

## 6 Videokompression

### 6.1 Fernseh-Charakteristika

### 6.2 H.261

### 6.3 MPEG-1 und MPEG-2

### 6.4 Berechnung von Differenzbildern

### 6.5 Ausblick auf MPEG-4

#### Fernseh-Charakteristika:

Farbnorm:	PAL	Secam	NTSC
Bilder pro s:	25	25	29,97
Zeilenzahl pro Bild:	625	625	525
Zeilendauer:	64 $\mu$ s	64 $\mu$ s	63,492 $\mu$ s
Vertikalfrequenz:	50 Hz	50 Hz	59,94 Hz
Verhältnis Breite zu Höhe:	4:3	4:3	4:3
Kanalbreite:	7 MHz	8 MHz	6 MHz
Videobandbreite:	5 MHz	6 MHz	4 MHz
Leistungsverhältnis Bild zu Ton:	5:1	8:1	4:1

#### Abkürzungen:

PAL = Phase Alternating Line

Secam = Séquentielle Couleur avec Mémoire

NTSC = National Television Systems Committee

Zum Vergleich: Zahl der Bilder beim Tonfilm: 24

#### Bemerkungen:

- (i) Transformation Tonfilm nach PAL: Keine, der Film wird mit einer leicht erhöhten Geschwindigkeit abgespielt.
- (ii) Transformation Tonfilm nach NTSC: Abwechselnd werden aus einem Bild drei Halbbilder und zwei Halbbilder erzeugt, so werden aus 24 Bildern 60 Halbbilder.
- (iii) Transformation NTSC nach PAL: Hinzufügen von 100 Zeilen je Bild und Streichen jeden 6-ten Halbbildes.

## Einige Zahlen:

Hier nur Variante mit 625 Zeilen:

Dauer eines Bildes: 40 ms

Dauer einer Zeile: 64  $\mu$ s

Strahlrückführung: 12  $\mu$ s (horizontale Austastlücke)

Vertikalaustastlücke: 1,6 ms pro Halbbild, daher nur  
575 Zeilen nutzbar, denn  
 $(625 - 575) * 64 \mu\text{s} = 3200 \mu\text{s}$ .

Die obigen Zahlen sind die Zahlen für Studiogeräte, zur Ausstrahlung an den Endverbraucher gelangen etwa 385 Zeilen, damit ist die Spaltenzahl  $4/3 * 385 \approx 514$ .

Man berechnet hieraus eine Videogrenzfrequenz von  $(514/2) / (52 * 10^{-6}) \text{ Hz} = 4,94 * 10^6 \text{ Hz}$ . Dieser Wert korreliert recht gut zur Videobandbreite von 5 MHz.

## H.261 (März 1993):

Voller Name: VIDEO CODEC FOR AUDIOVISUAL SERVICES AT  $p \times 64$  kbits

Bemerkung:  $p$  kann die Werte von 1 bis 30 annehmen.

Entwicklungszeit: 1984 – 1990

### Bemerkungen:

- (i) Der Standard H.261 unterstützt Bildtelefonie und Videokonferenzen.
- (ii) Die Bildformate CIF und QCIF werden unterstützt.
- (iii) Die bei der Bewegungsvorhersage betrachtete Einheit ist der Makroblock.
- (iv) Das verwendete Farbmodell ist YCrCb in der Variante 4 : 2 : 0.
- (v) Die benutzte Transformation ist die diskrete Cosinus-Transformation für 8 x 8 – Matrizen.
- (vi) Eine Glättungsoperation wird notwendig wegen der Integration von alten Bildteilen in neue Bilder.
- (vii) Die übertragenen Daten werden durch einen BCH (511, 493)-Code mit dem Generatorpolynom  $(x^9 + x^4 + 1) * (x^9 + x^6 + x^4 + x^3 + 1)$  geschützt. Die Berücksichtigung des Fehlercodes im Decoder ist optional.

**Hierarchie:**

**Bildfolge:**

**Bild** (= 3 oder 12 Blockgruppen)

**Blockgruppe** (= 33 Makroblöcke)

**Makroblock** (= 6 Blöcke)

**Block** (= 8 x 8 – Pixelwerte)

1	2	3	4	5	6	7	8	9	10	11
12	13	14	15	16	17	18	19	20	21	22
23	24	25	26	27	28	29	30	31	32	33

