2.2 Kompressionsverfahren für Standbilder

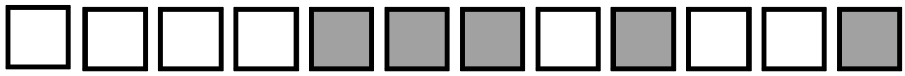
2.2.1 Fax-Kompression

Standards in der Telekommunikation werden von der International Telecommunications Union (ITU-T) entwikkelt (früher: CCITT = Commitée Consultatif International de Téléphonie et Télégraphie).

Der Standard für verlustlose Kompression von Faksimile-Anwendungen war einer der ersten Standards zur Bildkompression.

Bilder werden von Faxalgorithmen der Gruppe 3 als **bitonal** betrachtet, d.h. jedes Pixel kann entweder schwarz oder weiß sein. So kann jedes Pixel mit einem Bit repräsentiert werden. In jedem bitonalen Bild gibt es in der Regel größere Bitfolgen (Runs), die entweder vollkommen schwarz oder vollkommen weiß sind.

Beispiel:



Lauflängencodierung: 4w 3s 1w 1s 2w 1s

Fax-Standards der ITU (vormals CCITT)

Standard T.4 erstmals verabschiedet 1980, überarbeitet 1984 und 1988 (Fax Gruppe 3) für *fehleranfällige* Leitungen, insbesondere Telefonleitungen.

A4-Dokumente in schwarz/ weiß.

100	dots per inch (DPPI) oder 3,85 Zeilen/mm vertikal
1728	Pixels pro Zeile

Ziel:
Übertragung einer Seite in einer Minute bei 4800 bit/s

Standard T.6 erstmals verabschiedet 1984 (Fax Gruppe 4) für *fehlerfreie* Leitungen oder digitale Speicherung.

Standards für die Fax-Kompression

Fax Gruppe 3, ITU-T Recommendation T.4:

Kodierungsansatz 1: Modified Huffman Code (MH)

- Bild wird als Folge von Pixel-Zeilen betrachtet.
- Für jede Zeile wird eine Lauflängencodierung ermittelt.
- Die (Lauflänge,Wert)-Beschreibung wird dann Huffman-kodiert, mit standardisierter, festgeschriebener Tabelle
- Dabei werden schwarze und weiße Runs mit separaten Huffman-Codes behandelt, da die Auftrittswahrscheinlichkeiten der Lauflängen recht unterschiedlich sind.
- Zur Fehlererkennung wird nach jeder Zeile ein EOL (end-of-line=)-Codewort eingefügt. Dies ermöglicht das Wiederaufsetzen nach Bitübertragungsfehlern.

Fax-Kompression

Kodierungsansatz 2: Modified Read (MR) Code

- Pixelwerte der vorangehenden Zeile werden für die Vorhersage der aktuellen Zeile verwendet.
- Anschließend wird eine Lauflängencodierung und ein statischer Huffman-Code wie bei MH eingesetzt.
- Auch hier wird ein EOL-Codewort verwendet.

Die Kodierungsansätze MH und MR werden regelmäßig abgewechselt, um die Fortpflanzung von Fehlern zu vermeiden.

Huffman-Tabelle für Fax Gruppe 3 (Ausschnitt)

White run length	Code word	Black run length	Code word
0	00110101	0	0000110111
1	000111	1	010
2	0111	2	11
3	1000	3	10
4	1011	4	011
5	1100	5	0011
6	1110	6	0010
7	1111	7	00011
8	10011	8	000101
9	10100	9	000100
10	00111	10	0000100
11	01000	11	0000101
12	001000	12	0000111
13	000011	13	00000100
14	110100	14	00000111
15	110101	15	000011000
16	101010	16	0000010111
17	101011	17	0000011000
18	0100111	18	0000001000
19	0001100	19	00001100111
20	0001000	20	00001101000

Fax Gruppe 4

ITU-T Rec. T.6

Kodierungstechnik: Modified Modified Read Code (MMR)

- Vereinfachung des MR-Codes; es werden keine Fehlererkennungsmechanismen mehr eingesetzt, um die Kompressionsrate zu erhöhen.

Kompressionsraten:

	Geschäftsdokumente
Gruppe 3:	20:1
Gruppe 4:	50:1

Für Fotos u.ä. ist die Kompressionsrate schlecht, weil dort nur kurze Lauflängen vorkommen. Eine adaptive arithmetische Codierung wäre beispielsweise besser geeignet.

2.2.2 Block Truncation Coding (BTC)

Ein einfaches Verfahren für *Graustufen*-Bilder. Man nimmt an, daß jedes Pixel im Original mit einem Grauwert von 0 (schwarz) bis 255 (weiß) beschrieben ist.

Algorithmus BTC

1. Zerlege das Bild in Blöcke der Größe n x m Pixel
2. Berechne Mittelwert und Standardabweichung der Pixelwerte für jeden Block

$$\mu = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} Y_{i,j}$$

$$\sigma = \sqrt{\frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} Y^2_{i,j} - \mu^2}$$

mit $Y_{i,j}$ = Helligkeit (Graustufe) des Pixels.

3. Weise dem Block eine Bitmatrix der Größe n x m nach folgender Regel zu:

$$B_{i,j} = \begin{cases} 1 \ldots falls \;\; Y_{i,j} \leq \mu \\ \\ 0 \ldots sonst \end{cases}$$

4. Berechne zwei Graustufenwerte für den Block, *a* für den Wert der dunkleren Pixel, *b* für den Wert der helleren Pixel:

$$a = \mu - \sigma\sqrt{p/q}$$

$$b = \mu + \sigma\sqrt{q/p}$$

p ist dabei die Anzahl der Pixel, die heller als der Mittelwert μ sind, *q* die Anzahl der dunkleren Pixel.

5. Ausgabe: (Bitmatrix, *a*, *b*) für jeden Block

Dekompression bei BTC

Für jeden Block werden die Graustufen der einzelnen Pixel wie folgt berechnet:

$$Y'_{i,j} = \begin{cases} a \ldots falls \;\; B_{i,j} = 1 \\ b \ldots sonst \end{cases}$$

Beispiel für die Kompressionsrate

Blockgröße:	4x4
Original (Grauwerte)	1 Byte pro Pixel
codierte Darstellung:	Bitmatrix mit 16 Bits + 2x8 Bits für *a* und *b*

⇨Reduktion von 16 Bytes auf 4 Bytes

2.2.3 Color Cell Compression

Ein Verfahren zur Kompression von Farbbildern. Im Prinzip könnte man BTC für jeden der drei Farbkanäle RGB getrennt anwenden. Aber CCC liefert qualitativ bessere Ergebnisse.

Algorithmus CCC

1. Zerlege das Bild in Blöcke der Größe m x n Pixels.

2. Berechne für jedes Farbpixel die Helligkeit gemäß:

$$Y = 0.3P_{red} + 0.59P_{green} + 0.11P_{blue}$$

Y=0 entspricht weiß, Y=1 entspricht schwarz

3. Für c = red, green, blue berechne die mittleren Farbwerte der Pixel gemäß

$$a_c = \frac{1}{q} \sum_{Y_{i,j} \le \mu} P_{c,i,j} \,, \qquad b_c = \frac{1}{p} \sum_{Y_{i,j} \succ \mu} P_{c,i,j}$$

p ist die Anzahl der Pixel, die heller als der Mittelwert sind, q die Anzahl der dunkleren Pixel.

4. Weise dem Block eine Bitmatrix der Größe n x m nach folgender Regel zu:

$$B_{i,j} = \begin{cases} 1 \ldots falls \;\; Y_{i,\,j} \leq \mu \\ \\ 0 \ldots sonst \end{cases}$$

5. Die Werte $a = (a_{red}, a_{green}, a_{blue})$ und $b = (b_{red}, b_{green}, b_{blue})$ werden in eine Farbtabelle abgebildet, gemäß einem Nähe-Maß. Es ergeben sich die Werte a' und b' als Indizes der Farbtabelle (Color Looking Table, CLUT).

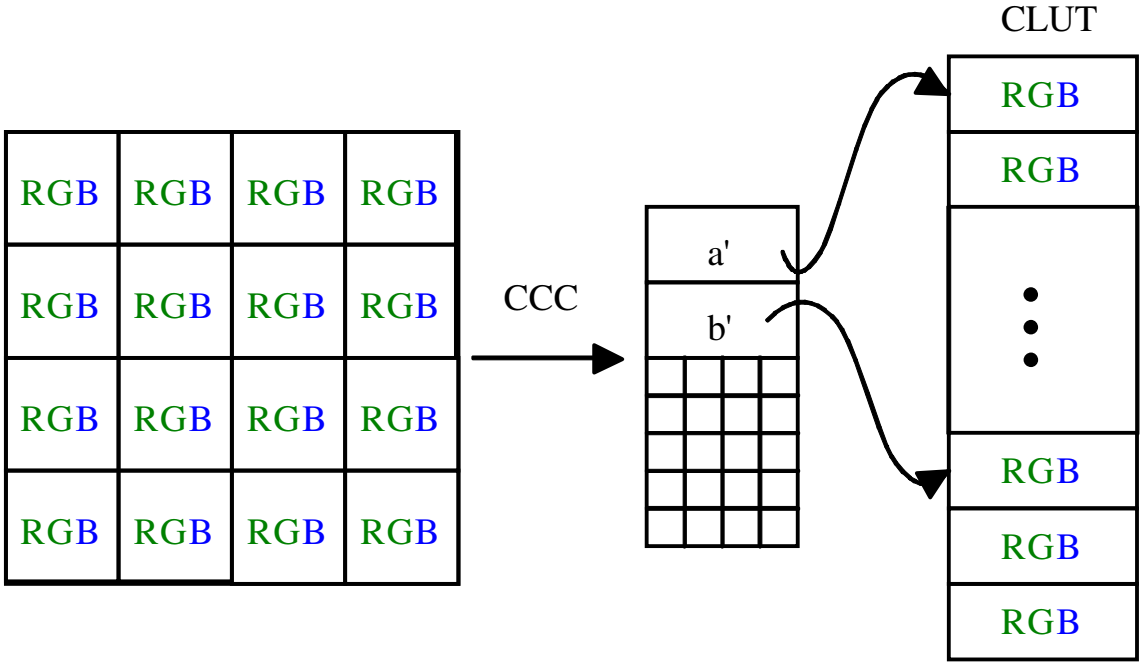
6. Ausgabe: (Bitmatrix, a', b') für jeden Block

Dekompression bei CCC

Die Dekompression funktioniert für jeden Block wie folgt:

$$P'_{i,j} = \begin{cases} CLUT[a'] \ldots falls\, B_{i,j} = 1 \\ CLUT[b'] \ldots sonst \end{cases}$$

