

# 8. Bewegtbilder

8.1 Bewegungswahrnehmung

8.2 Videokompression

insbesondere MPEG-1 und MPEG-2

8.3 Videodatenformate



Literatur:

Arne Heyna/Marc Briede/Ulrich Schmidt: Datenformate im Medienbereich, Fachbuchverlag Leipzig 2003

John Watkinson: The MPEG Handbook, Focal Press 2001

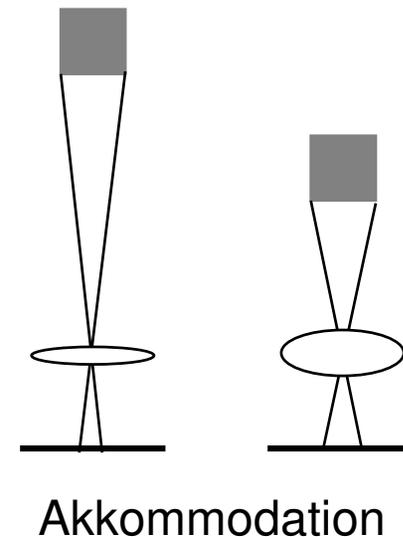
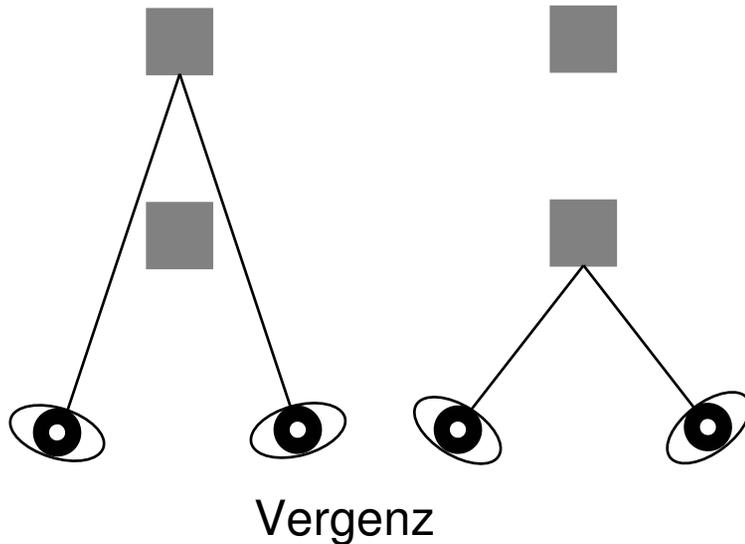
Iain E.G. Richardson: H.264 and MPEG-4 Video Compression, Wiley 2003

# Bewegungswahrnehmung (1)

- Psychologische Faktoren:
  - Bewegungswahrnehmung ist eine komplexe Berechnungs- und Bewertungsleistung des Gehirns
  - Physikalisch „falsche“ Wahrnehmung durch Unterdrückung von Wahrnehmungen im Gehirn möglich
    - » Beispiel: Von bewegtem Objekt herunterfallendes Objekt
- Bewegungseindruck durch Betrachten von Bildfolgen
  - Grundprinzip bereits mit einfachen mechanischen Geräten nutzbar
  - Lumière 1895: Cinematograph
  - Maß: Bilder/Sekunde (*frames per second, fps*)
  - Physiologische Grenze: 50 – 60 Bilder/Sekunde (Hz)
    - » z.B. bei 100 Bildern/Sekunde keine Zwischenstufen mehr durch das Auge auflösbar
  - Psychologische Grenze: 25 – 30 Bilder/Sekunde (Hz)
    - » unter 50 Bildern/Sekunde aber sehr anstrengend
  - Koordination Bewegung-Wahrnehmung benötigt mindestens 5 Hz

# Bewegungswahrnehmung (2)

- Physiologische Faktoren:
  - Gegenseitige Beeinflussung benachbarter Lichtsinneszellen auf der Netzhaut (Verschaltung)
  - Nachführung der Augen, um ein bewegtes Objekt auf die Fovea (Gelber Fleck, Stelle der besten Sehleistung) zu fokussieren (foveale Objektverfolgung)
  - Anpassung an veränderliche Entfernung des Objekts
    - » Vergenz
    - » Akkommodation



# Bewegungssimulation durch Bildfolgen

- Natürlichkeit des Bewegungseindrucks („Immersion“) ist bei klassischen Ausgabegeräten begrenzt:
  - Fehlende Beschleunigungswahrnehmung
    - » Übelkeit wegen inkonsistenter Signal-Information
  - Fehlende Akkommodation und Vergenz
    - » Bewegungen in Richtung zum Betrachter und vom Betrachter weg
    - » Alle Objekte in gleicher Entfernung dargestellt/aufgenommen
- Stereoskopische Ausgabegeräte:
  - Kopplung von Akkommodation und Vergenz nicht ausreichend
  - Belastende Präsentationstechnik (z.B. Shutterbrillen)

# 8. Bewegtbilder

- 8.1 Bewegungswahrnehmung
- 8.2 Videokompression  
insbesondere MPEG-1 und MPEG-2
- 8.3 Videodatenformate