**Anordnung der Makroblöcke in einer Blockgruppe**

**CIF:**

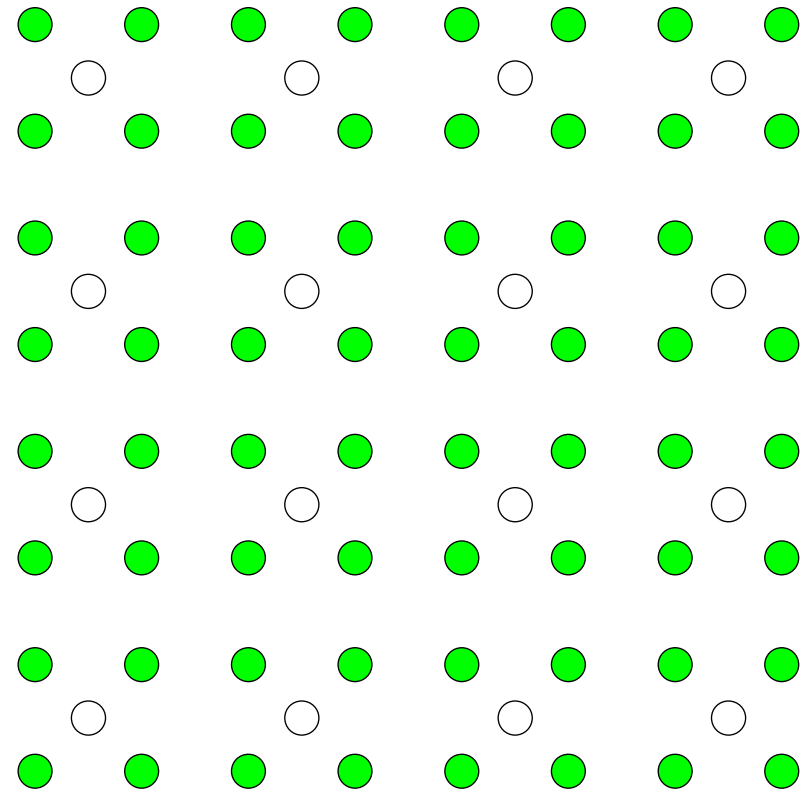
1	2
3	4
5	6
7	8
9	10
11	12

**QCIF:**



1
3
5

**Anordnung der Blockgruppen für CIF und QCIF**

**Reduktion der Chrominanz 4 : 2 : 0 :**



**Verhältnis der Farbinformationen:**

-  **Luminanz**
-  **Chrominanz (zweimal)**

## Formate:

### Common Intermediate Format:

	CIF	QCIF
Luminanz (Y):	352 x 288	176 x 144
Chrominanz (Cr):	180 x 144	90 x 72
Chrominanz (Cb):	180 x 144	90 x 72
Bildfrequenz:	29,97 Hz	29,97 Hz

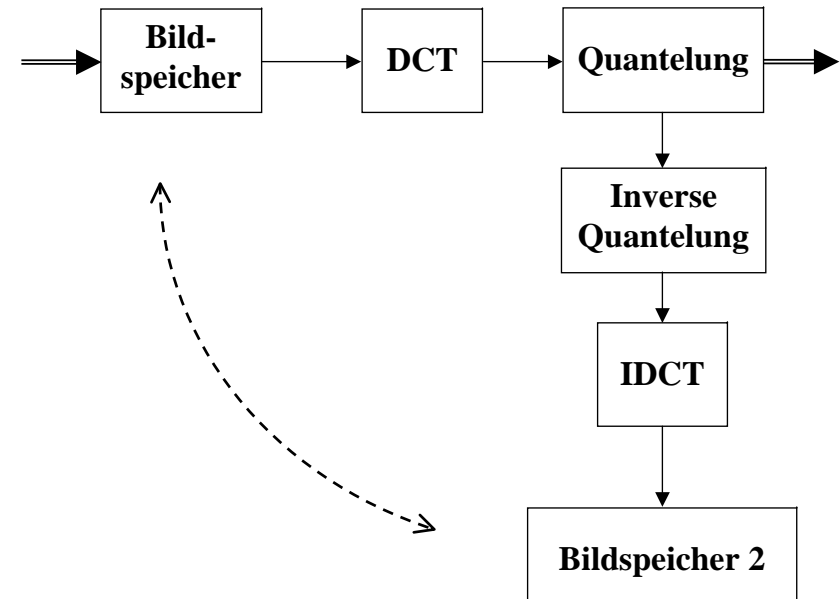
### Glättungsoperation:

Die Berücksichtigung von Kanten bei der Differenzbildung zur Bewegungskompensation kann zu großen Werten bei den Hochfrequenzkomponenten in der Bildtransformierten führen. Daher nutzt man in diesem Fall ein- oder zweidimensionale Glättungsoperationen mit den eindimensionalen Parametern  $1/4$ ,  $1/2$ ,  $1/4$ . Die Glättung erfolgt nur innerhalb von  $8 \times 8$  - Pixelblöcken.

Beispiel:    Original:        110 218 116 112  
                              108 210 110 114  
                              110 218 210 112  
                              112 108 110 116

                  Zeilenglättung: 110 166 141 112  
                                      108 160 136 114  
                                      110 189 188 112  
                                      112 110 111 116

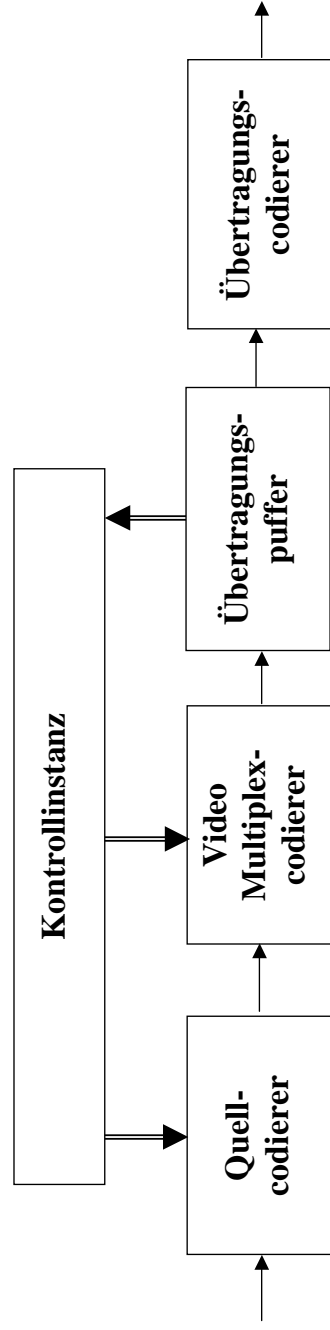
                  Spaltenglättung: 110 166 141 112  
                                      109 169 150 113  
                                      110 162 156 114  
                                      112 110 111 116



### Überblick über Codierung

**Bemerkung:** Um die temporale Redundanz einer Bildfolge zu nutzen, wird versucht, zu einem Teilbild eines neuen Bildes ein fast gleichwertiges Teilbild im Bildspeicher 2 der rekonstruierten Bilder zu finden.