Funktionsweise von CCC

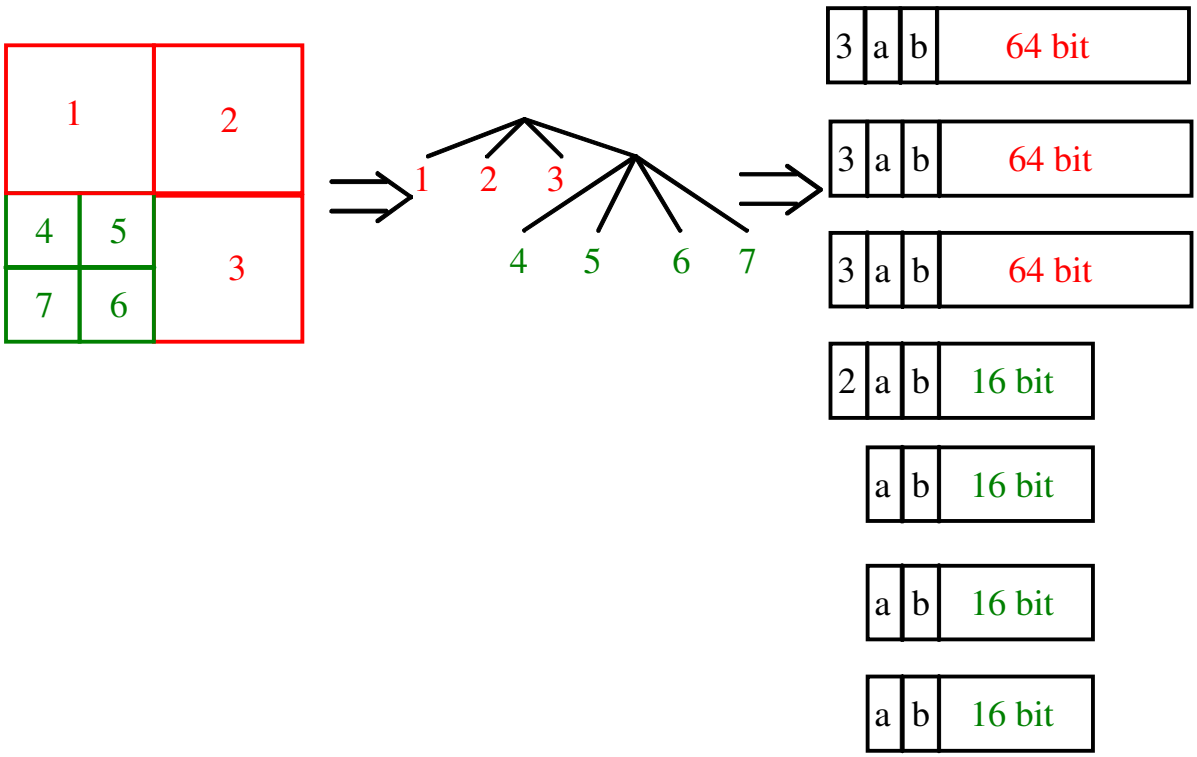


Extended Color Cell Compression (XCCC)

Eine Erweiterung von CCC zur weiteren Verbesserung der Kompressionsrate.

Idee

Hierarchie von Blockgrößen. Man versucht zunächst, einen großen Block mit CCC zu codieren. Wenn die Abweichung der tatsächlichen Farbwerte im Block von a‘ oder b‘ größer ist als ein vorgegebener Schwellenwert, wird der Block in vier Teilblöcke zerlegt. Der Algorithmus arbeitet rekursiv.



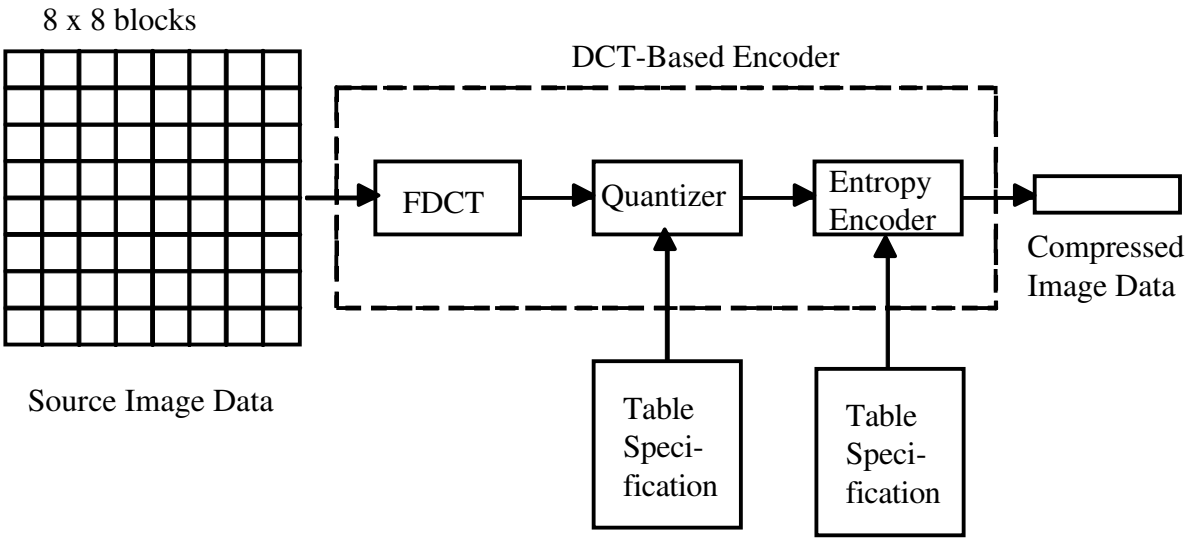
JPEG

Die Joint Photographic Experts Group (Komitee der ISO) hat einen sehr effizienten Kompressionsalgorithmus entwickelt, der nach ihr benannt ist. JPEG hat beispielsweise im WWW eine sehr weite Verbreitung gefunden.

Kompression in vier Schritten:

1. Bildaufbereitung
2. Diskrete Cosinus-Transformation (DCT)
3. Quantisierung
4. Entropie-Kodierung

Der DCT-basierte Kodierungsprozeß



JPEG "Baseline" Mode

Ein Kompressionsverfahren auf der Basis einer Transformation aus der Zeitdomäne in die Frequenzdomäne

Bildtransformation

FDCT (Forward Discrete Cosine Transformation). Sehr ähnlich der Fourier-Transformation. Wird auf jeden Block von 8x8 Pixeln einzeln angewendet.

$$S_{vu} = \frac{1}{4} C_u C_v \sum_{x=0}^{7} \sum_{y=0}^{7} s_{yx} \cos\frac{(2x+1)u\pi}{16} \cos\frac{(2y+1)v\pi}{16}$$

mit

$$C_u, C_v = \frac{1}{\sqrt{2}} \text{ für u, v = 0;}$$

$$C_u, C_v = 1 \text{ sonst}$$

64 mal anwenden, ergibt 64 Koeffizienten in der Frequenzdomäne

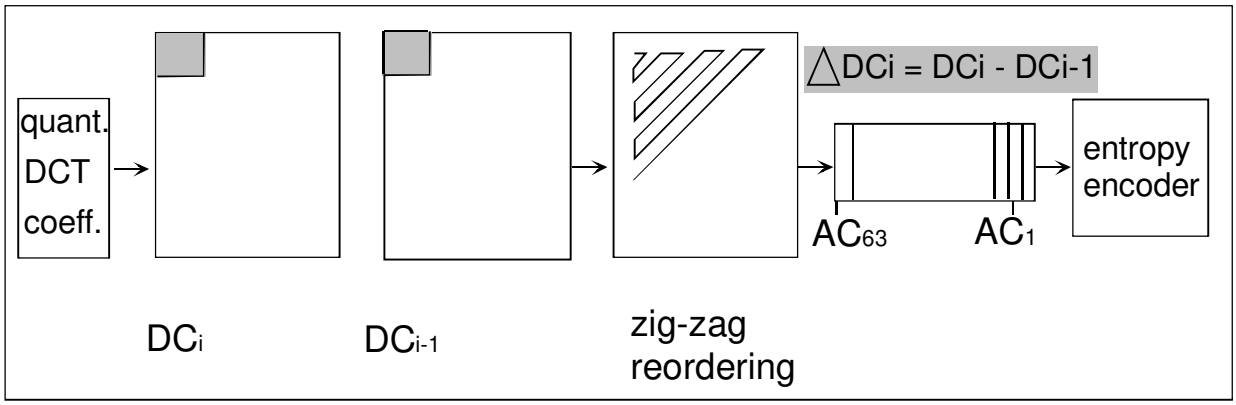
Quantisierung

Anzahl der Quantisierungsstufen je DCT-Koeffizient einzeln einstellbar (Q-Faktor). Dies geschieht durch Angabe einer Quantisierungstabelle.

Entropiekodierung

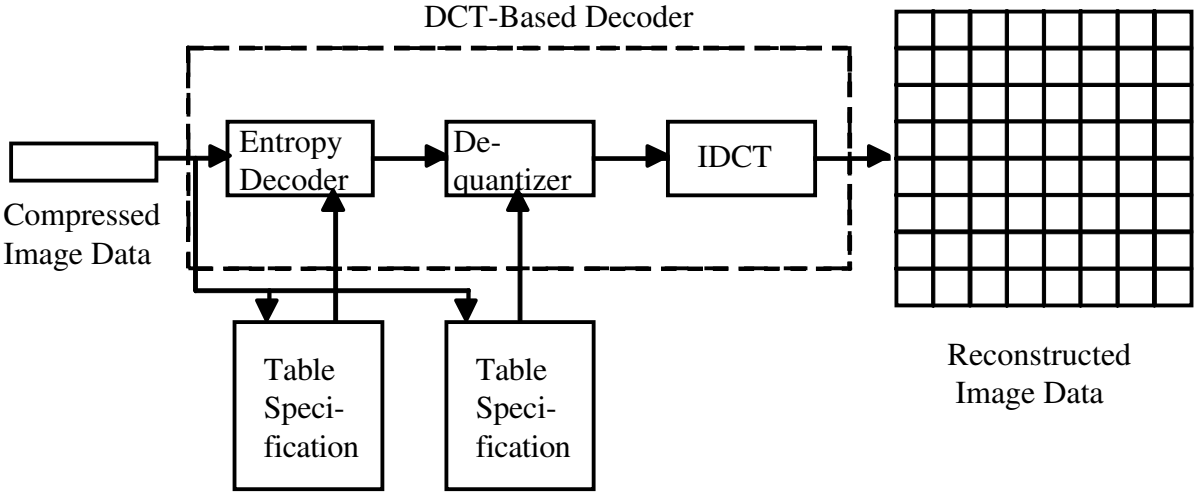
1. DC-Koeffizient sehr wichtig (Grundfarbton). Codierung durch Differenzbildung zu vorherigem DC-Koeffizienten.
2. AC-Koeffizienten, Abarbeitung in „Zick-Zack“-Reihenfolge, Lauflängenkodierung
3. Huffman-Codierung

