4-Codec zeichnet sich durch qualitativ hochwertige Videodarstellungen aus, er ist extrem schnell und benötigt weniger Speicherbedarf als vergleichbare *Codecs* wie *XviD* oder *DivX*. Angegeben wird eine Datenreduktion auf etwa 15 % des ursprünglichen Wertes.

HDX4 unterstützt alle gängigen Bildschirmauflösungen von 32 x 32 bis hin zu HDTV mit 1.920 x 1.080 Bildpunkten. neben dem PAL-Fernsehstandard wird auch NTSC und HDTV unterstützt sowie die Auflösungen für das CIF-Darstellformat und für QCIF.

Wiedergabeseitig unterstützt HDX4 alle Player, die die Schnittstellen von Video for Windows (VfW) und DirectShow benutzen, also auch den Windows Media Player. Für die Wiedergabe stellt HDX4 einige Funktionen für die Synchronität von Bild und Ton und zur Verbesserung der Bildqualität bereit.

Indeo,
Intel video

Das von Intel entwickelte Indeo ist ein *Kompressionsverfahren* für Video- und Tonsignale. Es arbeitet auf Softwarebasis und wurde für den i750-Prozessor entwickelt. Indeo hat wie der *Cinepak-Codec* eine Farbtiefe von 24 Bit und eine maximale Auflösung von 320 x 240 Pixeln. Das in Indeo benutzte Komprimierungsverfahren arbeitet asymmetrisch und ist verlustbehaftet. Die *Kompressionsrate* liegt bei etwa 7:1.

Die Datei-Extension ist *.avi; die Dateien können mit Video for Windows und der Medienwiedergabe abgespielt werden.

I-frame
Intra-Frame

Intra-Frame ist ein Rahmenformat von *MPEG* mit dem der Bildbereich von Digitalvideos in Makroblöcke geteilt wird. Die Intra-Frames sind vollständig gespeicherte Einzelbilder und berücksichtigen die Bild-Redundanz zwischen den Zeilen. Sie werden aus der *Luminanz* und dem *Farb-Subsampling* gebildet und haben eine Größe von 8 x 8 Pixeln. Die einzelnen Makroblöcke enthalten Informationen über das Einzelbild und den zeitlichen Bezug zu den folgenden Bildern. Nur jedes 10. bis 15. Bild wird als Intra-Frame gespeichert. Bevor die Intra-

Frames komprimiert werden, werden sie mittels *DCT-Transformation* in die Frequenzebene transformiert und anschließend quantisiert. Die *Kompression* in JPEG erfolgt mittels Huffman-Codierung.

Intra-Frames haben Dateigrößen von etwa 100 kByte, die um Faktoren von etwa 10:1 komprimiert werden. Aus den Intra-Frames werden die *Predicted-Frames* und die *Bidirectional-Frames* abgeleitet.

In Abhängigkeit vom Filminhalt erscheinen Intra-Frames in der Regel zweimal innerhalb von 30 Frames.

JND, just noticeable difference

Die mathematische Qualitätsbewertung von *Video-Codecs* erfasst nicht alle vom menschlichen Wahrnehmungsvermögen erkennbaren Qualitätsmerkmale der Videodarstellung von Video-Codecs. Aus diesem Grund gibt es eine psychovisuelle Bewertung, die sich an der menschlichen Wahrnehmung orientiert und die kleine gerade noch erkennbare Unterschiede zur Bewertung heranzieht. Dieses Qualitätsverfahren heißt Just Noticeable Difference (JND) und arbeitet mit einem JND-Index.

Ein JND ist definiert als die Wahrscheinlichkeit mit der ein Zuschauer mit 75 %-iger Sicherheit Unterschiede zwischen dem Originalvideo und dem dekomprimierten Video erkennt. Höhere JND-Werte gehen bei 4 JND bereits gegen 100 %.

JVT, joint video team

Das Joint Video Team (JVT) befasst sich u.a. mit dem technischen Design von *Video-Codecs* wie dem für *H.264/AVC*. Das JVT besteht aus Experten der Video Coding Experts Group der ITU sowie der Motion Pictures Expert Group (*MPEG*), die wiederum zur International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) gehört.

<http://www.itu.int/ITU-T/studygroups/com16/jvt/>

Kompression (compression)

Das Prinzip der Datenkompression von Text-, Grafik- und Bewegtbilddaten beruht auf der Eliminierung redundanter Zeichen bzw. einer dynamischen Zuordnung von Datenbits in Abhängigkeit oder Häufigkeit eines Zeichens (Character). Im Gegensatz

dazu werden bei der Sprachkompression effiziente Codierverfahren bei den Codes eingesetzt.

Ausgefeilte Kompressionsalgorithmen ermöglichen die Datenkompression von beliebigen Datenströmen, auch wenn es sich um Computerprogramme oder Zufallsfolgen handelt.

Ziel der Datenkompression ist es, die Dateigrößen in ihrem Umfang zu reduzieren, um dadurch einen schnelleren Datendurchsatz zu erreichen und mit weniger Speicherbedarf auszukommen.

In der Datenkompression unterscheidet man zwischen echter Kompression, verlustfreier Kompression und *verlustbehafteter Kompression*. Die Verfahren setzen auf unterschiedlichen Algorithmen auf und erfassen häufig nur Zustandsänderungen gegenüber dem vorherigen Zustand.

Besondere Bedeutung finden die Kompressionsverfahren bei der Bildkompression (Grafiken, Fotos), bei Audio sowie bei Bewegtbildern (Video). Bei der Kompression von Bildern und Grafiken setzt man auf die Redundanz von Farben, Flächen und Linien, bei Audio auf die Irrelevanz von Tonfrequenzen und Pegeln und bei der *Videokompression* auf die Unterschiede zwischen aufeinander folgenden Bildern.

Bekannte Kompressionsverfahren für Bilder und Grafiken sind die Lauflängencodierung, TIFF und JPEG, für die Sprachkompression GSM 06.10, das A-Law-Verfahren nach H.321 und ADPCM, für die Audiokompression *MPEG*, MP3, ATRAC und Advanced Audio Coding (AAC) und für die Video- und Bewegtbildkompression *H.261* und *MPEG*.

Verfahren für die Textkompression sind die Lauflängencodierung (RLE), das Word-Coding, die Huffman-Codierung und der LZW-Algorithmus.

In der Modemtechnik wird zur Datenkompression häufig das Kompressionsverfahren V.42bis eingesetzt, das als Vorgänger MNP4 und MNP5 hat.

Kompressionsrate
(*compression rate*)

Die Kompressionsrate spiegelt das Verhältnis der Dateigröße wider, das zwischen einem komprimierten und einer unkomprimierten Signal besteht. Das Verhältnis kann als solches angegeben werden, beispielsweise 1:100, und besagt, dass die

Dateigröße nach der *Kompression* einhundertstel der vorherigen Dateigröße beträgt. Statt der Verhältnisangabe kann die Kompressionsrate auch als Prozentsatz angegeben werden, beispielsweise 94 %. Das bedeutet, dass die Dateigröße der komprimierten Datei um 94 % gegenüber der Originaldatei verkleinert wurde. Die Kompressionsrate ist wesentlich von der Art der Kompression abhängig. Eine verlustfreie Kompression wird immer eine geringere Kompressionsrate haben, als eine *verlustbehaftete Kompression*. Des Weiteren ist die Kompressionsrate von der Art der Medien abhängig: Daten haben eine geringere Kompression als beispielsweise Audio oder Video. Der Spielraum den Kompressionsraten umfassen können liegt zwischen 1:1,.. und 1:1.000 und höher wie beispielsweise bei der Kompression von Videosequenzen nach *H.261*.