### Kontrolle des Videostroms:



### Methoden der Anpassung:

- (i) Bis zu drei Bildern können zwischen zwei gesendeten Bildern ausgelassen werden.
- (ii) Anpassung des Bitaufkommens durch Veränderung der Quantisierungswerte.
- (iii) Anpassung des Bitaufkommens durch Veränderung des Schwellwertes zur Versendung von Differenzblöcken im Interframe-Modus.

### MPEG:

**MPEG = Moving Picture Experts Group**

**Offiziell: ISO/IEC JTC1/SC29/WG11**

### Standards:

**MPEG-1: ISO/IEC 11172-1 - 11172-5**  
Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s.

**MPEG-2: ISO/IEC 13818-1 - 13818-11 (w/o 13818-8)**  
Generic coding of moving pictures and associated audio information.

**MPEG-4: ISO/IEC 14496-1 - 14496-19 (w/o 14496-17)**  
Coding of audio-visual objects.

**MPEG-7: ISO/IEC 15938-1 - 15938-11**  
Multimedia content description interface.

**MPEG-21: ISO/IEC 21000-1 - 21000-16 (w/o 21000-4, 21000-8, 21000-13 - 21000-15)**  
Multimedia framework.

**Datenrate für Bildsequenz:**

$$\begin{aligned} \text{Bildgröße} * \text{Farbtiefe} * \text{Abspielrate} &= \\ 360 * 240 * 24 * 30 &= 62.208.000 \text{ Bit/s} \end{aligned}$$

**Datenrate einer Audio-CD:**

$$\begin{aligned} \text{Abtastrate} * \text{Abtastgröße} * \text{Kanalzahl} &= \\ 44100 * 16 * 2 &= 1.411.200 \text{ Bit/s} \end{aligned}$$

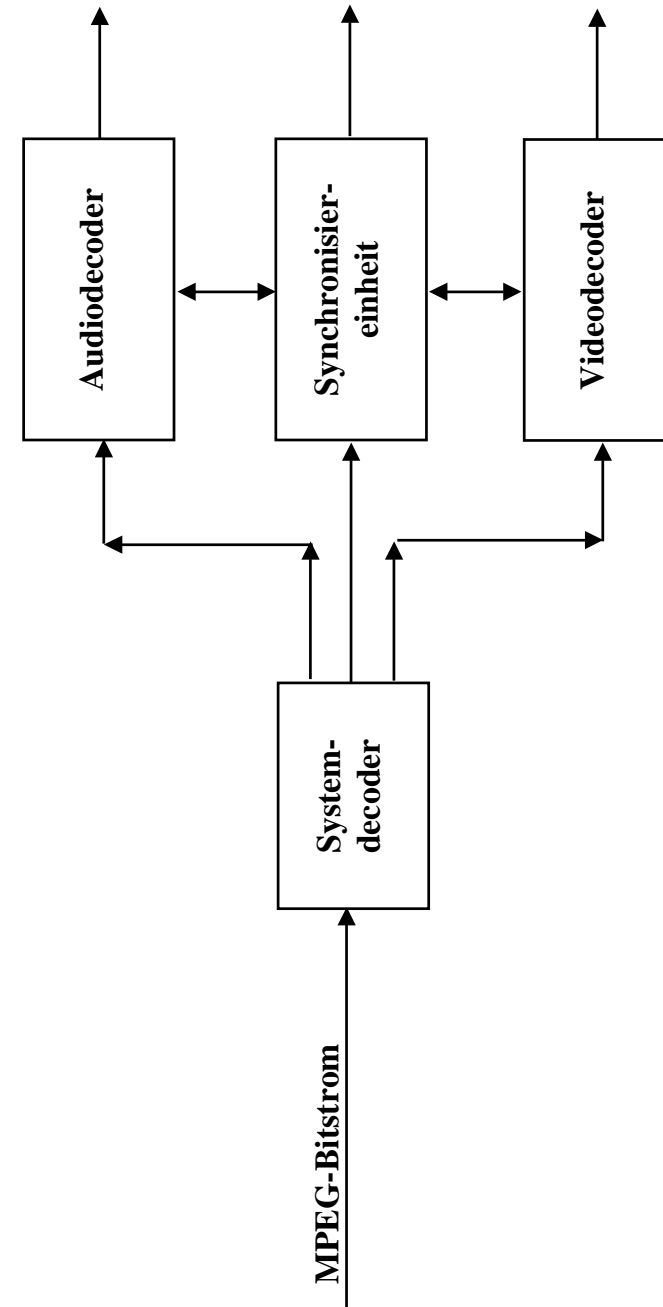
**Erstrebte Kompressionsrate: 44,08**

**Dies betrifft nur die Bildfolge, für den Ton wird zusätzliche Kanalkapazität benötigt.**

**Eine Kompressionsrate von 44 ist erreichbar, wenn man bedenkt, daß JPEG einen Kompressionsfaktor von 20 bei angemessener Qualität erlaubt. Einen Faktor 2 bis 3 sollte die Berücksichtigung der Ähnlichkeit aufeinanderfolgender Bilder liefern.**

**Nutzt man die doppelte Abspielrate einer CD, dann läßt sich auf einer CD ein halbstündiger Film im Kleinformat speichern.**

**Überblick über MPEG-1:**



## Forderungen an MPEG-1:

- (i) Gefordert wird VHS-Qualität bei einer Bandbreite von 1,5 Mbit/s.
- (ii) Asymmetrie bevorzugt: schnellere und häufigere Dekompression.
- (ii) Direkter Zugriff über Zugriffspunkte (ohne Fremdreferenz der Kodierung) innerhalb 0,5 s.
- (iv) Schneller Vor/Rücklauf über besondere Zugriffspunkte.
- (v) Reverse Playback.
- (vi) Audio-Video-Synchronisation.
- (vii) Robustheit, Fehlertoleranz.
- (viii) Codier- und Decodierverzögerung unter 150 ms.
- (ix) Editierbarkeit auch komprimierter Ströme.
- (x) Unterstützung vieler Bildgrößen.