Quantisierung und Entropie-Codierung



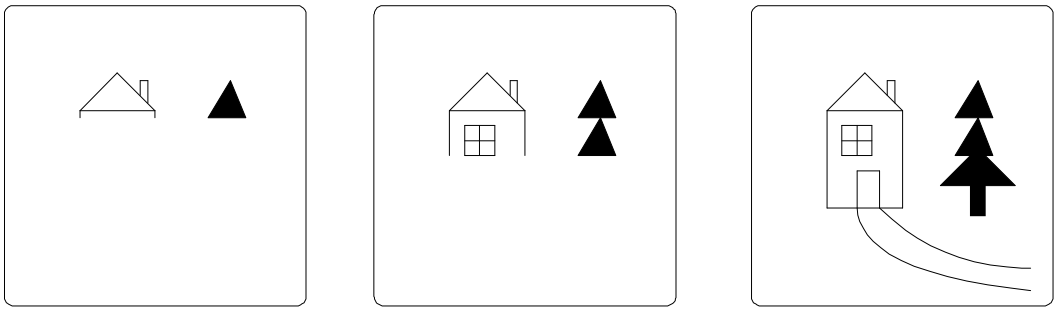
Das Zickzack-Auslesen der Koeffizienten ist besser als ein zeilenweises Auslesen. Grund: Nach der Quantisierung kommen in der rechten unteren Dreiecksmatrix viele gleiche Werte nacheinander vor, darunter viele Nullen (Amplituden mit zunehmender Frequenz entsprechen schärfer werdenden Kanten). Die Zickzack-Anordnung maximiert die Lauflängen für die anschließende Entropie-Codierung (Lauflängencodierung, gefolgt von Huffman-Codierung)

JPEG Decoder

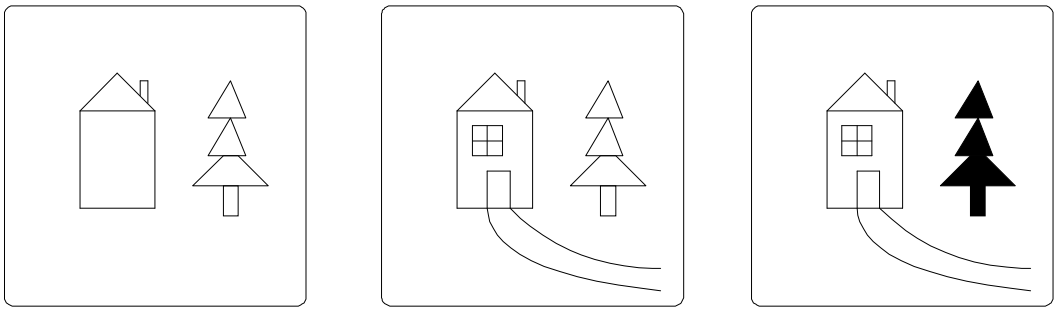


Verschiedene Modi in JPEG

JPEG Sequential Mode



JPEG Progressive Mode



Quantisierungsfaktor und Bildqualität

Bildbeispiel 1

Schloß, Original

Schloß, Q=6

Schloß, Q=12

Schloß, Q=20

Bildbeispiel 2

Blumen, Original

Blumen, Q=6

Blumen, Q=12

Blumen, Q=20

2.2.4 Bildkompression mit Fraktalen

Theorie der Fraktale = Theorie der **Selbstähnlichkeit**. Mathematisch beschreibbar.

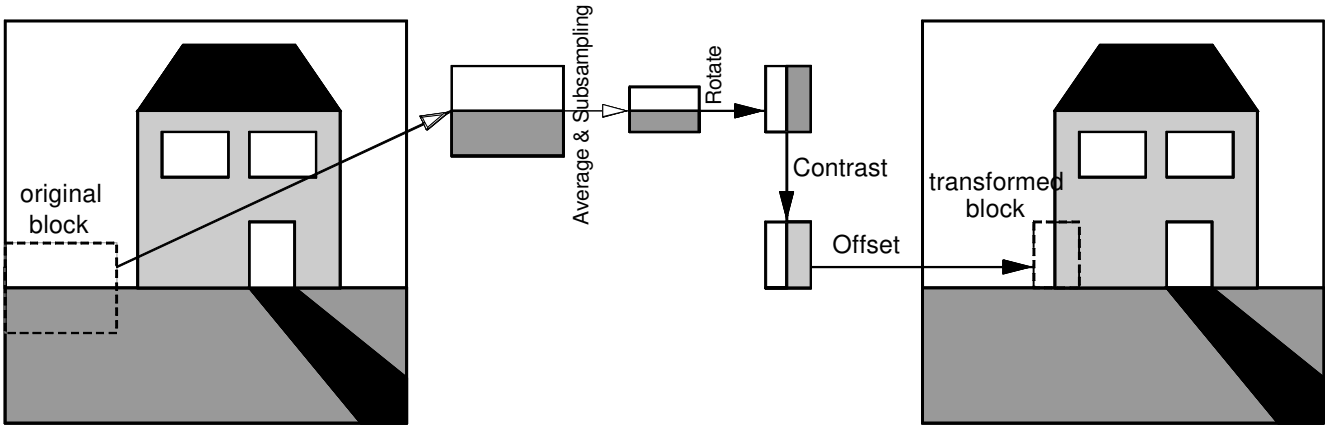
Beispiel: Küstenlinie einer Insel

Idee für die Bildkompression:

- Häufig ist ein Ausschnitt aus einem Bild einem anderen Ausschnitt ähnlich. Genauer: Er kann unter Anwendung von einfachen mathematischen Operationen aus dem anderen Ausschnitt errechnet werden (Translation, Rotation, Skalierung)
- Codierung: Ausgabe des ersten Bildausschnitts sowie für die ähnlichen Bildausschnitte Ausgabe der Transformationsoperationen.

Bildkompression mit Fraktalen

Beispiel



Literatur:

M.F. Barnsley, L.P. Hurd:
Bildkompression mit Fraktalen
Vieweg-Verlag, 1996

2.3 Kompressionsverfahren für Video

2.3.1 MPEG

Moving Picture Experts Group (Komitee der ISO).
MPEG konzentriert sich nicht nur auf Videodaten, sondern behandelt auch die damit verbundenen Audioströme. Ziel von MPEG-I: ein Videosignal soll samt Audio bei annehmbarer Qualität mit 1,5 Mbit/s komprimiert werden. (Datenrate einer T1-Verbindung in den USA).

Ähnlich wie JPEG sollte MPEG ein möglichst allgemeiner Standard werden, der viele Anwendungen unterstützt:

- Generalität
- Flexibilität
- Effizienz

Entwurfsziele des MPEG-1-Videokompressionsalgorithmus

- Wahlfreier Zugriff innerhalb von 1/2 Sekunde ohne signifikanten Qualitätsverlust
- schnelle Vorwärts-/Rückwärtssuche
- Bildfolge kann rückwärts abgespielt werden
- gute Editierbarkeit

MPEG - Videokodierung

Vier Frame-Typen in MPEG-1:

I-Frame (Intra Picture)

Intra-codiertes Vollbild, sehr ähnlich dem JPEG-Standbild, codiert mit DCT, Quantisierung, Lauflänge und Huffman

P-Bild (Predicted Picture)

Bezug auf vorherige I- und P-Bilder.
Delta-Codierung, DPCM-kodierte Makroblöcke
Bewegungsvektoren möglich

B-Bild (Interpolated Picture)

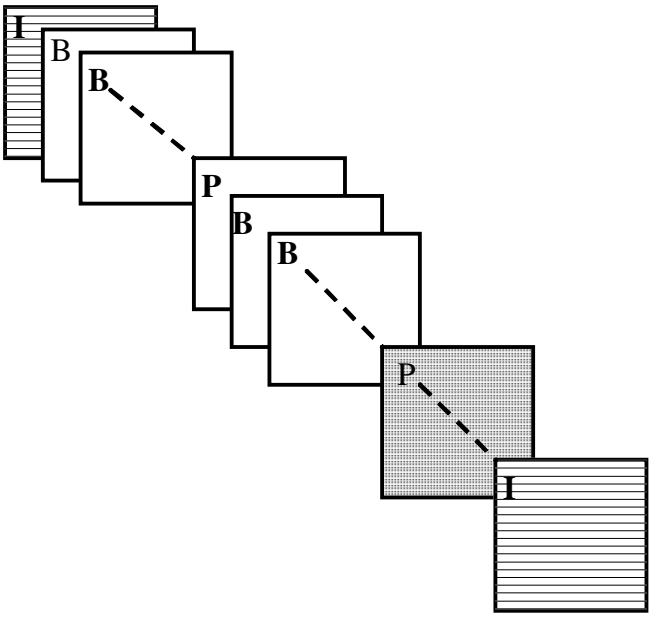
"bidirectionally predictive coded pictures", Bezug auch zu nachfolgenden Bildern
auch Interpolation zwischen Makroblöcken möglich