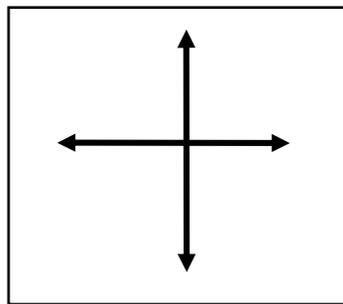
Literatur:

Arne Heyna/Marc Briede/Ulrich Schmidt: Datenformate im Medienbereich,  
Fachbuchverlag Leipzig 2003

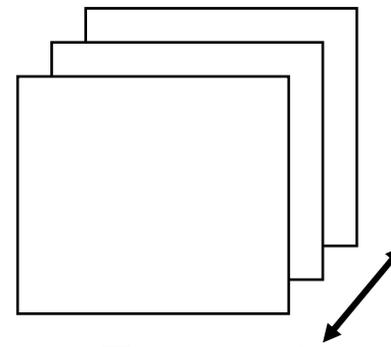
John Watkinson: The MPEG Handbook, Focal Press 2001

# Ansatzpunkte zur Video-Kompression

- Videodaten haben vier Dimensionen:
  - Zwei Bilddimensionen
  - Eigenschaften der Pixel (Helligkeit, Farbe)
  - Zeitachse
- Kompressionsansätze:
  - *Spatial* oder *intra-coding*: Redundanz aus einem Bild entfernen
    - » DCT, DWT, Vektorquantisierung, Konturbasierte Kodierung
  - *Temporal* oder *inter-coding*: Redundanz zwischen Bildern entfernen
    - » Differenzcodierung, Bewegungskompensation



Spatial



Temporal

# Konzept: Vektorquantisierung

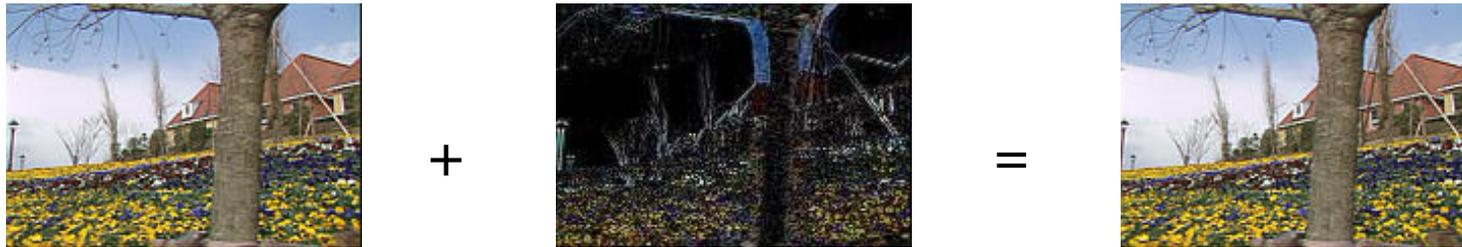
- Idee:
  - Bild aufteilen in Blöcke, z.B. 4 x 4 Pixel
  - Suche nach Ähnlichkeiten zwischen den Blöcken
  - Ähnliche Blöcke durch einen „Durchschnittsblock“ ersetzen
  - Palette für Bildblöcke, d.h. Kodierung durch Index
- Verwendung in Codecs:
  - Indeo, Cinepak
- Langsame Codierung (Spezial-Hardware)
- Schnelle Decodierung
- In Kompression und Bildqualität nicht besser als DCT und DWT

*Codec = Coder/Decoder*

# Konzept: Konturbasierte Kodierung

- Idee:
  - Bild trennen in *Konturen* und *Texturen*
  - Konturen z.B. durch Beziér-Kurven beschreiben
  - Texturen z.B. nach DCT kodieren
- Verwendung:
  - Ansatzweise in MPEG-4
- Vermeidet Darstellungsprobleme an Kanten
- Problem: Finden der Konturen in gegebenem Bild
  - Forschungsthema

# Konzept: Differenzkodierung (*frame differencing*)

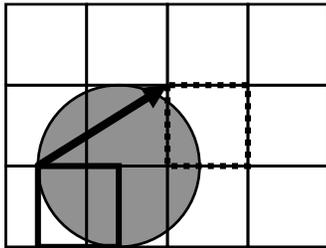


- In den meisten Fällen unterscheiden sich aufeinander folgende Bilder nur in Details
- Idee:
  - Startbild (und regelmäßig weitere *key frames*) intracodiert übertragen
  - Differenz zum nächsten Bild als Bild auffassen und komprimieren
    - » Z.B. mit DCT und anschließender Entropiecodierung
    - » Viele niedrige Werte, also hoher Kompressionsfaktor möglich

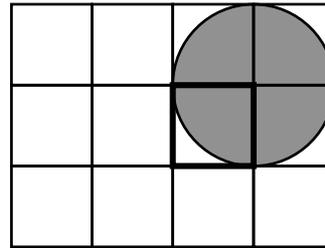
# Konzept: Bewegungskompensation (*motion compensation*)

- Idee:
  - Bewegungen von Objekten zwischen aufeinander folgenden Bildern identifizieren
  - Für Teilbilder übertragen:
    - » Differenzbild
    - » Verschiebungsvektor
- Verwendung u.a.:
  - MPEG-1, -2 und -4, H.261-H.264
- Problem: Algorithmen zur Bewegungsabschätzung (*motion estimation*)
  - *block matching*
  - *gradient matching*
  - *phase correlation*