Folgenden Parametern ist zu genügen:

Zahl der Pixel pro Zeile	$\leq 720$
Zahl der Zeilen	$\leq 576$
Zahl der Makroblöcke pro Bild	$\leq 396$
Zahl der Makroblöcke pro Sekunde	$\leq 396 * 25 = 330 * 30$
Bildrate	$\leq 30$ Bilder/s
Bitrate	$\leq 1,86$ Mbit/s
Decoder Puffer	$\leq 376.832$ Bit

**Bemerkung:** Es ist nicht vorgesehen, daß alle Parameter gleichzeitig ihren Höchstwert haben; denn z. B.  $720 * 576 / (16 * 16) = 1.620$ , dies ist größer als 396.

## Vier Arten von Bildern:

- I-Bild** (= Intracoded Picture)
- P-Bild** (= Predictive Picture)
- B-Bild** (= Bidirectional Picture)
- D-Bild** (= DC coded Picture)

## Bemerkungen:

- (i) Verschiedene Teile eines P-Bildes können sich auf verschiedene vorhergehende Bilder beziehen. Auch können sich verschiedene Teile eines B-Bildes auf verschiedene Referenzbilder beziehen.
- (ii) Man unterscheidet zwischen offenen und geschlossenen Bildgruppen.
- (iii) Man unterscheidet zwischen der Codierfolge und der Abspielfolge.

## Beispiel einer Bildgruppe:

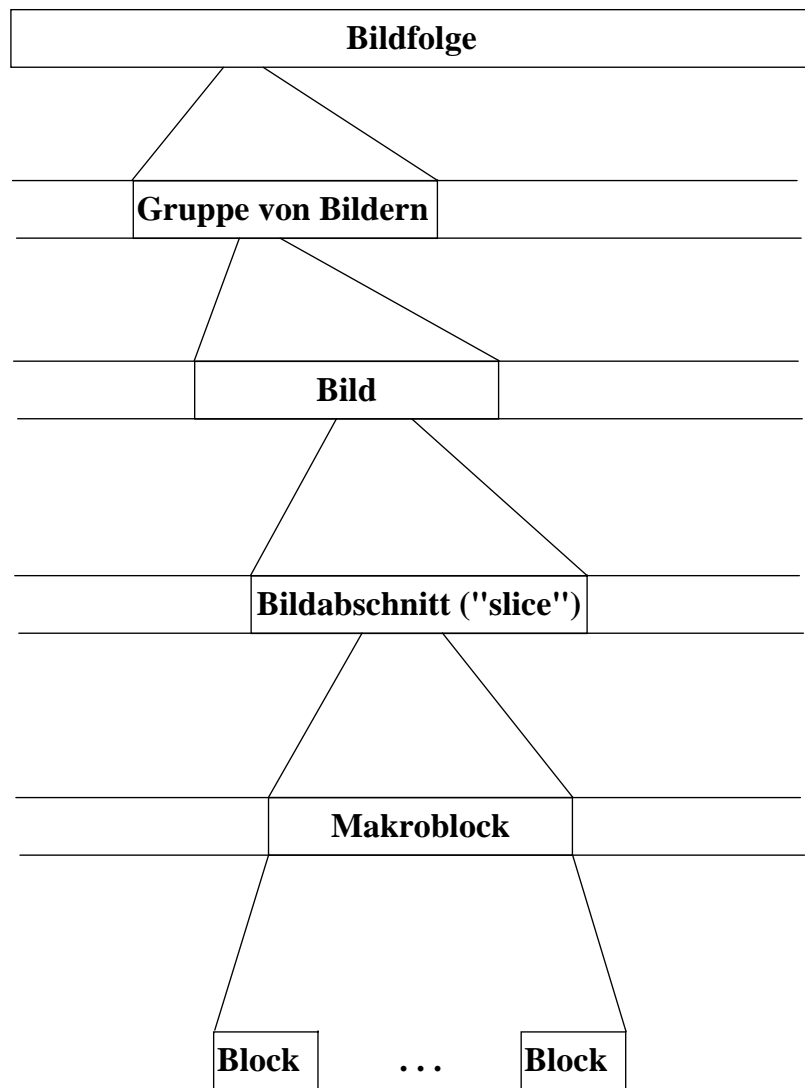
Reihenfolge bei der Wiedergabe:

$I_1 \ B_2 \ B_3 \ B_4 \ P_5 \ B_6 \ B_7 \ B_8 \ P_9$

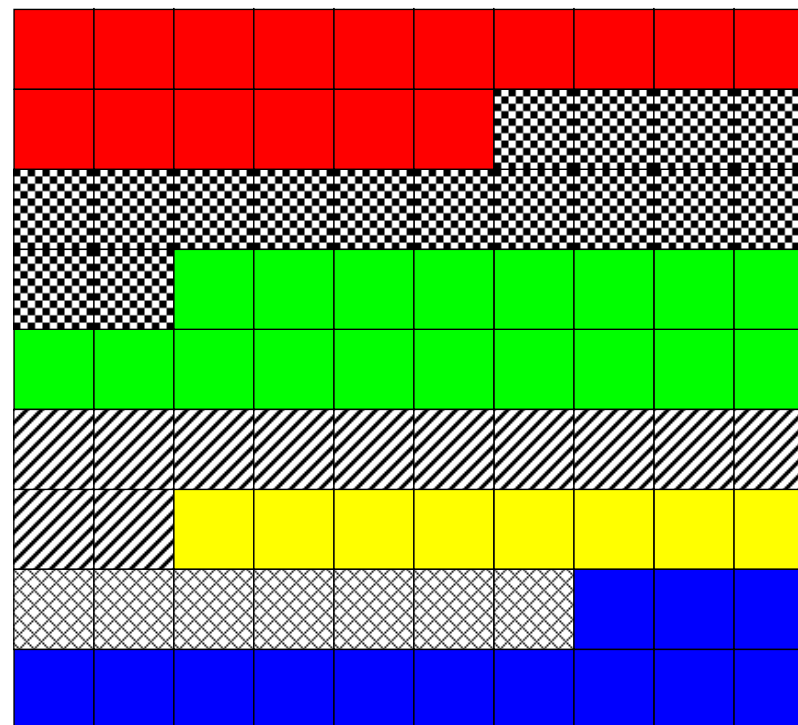
Reihenfolge beim Übertragen:

$I_1 \ P_5 \ B_2 \ B_3 \ B_4 \ P_9 \ B_6 \ B_7 \ B_8$

## MPEG-1 Hierarchie:



## Bildabschnitte:



## Bemerkungen:

- (i) Ein Bildabschnitt besteht aus einer beliebigen Zahl von Makroblöcken.
- (ii) Die Makroblöcke in einem Bildabschnitt sind unabhängig von den Makroblöcken in anderen Bildabschnitten codiert. Daher dienen Bildabschnitte als Aufsetzpunkte, falls Bildteile bei der Übertragung beschädigt wurden.

## 6-Stufen-Hierarchie:

### Bildfolge

**Gruppe von Bildern**  
(Einheit für den direkten Zugriff)

**Bild**  
(Codiereinheit)

**Bildabschnitt**  
(Einheit zur Resynchronisation)

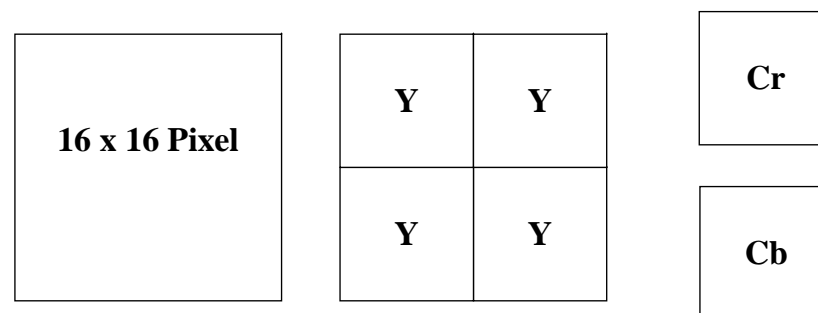
**Makroblock**  
(Einheit bei Differenzbildung)

**Block**  
(Einheit für Cosinustransformation)

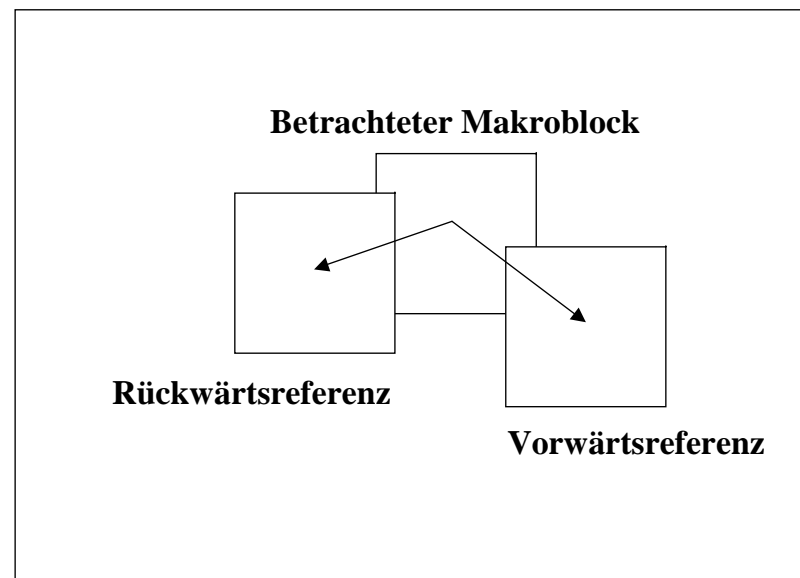
### Source Input Format:

	SIF-525	SIF-625
<b>Luminanz (Y):</b>	<b>360 x 242</b>	<b>360 x 288</b>
<b>Significant Pel Area:</b>	<b>352 x 240</b>	<b>352 x 288</b>
<b>Chrominanz (Cr, Cb):</b>	<b>180 x 121</b>	<b>180 x 144</b>
<b>Significant Pel Area:</b>	<b>176 x 120</b>	<b>176 x 144</b>
<b>Bildfrequenz:</b>	<b>29,97</b>	<b>25</b>