D-Bild

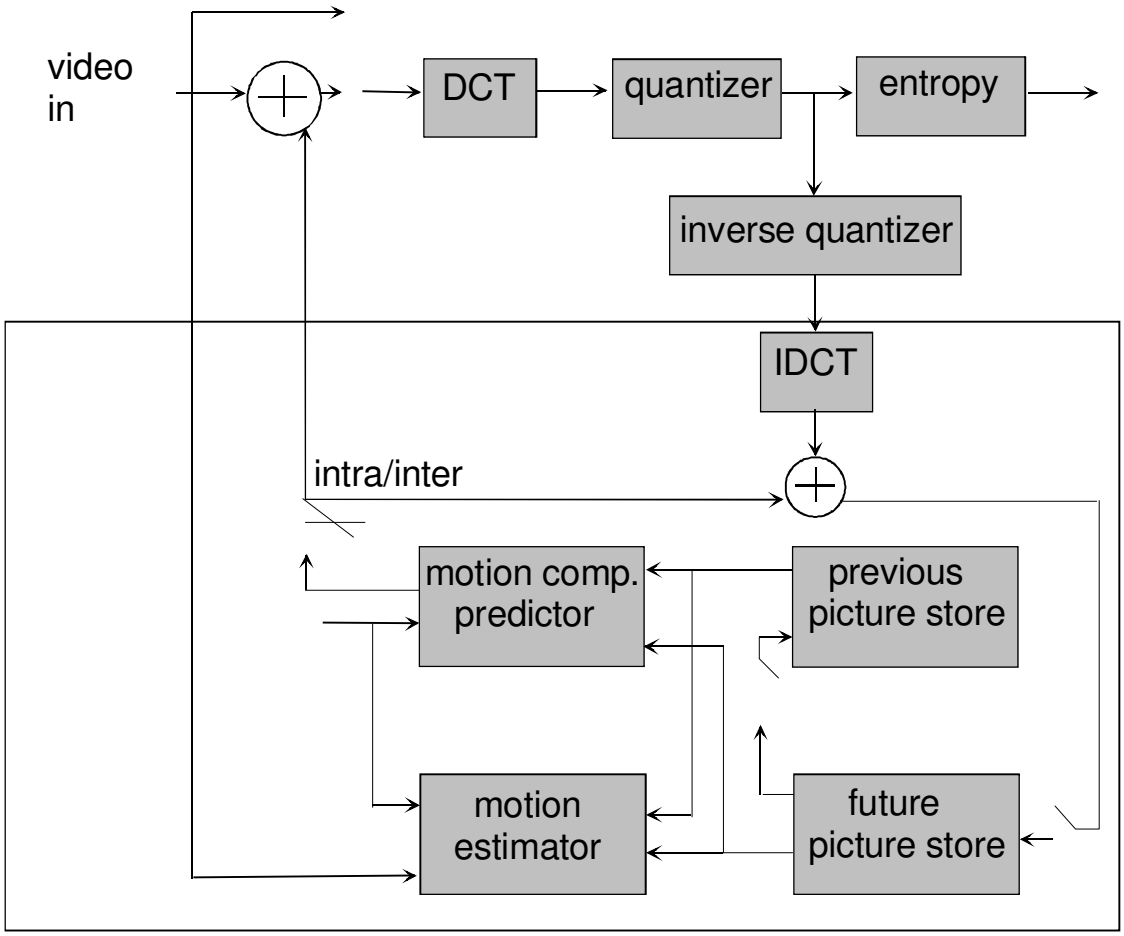
"DC coded picture", nur DC-Koeffizienten der Blöcke (linke obere Ecke der Koeffizienten-Matrix) für schnellen Vorlauf

„Group of Pictures“ in MPEG

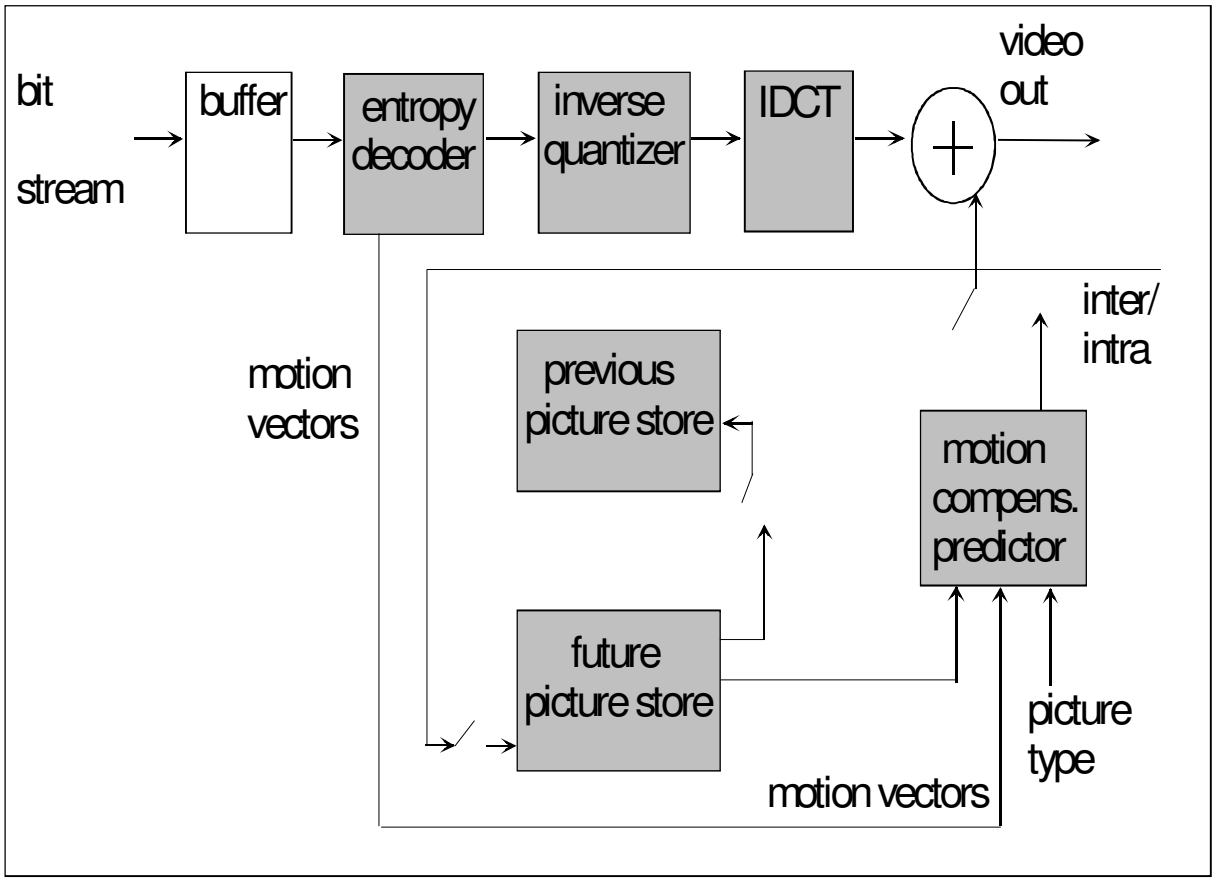
Die Folge von I, P und B-Bildern ist nicht standardisiert, sondern kann von der Anwendung bestimmt werden. So lassen sich Bildqualität und Kompressionsrate anwendungsabhängig wählen.



MPEG-Encoder

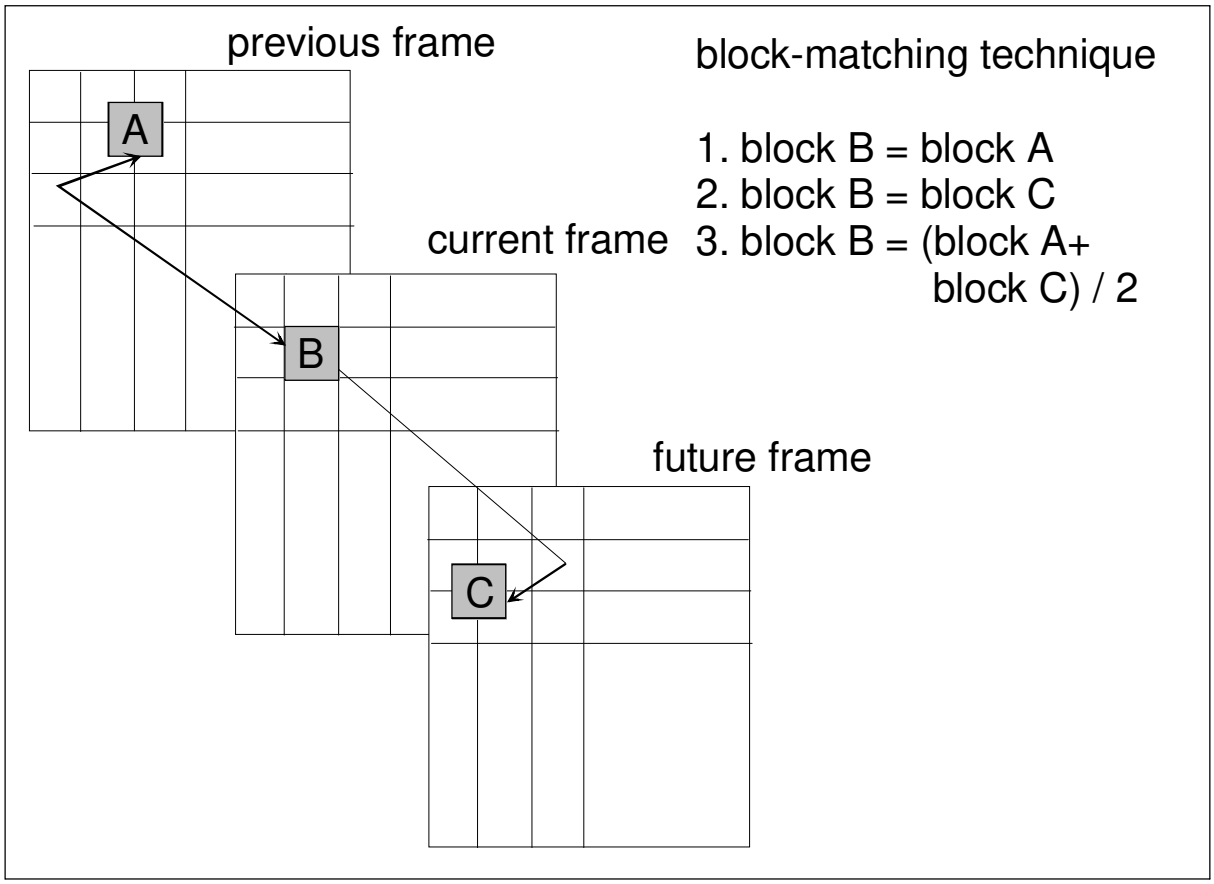


MPEG-Decoder



Zeitliche Redundanz und Bewegungsvektoren

"Motion Compensated Interpolation"



Der Suchradius (search range) kann bei der MPEG-Codierung als Parameter eingestellt werden. Je größer der Suchradius, desto besser das Kompressionspotential, aber desto länger auch die Laufzeit.

MPEG-1 “Constraint Parameter Set“

Im Prinzip ist der MPEG-Standard sehr flexibel, was aber die Implementierung standardkonformer Decoder erschwert. Deshalb wurde ein „Constraint Parameter Set“ standardisiert. Jeder MPEG-Encoder und -Decoder muß diesen Einschränkungen entsprechend operieren können, auch Hardwareimplementierungen.

horizontal size	**<= 720 pixels**
vertical size	**<= 576 pixels**
total number of macro-blocks/picture	**<= 396**
total number of macroblocks/s	**<= 396*25 = 330*30**
picture rate	**<= 30 frames/s**
bit rate	**<= 1.86 Mbits/s**
decoder buffer	**<= 376832 bits**

MPEG-2

MPEG-2 erweitert MPEG-1 für höhere Bandbreiten und besser Bildqualität, bis hin zu HDTV. Es wurde gemeinsam von ISO und ITU entwickelt (H.262).

Erstmals sieht MPEG-2 auch skalierbare Datenströme vor, die zur Bedienung von Empfängern verschiedener Leistungsklassen effizient eingesetzt werden können.