# Block Matching



Referenzframe  $N$

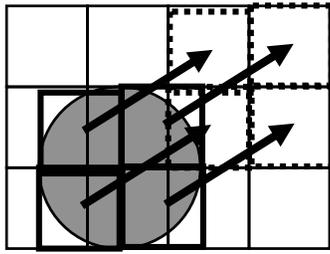


Zielframe  $N+1$

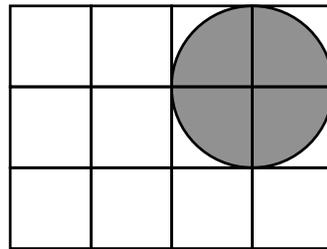
- Referenzframe und Zielframe (aktueller Frame)
  - Referenzframe = vorheriges Bild
- Einteilung des Bildes in Blöcke
- Für jeden Block des Zielframes:
  - Suche nach „best match“ im Referenzframe
    - » z.B. mittlere quadratische Abweichung oder mittlere Differenz
  - Speichern des Verschiebungsvektors
- Algorithmusbeschleunigung:
  - Hierarchische Suche zunächst auf vergrößertem Bild

# Differenzbilder

- Auch bei relativ schlechtem Block Matching werden die Differenzbilder (*residual error pictures*) relativ einfach und damit klein.



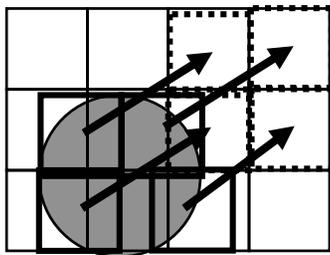
Referenzframe  $N$



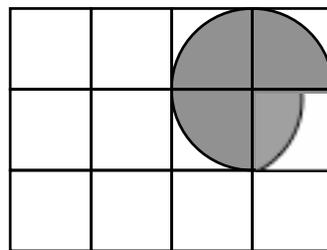
Zielframe  $N+1$



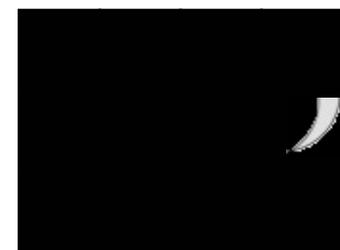
Differenzframe



Referenzframe  $N$



Zielframe  $N+1$



Differenzframe

# 8. Bewegtbilder

8.1 Bewegungswahrnehmung

8.2 Videokompression

insbesondere MPEG-1 und MPEG-2



8.3 Videodatenformate

Literatur:

Arne Heyna/Marc Briede/Ulrich Schmidt: Datenformate im Medienbereich,  
Fachbuchverlag Leipzig 2003

John Watkinson: The MPEG Handbook, Focal Press 2001

# MPEG: Übersicht

- MPEG = Moving Pictures Experts Group
  - Expertengruppe bei der ISO, Standards für Bewegtbild-Kompression
  - Benutzt konsequent JPEG-Standards
  - Ansatz: Nur Decodierung spezifiziert, viele Encoder möglich
- MPEG-1 (ISO 11172, 1992)
  - Video und Audio mit der Datenrate einer Audio-CD (1,8 Mbit/s, davon 1,25 Mbit/s Video + zwei Audio-Kanäle) (--> Video-CD)
  - Auflösung: CIF (bei PAL 352 x 288)
- MPEG-2 (ISO-13818 und ITU Rec. H.262, 1993)
  - Hohe Bandbreite zwischen 2 und 80 Mbit/s, skalierbare Qualität
  - Bis zu 5 Audio-Kanäle
- MPEG-4 (ISO 14496, 2000)
  - Unregelmäßig geformte Objekte, Animationen, Interaktion
- Weitere MPEG-Standards abgeschlossen und in Vorbereitung:
  - MPEG-7, MPEG-21 (sind aber KEINE Kompressionsverfahren)

# MPEG-2: Profiles und Levels

		Profiles					
		Simple	Main	4:2:2	SNR	Spatial	High
Levels	High (HDTV 16:9)		4:2:0 1920 x 1152 90 Mb/s				4:2:0/2 1920 x 1152 100 Mb/s
	High 1440 (HDTV 4:3)		4:2:0 1440 x 1152 60 Mb/s			4:2:0 1440 x 1152 60 Mb/s	4:2:0/2 1440 x 1152 80 Mb/s
	Main	4:2:0 720 x 576 15 Mb/s	4:2:0 720 x 576 15 Mb/s	4:2:2 720 x 608 50 Mb/s	4:2:0 720 x 576 15 Mb/s		4:2:0/w 720 x 576 20 Mb/s
	Low		4:2:0 352 x 288 4 Mb/s		4:2:0 352 x 288 4 Mb/s		

SNR = Signal Noise Ratio, HDTV = High Definition TV

Schreibweise z.B.: MP@ML

# Frametypen in MPEG

- Intraframes (I-Frames)
  - Vollständige Bilddaten, nur innerhalb des Bildes komprimiert (ca. 92 kB, 7:1)
  - Ca. jedes 15. Frame ist ein I-Frame
- Predicted Frames (P-Frames)
  - Bewegungskompensation und Differenzbildung (ca. 32 kB, 20:1)
  - Typischerweise 3 P-Frames zwischen zwei I-Frames
- Bidirectionally Predicted Frames (B-Frames)
  - Bewegungskompensation unter Berücksichtigung von nachfolgendem und vorausgehendem I- oder P-Frame (ca. 13 kB, 50:1)
  - Typischerweise 2-3 B-Frames zwischen zwei P-Frames