## Struktur eines Makroblocks



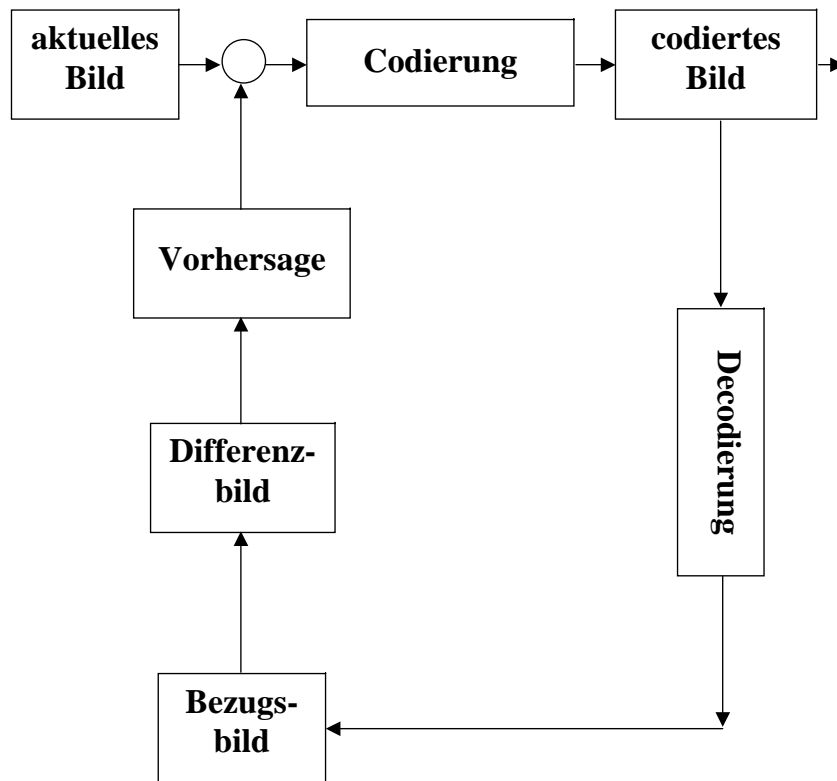
## B-Bild



## Bestimmung von Bewegungsvektoren



## Differenzcodierung:



**Bemerkung:** Statt nur eines Bezugsbildes kann man mehrere Bezugsbilder haben. Auch muß das Bezugsbild nicht unbedingt das letzte Bild sein.

## Gütemaße:

Als Abstandsmaß für zwei  $N \times N$ -Blöcke aus den Bildern **A** und **B** eignet sich der mittlere quadratische Fehler MSE.

$$MSE = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (A_{ij} - B_{ij})^2$$

**Beispiel:**

Gegeben seien ein 5x5-Bild und ein 3x3-Block:

Bild =

2	3	2	4	5
6	4	2	3	2
5	5	3	3	3
4	4	2	4	3
4	6	4	5	5

Block =

3	3	2
2	4	2
4	5	3

Man erhält die folgenden MSE-Werte:

(0, 0): 2  
(0, 1): 2,111  
(0, 2): 1,889  
(1, 0): 2,556  
(1, 1): 2,667  
(1, 2): 1  
(2, 0): 1,667  
(2, 1): 2,889  
(2, 2): 0,667

Man nutzt auch andere Abstandsmaße, z. B. die Summe der absoluten Differenzen SAD:

$$\text{SAD} = \sum_{i=1}^N \sum_{j=1}^N |A_{ij} - B_{ij}|$$

Für obiges Beispiel erhält man diese Werte:

(0, 0): 6  
(0, 1): 11  
(0, 2): 9  
(1, 0): 11  
(1, 1): 12  
(1, 2): 7  
(2, 0): 9  
(2, 1): 14  
(2, 2): 4

Das Minimum liegt auch hier beim Indexpaar 2, 2.

## MPEG-2:

MPEG-2 ist das Kompressionsverfahren für das digitale Fernsehen bis zum HDTV (1920 x 1080 x 30) und die Videoaufzeichnung für die DVD. Es beherrscht auch den "surround sound" (5 Kanäle).

Änderungen gegenüber MPEG-1:

Farbabtastung: 4:2:0, 4:2:2, 4:4:4.

Bildabschnitte: Ein Bildabschnitt darf nur Makroblöcke einer Makroblockzeile enthalten.

Quantelung: MPEG-1 kennt nur die Genauigkeitsstufe 8 Bit für DC-Koeffizienten; MPEG-2 ermöglicht auch Genauigkeiten von 9 und 10 Bit.

Ersatzbewegungsvektoren: Falls ein intracodierter Makroblock verloren geht, verfügt man eventuell über einen Ersatzblock.

MPEG-2 beherrscht das Zeilensprungverfahren und die zugehörige Halbbildcodierung.

Anpassung des Bildseitenformats bei der Darstellung, z. B. von 16:9 nach 4:3.