MPEG-2 Video-Profil

	Simple profile no B frames not scalable	Main profile B frames not scalable	SNR scalable profile B frames SNR scalable	Spatially scalable profile B frames SNR scalable	High profile B frames spatial or SNR scalable
High level 1920x1152x60		≤80 Mbits/s			≤100 Mbits/s
High-1440 level 1440x1152x60		≤60 Mbits/s		≤60 Mbits/s	≤80 Mbits/s
Main level 720x576x30	≤15 Mbits/s	≤15 Mbits/s	≤15 Mbits/s		≤20 Mbits/s
Low level 352x288x30		≤4 Mbits/s	≤4 Mbits/s		

Skalierbare Codierung = Codierung in Schichten (layers)

base layer: niedrigste Qualität

enhancement layers: Nachbesserung der Qualität

SNR scaling: enhancement layer enthält beispielsweise Chroma-Daten in höherer Auflösung

Spatial scaling: Erhöhung der Pixeldichte

MPEG-4

Ursprünglich planten ISO und ITU einen Standard MPEG-3 für HDTV-Qualität bei sehr hohen Datenraten. Dieser wurde inzwischen in MPEG-2 integriert und somit hinfällig.

Für MPEG-4 war zunächst an Video für sehr niedrige Bandbreiten gedacht (z.B. Multimedia für Mobilfunk). Inzwischen wird nach einer radikalen Kehrtwendung im Normungsgremium an einer Einzelobjekt-Codierung gearbeitet. Das Bild wird in einzelne Objekte segmentiert. Für jedes Objekt kann die Codierungstechnik getrennt gewählt werden. Die objektorientierte Codierung eröffnet zugleich völlig neue Möglichkeiten zur Weiterverarbeitung beim Empfänger.

Die Verabschiedung von MPEG-4 als International Standard wird für Ende 1998 erwartet.

2.3.2 ITU Recommendation H.261

Auch bekannt als „p*64kbit/s“

- Ein Verfahren zur Videokodierung für audiovisuelle Dienste bei einer Bitübertragungsrate von p x 64 Kbit/s
- Gedacht für ISDN
- Für den Parameter p gilt:
  - p ist aus [1,30]
  - p klein (p=1, p=2): Bildtelefon (schlechte Qualität)
  - p groß (ab p=6): auch Videokonferenz (mittlere Qualität)
- heute gebräuchlich: p=2, p=6
- Intraframe-Coding: DCT-basiert (wie JPEG)
- Interframe-Coding: DPCM-basiert

Videocodierung in H.261

Die fundamentalen Ideen sind denen in JPEG und MPEG gleich. H.261 ist weniger flexibel, dafür aber leichter standardkonform implementierbar.

Es werden genau drei Komponenten codiert, nämlich Y, C_b und C_r, jeweils blockweise mit 8x8 Pixels pro Block. Die Helligkeit Y wird mit voller Auflösung codiert, die beiden Chroma-Komponenten jeweils mit halber Auflösung in horizontaler und vertikaler Richtung (4:2:0-Modell).

Es gibt nur zwei Bildformate:

CIF	(Common Intermediate Format):	352x 288
QCIF	(Quarter CIF):	176x144

Die wichtigsten Parameter von H.261

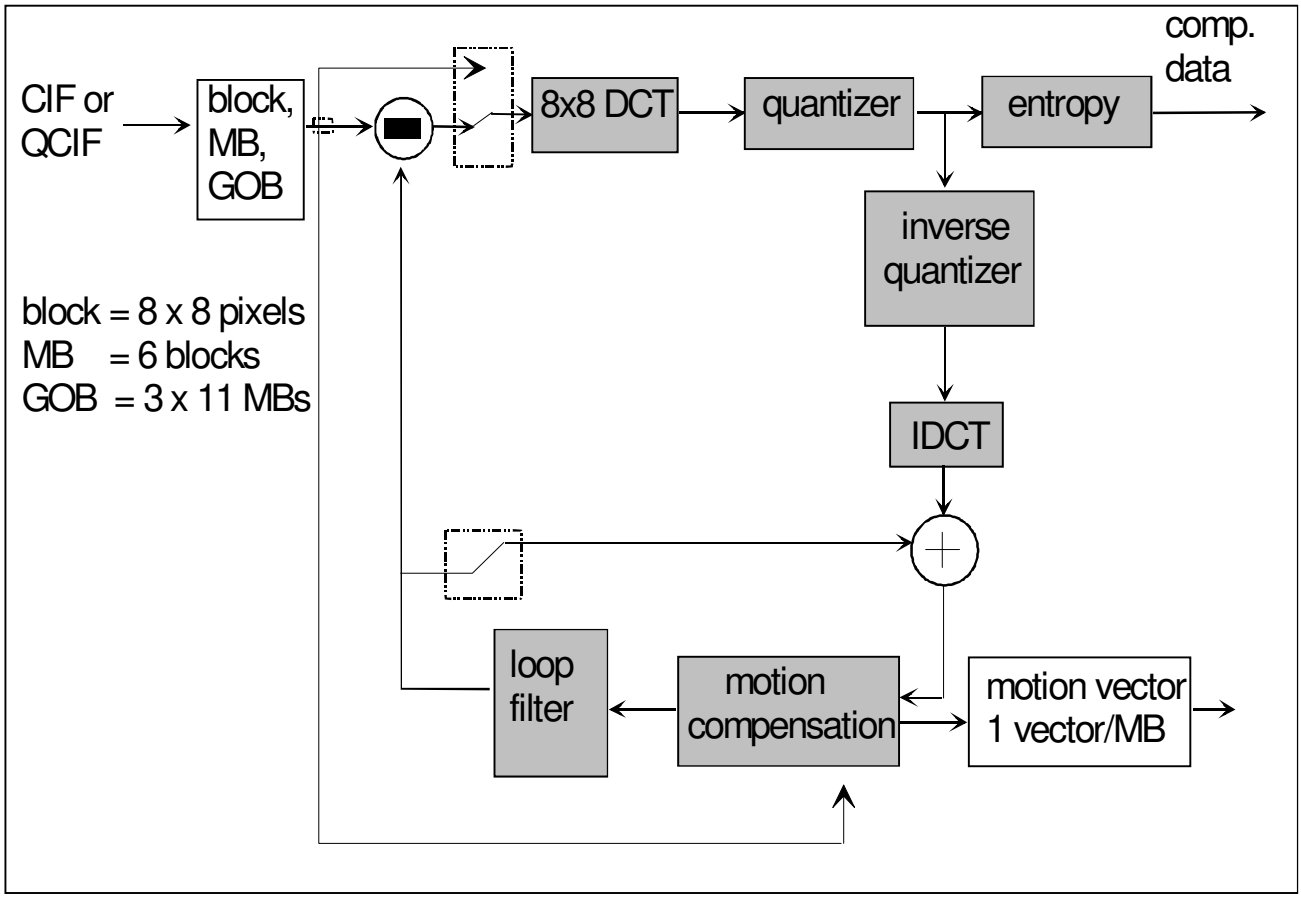
Die beiden Bildformate

	CIF (Breite x Höhe)	**QCIF (Breite x Höhe)**
Y	352 x 288	176 x 144
C_b	176 x 144	88 x 72
C_r	176 x 144	88 x 72

Die Hierarchie der Blockstrukturen

Strukturelement	**Beschreibung**
picture	one video picture (frame)
group of blocks	33 macro blocks
macro block	16 x 16 Y, 8 x 8 C_b, C_r
block	8 x 8 pixels (coding unit for DCT)

Die Funktionsweise eines H.261-Encoders



Status von H.261

Implementierungen in Hardware und Software verfügbar. Fast alle Bildtelefon-Hersteller haben inzwischen proprietäre Lösungen aufgegeben und auf H.261 umgestellt. PC-Lösungen sind auf dem Vormarsch.

2.4 Audio-Kompression

2.4.1 Grundlagen digitaler Audioströme

Audiosignale sind analoge Wellenverläufe. Die charakteristischen Eigenschaften werden im wesentlichen von der Frequenz (Tonhöhe) und von der Amplitude (Lautstärke) bestimmt.

Vor der weiteren Verarbeitung im Rechner müssen die analogen Wellenverläufe in digitale Signale transformiert werden. Dies geschieht durch **Abtastung**.

Abtastung

Der zum Abtastzeitpunkt vorliegende Momentan-Wert des Analogsignals wird der Analog-Digital-Umsetzung unterworfen. Praktisch wichtig ist vor allem die periodische Abtastung.

Abtasttheorem von Shannon und Raabe (1939)

Zur fehlerfreien Rekonstruktion des Signalverlaufs der abgetasteten Analogsignale ist eine Mindestabtasthäufigkeit (Abtastfrequenz f_A) erforderlich (bei periodischem Abtastzyklus).

Abtasttheorem

Eine Signalfunktion, die nur Frequenzen im Frequenzband B (bandgegrenztes Signal) enthält, wobei B gleichzeitig die höchste Signalfrequenz ist, wird durch ihre diskreten Amplitudenwerte im Zeitabstand

$$t_0 = \frac{1}{2} B$$

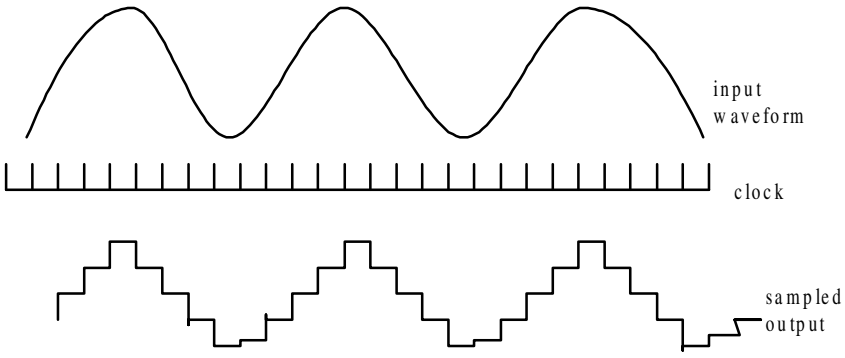
vollständig bestimmt.

Andere Formulierung:
Die Abtastfrequenz f_A muß doppelt so hoch sein wie die höchste im abzutastenden Signal vorkommende Frequenz f_S

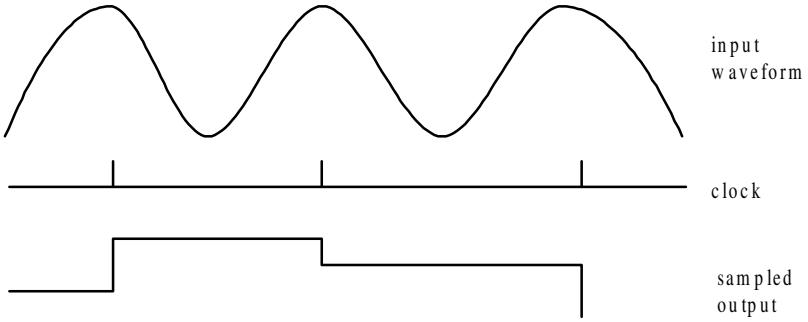
$$f_A = 2\, f_S$$

Abtastung und Quantisierung sind voneinander unabhängig zu betrachten. Eine exakte Rekonstruktion des Zeitverlaufs (bzw. des Frequenzspektrums) sagt nichts über den Fehlergrad bei der Signalwertdiskretisierung (Quantisierung) aus.