Luminanz
(*Y, luminance*)

Die Luminanz (Y) ist die Leuchtdichte einer Videodarstellung. Sie ist ein Maß für den Helligkeitseindruck. Da das menschliche Auge Helligkeitsunterschiede wesentlich besser auflöst als Farbnuancen, bestimmt die Luminanz den Schärfeeindruck des Videobildes. Das Luminanzsignal wird daher in allen Fernseh- und Video-Standards mit wesentlich höherer Bandbreite übertragen als der Farbanteil, die so genannte *Chrominanz* (C). In Farbfernsehsystemen setzt sich das Luminanzsignal aus der Addition von Farbanteilen der Primärfarben Rot (R), Grün (G) und Blau (B) zusammen. Diese Farbanteile sind durch die Augenempfindlichkeit bedingt, die im Blau- und Rotbereich wesentlich unempfindlicher ist als im Grünbereich: $Y = 0,299(R) + 0,587(G) + 0,114(B)$.

MJPEG,
motion JPEG

Motion JPEG ist ein *Kompressionsverfahren* für *Codecs*, bei dem jedes Einzelbild einer Videosequenz mit JPEG komprimiert wird. Es ist somit eine Umsetzung von JPEG auf Bewegtbildsequenzen. Dem Verfahren nach werden bei MJEP zuerst die Farbwerte mit *Farb-Subsampling* reduziert, da das menschliche Auge Helligkeitsunterschiede besser auflösen kann. Danach werden die Bilddaten mittels *DCT-Transformation* in die Frequenzebene transferiert und mit der Huffman-

Codierung und der Lauflängencodierung in ihren Bilddaten komprimiert. Zuletzt werden die bereits komprimierten Daten mit einer verlustfreien Kompression weiter reduziert.

MJPEG ist somit ein Verfahren, bei dem die Inhalte der Einzelbilder reduziert werden, und nicht wie bei *MPEG* die Unterschiede zwischen aufeinander folgenden Bildern.

Mit der MJPEG-Technik können Videodatenströme auf ca. 30 Mbit/s reduziert werden.

MP4-Dateiformat
(*MP4 data format*)

MP4 ist ein multimediales Container-Format für *MPEG-4*-Dateien. In einem solchen Container-Format können verschiedene multimediale Streams in einer einzelnen Datei kombiniert werden, so beispielsweise Audio und Video. Die bekanntesten multimedialen Container-Formate sind u.a. MP4, das AVI-Dateiformat, *MPEG*, Matroska und Quicktime. MP4 ist das von MPEG definierte und standardisierte Container-Format. Es kann für Streaming eingesetzt werden und unterstützt die verschiedensten multimedialen Inhalte wie Video, Audio, Bilder, 2D- und 3D-Animationen, Kapitelbeschreibungen, Untertitel, interaktive Anwendungen, DVD-ähnliche Menüs und einiges mehr. Zur Generierung einer MP4-Datei werden Audio- und *Videodateiformate* basierend auf MPEG-4 benutzt wie *XivD*, *DivX*, *Nero Digital*, Advanced Audio Coding (AAC), *H.264/AVC*, Quicktime, MP3, CELP, TwinVQ und andere. Diese Video- und Audiodateien werden mit entsprechenden Tools wie dem mp4creator oder dem “*3ivx mp4 muxer*” und vielen anderen in dem MP4-Container kombiniert. Ebenso andere Grafikdateien, die Kapitelbeschreibungen und die Untertitel.

MPEG, moving picture experts group
(*MPEG-Standard*)

Moving Picture Experts Group (MPEG) steht gleichermaßen für ein Standardisierungs-Gremium und für mehrere asymmetrische Codierverfahren sowie für verschiedene Dateiformate für die Bewegtbild- und Audiokompression. Die *MPEG-Kompression* wird für Digitalvideo auf Computern, für Streaming-Media und für die Übertragung von digitalen Fernsehbildern eingesetzt. 1988 begannen die ISO

und das IEC einen Standard für die Video- und Audiokompression zu entwickeln, der 1992 in der ersten Version als *MPEG-1* unter ISO 11172 verabschiedet wurde. Die Audiokompression wird u.a. in Compact Discs (CD), DVDs und im Digital Audio Broadcast (DAB) eingesetzt.

Bei MPEG handelt es sich um eine *verlustbehaftete Kompression*, das bedeutet, dass bei der Decodierung das Originalbild nicht mehr exakt generiert werden kann. Darüber hinaus ist die Codierung technisch sehr viel aufwändiger als die Decodierung.

MPEG-1 wurde 1992 verabschiedet und ist ein integrierter Standard für die Codierung von Bewegtbildern mit zugehörigem Audio zur Speicherung auf CDs. Der

Benutzeroberfläche	Bildauflösung	Übertragungsbandbreite (1 Kanal Halbduplex)	Übertragungsbandbreite/Videokonferenz (4 Teilnehmer)
Video (MPEG-1 komprimiert)	352 x 288	1,15 Mbit/s	13,8 Mbit/s
Video (MPEG-2 komprimiert)	720 x 576	4 Mbit/s	48 Mbit/s
Video (MPEG-3 komprimiert) (HDTV)	1920 x 1080	20 Mbit/s	240 Mbit/s
Video (MPEG-4 komprimiert) (Videophone)	176 x 144	0,064 Mbit/s	0,768 Mbit/s
ASCII-basierende Darstellung	40 x 80 Zeichen	0,0096 Mbit/s- 0,0144 Mbit/s	—
Graphische Benutzer- oberfläche im LAN-Verbund	800 x 600	* Spitzenlasten: bis zu 4 Mbit/s durchschnittliche Last: 5-50 kbit/s	—
* Laden von MS-Windows-Applikationen wie "Word für Windows" oder "Excel" Dauer der Lastspitze: ca. 5 Sekunden			

Übertragungsbandbreite von
Grafik-Files

Standard ist in einen Videoteil und einen Audioteil untergliedert. MPEG-1-Audio sieht die Codierung von Mono- und Stereosignale vor. MPEG-1 entspricht in der Video-Wiedergabequalität der Auflösung von VHS und eignet sich für multimediale Präsentationen auf CD-ROM oder CD-I. Für die Übertragung von Bewegtbild in Standard-Fernsehqualität und Studioqualität mit mehreren Tonkanälen wurde Ende 1994 MPEG-2 standardisiert. MPEG-2 zeichnet sich durch seine hohe Flexibilität in der Audiokompression aus, unterstützt Audio für höchste Ansprüche und kann theoretisch einen Raumklang aus bis zu 48 Kanälen bilden.

MPEG-4 kam 1999 auf den Markt und legt im Gegensatz zu MPEG-1 und MPEG-2 kein einheitliches *Kompressionsverfahren* fest, sondern definiert eine Sammlung von Audio- und Video-Kompressionsverfahren. Es bildet ein Rahmenwerk das die Integration beliebiger Medienobjekte unterstützt, ebenso wie die Interaktion zwischen menschlichen Benutzern und der audiovisuellen Welt.

MPEG-7 wurde 2001 standardisiert, um multimediale Inhalte zu beschreiben und ein Ordnungssystem für die vielen MPEG-Varianten zu schaffen. MPEG-7 berücksichtigt neben den audiovisuellen Inhalten auch Textinformationen und wird von der Description Definition Language (DDL) unterstützt. Mit XML-Anwendungen soll ein Multimedia Description Scheme (MDS) entwickelt werden, mit dem sich Audio- und Video-Inhalte mittels Schlüsselwörtern identifizieren lassen.

