



Videocodierstandards H.120, H.261, MPEG-1, MPEG-2/H.262, H.263 und MPEG-4

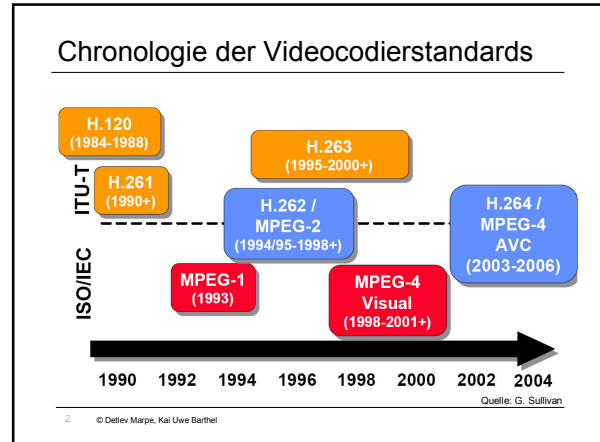
Dr. Detlev Marpe
Fraunhofer Institut für Nachrichtentechnik HHI



Fachhochschule für Technik und Wirtschaft Berlin
University of Applied Sciences



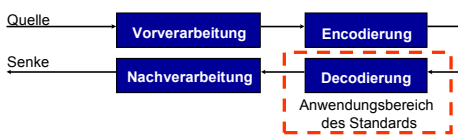
Fraunhofer Institut für Nachrichtentechnik
Heinrich-Hertz-Institut



Was wird standardisiert?

Quelle

Senke



- Nur die **Syntax** (Aufbau und Form) des Bitstroms und der **Decoder** werden standardisiert
 - Erlaubt implementationsspezifische Optimierung, z. B. bezüglich Komplexität, Qualität und Bitrate
 - D. h. insbesondere: **Standard garantiert keine Qualität, sondern lediglich Interoperabilität**
- Definierte **Konformitätspunkte** (Profile@Level) schränken Auswahl von Tools bzw. Bitstrom-Parameter (z.B. Videoformat, Bitrate, ...) ein

ITU-T H.120

Jahr: 1984 (version1), 1988 (version2)
Format: PAL und NTSC

Methoden: v1: Conditional Replenishment, DPCM (örtlich), lineare Quantisierung, VLC
 v2: Bewegungskompensation und Hintergrund-Prädiktion

Rate: 1544 (NTSC) bzw. 2048 (PAL) kbit/s

Anwendung: Bildtelefonie, Videokonferenz

ITU-T H.261

Jahr: 1990
Format: QCIF + CIF

Methoden: MB-basierte Bewegungskompensation
 Pixel-genaue Bewegungsparameter
 Intra / Interframecodierung
 DCT-basierte Codierung

Rate: 1x64 – px64 (64-2048) kbit/s

Anwendung: Bildtelefonie, Videokonferenz via ISDN

ISO/IEC MPEG-1 (Video)

Jahr: 1993
Format: Jede Auflösung < 4096 x 4096, nur progressiv; typische Auflösung SIF/CIF

Methoden: Bewegungskompensation mit **Halb-Pixel Genauigkeit, bidirektionale Prädiktion**, Intra- / Interkodierung mit P- und B-Bildern, DCT-Codierung, Quant.-Gewichtungsmatrizen

Rate: typisch: 1 – 2 Mbit/s inkl. Audio (MPEG-1 Layer3 = **MP3**) => Video in VHS-Qualität für SIF/CIF annähernd wahlfreier Zugriff, schneller Vor- und Rücklauf (nur über I-Bilder)

Merkmale: Speicherung von Video auf Video-CD, Festplatte, etc. (weitgehend durch MPEG-2 ersetzt)

MPEG-1/2: I, P und B Bildtypen

Group of Pictures (GOP)

Codierreihenfolge 1 3 2 5 4 6 8 7 10 9 ...

I = Intra-bilder, intra-codiert ohne Rückgriff auf andere Bilder
P = aus vorangegangenen I- oder P-Bildern prädiert
B = bidirektionale prädierte Bilder aus zwei aufeinander folgenden I- oder P-Bildern (zeitlich nachfolgende Referenzbilder müssen vor B-Bild en-/decodiert werden)

7 © Detlev Marpe, Kai Uwe Barthel

Eigenschaften der Bildtypen

- I Bilder**
 - Ermöglichen nahezu **wahlweisen Zugriff**
 - Decodierung kann nur mit einem I-Bild beginnen
 - Z.B. in TV-Übertragung alle 500 ms ein I-Bild
- P Bilder**
 - Ermöglichen i.a. **höhere Codiereffizienz**
 - Können nicht für sich alleine decodiert werden
- B Bilder**
 - Können **Codiereffizienz weiter verbessern**
 - Führen aber zu einer **zeitlichen Latenz** bei der Decodierung (kein Problem in TV-Übertragung, aber unzulässig für Anwendung in Videokommunikation, wie z.B. Videokonferenz)
 - Können zur **zeitlichen Skalierbarkeit** verwendet werden