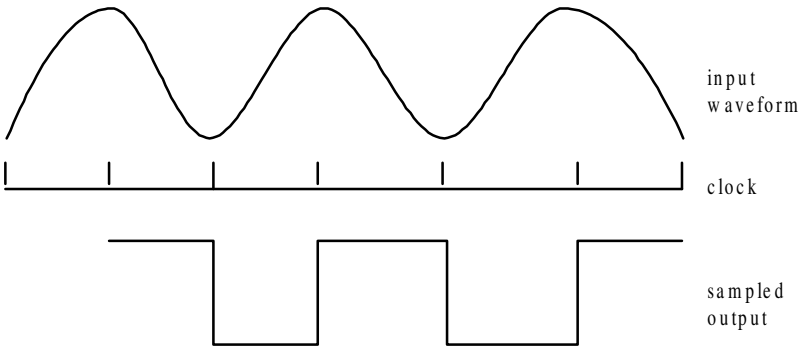
Verschiedene Abtastfrequenzen



(a) Sampling rate is much higher than signal frequency



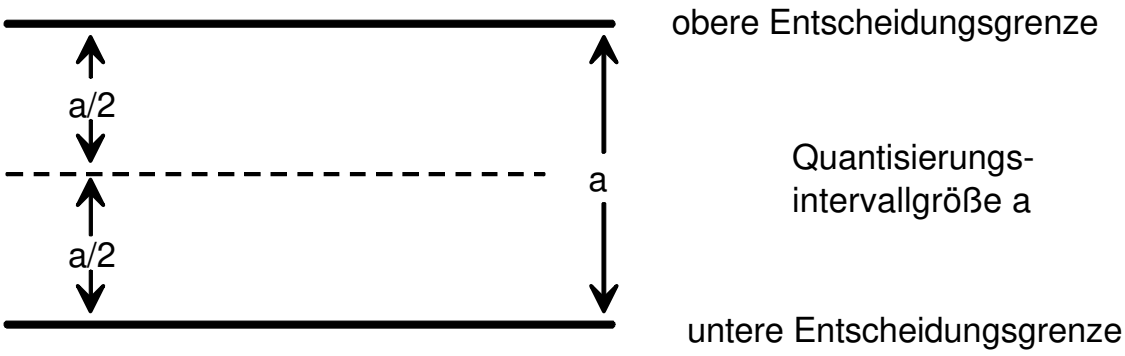
(b) Sampling rate is lower than signal frequency



(c) Sampling rate is at the Nyquist limit

Quantisierung

Der gesamte Wertebereich des Analogsignals wird in eine endliche Anzahl von Intervallen Quantisierungsintervallen) eingeteilt, denen jeweils ein fester diskreter Wert zugeordnet wird. Da alle in ein Quantisierungsintervall fallenden Analogwerte **demselben** diskreten Wert zugeordnet werden, entsteht ein Quantisierungsfehler.



Quantisierungintervall für die Zuordnung eines diskreten Wertes zu allen zwischen + a/2 und - a/2 liegenden Werten einer Analogdarstellung.

Rückwandlung:

Beim Empfänger wird ein Analogwert rückgewonnen (Digital-Analog-Umsetzung), der dem in der Mitte des Quantisierungsintervalls liegenden Analogwert entspricht.

Maximaler Quantisierungsfehler: **a/2**

Codierung

Die Quantisierungsintervalle werden durch die Zuordnung eines - im Prinzip frei wählbaren - (Binär-) Codes gekennzeichnet und unterschieden. Anstelle des ursprünglichen Analogsignals wird die - mit dem Quantisierungsfehler behaftete - digitale Darstellung übertragen.

Im einfachsten Fall wird ein reiner Binärcode Darstellung als Binärzahl) als Codierung des diskreten Digitalwertes gewählt.

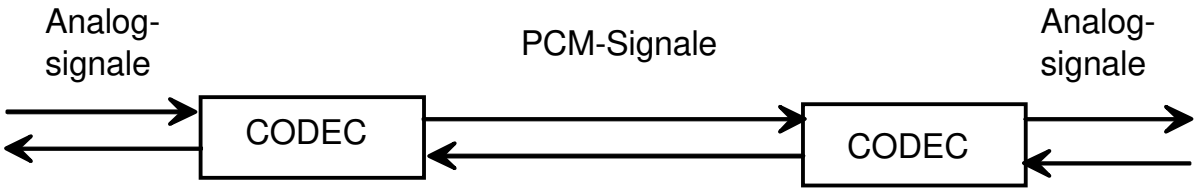
2.4.2 Pulse Code Modulation

Die Zusammenfassung der Schritte

Abtastung
Quantisierung
Codierung

und die Darstellung der gewonnenen Codewörter als digitale Basisbandsignale am Ausgang des PCM-A/D-Umsetzers ist Grundlage der in großem Umfang eingesetzen **PCM-Technik**.

Die A/D-Umsetzung (Abtastung/Quantisierung) und Codierung sowie die Rückkonvertierung erfolgt im sogenannten **CODEC (*Co*dierer/*D*ecodierer)**.



PCM-Fernspechkanal

Ausgangspunkt: Analoger CCITT-Fernsprechkanal
Frequenzlage 300-3400 Hz
Bandbreite 3100 Hz

Abtastfrequenz: $\mathbf{f_A = 8\ kHz}$

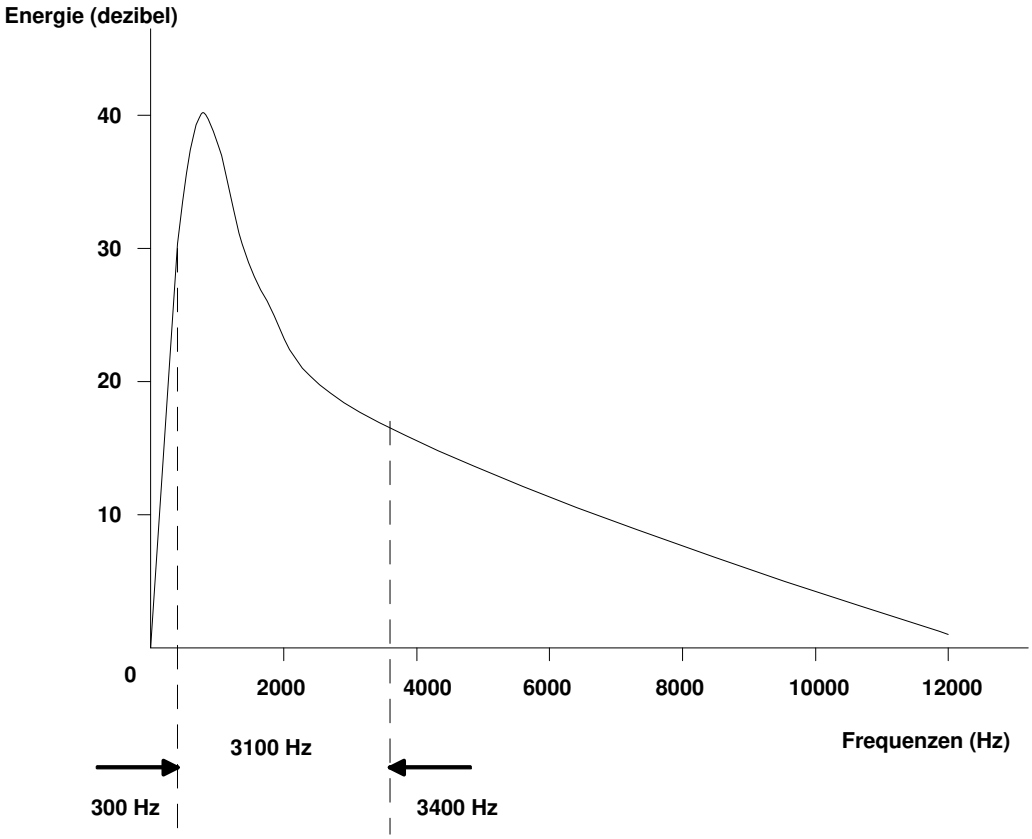
Abtastperiode: $\mathbf{T_A = 1/\ f_A = 1/8000\ Hz = 125\mu s}$

Die vom CCITT gewählte Abtastfrequenz ist höher als nach Shannon-Abtasttheorem erforderlich: 3400Hz obere Bandgrenze ergibt 6800 Hz Abtastfrequenz.
Für die höhere Abtastfrequenz gibt es technische Gründe (Filtereinfluß, Kanaltrennung usw.).

Frequenzspektrum eines Signals

Bandbegrenztes Signal

Signale können ein „natürlich“ begrenztes - kontinuierliches - Frequenzspektrum umfassen oder durch technische Mittel auf einen Ausschnitt ihres Spektrums begrenzt werden (Filter).



Kontinuierliches Frequenzspektrum der menschlichen Stimme und Bandbreite des CCITT-Standardtelefonkanals.

Amplitudenquantisierung

Die Zahl der benötigten Quantisierungsintervalle wird bei der akustischen Sprachkommunikation (Fernsprechen) durch den Grad der Silbenverständlichkeit beim Empfänger bestimmt.

Mit „Sicherheitszuschlag“ wurden vom CCITT 256 Quantisierungsintervalle genormt.

Bei binärer Codierung ergibt dies 8 Bits pro Abtastung.