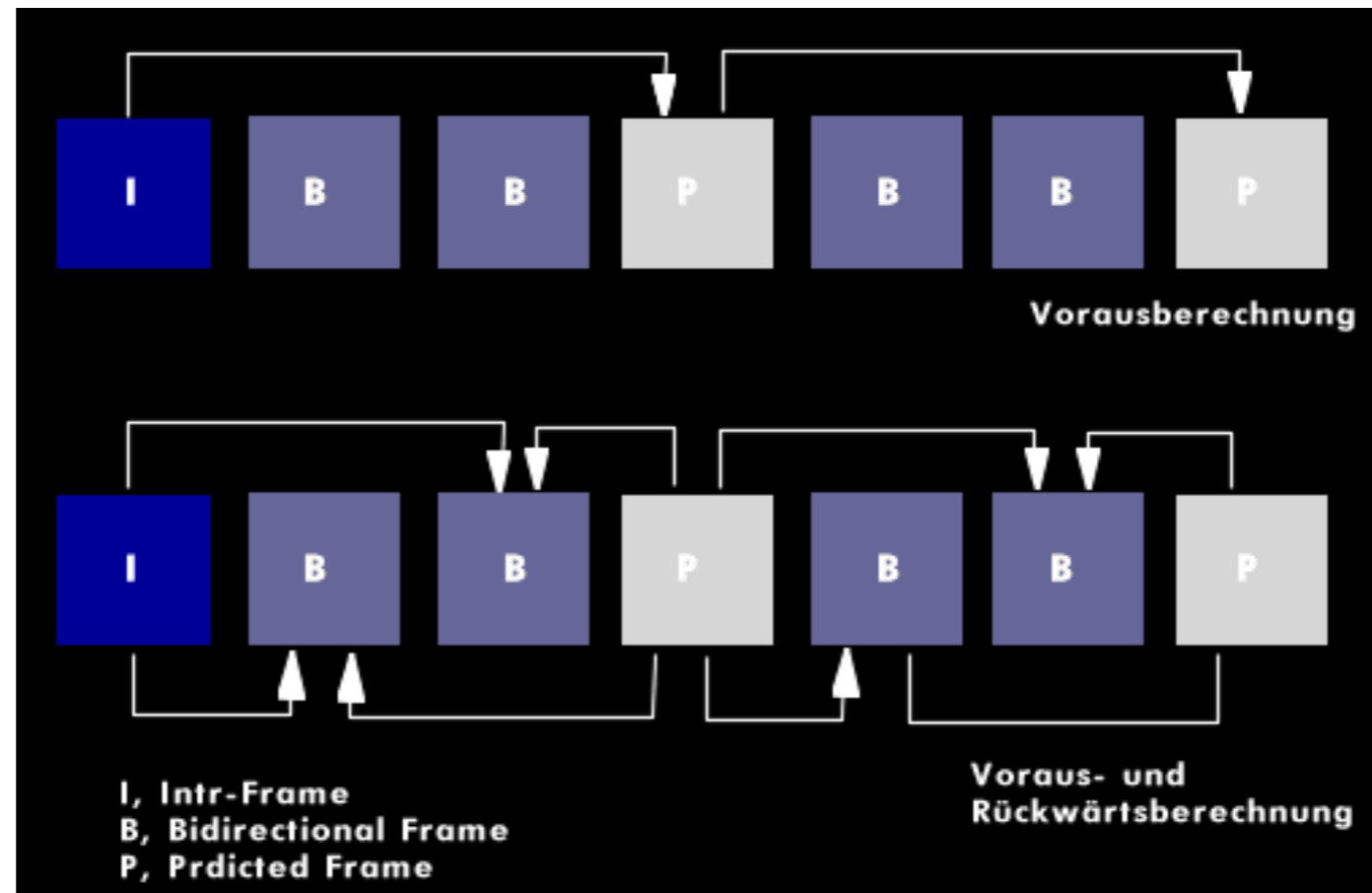
MPEG-21 unterstützt den multimedialen Nachrichtenaustausch. Die vorherige Abgrenzung in Bezug auf fest strukturierte Benutzergruppen wie Videostudios öffnet sich damit zur Konsumelektronik. Der Endverbraucher hat damit die Möglichkeit, persönliche Fotos, Audio- und Videosequenzen mit Fremdproduktionen zu kombinieren. Auch die Weiterverarbeitung kann Online erfolgen, ebenso können die semiprofessionellen Aufnahmen und Videoclips kommerziellen Anbietern angeboten werden.

<http://www.mpeg.org>

MPEG-Kompression (*MPEG compression*)

Der *MPEG-Kompression* liegt die Tatsache zugrunde, dass bei einem Videobild bis zu 95 % der digitalen Daten aus Wiederholungen bestehen und ohne merklichen

MPEG-Bildtypen



Qualitätsverlust komprimierbar sind. MPEG setzt auf Ähnlichkeiten in aufeinander folgenden Videobildern, es berücksichtigt die Tatsache, dass ein Videobild hauptsächlich aus Flächen und Linien bzw. Kanten besteht und dass die Helligkeitsverteilung in Videobildern über mehrere Bilder vergleichbar ist.

Aus diesen Annahmen resultiert der Algorithmus, in den Voraussagen, so genannte Prädiktionen, auf folgende Videobilder implementiert sind. Flächen, Kanten und Helligkeitswerte weichen in der Prädiktion von dem ursprünglichen Bild nur geringfügig ab, was zu einer erheblichen Datenreduktion führt. Weichen die vorausgesagten Bilder allerdings erheblich von der Vorhersage ab, ist eine datenintensivere Korrektur erforderlich.

MPEG arbeitet mit analogem Video im *YUV-Farbmodell*. Nach der Digitalisierung und dem *Farb-Subsampling* werden die Bilder in 8x8 Pixel große Blöcke aufgeteilt. Mittels *DCT-Transformation* werden die Helligkeits- oder Farbwerte in Frequenzparameter umgesetzt und in einem 8x8-Block gespeichert. Dabei wird in dem linken oberen Block der Mittelwert der Helligkeit gespeichert, in den anderen das Frequenzspektrum des Bildinhaltes, wobei die Detailauflösung der einzelnen Datenblöcke ständig ansteigt, bis sie in dem rechten unteren Datenblock ihr

Maximum erreicht. Eine nachgeschaltete Quantisierung der Daten dient der weiteren Datenreduktion.

Da aufeinander folgende Videobilder sich sehr ähnlich sind, teilt man sie in mehrere Makroblöcke auf, den so genannten *Group of Pictures* (GOP). Innerhalb dieser Bildgruppen werden die Bilder nacheinander komprimiert, ausgehend von dem datenintensiven Anfangsbild der Gruppe, gefolgt von hoch komprimierten Bilder und Bildern, die auf einer Vorhersage basieren.

MPEG-1

Der 1992 verabschiedete *MPEG-Standard* MPEG-1 ist ein integrierter Standard zur Speicherung von Video und Audio auf einer Compact Disc (CD) mit einer Datenrate von 1,5 Mbit/s. MPEG-1 arbeitet mit einer Auflösung von 352 x 288 Bildpunkten und einer Bildwiederholfrequenz von 25 Hz, was dem europäischen CIF-Darstellformat und dem SIF-Darstellformat entspricht. Die Auflösung ergibt sich aus dem Qualitätsanspruch, der bei MPEG-1 dem VHS-System entsprechen soll. Da Fernsehen mit zwei Halbbildern arbeitet, die zusammen ein Vollbild bilden, geht man bei der vertikalen Auflösung von der eines Halbbildes aus. Beim PAL-Fernsehstandard sind von den 625 Zeilen 575 sichtbar und damit pro Halbbild 288. Bei der Horizontalen ergeben sich 352 Pixel.

Die *MPEG-Kompression* für Video verwendet eine komplexe Codierung, bei der zeitliche und inhaltliche Bildredundanzen für die *Datenkompression* ausgenutzt werden. Mit MPEG-1 können *Videosignale* so komprimiert werden, dass sie bei einer Auflösung von 352 x 288 Bildpunkten und einer Bildwiederholfrequenz von 25 Bildern pro Sekunde, mit Datenraten zwischen 1,2 Mbit/s und 3 Mbit/s übertragen werden. Dies entspricht einer Kompression von etwa 20:1.