8 © Detlev Marpe, Kai Uwe Barthel

MPEG-1/2/4: Sub-Pixel genaue Verschiebungen

Erzeugung von Sub-Pixel Information durch Interpolation; MPEG-1/2: Halb-Pixel Bildpunkte werden **bilinear** aus Ganzpixel-Werten generiert
Sub-Pixel Verschiebung wird häufig nur als Verfeinerung der Verschiebungen auf vorheriger Stufe geschätzt; *Beispiel:* G → e → □

9 © Detlev Marpe, Kai Uwe Barthel

ISO/IEC MPEG-2 / ITU-T H.262

Jahr: 1994 (gemeinsam von ISO/IEC und ITU-T)
Format: Zeilensprung und Progressiv (Halb-, Vollbild)

Methoden: alle MPEG-1 Methoden + Unterstützung für Halbbilder (Frame-/Field-Codiermodi) und Skalierbarkeit

Rate: ausgelegt für 2-20 Mbit/s (2-5 Mbit/s: DVB; 4-10 Mbit/s: DVD; ~20 Mbit/s HDTV)

Anwendung: **der** Standard für Digitales Fernsehen und DVD in Standard-Auflösung (SDTV: PAL/NTSC); **der meistbenutzte Videocodierstandard weltweit**

10 © Detlev Marpe, Kai Uwe Barthel

ITU-T H.263

Jahr: 1995 (v1), 1998 (v2:"+"), 2000 (v3:"++")
Format: alle progressiven Formate (ab v2)

Methoden: **v1** "baseline: Verbesserungen gegenüber H.261: Halb-Pixel Bewegungskomp., verbesserte Entropiecodierung, etc.
v2/v3: viele **optionale Modi** zur Verbesserung der Codiereffizienz, des Fehlerschutzes, der Skalierbarkeit, ...

Rate: > 10 kbit/s; typisch für niedrige Bitraten: **halbierte Rate v. H.261** bei gleicher Qualität

Anwendung: Videokonferenz, Bildtelefonie, Internet (Streaming), Mobilkommunikation

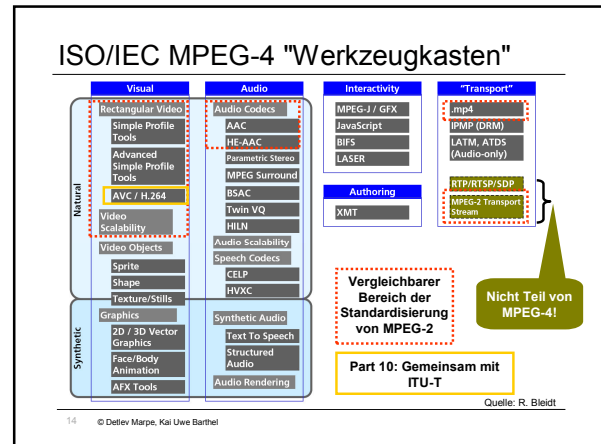
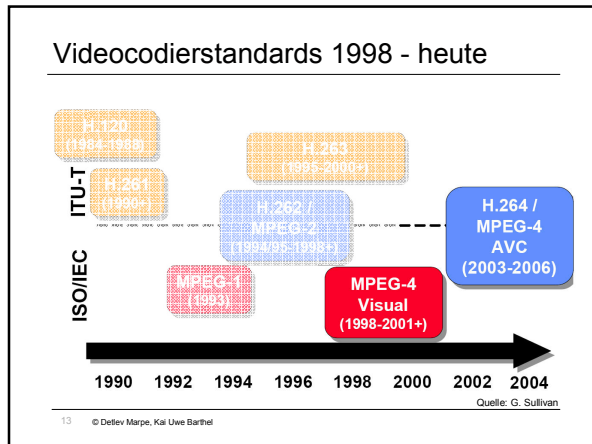
11 © Detlev Marpe, Kai Uwe Barthel