Die Übertragungsgeschwindigkeit (Bitrate) für einen digitalisierten PCM-Fensprechkanal ist demnach

Bitrate	**= Abtastfrequenz**	**x**	**Codewortlänge**
kbit/s	**= 8000/s**	**x**	**8 bit**

= **64 kbit/s**

Ungleichförmige Quantisierung

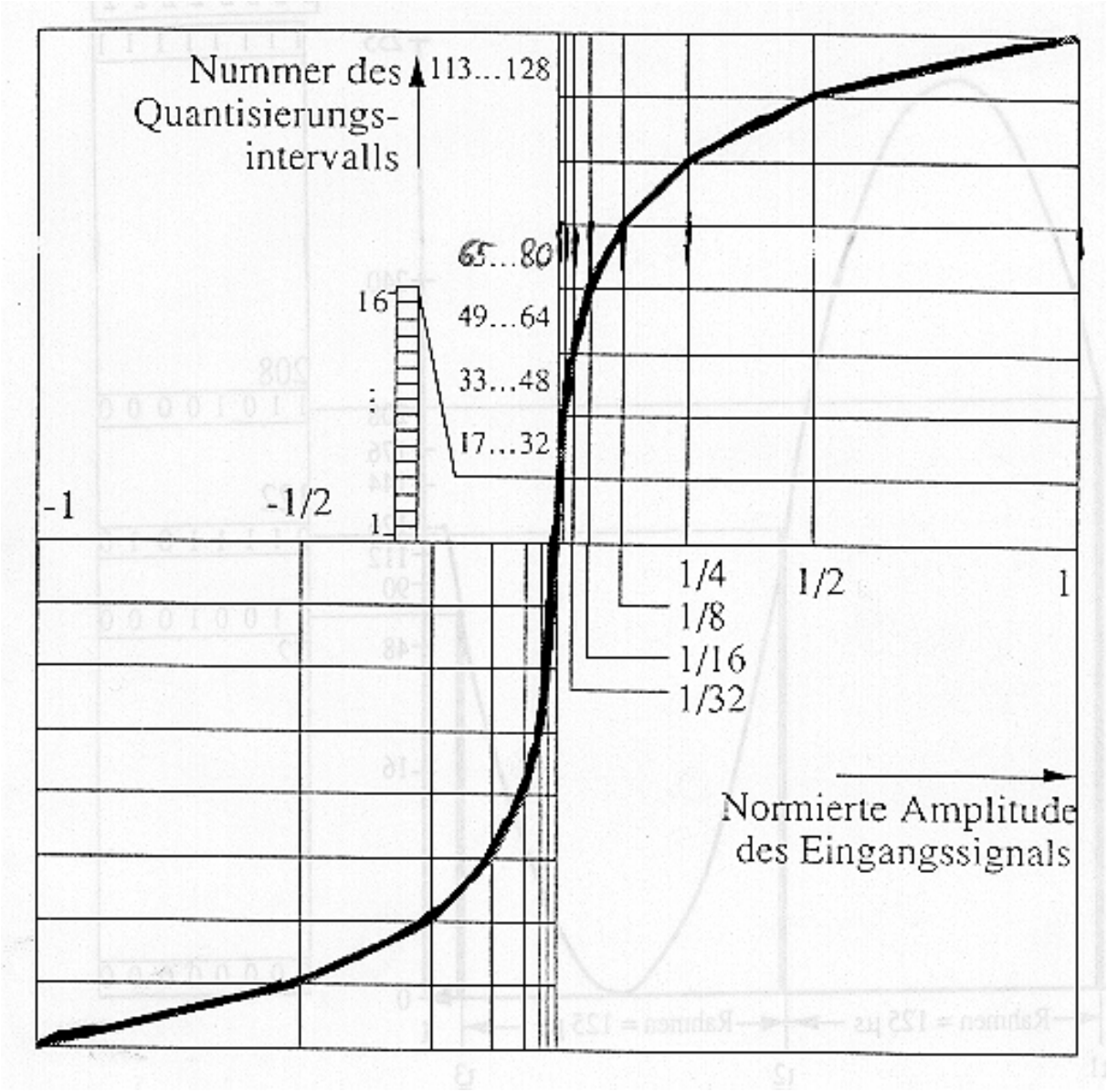
Bei gleichförmiger Quantisierung sind alle Intervalle gleich groß und vom Momentanwert des Signals unabhängig. Quantisierungsfehler machen sich bei gleichförmiger Quantisierung bei kleinen Signalwerten sehr stark bemerkbar (Quantisierungsrauschen).

Bei ungleichförmiger Quantisierung sind die Quantisierungsintervalle bei großer Signalamplitude größer und bei kleiner Amplitude kleiner.

Die ungleichförmige Intervallgröße wird durch einen dem Quantisierer vorgeschalteten (Signal-) *Kompressor* erzielt. Auf der Empfangsseite wird in inverser Funktion ein *Expander* eingesetzt. Er dient zur Wiederherstellung der ursprünglichen Größenverteilung der Signale (Dynamik der Signale).

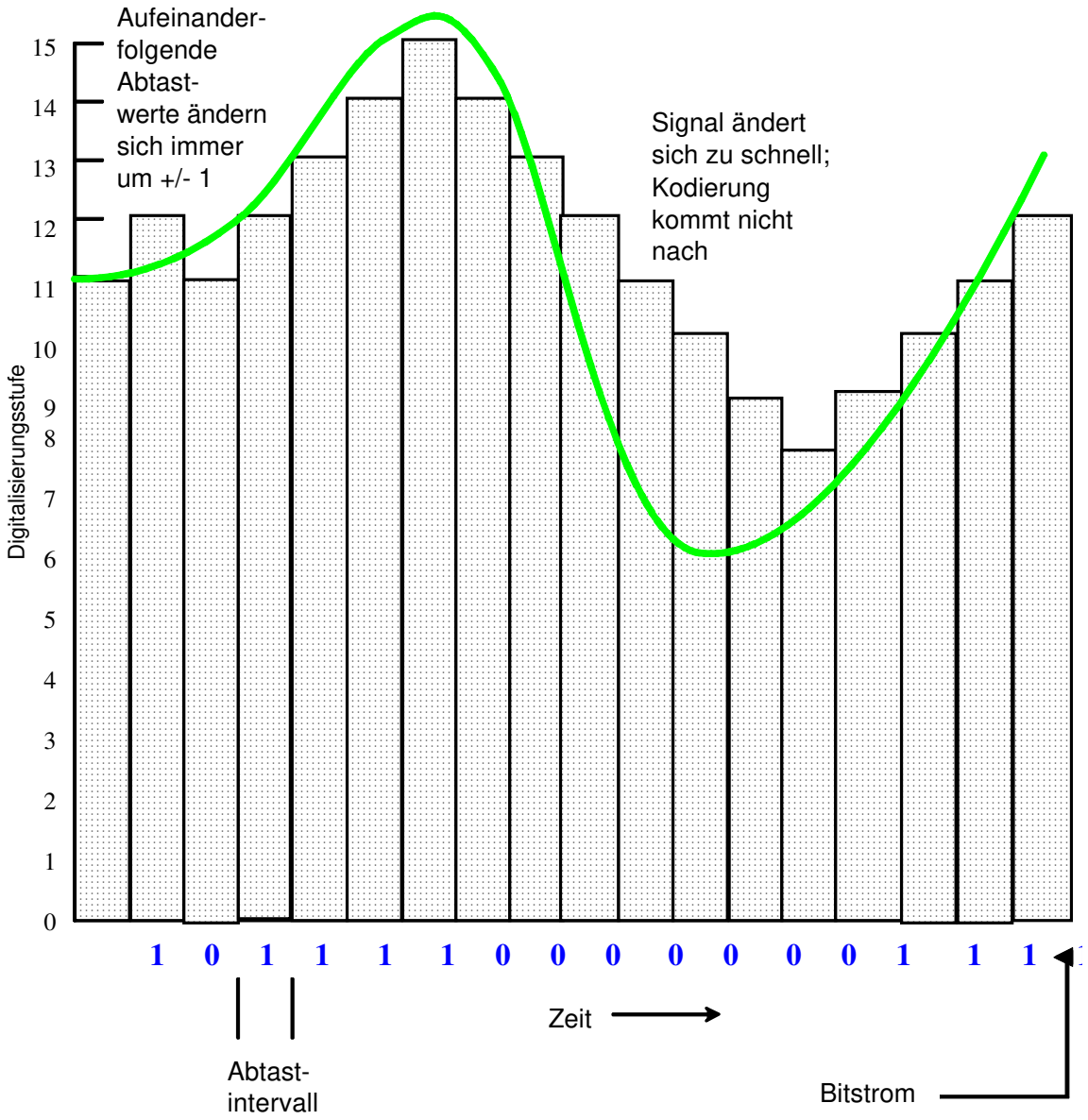
Als Kompressionskennlinien werden logarithmische Kennlinien verwendet, die schaltungstechnisch durch lineare Teilstücke approximiert werden.

13-Segment-Kompressorkennlinie



Delta-Modulation

Statt der Absolutwerte der Amplitude werden die Wert-differenzen codiert.



Codierung:

1 = steigendes Signal

0 = fallendes Signal

Differentielle PCM (DPCM)

Differenz zweier PCM-Werte wird codiert. Ergibt kleinere Werte, geringere Bitrate, aber Quantisierungsfehler bei schnellen Signalschwankungen.

Adaptive DPCM (ADPCM)

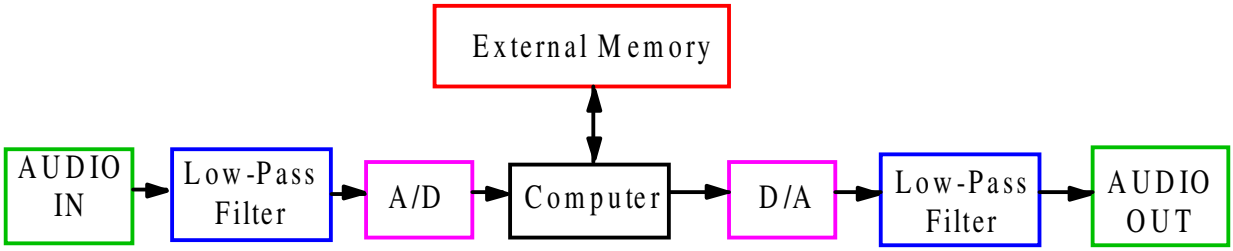
Die Quantisierungstabelle ändert sich mit den tatsächlichen Signalschwankungen. Der Empfänger kann diese Tabellenänderung dynamisch nachvollziehen. Bei kleinen Signalschwankungen wird mit weniger bit/s übertragen als bei großen Signalschwankungen. Dadurch kann beispielsweise eine Kompression von 1,4 Mbit/s auf 0,2 Mbit/s bei vergleichbarer Qualität möglich werden.