## Beispiele zu Chromatizitätskoordinaten für Primärfarben:

### ITU-R BT.709

	x	y
Grün	0,300	0,600
Blau	0,150	0,060
Rot	0,640	0,330
Weißpunkt	0,3127	0,3290

### ITU-R BT.470-2 System B, G

	x	y
Grün	0,29	0,60
Blau	0,15	0,06
Rot	0,64	0,33
Weißpunkt	0,313	0,329

### SMPTE 240M (1987)

	x	y
Grün	0,310	0,595
Blau	0,155	0,070
Rot	0,630	0,340
Weißpunkt	0,3127	0,3290

## Farben nach ITU-R BT.709

$$E_Y = 0,7154 E_G + 0,0721 E_B + 0,2125 E_R$$

$$E_{PB} = -0,386 E_G + 0,500 E_B - 0,115 E_R$$

$$E_{PR} = -0,454 E_G - 0,046 E_B + 0,500 E_R$$

$$(0 \leq E_G, E_B, E_R \leq 1)$$

## Farben nach ITU-R BT.470-2 System B, G

$$E_Y = 0,587 E_G + 0,114 E_B + 0,299 E_R$$

$$E_{PB} = -0,331 E_G + 0,500 E_B - 0,169 E_R$$

$$E_{PR} = -0,419 E_G - 0,081 E_B + 0,500 E_R$$

## Farben nach SMPTE 240M (1987)

$$E_Y = 0,701 E_G + 0,087 E_B + 0,212 E_R$$

$$E_{PB} = -0,384 E_G + 0,500 E_B - 0,116 E_R$$

$$E_{PR} = -0,445 E_G - 0,055 E_B + 0,500 E_R$$

## Farbmodell:

$$Y = (219 * E_Y) + 16$$

$$Cb = (224 * E_{PB}) + 128$$

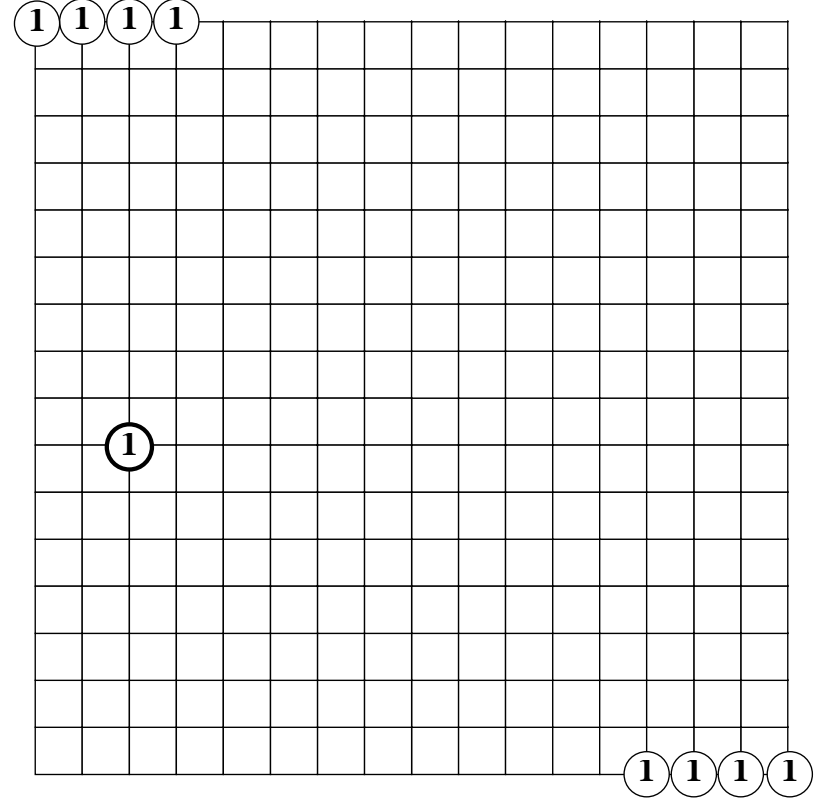
$$CR = (224 * E_{PR}) + 128$$

**Profiles and Levels in MPEG-2, obere Grenzwerte der Abtastdichte:**

Level	Simple	Profile				
		Main	SNR	Spatial	High	
HIGH		1920H	1920H	1440H	1920H	
		1152V	1152V		1152V	
		60 Hz	60 Hz		60 Hz	
HIGH 1440		1440H	1440H	1440H	1440H	
		1152V	1152V	1152V	1152V	
		60 Hz	60 Hz	60 Hz	60 Hz	
MAIN	720H 576V 30 Hz	720H	720H	720H	720H	
		576V	576V		576V	576V
		30 Hz	30 Hz		30 Hz	30 Hz
LOW		352H	352H			
		288V	288V			
		30 Hz	30 Hz			

**Suchstrategie:**

**Vollsuche:**



**Nachteil der Vollsuche ist der hohe Aufwand, daher lohnt es sich, Schnellsuchverfahren einzusetzen.**



## Ausblick auf MPEG-4:

### Anforderungen an MPEG-4:

eigenständige audiovisuelle Objekte,  
inhaltsbasierter Datenzugriff,  
flexible Interaktivität.

Der Szenenaufbau in MPEG-4 ist objektorientiert, so lässt sich z. B. eine Szene komponieren aus einem Standbild als Hintergrundobjekt, mehreren handelnde Personen ohne Hintergrund als Bewegtbildobjekte, zugeordnete Stimmen als Audioobjekte.

Zusammenführen einzelner audiovisueller Objekte zwecks Komposition zu Szenenfolgen mit unabhängigen Plazierungen und Transformationen der Teilobjekte, Animationen mit frei wählbaren Betrachtungsstandpunkten werden unterstützt.

In MPEG-4 wurde der VRML (Virtual Reality Modeling Language) Standard übernommen. Bei VRML handelt es sich um eine Sprache zur Beschreibung von 3D-Welten, die durch einfache geometrische Objekte, wie Kugel, Würfel, u.s.w. zusammengestellt werden. In VRML ist ein Ereignismodell integriert, so lassen sich z. B. Rotationen von Objekten über Ereignisse steuern.