Neben der *Videokompression* bietet MPEG-1-Audio zwei Audiokanäle für die je nach MPEG-1-Variante Bitströme von 96 kbit/s, 128 kbit/s und 192 kbit/s zur Verfügung stehen.

Die komprimierten Video- und Audiodatenströme von MPEG-1 werden zur Übertragung gemultiplext und mit Synchronisationssignalen versehen.

MPEG-2
MPEG-2

Der 1994 verabschiedete *MPEG-Standard* MPEG-2 eignet sich für die Übertragung von Bewegtbild in Standard-Fernsehqualität und Studioqualität mit mehreren Tonkanälen. Diese Fernsehstandards sind im ITU-Dateiformat *BT.601* spezifiziert. MPEG-2 baut auf *MPEG-1* auf und berücksichtigt das Zeilensprungverfahren. Von der Auflösung her reicht der Videoteil von MPEG-2 von niedriger Fernsehqualität über die Standard-Fernsehqualität bis hin zu HDTV. MPEG-2 verfügt über mehrere Profile, von denen einige skalierbar sind; das bedeutet, dass aus vorhandenen Video-Datenströmen Videosequenzen mit unterschiedlicher Bildqualität erzeugt werden können. Die verschiedenen Profile repräsentieren Videodarstellungen mit Auflösungen zwischen 352 x 288 (Low), 720 x 576 (Main) und HDTV mit 1.440 x 1.152 bzw. 1.920 x 1.152 (High) Pixeln bei einem Bildseitenverhältnis von 16:9. Um Studio-Aufnahmen in hoher Qualität komprimieren zu können, verfügt MPEG-2 mit dem 422-Profile über ein *Farb-Subsampling* von 4:2:2. Das Main-Profile und der Main-Level gehören zu den verbreiteten Profilen, die vor allem in der DVDs und in Digital-TV eingesetzt werden. Das High-Profile zielt in Kombination mit dem High-Level auf HDTV ab. Die resultierenden Datenraten liegen zwischen 4 Mbit/s und 166 Mbit/s und werden um Faktoren zwischen 10:1 und 40:1 komprimiert, so dass im Falle des 166-Mbit/s-Stromes ein MPEG-2-Strom von 4 Mbits entsteht.

MPEG-2 unterstützt korrespondierende Audiosequenzen, die bei der Übertragung

Level	Auflösung Bildfrequenz	Pixel/s	max. Bitrate Mbit/s	Vergleichbar mit
Low	352 x 240 x 30	3,05	4	CIF-Dateiformat
Main	720 x 480 x 30	10,4	15	CCIR-Dateiformat
High 1440	1440 x 1152 x 30	47,99	60	Consumer HDTV
High	1920 x 1080 x 30	62,7	80	SMPTE

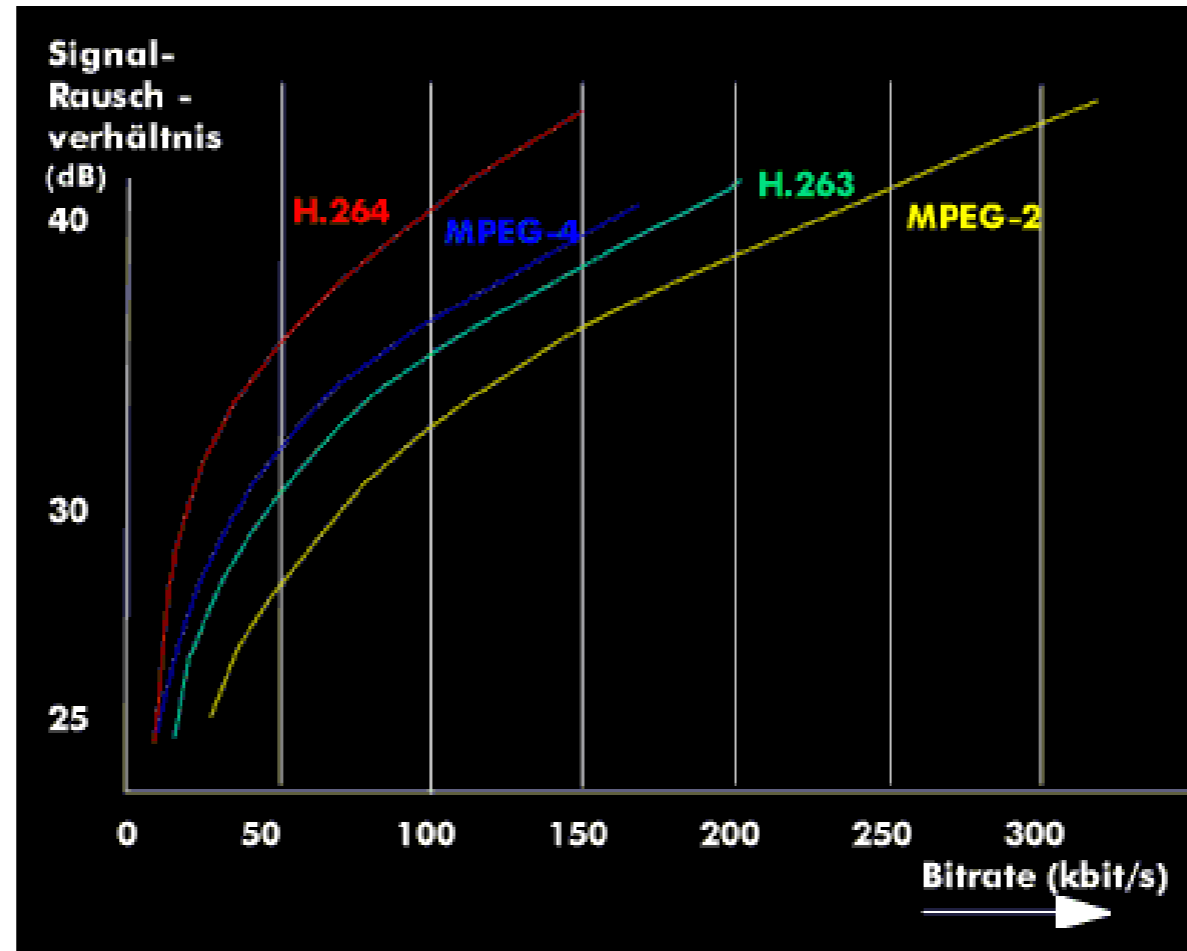
mit den Videosequenzen synchronisiert werden. MPEG-2-Audio ist kompatibel zu MPEG-1-Audio und

MPEG-2-Profile

bietet zusätzlich noch *AAC-Kompression*.
 MPEG-2-Audio ist wie MPEG-1-Audio auf die *verlustbehaftete Kompression* von qualitativ hochwertigen Audiosignalen optimiert. Neben den auch in MPEG-1-Audio verwendeten Abtastraten von 32 kHz, 44,1 kHz und 48 kHz sind bei MPEG-2-Audio auch Abtastraten von 8 kHz, 11 kHz, 16 kHz, 22 kHz, 24 kHz und 96 kHz möglich, die in der Sprachkompression in ISDN, bei G.721 oder in der Online-Distribution, der Kino- und Audio-Studiotechnik eingesetzt werden.
 MPEG-2-Audio unterstützt HDTV, Digital Theatre Sound (DTS) und Dolby-Digital mit bis zu 7 Tonkanälen und Tieftonkanal (7.1) mit Tiefenanhebung (LFE). Die Kanalcodierung ist sehr flexibel, wobei bis zu 48 Tonkanäle zu einem Raumklang kombiniert werden können. Zusätzlich zu den Möglichkeiten von MPEG-1-Audio - Mono (1.0), Stereo (2.0) und Joint-Stereo - bietet MPEG-2 Intensity Stereo Coding (ISC), Phantom Coding of Center (PCC), Dynamic Transmission Channel Switching, Dynamic Cross Talk und Adaptive Multichannel Prediction.
 MPEG-2 ist auch für Client-Server-Anwendungen ausgelegt und kann über heterogene Netze verteilt werden. MPEG-2 ist die Basistechnologie für DVDs; es wird aber auch in Digital Video Broadcasting (DVB) eingesetzt und in Settop-Boxen. Durch die Erweiterung der MPEG-2-Spezifikationen hin zu HDTV, konnte MPEG-3, das ursprünglich für hochauflösendes Fernsehen (HDTV) konzipiert wurde, zugunsten MPEG-2 aufgegeben werden.