Darstellungsreihenfolge:

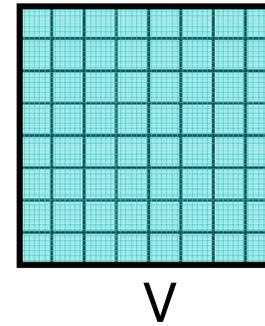
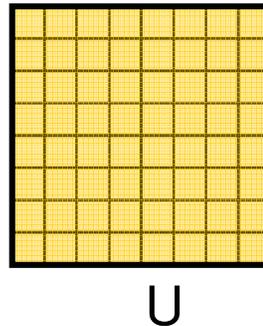
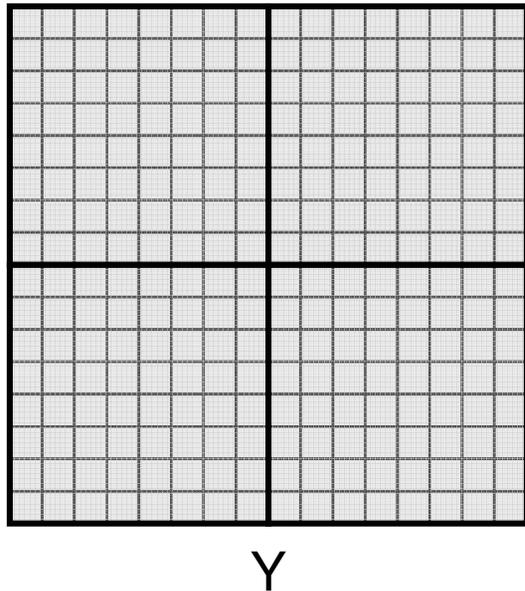


Group of Pictures (GOP)

Übertragungsreihenfolge:

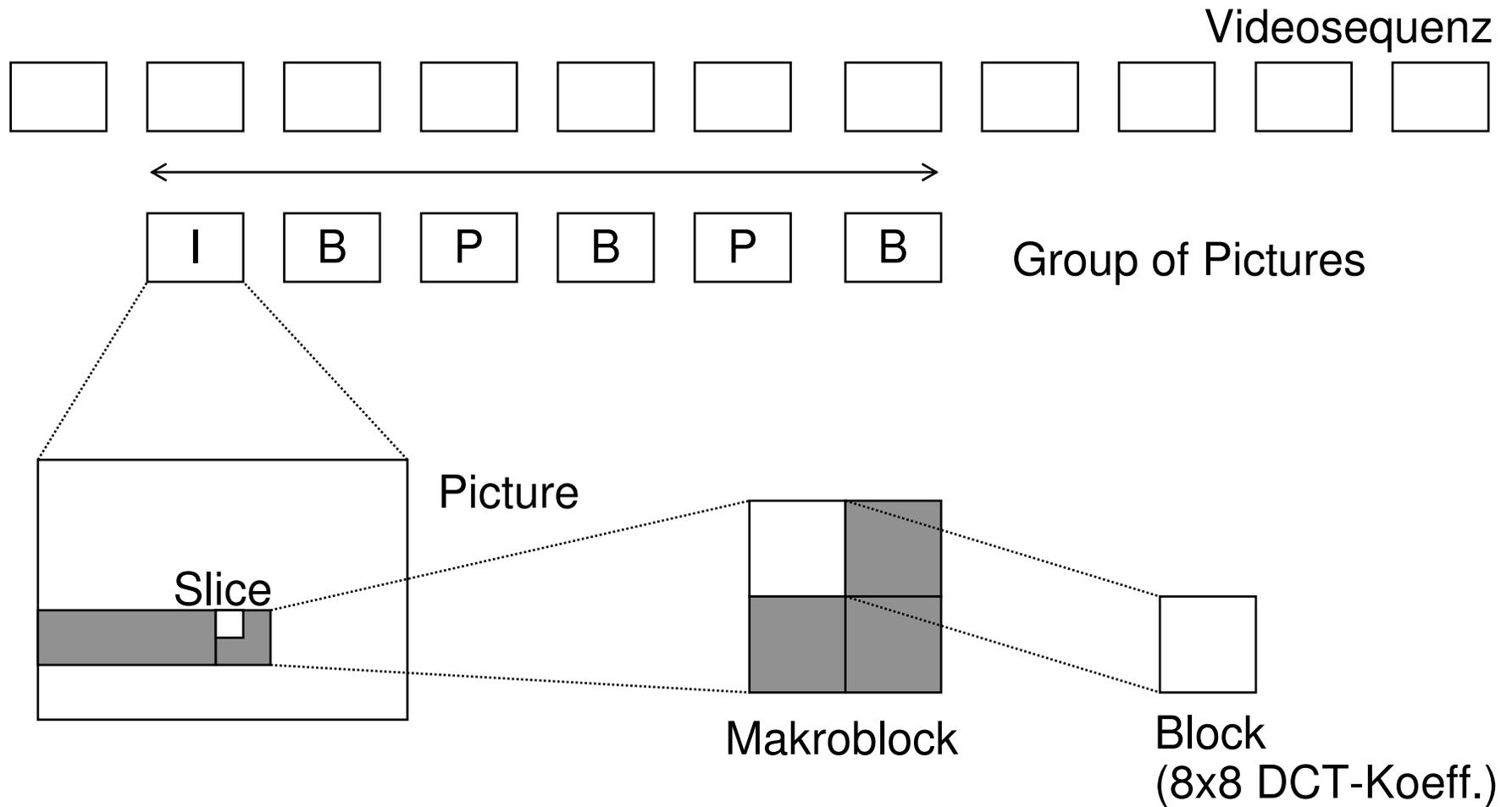


# Makroblöcke in MPEG-1/2

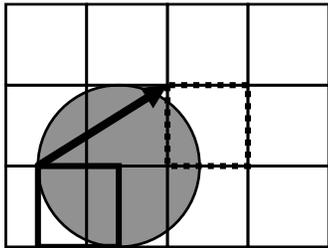


- Makro-Blockeinteilung so gewählt, dass
  - Vielfache von 8x8-Blöcken
  - Kompatibel mit Chroma-Subsampling
- Typisch: 16x16-Pixel Makroblöcke

