



Grundlagen der Videocodierung

Dr. Detlev Marpe
Fraunhofer Institut
für Nachrichtentechnik HHI



Fachhochschule für Technik
und Wirtschaft Berlin
University of Applied Sciences



Fraunhofer
Institut
Nachrichtentechnik
Heinrich-Hertz-Institut

Beispiel: PAL (4:2:0) über DVB-T

Quelldatenrate = $720 \cdot 576 \cdot 3/2 = 622.080$ Bytes/frame
 25 Hz Vollbild- bzw. 50 Hz Halbbild-Wiederholrate:
 $\Rightarrow 25 \cdot 622.080$ Bytes/s = $15.552.000$ Bytes/s
 $= 124.416.000$ bits/s \approx **124 Mbit/s**

Mittlere Videodatenrate für DVB-T: **2-5 Mbit/s**
 \Rightarrow Erforderliche **Kompressionsfaktoren** \approx **25 : 1 bis 62 : 1**
 d.h. 5-12 mal höherer Kompressionsfaktor gegenüber DV

Notwendig:

- Abstriche bei der Rekonstruktionsqualität im Vergleich zu DV
- Erhöhung der Codiereffizienz durch **Ausnutzung zeitlicher Redundanzen**

2 © Detlev Marpe, Kai Uwe Barthel

Ausnutzung der zeitlichen Redundanz


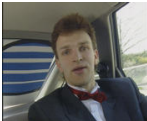



-

=


Bild zum Zeitpunkt $t-1$ Bild zum Zeitpunkt t kontrastverstärktes Differenzbild

- Zeitlich aufeinander folgende Bilder weisen große Ähnlichkeiten auf
- Redundante Information soll nicht mehrfach codiert werden
- Redundanzreduktion durch **zeitliche Prädiktion** von Bildpunkten

3 © Detlev Marpe, Kai Uwe Barthel

Zeitliche Prädiktion: Conditional Replenishment



Decodiertes Bild zum Zeitpunkt $t-1$




Bild zum Zeitpunkt t (zu codierendes Bild)

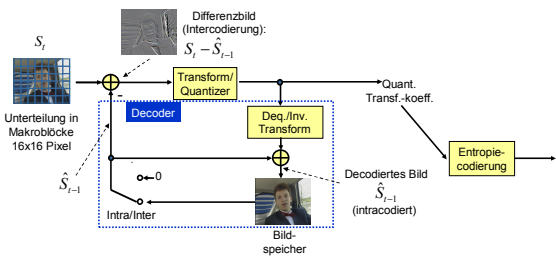
Blaue Blöcke: zeitlich prädiert, d.h. keine Übertragung von Bildinformation von $t-1 \rightarrow t$
 Rote Blöcke: Bildinformation in Bild t wird (verlustbehaftet) neu übertragen

Annahme: viele Bereiche des Bildes bleiben zeitlich konstant

- Einteilung des Bildes in Blöcke (16 x 16 Pixel)
- Prädiktion erfolgt blockweise auf Basis decodierter Bilder ("inter-codiert")
 - Ersetze decodierten Block aus Bild t mit örtlich korrespondierendem Block aus Bild $t-1$ (alle blauen Blöcke)
- Wenn Energie des Prädiktionsfehlers zu groß (rote Blöcke), übertrage die Bildinformation eines kompletten Blocks neu ("intra-codiert"), z.B., mit örtlicher Prädiktions- oder Transformationscodierung (à JPEG)