MPEG-4 Bei *MPEG-4* handelt es sich um einen von MPEG im Jahre 1998 entwickelten und von ISO/IEC verabschiedeten Standard. Die rückwärtskompatible Version 2 von *MPEG-4* wurde im Jahre 2000 verabschiedet. MPEG-4 befasst sich mit Video, interaktiven audiovisuellen Anwendungen und interaktivem Multimedia.
MPEG-4 legt im Gegensatz zu *MPEG-1* und *MPEG-2* kein einheitliches *Kompressionsverfahren* fest, sondern definiert eine Sammlung von Audio- und Video-Kompressionsverfahren. Es bildet ein Rahmenwerk, das die Integration beliebiger Medienobjekte unterstützt, ebenso wie die Interaktion zwischen menschlichen Benutzern und der audiovisuellen Welt. Darüber hinaus definiert

Die verschiedenen Video-Codecs im Vergleich



MPEG-4 eine Sprachcodierung mit niedrigen Datenraten, audiovisuelle Objekte, die individuell komprimiert und zu Audiosequenzen zusammengestellt werden können, Mehrkanalton, Animation von Objekten der virtuellen Realität, das Einfügen von Einzelbildern und einiges mehr. Mit MPEG-4 kann eine Szene in audiovisuelle Objekte aufgeteilt werden. Die Objekte können

separat mit einer speziellen Beschreibungssprache beschrieben und beim Benutzer durch Interaktion verändert und neu zusammengefügt werden. Die Komposition von multimedialen Präsentationen kann somit beim Empfänger erfolgen. Zu diesem Zweck ist in MPEG-4 ein binäres Format für die Beschreibung von Szenen definiert, das Binary Format for Scene Description (BIFS), das auf einer Erweiterung der Virtual Reality Modelling Language (VRML) basiert. MPEG-4 hat viele unterschiedliche Profile wie die Visual Profiles, Audio Profiles, Graphics Profiles, Scene Graph Profiles, MPEG-Profiles und die objektbeschreibenden Profile. Mit diesen Profiles werden grafische Elemente in Form von einfachen Linien oder komplexen Polygonen unterstützt, Körper und Flächen, 2D- und 3D-Modelle, reale, synthetische und Video-Objekte und verschiedene Audiokompressionen wie CELP, HVXC, HILN, TwinVQ und die AAC-Kompression. Das generische Interface zwischen der Anwendung und den Transportnetzen bildet

bei MPEG-4 das Delivery Multimedia Integration Framework (DMIF) mit die DMIF-Schnittstellen zum Netzwerk (*DNI*) und zur Applikation (*DAI*).

Die Datenraten von MPEG-4 liegen bei Video zwischen 5 kbit/s und 4 Mbit/s, also wie bei MPEG-1 und MPEG-2. Darüber hinaus unterstützt MPEG-4 diverse *Videodateiformate* mit und ohne Zeilensprungverfahren vom CIF-Darstellformat bis hin zu HDTV. Bei der Audiokompression von MPEG-4 werden mehrere effiziente Codierverfahren eingesetzt, die speziell für Sprache, Musik oder synthetische Klänge optimiert sind. Dazu gehören die AAC-Kompression, CELP und HVXC für Sprachkompression und HILN für Musik.

Nero Digital

Nero Digital ist ein *Video-Codec* für Audio- und *Videokompressionen* von hoher Qualität, der standardkonform mit *MPEG-4* arbeitet. Bei der Audiokompression von Nero handelt es sich um die effiziente *AAC-Kompression* mit Mehrkanalton mit 5.1 Kanälen. Nero hat sehr hohe *Kompressionsraten* und -faktoren und kann außerdem andere MPEG-4-Implementierungen wie *XviD* oder *3ivx* decodieren. Die MPEG-4-Videos von Nero Digital werden in MP4-Containern gespeichert, die mit Kapitelinformationen und Untertiteln erstellt werden können.

Nero Digital steht im Wettbewerb mit *XviD* und *DivX* und wird je nach Tester und Version mit vergleichbarer oder besserer Qualität bewertet. Wie die anderen Video-Codexs zielt auch Nero Digital auf die Unterhaltungselektronik mit Camcordern, DVD-Technik, Smartphones usw.

<http://www.nero.com>

P-frame
Predicted-Frame
predicted frame

P-Frames sind vorhergesagte Bilder in der *MPEG-Kompression*, basierend auf der Prädiktion von vorausgegangenen *Intra-Frames* und Predicted-Frames. Predicted-Frames haben eine höhere *Kompressionsrate* als die Intra-Frames, die bei 20:1 liegen kann. Entsprechend geringer ist das komprimierte Datenvolumen eines Predicted-Bildes.

Im Predicted-Frame wird die Tatsache berücksichtigt, dass sich Bilddetails nur geringfügig verschieben. Die Verschiebung wird als Bewegungsvektor gespeichert.

PES, packetized elementary stream

Ein Formatierungsverfahren in *MPEG-2*. Nach der Digitalisierung und *Kompression* des Datenstromes in *MPEG-2*, wird dieser paketiert und gemultiplext in einen Programm- oder Transport-Stream.

SMP-Kompression
SMP, software motion picture

Software Motion Picture (SMP) ist ein von Digital Equipment (DEC) entwickeltes *Kompressionsverfahren* für Bewegtbild und Digitalvideo. Der softwarebasierte Kompressionsalgorithmus arbeitet verlustbehaftet und ist auf allen Windows-Plattformen lauffähig. Software Motion Picture wurde für einfache Netzanwendungen entwickelt, die Qualität ist mit einer Auflösung von 320 x 240 Bildpunkten, einer Farbtiefe von 8 Bit und einer Bildwiederholrate von 15 Bildern/s den Kommunikationsanwendungen angepasst. Durch die Kompression um den Faktor von etwa 9:1 werden Datenraten von ca. 1 Mbit/s für Video und 100 kbit/s für Audio erzielt.

Die SMP-Kompression arbeitet blockweise, wobei zuerst die Werte für die *Luminanz* und die *Chrominanz* mittels *Farb-Subsampling* ermittelt werden. Anschließend werden die 16-Bit-Datenblöcke für die Chrominanz durch Mittelwertbildung der Farbwerte von nebeneinander liegenden Bildpunkten auf 8 Bit halbiert. Der Luminanzanteil wird in Blöcke unterteilt, die aus 4x4-Bildpunkten bestehen. Über die Mittelung der Grauwerte wird der Datenumfang der Blöcke reduziert.