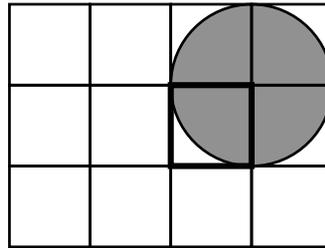
# Struktur des MPEG-2 Videodatenstroms



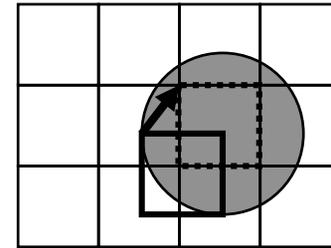
# Bidirektionale Bewegungskompensation



Referenzframe  $N$



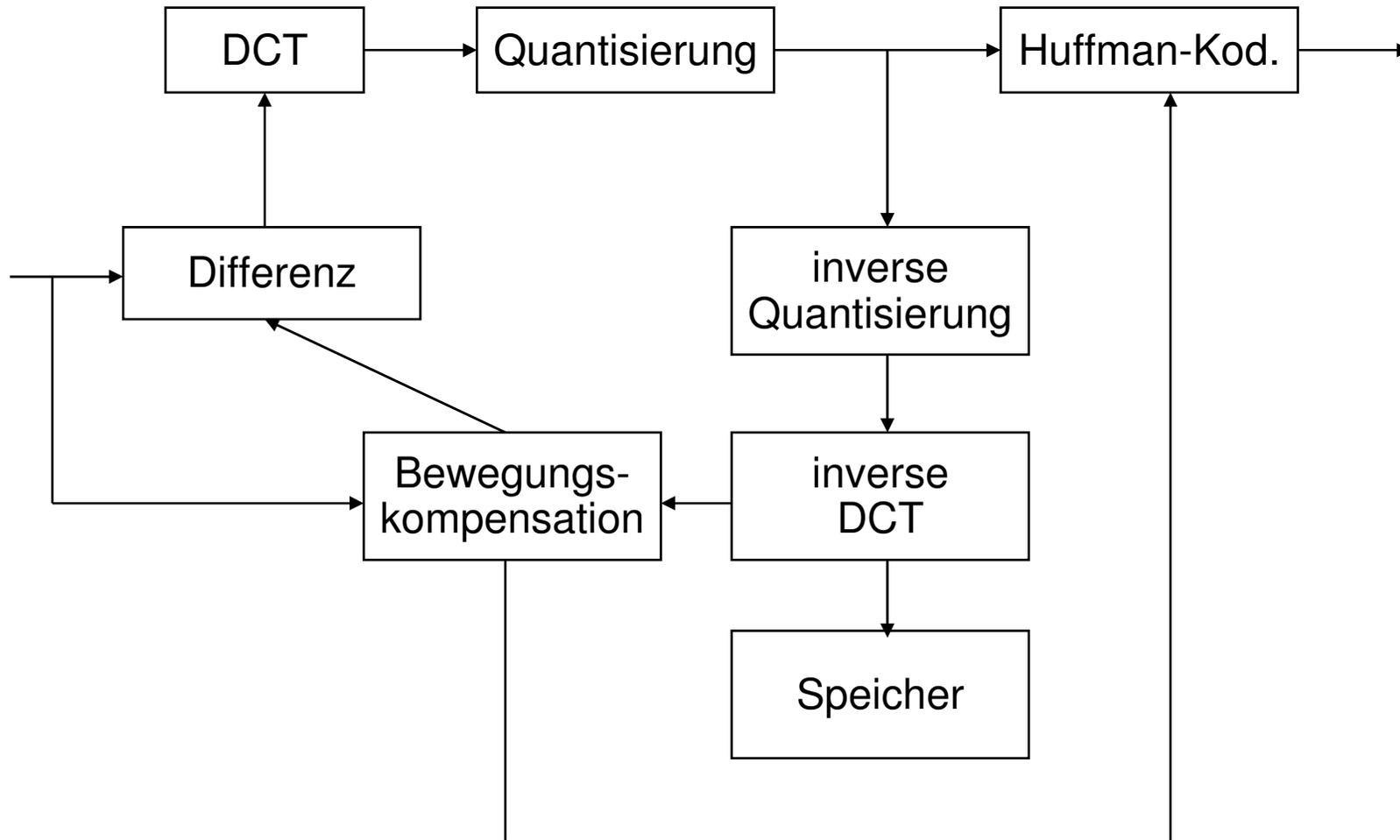
Zielframe  $N+1$



Referenzframe  $N+2$

- Das Zielframe soll im Decoder zwischen zwei anderen Frames interpoliert werden.
  - Bestimmung von zwei Verschiebungsvektoren
  - Differenzkodierung im Vergleich zum Durchschnitt der Darstellung des Makroblocks in den beiden Referenzbildern
- Ermöglicht es, mehr Information durch Verschiebungsvektoren zu kodieren als bei unidirektionaler Bewegungskompensation
- Nur in MPEG-2, nicht in MPEG-1

# Schema der P- und B-Frame-Kodierung



# Frame-Kompression in MPEG 1/2

- I-Frames:
  - JPEG-Kompression, d.h.
  - DCT, Quantisierung, Lauflängencodierung, Entropiecodierung
- B- und P-Frames:
  - Differenzen werden wie Bilder behandelt und nach dem gleichen Schema komprimiert
- Vereinfachung: Standardisierte Quantisierungstabellen

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	28	46	56	69
27	29	35	38	46	56	69	83

Luminanz

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

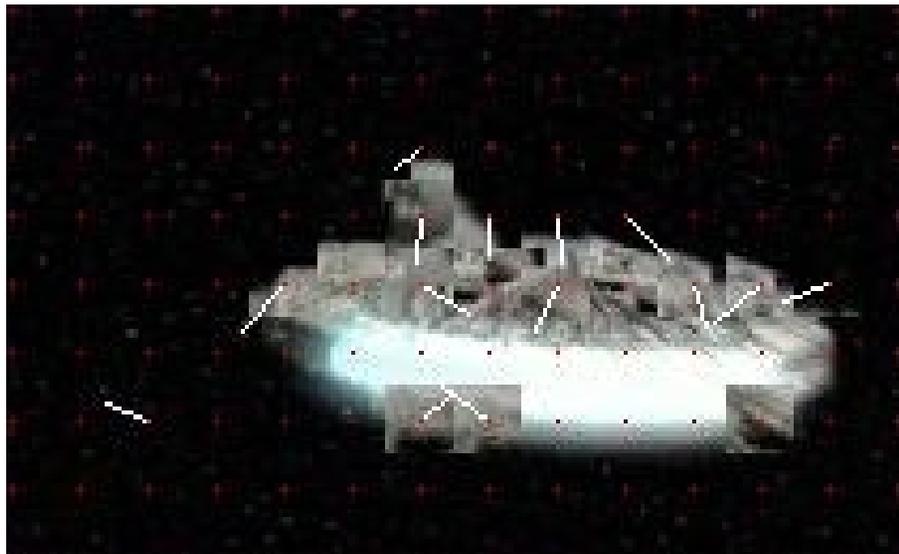
Chrominanz

# Symmetrische und unsymmetrische Verfahren

- Symmetrisch:
  - Aufwand für Codierung und Decodierung vergleichbar
  - Relativ geringe Kompression
  - Z.B. DV-Standard für Digital Video (Chroma-Subsampling + Intracodierung)
- Unsymmetrisch:
  - Codierung wesentlich aufwändiger als Decodierung
  - Hohe Kompression erreichbar
  - Qualität der Kompression oft abhängig von investiertem Aufwand
  - Z.B. MPEG-Kompression

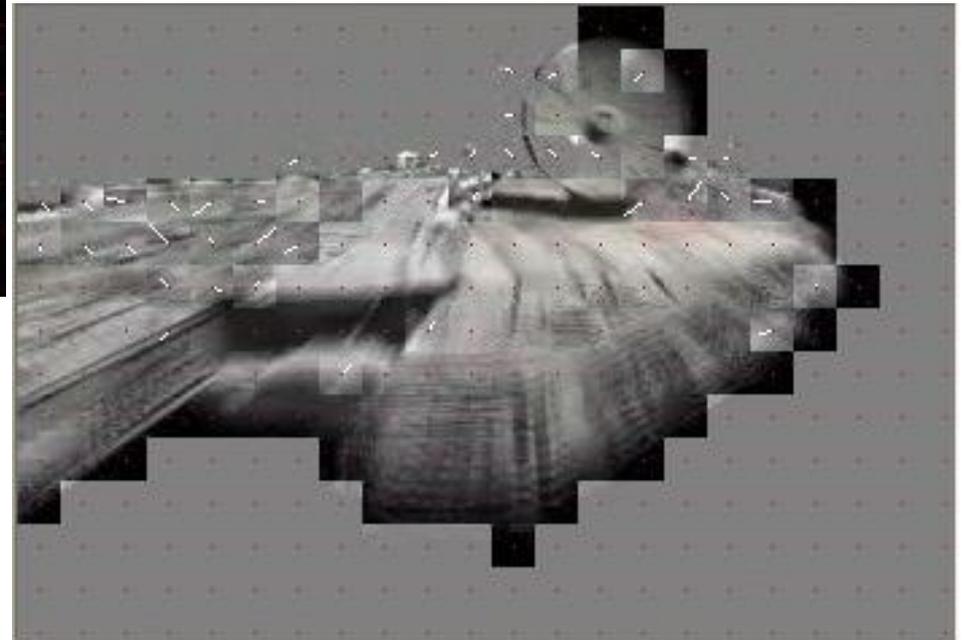
# Beispiel: MPEG Video Analyse

- Demonstrationssoftware "VCDemo", siehe:
  - <http://www-it.et.tudelft.nl/~inald/vcdemo/>