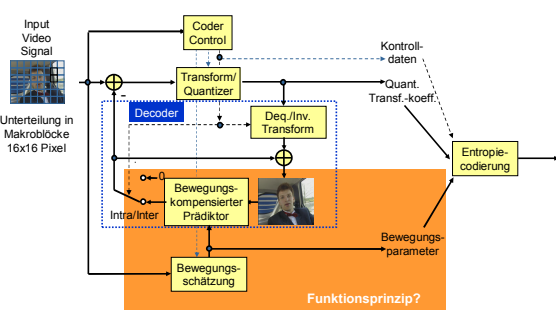
4 © Detlev Marpe, Kai Uwe Barthel

Zeitliche Prädiktion: Differenzbildcodierung



5 © Detlev Marpe, Kai Uwe Barthel

Zeitliche Prädiktion: Bewegungskompensation



6 © Detlev Marpe, Kai Uwe Barthel

Bewegungskompensation mit verschobenen Blöcken

Verschiebungsvektor = Bewegungsvektor

Suchbereich

Schrittweite

Vorgängerbild

Referenzbereich

aktuelles Bild

7 © Detlev Marpe, Kai Uwe Barthel

Bewegungsschätzung

- Blockverschiebung wird durch einen Bewegungsvektor $\mathbf{v} = (\Delta m, \Delta n)$ pro Block charakterisiert, der als Nebeninformation codiert wird
- Problem:** Bestimmung der "Bewegung" mittels *Bewegungsschätzung*
- Einfachste Realisierung: Blockmatching**
 - Berechne den Prädiktionsfehler unter Verwendung von verschiedenen Kandidaten-Vektoren \mathbf{v}_i innerhalb eines Suchbereichs
 - Finde den besten "Match", d.h. wähle den Vektor $\mathbf{v}_i = (\Delta m_i, \Delta n_i)$ mit kleinstem mittleren quadr. Fehler MSE:

$$MSE_{\Delta m_i, \Delta n_i} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (S_i(m, n) - S_{i-1}(m + \Delta m_i, n + \Delta n_i))^2$$

- Rechenaufwand ist proportional zur Anzahl der Suchpositionen
- Es gibt Suchstrategien zur schnellen Bewegungsschätzung

8 © Detlev Marpe, Kai Uwe Barthel

Prädiktionsfehler

aktuelles Bild A - Prädiktionsbild B mit decodiertem letztem Bild = Prädiktionsfehler A - B (Differenzbild)

aktuelles Bild A - Prädiktionsbild C mit verschobenen Blöcken aus letztem decodierten Bild = Prädiktionsfehler A - C (bewegungskompensiert)

9 © Detlev Marpe, Kai Uwe Barthel

Hybride Videocodierung: Basis aller modernen Standards

Input Video Signal

Coder Control

Transform/Quantizer

Decoder

Deq./Inv. Transform

Bewegungs-kompensierter Prädiktor

Bewegungsschätzung

Entropie-codierung

Kontroll-daten

Quant. Transf.-koeff.

Bewegungs-parameter

Unterteilung in Makroblöcke 16x16 Pixel

Hybride Codierung = Bewegungskompensierte Prädiktion + Transformationscodierung

10 © Detlev Marpe, Kai Uwe Barthel

Beispiel: ITU-T H.261

- ITU-T H.261 (ex-CCITT): der erste erfolgreiche, weit verbreitete Videocodierstandard [Version 1: 1990]
- Anwendungen: Videotelefonie über ISDN mit $p \times 64$ kbit/s
- Formate: CIF und QCIF, jeweils mit 4:2:0 Sampling
- Bildfolgefrequenzen zwischen 8 1/3 und 30 Hz
- Bitraten: 64 - 2048 kbit/sec
- Codierelemente:
 - 16x16 Makroblock-basierte Bewegungskompensation
 - 8x8 DCT
 - Skalare Quantisierung
 - Zig-Zag Scan
 - Laufängencodierung (RLE) und Codes variabler Länge (VLC)

11 © Detlev Marpe, Kai Uwe Barthel

H.261 Intracodierung

Makroblock (MB): 16x16 Y, 8x8 Cb, Cr

DCT

Quant

Zig-zag

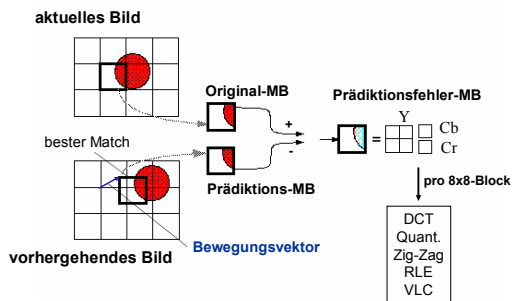
RLE

Huffman

01101...

12 © Detlev Marpe, Kai Uwe Barthel

H.261 Interkodierung



13 © Detlev Marpe, Kai Uwe Barthel

H.261: Zusammenfassung

Erster standardisierter hybrider Videocoder:

- ❑ Bewegungskompensierte Prädiktion für Makroblöcke der Größe 16x16, ganzzahlige Verschiebungen auf Pixelraster
- ❑ Wahlweise Intraframe- oder Interframe-Codierung für jeden Makroblock
- ❑ 8x8 DCT
- ❑ Lineare, ungewichtete Quantisierung der Koeffizienten
- ❑ Prädiktive Codierung der Bewegungsvektorkomponenten und der Gleichanteil (DC) Koeffizienten
- ❑ Lauflängencodierung der AC-Koeffizienten mit Zigzag-Scan innerhalb jedes DCT-Blocks der Größe 8x8 Samples

14 © Detlev Marpe, Kai Uwe Barthel