Sorenson-Codec

Der Sorenson-Codec, auch bekannt als Sorenson-Video-Codec ist ein *Video-Codec*, der von der Firma Sorenson Media entwickelt wurde und in Apple's Quicktime, Adobes Flash-Player und dem MPlayer für Unix und Linux eingesetzt wird. Die erste Version des Sorenson-Codecs wurde 1998 vorgestellt, gefolgt von Weiterentwicklungen in den Jahren 1999, 2001 und 2004. Die Spezifikationen des Codecs wurden nicht veröffentlicht. Es ist allerdings bekannt, dass der Sorenson-Codec auf H.264 basiert und FFmpeg unterstützt, wodurch die Videos auf allen Plattformen abgespielt werden können, die FFmpeg unterstützen. Die neueren Versionen sollen auf H.263 basieren.

SSIM
structural similarity index

SSIM ist ein Methode zur objektiven Qualitätsbewertung von *Video-Codexs*. Bei diesem Verfahren werden die Helligkeit und die *Chrominanz* der Vollbilder analysiert und das Helligkeits- und Farbempfinden des menschlichen Auges berücksichtigt. Da das menschliche Auge Helligkeitssignale wesentlich besser auflösen kann als Farbsignale, gehen in die SSIM-Bewertung 80 % des Luminanzsignals und jeweils 10 % der *Farbdifferenzsignale* ein.

Theora-Codec

Der Theora-Codec ist ein Video-Codec, der von der Vereinigung Xiph.Org entwickelt wurde und auf der ON2-Technologie des VP3-Codecs basiert. Nach der Zielsetzung von Theora sollte der *Video-Codec* vergleichbare Eigenschaften wie Video unter *MPEG-4* haben, wie *Xvid* oder *DivX*, und ähnlich sein wie andere *Videokompressionen*, die mit niedrigen Bitraten arbeiten; RealVideo oder Windows Media Video (WMV).

Während VP3 eine patentierte Technologie ist, die lizenziert werden muss, kann das darauf basierende Theora von Jedermann genutzt werden.

Theora ist eine verlustbehaftete Kompression, abgeleitet vom VP3-Codec von ON2. Das komprimierte Videoformat kann in einem geeigneten Container gespeichert werde, so beispielsweise in einem Ogg-Container. Es wird regelmäßig mit Vorbis Audio kombiniert. Das aus dem Ogg-Container, Theora-Video und dem Vorbis-Audio entstehende Format kann lizenzfrei als Multimedia-Format genutzt werden. Dagegen sind andere Multimedia-Formate wie MPEG-4 Video oder MP3 patentiert und lizenzierungspflichtig und werden kommerziell eingesetzt.

Wie alle anderen Bild- und Videokompressionen arbeitet Theora mit Farb-Subsampling. Es benutzt eine Kompressionstechnik die blockweise arbeitet und 8x8-Pixelblöcke über eine DCT-Transformation umwandelt, vergleichbar *MPEG* in den verschiedenen Ausführungen.

TS
transport stream

Einer von zwei Datenströmen, die von *MPEG-2*-Systemen vorgegeben werden. Der Transport Stream (TS) besteht aus 188-Byte-Paketen und kann mehrere Programme enthalten.

Die Empfehlungen für den Transport Stream sind in der ISO/ICE Empfehlung 13818-1 definiert und Grundlage von DVD.

VC-1,
video codec 1

Der *Video-Codec* VC-1 von Microsoft basiert auf dem Windows Media Video (*WMV9*). Es handelt sich dabei um einen Video-Codec mit *DCT-Transformation* wie sie auch in *H.261, H.263, H.264/AVC, MPEG-2* und *MPEG-4* benutzt wird. Video-Codex wie der VC-1 sind für Konsumerprodukte wie Settop-Boxen (STB), HDTV, HD-DVDs und Blu-Ray-Disc konzipiert und werden von diesen Technologien unterstützt. VC-1, das auch in HDTV eingesetzt wird und sich durch gute *Kompressionsraten* auszeichnet, kann gleichermaßen für HDTV-Darstellungen im Interlace mit 1.920 x 1.080 Pixel benutzt werden, was mit 1.080i bezeichnet wird, als auch für Darstellungen mit progressive Scan mit 1.280 x 720 Pixel, bezeichnet als 720p. VC-1 wurde 2006 von der Society of Motion Picture and Television Engineers (SMPTE) als offener Standard verabschiedet. Neben der Standardversion gibt es mit WMVA und WVC1 zwei weitere Versionen mit unterschiedlichen Profilen.

**Verlustbehaftete
Kompression**
lossy data compression

Die verlustbehaftete *Kompression* orientiert sich an der Wahrnehmungsfähigkeit des Betrachters oder des Zuhörers. So können aufgrund der menschlichen Gehirntätigkeit und der Auflösungsmöglichkeit von Farben, Formen und Tönen bei Grafiken, Audios oder Videos Informationen verändert oder entfernt werden, ohne dass sich die Wahrnehmung wesentlich ändert und der subjektive Eindruck verschlechtert. Bei der Dekompression einer verlustbehafteten Kompressionsdatei kann nicht mehr das exakte Originalbild regeneriert werden. Die Abweichungen hängen vom Kompressionsgrad und vom gewählten Kompressionsverfahren ab. Das bekannteste verlustbehaftete Kompressionsverfahren für Bilder, Fotos und Grafiken ist JPEG.

Video-Codec
video codec

Video-Codex sind Funktionseinheiten in denen die *Videokompression* hard- oder softwaremäßig umgesetzt wird. Sie werden überall dort eingesetzt, wo ein

Videosignal über Datennetze übertragen wird. Video-Codecs reduzieren die enorme Datenmenge der Videosignale, damit diese mit möglichst geringer Datenrate und guter Qualität übertragen und empfangen werden können. Senderseitig eingesetzt codieren sie die Videosignale durch *Kompression*, empfangsseitig decodieren sie das komprimierte Video in Encodern unter Benutzung eines Kompressionsalgorithmus.

Ziel aller Video-Codec-Entwicklungen ist die Reduzierung der Datenmenge bei gleicher oder verbesserter Videoqualität.

Viele Video-Codecs setzen auf *verlustbehaftete Kompression* von *MPEG*, vorwiegend *MPEG-2* und *MPEG-4*. Es handelt sich aber immer um entwicklungspezifische Implementierungen von Kompressionstechniken. Daher gibt viele firmenspezifische, lizenzpflichtige Video-Codecs, ebenso aber auch diverse lizenzfreie Open-Source-Produkte.

Einer unter vielen Parametern zur Bewertung von Video-Codecs ist die Effizienz der verlustbehafteten Codecs, die sich aus in effektiven Anzahl der Bits pro Pixel (bpp) ausdrückt und von der Bitrate abhängt. Soll beispielsweise ein Videosignal nach dem PAL-Fernsehstandard mit 720 x 575 Bildpunkte und 25 Vollbildern mit einer Datenrate von 1 Mbit/s übertragen werden, dann ergibt sich die effektive Bit-pro-Pixel-Rate zu 0,097. Diese Effizienz muss der Video-Codec realisieren. Das bedeutet beispielsweise, dass etwa 10 Pixel durch ein Bit dargestellt werden.

Die Qualitätsbewertung erfolgt durch mathematische Verfahren bei denen der Signal-Rauschanteil (PSNR) des Originalbildes mit dem dekomprimierten Bild in Relation gesetzt wird. Eine weitere Methode zur objektiven Bewertung ist der

Structural Similarity Index (*SSIM*).

Der Einsatzbereich von Video-Codes reicht von der Übertragung von Fernsehbildern mit hoher und höchster Auflösung,

$$\text{Bit pro Pixel} = \frac{\text{Bitrate}}{\text{Auflösung (H)} \times \text{Auflösung (V)} \times \text{Bildfrequenz}}$$

bpp, bit per pixel
 bps, bitrate
 Fps, frame per second
 Bildfrequenz

Ermittlung der Effizienz eines Video-Codecs

HDTV, über Bildtelelefon und Videokonferenzen bis hin zu Darstellungen auf dem Handy.