Bewegungskompensation

Inter-Frame Codierung



# 8. Bewegtbilder

8.1 Bewegungswahrnehmung

8.2 Videokompression

insbesondere MPEG-1 und MPEG-2

8.3 Videodatenformate



Literatur:

Arne Heyna/Marc Briede/Ulrich Schmidt: Datenformate im Medienbereich,  
Fachbuchverlag Leipzig 2003

John Watkinson: The MPEG Handbook, Focal Press 2001

# Was ist mit MPEG-3 passiert?

- Ursprünglicher Plan:
  - MPEG-3 als Erweiterung von MPEG-2 für HDTV
  - Wurde von MPEG-2 vollständig abgedeckt
  - Auflösung der MPEG-3 Aktivitäten
- Es gibt keinen MPEG-3 Standard!
- Parallel neue Entwicklung gestartet: MPEG-4
  - Start 1993
  - 1995: H.263 (siehe später) als Basis für Videocodierung gewählt
  - 1999: MPEG-4 Visual Standard publiziert, Ergänzungen 2002
  - 2003: H.264/MPEG-4 Part 10 (siehe später)

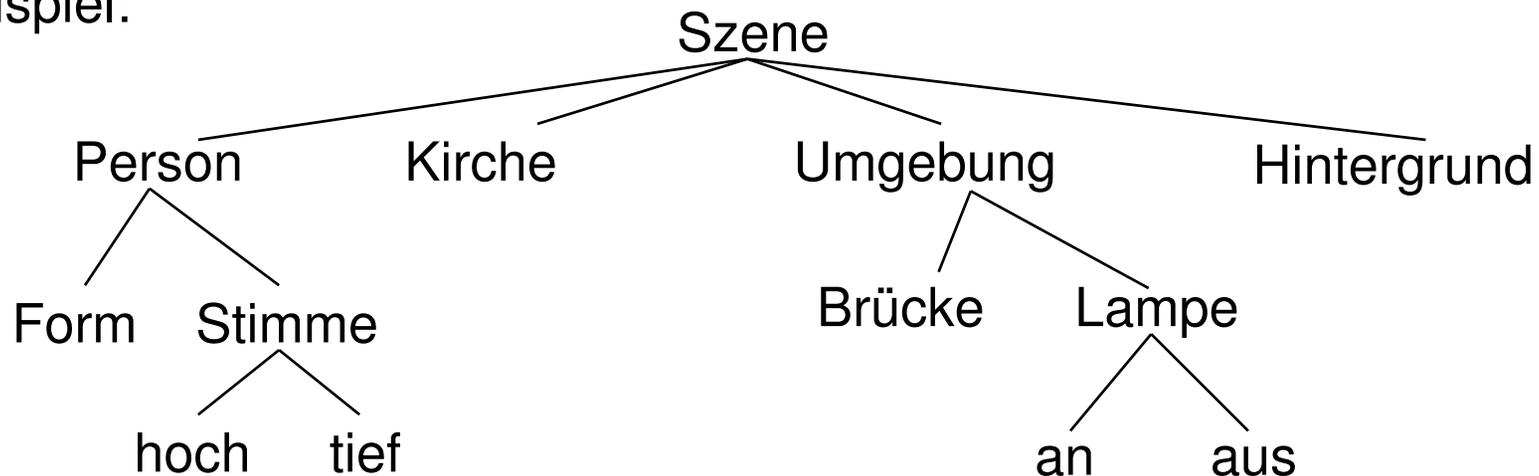
# MPEG-4 Visual (1999)

- Part 2 von ISO/IEC 14496 "Coding of Audio-Visual Objects"
  - Ca. 540 Seiten...
- Diverse "Coding tools"
  - Repräsentation verschiedener Datentypen für "Video-Objekte", siehe nächste Folie
- Kernstück = Video-Kompressionsalgorithmus, sehr ähnlich zu MPEG-2
  - Block-basiert, Bewegungskompensation, DCT, Quantisierung, Entropiecodierung
  - Verbesserungen der Codierungs-Effizienz u.a. durch
    - » Kleinere Blöcke für Bewegungskompensation (4x4)
    - » Intra-Frame prediction

# MPEG-4 Media Objects

- Media Objects
  - Beliebige audiovisuelle Datenformen, auch mit unregelmäßigen Grenzen
  - Z.B. Hintergründe, Video-Objekte (etwa Personen), Audio-Objekte, animierte Objekte (z.B. Avatare = Repräsentanten von Menschen in virtuellen Welten)
  - Synthetic Natural Hybrid Coding: Mischung aus künstlichen und abgetasteten Medienobjekten
  - Hierarchisch organisiert