Meilensteine der Videocodierung: 1988-1998

Bitrateneinsparung: 67%

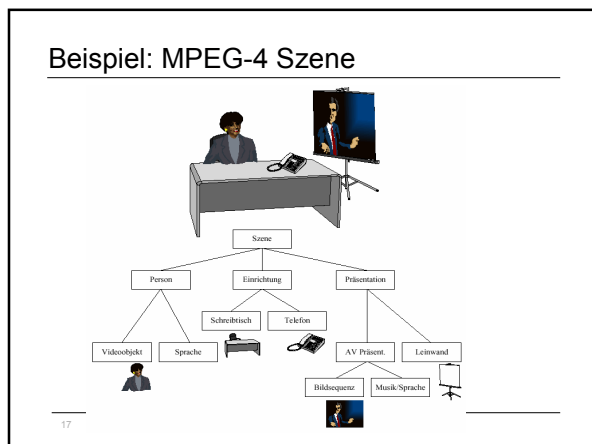
Foreman 10 Hz, QCIF 100 Frames codiert

12 © Detlev Marpe, Kai Uwe Barthel



- ### ISO/IEC MPEG-4 Visual
- Codierung von "rechteckigem" Video
 - Grundversion basiert auf H.263 Baseline → Simple Profile
 - Viertel-Pixel genaue Bewegungskompensation
 - Globale Bewegungskompensation
 - Sprites (zur Codierung von unveränderlichem Panorama-Hintergrund)
 - Neu: Codierung von "objekt-basiertem" Video
 - Definition von beliebig geformten Video Objekten (VO)
 - Codierung der Objektform ("Shape") per Alphamaske (binär- oder mehrwertig)
 - Shape-adaptive DCT zur Transformationscodierung
 - Neu: Mischung von natürlichen und synthetischen Medienobjekten in einer "Szene"

- ### Die MPEG-4 "Philosophie"
- Platzierung von unterschiedlichen „Medienobjekten“ in einer (virtuellen) Szene
 - Codierung kann auf unterschiedliche Objekttypen zugeschnitten werden
 - Text, 2D- oder 3D-Grafik, Standbilder, Video, etc.
 - Jedes Objekt wird in einem separaten "Elementarstrom" übertragen
 - Medien-Präsentation ist dynamisch
 - Anwendung von Transformationen zur Änderung der geometrischen oder akustischen Erscheinung
 - Interaktion und dynamische Änderung von Attributen der Medienobjekte
-
- Diagram illustrating the MPEG-4 philosophy, showing how objects are placed in a scene and how they are encoded and presented dynamically. Source: R. Bleid.



- ### MPEG-4 Szenenbeschreibung
- Binary Format for Scene description (BIFS)
 - Szenenbeschreibungsformat basiert auf der Virtual Reality Modeling Language (VRML 2)
 - Erweiterungen von BIFS gegenüber VRML:
 - Kompression der Szenenbeschreibung
 - Interaktion und Manipulation der multimedialen Szenen durch den Anwender
 - Streaming von Szenen, Update von Szenen (Dynamische Szenen)
 - Flexible Synchronisation der Objekte (über Object Descriptor)
 - "Light" Version von BIFS für Mobile Terminals: LAsER
 - XMT: eXtensible MPEG-4 Textual Format
 - Textuelle Szenenbeschreibung auf Basis von XML

HHI MPEG-4 Demo

Komposition und Echtzeit-Rendering von 2D- und 3D-Grafiken sowie Video und Standbilder

Demo in 3 Teilen:

- Klassische 3D Computergrafik-Objekte mit verschiedenen Texturierungen (interaktiv steuerbar)
- Komposition von Objekten (natürlich und synthetisch) in virtuellen Szenen mit Interaktionsmöglichkeiten
- Echtzeit-Rendering von vier CIF-formatigen Videos in einer Art Electronic Programming Guide (EPG)

19 © Detlev Marpe, Kai Uwe Barthel

MPEG-4 Kopf- und Körpermodelle

MPEG-4 Gittermodelle für Kopf und Körper

Synthese bzw. Animation z.B. mittels

- Facial Definition Parameter
- Facial Animation Parameter
- Facial Texture Coding

□ Beispiel: **Modell-basierte Videocodierung**

- Facial Animation Parameter werden vom Original (links) geschätzt und auf ein animiertes 3D-Modell (rechts) übertragen
- Modellparameter werden encodiert und mit **1-2 kbit/sec** übertragen
- Demomaterial: Dr. Peter Eisert, Fraunhofer HHI

20 © Detlev Marpe, Kai Uwe Barthel

MPEG Media Standards

Merkmalsbasierte Beschreibungssprache zur Klassifizierung von Medieninhalten
(z.B. zur Unterstützung von inhaltsbasierter Suche)

Verwaltung von Zugriffsrechten, Management von Medieninhalten, etc.
(z.B. zur dynamischen Anpassung von Inhalten)

"IT"
Content or Essence

AVC, AAC
MPEG-4
MPEG-2

Video, Audio,
Graphics, Interactivity

What is it?
Metadata
MPEG-7

Title, Author, Dates,
Dialogue, Thumbnails, Melody

How can I use it?
Rights
MPEG-21

Digital Item Identification
Rights Expression Language

Quelle: R. Bleidit

21 © Detlev Marpe, Kai Uwe Barthel