An Hardware-Codecs für Video sind zu nennen Motion-JPEG, Px64 nach *H.261* und *H.320*, und MPEG in den verschiedenen Varianten. An Software-Codecs sind zu nennen Cinepak, *Indeo*, *DivX*, Quicktime, Video for Windows (VfW), *XviD*, *3ivx*, *VC-1*, *Nero Digital*, *HDX4* und Windows Media Video (*WMV*).

Videodateiformat
video data format

Unter Videodateiformaten sind die Video-Darstellungsformate wie das CIF-Darstellformat, QCIF, FCIF und HDTV sowie die Dateiformate für die *Videokompression* und das Streaming zu verstehen. Hierzu gehören u.a. *MPEG*, *DivX*, Advanced Video Coding, *H.261*, *H.263*, *H.264/AVC*, QuickTime, das AVI-Dateiformat, FLV und *WMV-Dateiformat* u.a.

Darüber hinaus gibt es digitale, herstellerspezifische Videoformate, die in der Studioteknik eingesetzt werden. Die meisten digitalen Videoformate haben eine Auflösung von 720 x 576 Bildpunkten, andere entsprechen dem CCIR-Dateiformat und wieder andere sind für HDTV mit 1.920 x 1.080 und 1.280 x 720 Bildpunkten. Als *Kompression* verwenden die digitalen Videoformate *MPEG-2* und die *DCT-Transformation* mit *Intra-Frame-Codierung*, *Farb-Subsampling* und ein 16-Bit-Audiosampling. Bekannte Dateiformate sind MicroMV DVC-AM und IMX von Sony, DVC Pro25, DVC Pro50 und DVC Pro100 von Panasonic, High Definition Video (HDV) von Sony und JVC und AVCHD von Sony und Panasonic.

Videokompression
(video compression)

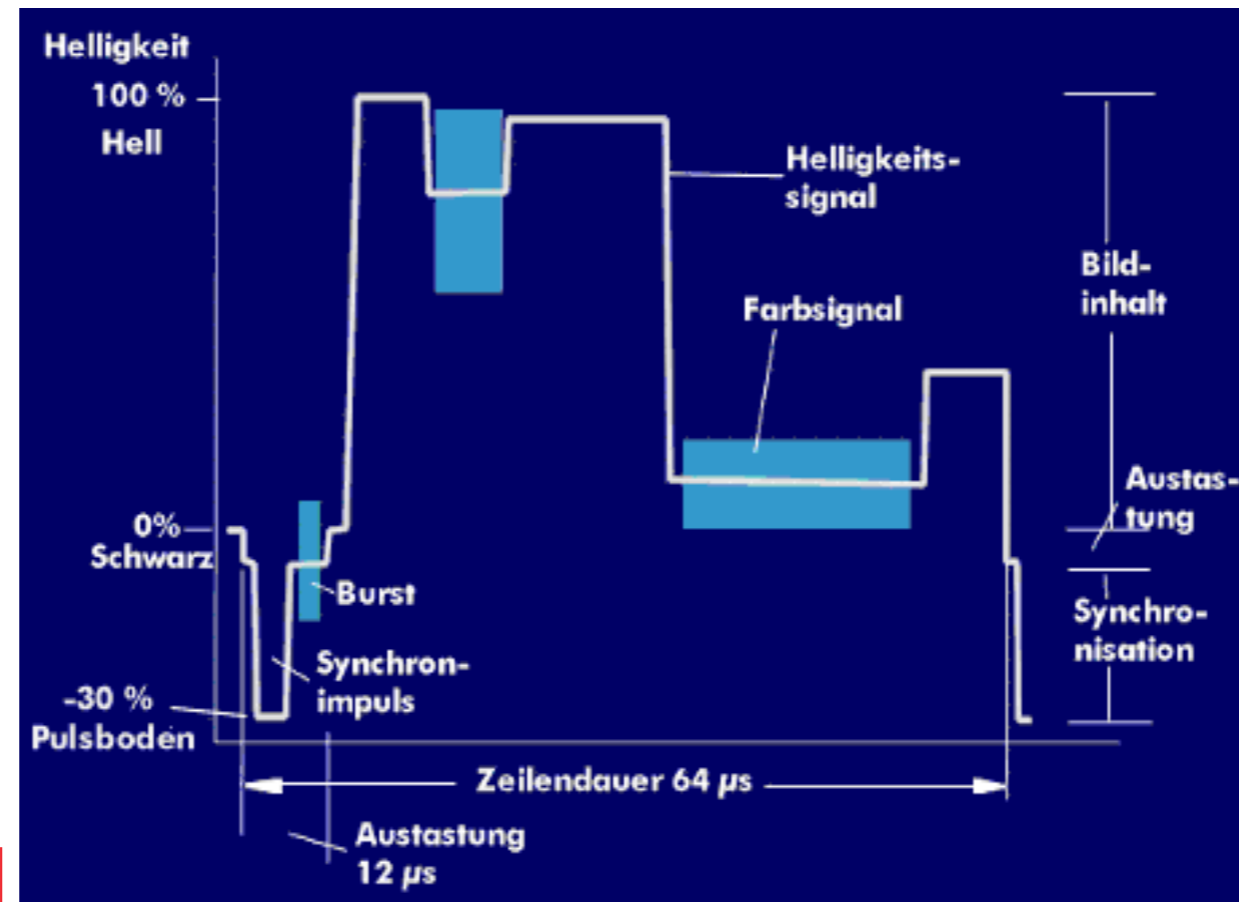
Die Videokompression ist eine *verlustbehaftete Kompression*, die sich die Trägheit und die unterschiedliche Augenempfindlichkeit für die Helligkeits- und Farbauflösung zunutze macht. Da aufeinanderfolgende Videobilder nur geringfügig unterschiedlich sind, kann man die Trägheit des Auges dadurch ausnutzen, dass man nur minimale Differenzen zwischen den aufeinander folgenden Bildern speichert.

Hinsichtlich der Sensibilität zwischen Helligkeit und Farbe nutzt man in der Videokompression das Farbmodell in dem diese Parameter einzeln vorhanden sind,

so wie beim *YUV-Farbmodell*. Da die Farbauflösung wesentlich geringer ist als die Auflösung der Helligkeitswerte, können die Farbdaten durch *Farb-Subsampling* bereits reduziert werden. Eine weitere Datenreduzierung erfährt die Farbinformation beispielsweise durch eine *DCT-Transformation*, durch erneute Quantisierung und durch die Bildergruppen, entsprechend den *Group of Pictures (GOP)* in der *MPEG-Kompression*.

Zwei internationale Standardisierungsinstitute, das ITU-T und ISO/IEC, entwickeln Standards für die Videokompresion. Bei der ITU-T heißen sie H-Empfehlungen und haben die Bezeichnungen *H.261*, *H.262*, *H.263* und *H.264/AVC*. Die ISO-IEC-Standards führen die Bezeichnungen *MPEG-x*: *MPEG-1*, *MPEG-2* und *MPEG-4*. Die meisten ITU-Empfehlungen sind für Echtzeitanwendungen wie Videokonferenzen und -telefonie, die ISO/IEC-Standards zielen auf die Videospeicherung, auf Broadcastvideo wie DVB-T, DVB-S und DVB-C und Video-Streaming über das Internet.