Beispiel:



# Szenenbeschreibung in MPEG-4

- BIFS (Binary Format for Scenes)
  - Basiert auf dem Standardformat für 3-dimensionale Szenen VRML
  - XML-basiertes Repräsentationsformat XMT, ähnlich zu SMIL  
Details zu diesen Technologien siehe später!
- Bäume von Medienobjekten sind dynamisch
  - Bestimmte Knoten können Objekte bewegen und modifizieren
    - » Z.B. abhängig vom Zeitverlauf
  - Interaktion mit Objekten
    - » Reaktion auf benutzererzeugte Ereignisse
    - » Verursacht Modifikation von Objekten
- Anwendungsbeispiele:
  - Interaktive Produktpräsentation im E-Commerce
  - Interaktives Video (z.B. Sprachversionen)
  - Virtuelle Konferenzen mit künstlich animierten Köpfen/Körpern und Möglichkeit zur Steuerung des „eigenen“ virtuellen Repräsentanten

# H.261 und H.263

- H.261: CCITT-Entwicklung (1984-1990)
- H.263: ITU-T (1996), Ersatz und Ergänzung von H.261
- Ziel: Videokonferenzen und Videotelefonie auf ISDN-Leitungen
  - Bandbreiten 64 kbit/s und Vielfache
- YUV-Farbmodell, Chroma-Subsampling 4:2:0
- Frames vom Typ CIF oder QCIF
  - CIF (Common Interchange Format)
    - » NTSC: 352 x 240 Pixel, PAL: 360 x 288 Pixel; kein Interlacing
    - » Chroma-Subsampling 4:2:0
    - » Datenrate 36,5 Mbit/s
  - QCIF (Quarter CIF)
    - » 176 x 144 Pixel, sonst wie CIF
- I-Frames und P-Frames wie in MPEG

# H.264

- Zusammenarbeit zwischen MPEG (Moving Pictures Expert Group der ISO) und der VCEG (Video Coding Experts Group der ITU)
  - VCEG entwickelte H.261 und H.263
  - Gemeinsames Joint Video Team (JVT)
- Resultat: Internationaler gemeinsamer Standard
  - H.264/MPEG-4 Part 10, von beiden Gremien publiziert
  - "Advanced Video Coding" (AVC)
  - Ziel: Bessere Kompressionsraten relativ zur Qualität
- Techniken (Auswahl):
  - Prädiktion kann auf Speicher von mehreren Bildern zurückgreifen (short- and long-term prediction)
  - Baum-strukturierte Bewegungskompensation mit variabler Blockgröße
  - Prädiktion für Bewegungsvektoren aus benachbarten Partitionen
  - Filter zur Beseitigung von Block-Artefakten
  - Spezielle Entropie-Codierungsverfahren

# AVI

- Audio Video Interleave
- Eingeführt von Microsoft mit *Video for Windows* (1991)
  - Wurde abgelöst von ASF (siehe nächste Folie)
  - Seit 2003 MS-Standard-Format für Video „Windows Media Video“ (WMV) (analog „Windows Media Audio“ (WMA))
- Basiert auf dem generischen „RIFF“-Format (Resource Interchange File Format)
  - Universelles „Container“-Format:
  - Kann unkomprimierte oder komprimierte Bitmap-Daten für Video enthalten
- Audio- und Video-Information in einer Datei, je genau ein Strom
- Einfach
  - Kaum Unterstützung zur Synchronisation der Spuren
  - Reines Heimanwenderformat, für lange Sequenzen ungeeignet

# Microsoft ASF

- Derzeitige Bedeutung des Akronyms „Advanced Systems Format“
  - Frühere Bedeutung u.a. „Active Streaming Format“
- Zweck:
  - Vielzahl von multimedialen Inhalten in geordneter Form speichern, abspielen und portionsweise über Netze übertragen („streamen“)
  - Offizieller Nachfolger von AVI
- Alternative zu Apple QuickTime
  - Allerdings Schwächen z.B. bei der Einbeziehung dateifremder Rohdaten (wichtig für Videoschnitt)

# DivX

- DivX;-)
  - Jérôme Rota („Gej“) u.a. entschlüsseln Microsofts MP43C32.dll (Microsofts Ansatz für MPEG-4, nicht Standard-konform) und publizieren Varianten davon
  - U.a. MP3 als Audioformat (statt WMA 2, wie von Microsoft vorgesehen)
- OpenDivX (oder DivX 4)
  - Echter MPEG4-Codec
  - Entwickelt von „Gej“s DivXNetworks, Universität Hannover, HHI Berlin
  - Basiert auf MPEG-4-Quellcode aus EU-Projekt
- XviD
  - OpenSource-Entwicklung auf der Basis von OpenDivX

# Und jetzt...

- Als kleiner Rückblick & Ausblick auf die kommenden Tage:
  - MPEG-Layer-III-Audiodatei
  - Singstimmen durch Sprachsynthese erzeugt
    - » Vielsprachig
- Quelle:
  - MBROLA Projekt
  - Siehe <http://tcts.fpms.ac.be/synthesis/mbrola/demo/xmas.mp3>