Gebräuchliche Sampling-Parameter

8 **kHz** Telefon-Standard, µ-law encoding, SUN Audio
32 **kHz** Digital Radio, NICAM, (DAT)
44,1 **kHz** CD
48 **kHz** Digital Audio Tape (DAT)

Sampling-Breite:

8 bit 256 Amplituden-Stufen (Sprache)
16 bit 65536 Amplituden-Stufen (HiFi)



Beispiel für Audio-Parameter

Signal:	20 kHz Bandbreite (HiFi)
Sampling Rate:	44,1 kHz
Sampling Breite:	15 bit

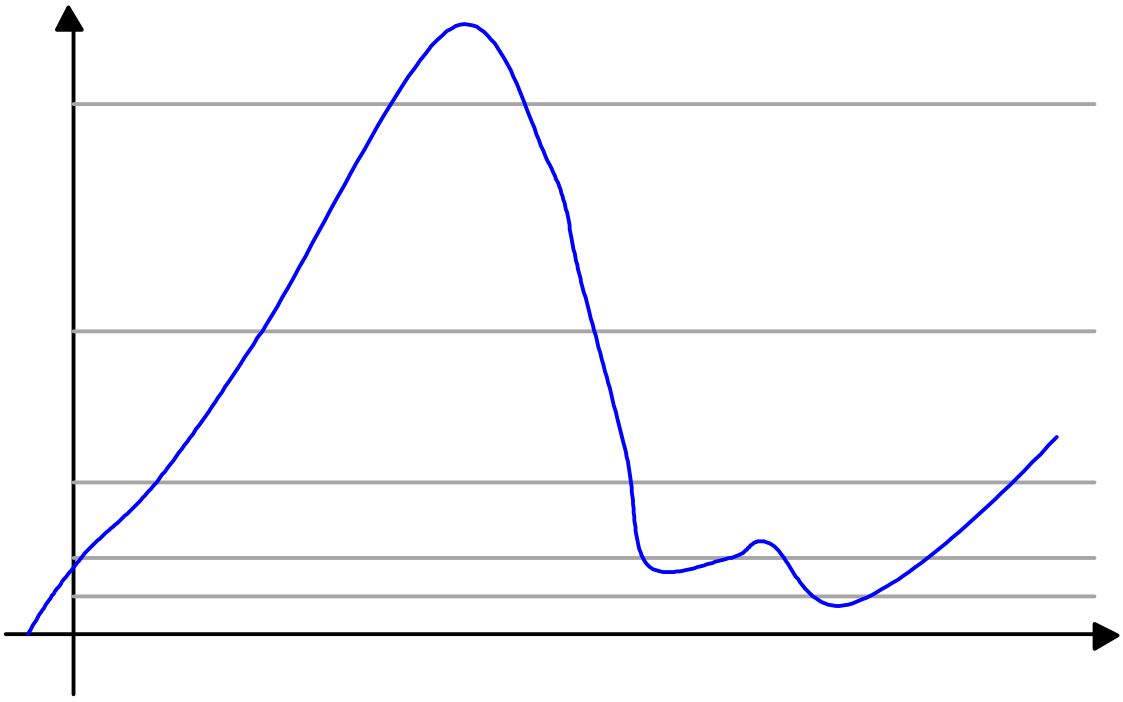
Datenrate:

mono	0,7 M bit/s
stereo	1,4 M bit/s

Das ist die Datenrate des CD-Players!

Beispiel für eine Audio-Codierung

Puls-Code-Modulation (PCM)
üblich: nichtlineare Quantisierung
A/D Wandler: 13 (14) bit
Abbildung auf 8 bit mit logarithmischer Kennlinie: A-law (µ-law)



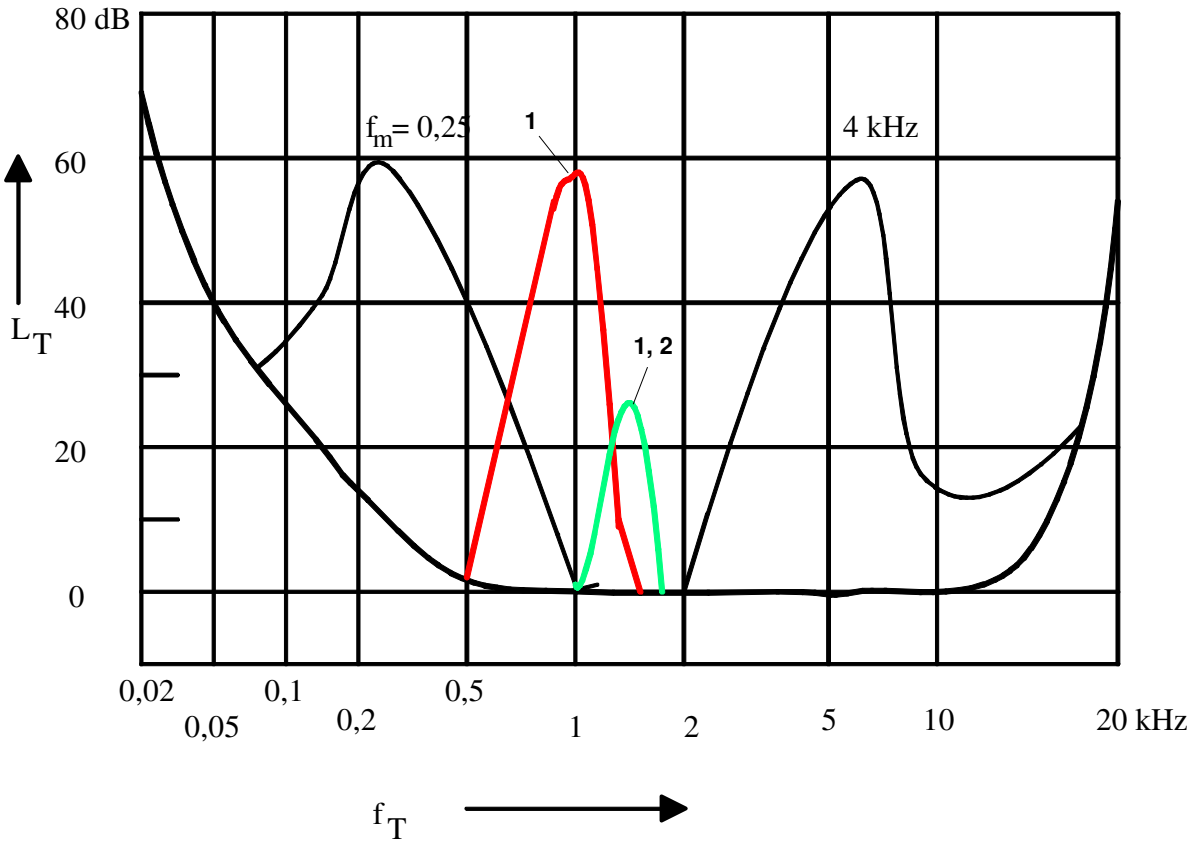
2.4.3 Psycho-akustische Modelle

"Irrelevanzreduktion"

Ausnutzung des Auflösungsvermögens des menschlichen Ohres

Verdeckungseffekt

- Mithörschwelle, frequenz- und pegelabhängig



Daten, die das Ohr sowieso nicht hören würde, werden schon an der Quelle herausgefiltert.

Beispiel: MPEG Audio

Merkmale

Kompression auf 32, 64, 96, 128 oder 192 kBit/s

Audiokanäle:

- Mono oder
- 2 unabhängige Kanäle oder
- "Joint Stereo"

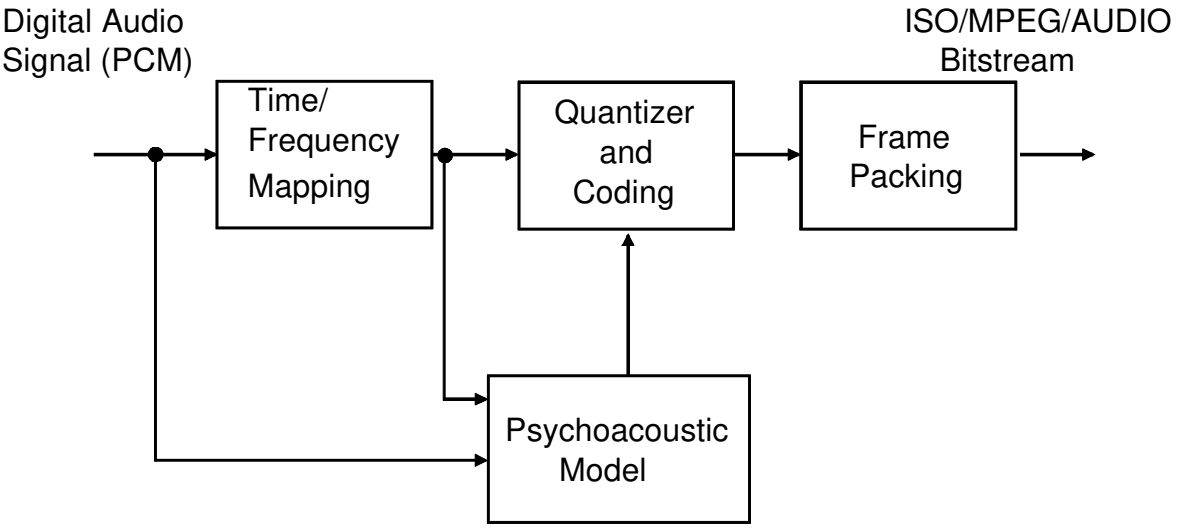
Verfahren

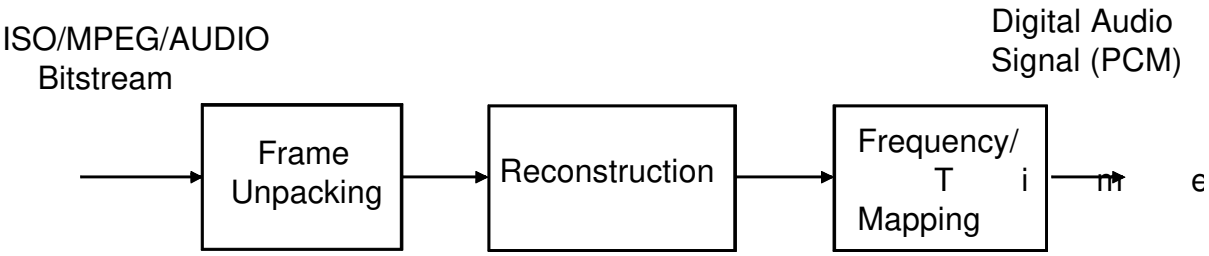
- Abtastrate 32 kHz, 44,1 kHz oder 48 kHz
- 16 Bits pro Abtastwert
- Verzögerung durch Codieren und Decodieren höchstens 80 ms bei 128 kbit/s
- **psychoakustisches Modell steuert die Quantisierung**

Zwei Verfahren in MPEG-Audio

MUSICAM	ASPEC
Masking Pattern Universal Subband Integrated Coding And Multiplexing	Advanced spectral Entropy Coding
Institut für Rundfunktechnik München	FhG Erlangen
Teilbandcodierung	dynamische Frequenz-bänder (überlappend) Entropiecodierung (Huffman)
einfacher Aufbau	sehr gute Ergebnisse bei niedrigen Bitraten

Funktionsweise eines MPEG-Audio-Encoders und -Decoders





Drei Schichten in MPEG Audio

I	Teilbandcodierung mit 32 Bändern nach MUSICAM • hohe Datenrate
II	Teilbandcodierung nach MUSICAM, aufwendigeres psycho-akustisches Modell bessere Bestimmung der Skalierungsfaktoren • mittlere Datenrate
III	Transformationscodierung mit max. 512 dynamischen Fenstern und Entropiecodierung nach ASPEC • niedrigste Datenrate