Videosignal
(video signal)



Videosignale finden ihren Ursprung im analogen Fernsehen. Der Anteil am Fernsehsignal, der die Bildinformationen enthält, heißt Video. Ausgehend von dieser klassischen analogen Videotechnik hat sich die Videotechnik mit der Digitalisierung andere Formate angenommen. Die klassische Videotechnik arbeitet mit

FBAS-Zeilensignal

einer Bildschirmdarstellung, die in der Horizontalen von links nach rechts und in der Vertikalen von oben nach unten erfolgt. Die horizontale Abtastung nennt man Zeile, die vertikale Abtastung Bild. Aus technischen Gründen und unter Ausnutzung der Trägheit des Auges unterscheidet man beim klassischen Video zwischen zwei Bildern, die als Halbbilder bezeichnet werden und zusammen ein Bild oder Vollbild ergeben. Die zwei Halbbilder unterscheiden sich dadurch, dass in einem Halbbild die geraden Zeilen (2, 4, 6, ..) und in dem folgenden Halbbild die ungeraden Zeilen (1, 3, 5, ...) dargestellt werden.

Legt man beide Halbbilder übereinander ist die Zeilenstruktur eines Vollbildes in der Horizontalen und Vertikalen. Diese trugen den damaligen technischen Möglichkeiten Rechnung, weil der Elektronenstrahl am Ende einer Zeile dunkelgesteuert wird und an den nächsten Zeilenanfang gestellt werden muss. Gleiches gilt für die Rückführung des Elektronenstrahls von unten nach oben, was während der

vertikalen Austastlücke erfolgt. Des Weiteren sind die Modulationstechniken für die Bildinformation Relikte eines technischen Entwicklungsstandes von vor über 50 Jahren. Dabei ist zu berücksichtigen, dass alle Video-Entwicklungen immer die Kompatibilität zwischen vorhandener und zukünftiger Technik berücksichtigen mussten. Das betraf im Besonderen die Einführung des Farbfernsehens. Neben dem Schwarz-Weiß-Signal,

Standard	CCIR 601 NTSC	CCIR 601 PAL/SECAM	CIF	QCIF
Luminanz- auflösung (Y)	720 x 480	768 x 576	352 x 288	176 x 144
Chrominanz- Auflösung (C)	360 x 480	384 x 576	176 x 144	88 x 72
Chroma- Subsampling	4 : 2 : 2	4 : 2 : 2	4 : 2 : 0	4 : 2 : 0
Halbbilder	60	50	30	30
Zeilensprung	ja	ja	nein	nein
Datenrate (Mbit/s)	167,6	165,9	36,5	9,1

*CCIR-Standards für
Digitalvideo*

das im Informationsgehalt nicht verändert werden konnte, musste zusätzlich ein Farbsignal übertragen werden, ohne dass Schwarz-Weiß-Fernseher durch das Farbsignal beeinträchtigt werden durfte. Deshalb gibt es das Videoformat mit zwei in der Bandbreite reduzierten *Farbdifferenzsignalen*, die die geringere Farbauflösung des menschlichen Auges berücksichtigen und sich im *YUV-Farbmodell* widerspiegeln. Bis auf das YUV-Signal gelten die Erläuterungen für alle Farbfernsehstandards PAL, NTSC und SECAM.

Das Videosignal besteht aus dem Schwarz-Weiß-Signal, das als *Luminanzsignal* bezeichnet wird, den Austast- und den Synchronimpulsen. Dieses Signal nennt man BAS-Signal, wobei "B" für Bild, "A" für Austastung und "S" für Synchronisation steht. Das entsprechende Farbsignal FBAS mit "F" für Farbe. Dieses Signal wird auch als Composite Video bezeichnet.

Neben dem Composite Video kann ein Videosignal auch aus den drei Farbkomponenten Rot (R), Grün (G) und Blau (B) bestehen. Man spricht dann von Komponentenvideo.

Ebenso ist es möglich das Helligkeitssignal (Y) und das komplette Farbsignal (Chrominanzsignal, C) getrennt aufzuzeichnen. Man spricht dann von separiertem Video oder von Y/C-Video.

WMF *Windows media format*

Windows Media Format (WMF) ist ein Microsoft-Konzept für multimediale Objekte und deren digitales Rechte management (DRM). Das WMF-Konzept besteht aus Windows Media Audio (WMA) für die Audiokompression, Windows Media Video (WMV) für die *Videokompression* und dem Windows Media Rights Manager (WMRM). Mit dem WMRM-Manager werden die WMF-Dateien mit einem Wasserzeichen für das digitale Rechte management versehen, wodurch die Wiedergabe auf anderen Personal Computern (PC) ausgeschlossen wird. Sie können allerdings auf Playern, die der Secure Digital Music Initiative (SDMI) entsprechen, abgespielt werden.

Zwischenzeitlich gibt es mehrere WMF-Versionen (WM7, WM8, WM9), bei denen die Wiedergabequalität und der Kopierschutz ständig verbessert wurden.

WMV
Windows media video

Windows Media Video (WMV) ist ein *Video-Codec* für die *Kompression* und Dekompression von *Videosignalen*. WMV hat hohe Kompressionsfaktoren und bietet CIF-Qualität bei Übertragungsraten von 250 kbit/s. Es ist ein Streaming-Format für die Wiedergabe von Audio und Video. Bei WMV handelt es sich um eine Serie an Audio- und Video-Codecs, die für spezielle Anwendungen optimiert wurden. Das WMV-Format unterstützt den Kopierschutz mittels digitalem Rechtemanagement (DRM).

Es gibt mehrere Versionen von Windows Media Video, deren Versionsbezeichnungen aber nicht eindeutig genutzt werden: WMV1, WMV3, WMV7 und WMV9, wobei die Versionen 1 und 7, sowie 3 und 9 identisch sind.

Die Version Windows Media Video 9 (WMV9) ist ein fortschrittlicher Video-Codec, der für das Streaming über IP-Netze optimale Vorteile bietet. WMV9 hat VC-1 implementiert und unterstützt die Interpolation von Halb- und Viertel-Pixeln, was sich unmittelbar auf die Bildqualität auswirkt, er reduziert *Artefakte*, die bei der

Profile	Level	Bitraten	Bildauflösung
Simple	Low	96 kbit/s	176 x 144/15 (QCIF)
	Medium	384 kbit/s	240 x 176/30
Main	Low	2 Mbit/s	352 x 288/15 (CIF)
	Medium	10 Mbit/s	320 x 240/24 (QVGA) 720 x 480/30 (480p) 720 x 576/25 (576p)
	High	20 Mbit/s	1.920 x 1.080/30 (1080p)
Advanced	L0	2 Mbit/s	352 x 288/30 (CIF)
	L1	10 Mbit/s	720 x 480/30 (NTSC-SD) 720 x 576/25 (PAL-SD)
	L2	20 Mbit/s	720 x 480/60 (480p) 1.280 x 720/30 (720p)
	L3	45 Mbit/s	1.920 x 1.080/24 (1080p) 1.920 x 1.080/30 (1080i) 1.280 x 720/60 (720p)
	L4	135 Mbit/s	1.920 x 1.080/60 (1080p) 2.048 x 1.536/24

Quantisierung entstehen, er kann sowohl mit *DCT*- als auch mit *IDCT*-Transformation arbeiten und unterstützt Blockgrößen von 8x8, 8x4, 4x8 und 4x4 Pixeln. Gegenüber *MPEG-2* bietet WMV9 eine höhere Codierungseffizienz und gegenüber *H.264/AVC* einen geringeren Rechenaufwand.

Profile und Darstellformate unter WMV9

WMV9 wurde auch für die HD-DVD und die Blu-Ray-Disc übernommen.
WMV ist das Hauptformat für Streaming-Media beim Windows Media Player. Der Media Player und der Codec können kostenlos bei Microsoft herunter geladen werden.

Die Extension von WMV-Dateien ist *.wmv.

XviD XviD ist ein *Video-Codec*, der als Open Source zur Verfügung steht und auf dem Kompressionsalgorithmus "MPEG-4 Part 2 Video Coding" basiert. XviD benutzt wie *DivX*, in dem es zu Konkurrenz steht, zur Codierung der Inhalte das Advanced Simple Profile (ASP) von MPEG-4.