### Objekte in MPEG-4:

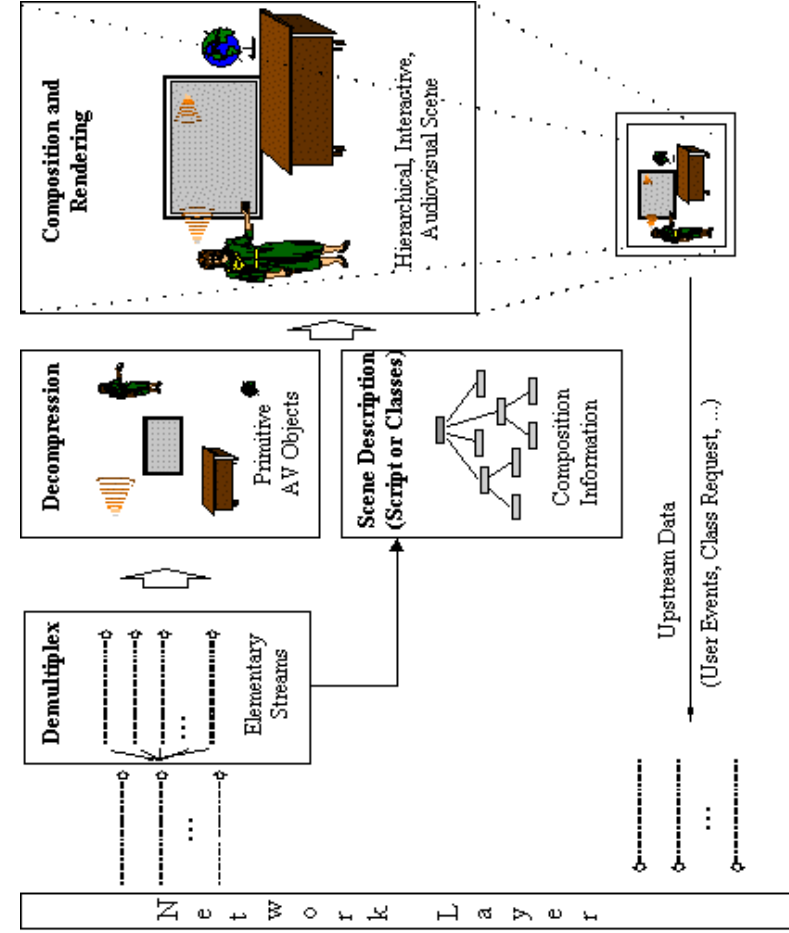


Bild aus Rob Koenen:  
Overview of the MPEG-4  
Standard

Zusammenführen von Objekten:

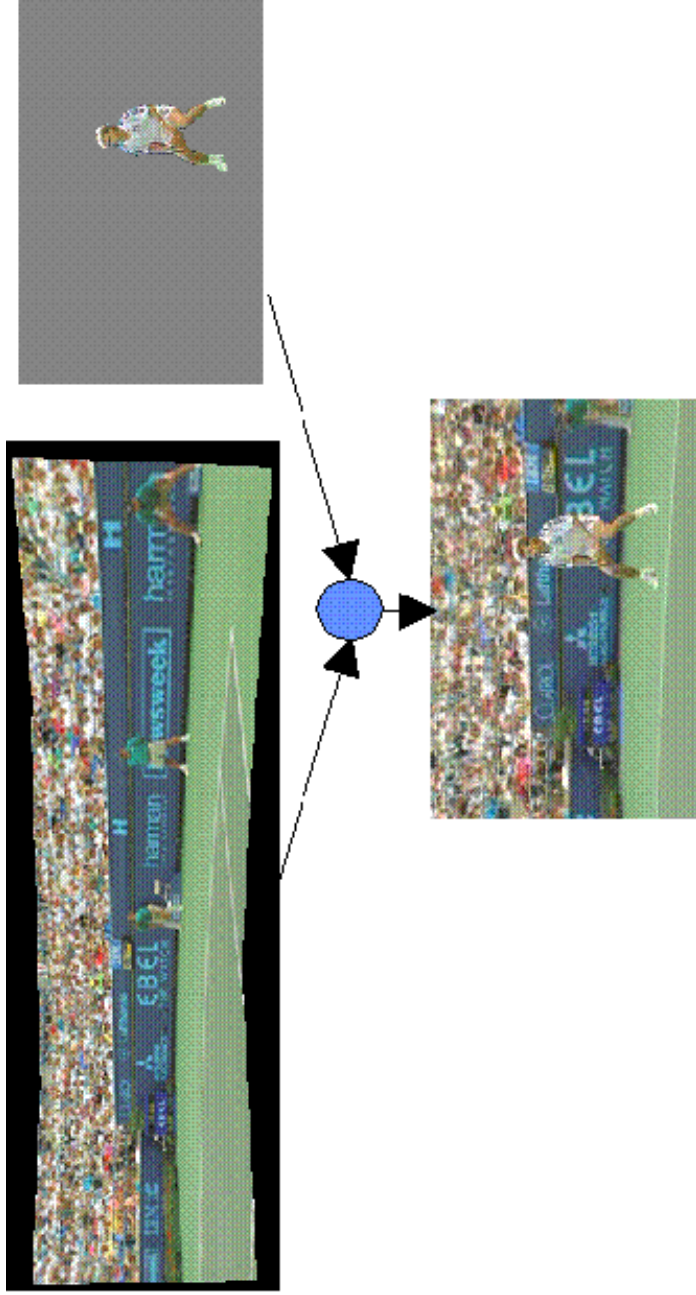


Bild aus Rob Koenen: Overview of the MPEG-4 Standard

### Bemerkungen zu MPEG-4, Teil 10 (H.264/AVC):

- (i) AVC wurde in den Jahren 2001 bis 2004 vom Joint Video Team, einem Zusammenschluß der Moving Picture Experts Group und der Video Coding Experts Group, entwickelt.
- (ii) Gegenüber den vorhergehenden Videostandards MPEG-2, MPEG-4 (2), H.261 und H.263 ist die Grundstruktur von AVC nicht verändert; es gibt aber in vielen Details erhebliche Fortschritte. Die Kompressionseffizienz gegenüber MPEG-2 wird etwa mit dem Faktor 2 veranschlagt.
- (iii) Es gelang die Diskrete Cosinus-Transformation durch Multiplikation mit ganzzahligen Matrizen hinreichend genau zu approximieren, hiermit ist ein Übergang von der Gleitkommarechnung zur Ganzzahlrechnung mit 16-Bit Zahlen gelungen.
- (iv) Durch Unterteilung der 16x16-Makroblöcke bis zu 4x4-Blöcken gelang eine bessere Anpassung an die Bildgegebenheiten.
- (v) Es wurden neue Profile High 10 mit Farbtiefen bis zu 12 Bit eingeführt.
- (vi) Es wurden adaptive Glättungsoperationen eingeführt, deren Wirkung auf den folgenden beiden Folien demonstriert wird. Die Beispiele stammen von Iain Richardson, [www.vcodex.com](http://www.vcodex.com).

**Effekt der Glättung der Blockübergänge (1):**

**Originalbild:**



**Bild ohne Glättung:**

**Effekt der Glättung der Blockübergänge (2):**

**Originalbild:**



**Bild mit Glättung:**