

Mehr Bildqualität mit weniger Bit



Quelle: Heinrich-Hertz-Institut

MPEG-4 etabliert sich als Nachfolger für MPEG-2

Digitales Video hat eine Reihe von Vorteilen:

Die Qualität hängt nicht mehr von den spezifischen Eigenschaften des Übertragungskanal ab, es gibt vielfältigere Methoden für Fehlerschutz und -korrektur, Speicherung und Kopieren ist ohne Qualitätsverlust mit unbeschränkter Lebensdauer möglich, die Methoden zur Bildbearbeitung und -manipulation sind vielseitiger, besser und kostengünstiger. Dem steht ein großer Nachteil entgegen: Die erforderliche Bandbreite ist bis zu 200-mal höher als bei analogen Video-Signalen. Schon zu Beginn der Digitalisierung von Video-Signalen hat man deshalb erkannt, dass kein Weg an der Videokompression vorbeiführt.

Größe eines Roh-Videos

Die Formel zur Berechnung der Größe eines Roh-Videos (RVG) lautet:

$$RVG = (f \cdot r \cdot c \cdot t) + A$$

- mit f: Bildgröße (frame size) in Bildpunkten (pixel), Höhe x Breite
- r: Bildwiederholrate (frame rate) in Bildern pro Sekunde (fps)
- c: Farbtiefe (colour depth) in Byte pro Pixel
- t: Länge des Videos in Sekunden
- A: Volumen des Audio-Soundtracks in Byte

Für ein zweistündiges Roh-Video in VGA-Auflösung ohne Ton ergibt sich also eine Größe von:

$$\begin{aligned}
 RVG &= \\
 &640 \cdot 480 \frac{\text{pixel}}{\text{frame}} \cdot 25 \frac{\text{frames}}{\text{s}} \cdot 3 \frac{\text{Byte}}{\text{pixel}} \cdot 7.200 \text{s} = \\
 &22.500 \frac{\text{KByte}}{\text{s}} \cdot 7.200 \text{s} = 162.000.000 \text{KByte} = \\
 &154,4952 \text{GByte}
 \end{aligned}$$

wobei 1 GByte = 1024 x 1024 x 1024 Byte sind. Es ist offensichtlich, dass Datenraten von über 20 MByte/s und Dateigrößen von mehr als 150 GByte die heute verfügbaren

Übertragungs- und Speichermedien weit überfordern. Das „Eindampfen“ des Datenstroms tut also Not!

Datenreduktion mit MPEG

Die Motion Pictures Experts Group (MPEG, gegr. 1989, sprich „empeg“) entwickelte einen CODEC (**C**oder/**D**ecoder oder auch Compressor/Decompressor) für die bitsparende Übertragung und Speicherung von Videos mit Ton. Anfang der 90er Jahre des vergangenen Jahrhunderts (also vor gut 15 Jahren) wurde der erste Entwurf eines Bildkompressionsalgorithmus veröffentlicht, der im August 1993 als Standard ISO/IEC 11172 (MPEG-1) verab-



**Bild 1: Qualitätsvergleich MPEG-4-2 gegen MPEG-4-10 (Version 2).
Ca. 1,5 Stunden Film passen in MPEG-4-Qualität damit auf eine normale
Daten-CD. (Quelle: Luke J. Crook, <http://www.balooga.com/mpeg4.php3>)**

schiedet wurde. Es folgte 1996 MPEG-2, das bis heute beim digitalen Fernsehen und der Filmspeicherung auf DVD eingesetzt wird. MPEG-4 (Anfang 1999) ist dessen Erweiterung, die bei höherer Kompressionsrate bessere Darstellungen ermöglicht.

Alle MPEG-Standards bauen aufeinander auf. Ihnen gemeinsam ist das Ausnutzen von Irrelevanz und Redundanz in einem Bild. Dabei wird das codierte Bild an die Eigenheiten des menschlichen Sehens angepasst, indem man ihm für das Auge nicht wahrnehmbare Bestandteile entzieht und die Ähnlichkeiten innerhalb eines Bildes oder in einer Folge von Bildern ausnutzt. Die Schritte des Verfahrens Chroma-Subsampling, zeitliche und räumliche Redundanzreduktion, Diskrete Cosinustransformation, Codierung, Prädiktion und Bewegungskompensation, Datenstrombildung, Fehlerkorrektur ... wurden bereits im „ELVjournal“ 1/05: „Satellitenempfang, Teil 3“ und „ELVjournal“ 2/05: „Satellitenempfang, Teil 4“ ausführlich beschrieben. Deshalb geht dieser Artikel nur auf die spezifischen Eigenheiten von MPEG-4 ein.

Der MPEG-2-Standard hat sich weltweit als Standard für digitales Fernsehen etabliert. Schätzungsweise 50 Millionen digitale Receiver (Set-Top-Boxen) arbeiten nach dem internationalen DVB-Standard (Digital Video Broadcast: Digitales Fernsehen), der wie ATSC (Advanced Television Systems Committee – digitales hochauflösendes Fernsehen in den USA), DVD (Digital Versatile Disc) und MP3 als Audiokompressionsstandard (MPEG-1, Layer 3) auf MPEG-2 beruht. Aber der technische Fortschritt hat neue Anwendungen ermöglicht: Vom hochauflösen-

den Fernsehen (HDTV: High-Definition TV) über IP-basiertes Videostreaming bis zum „Mäusekino“ auf dem Handy-Display (DVB-H) – stets sind die Datenraten für die zur Übertragung verfügbaren Kanäle zu groß. Damit hat die Stunde von MPEG-4 geschlagen.

MPEG-4 AVC/H.264

Der MPEG-4-Standard (AVC: Advanced Video Coding = H.264) zielt auf die möglichst effiziente Übertragung herkömmlicher A/V-Inhalte, strukturierter synthetischer 2D/3D-Grafiken und Audio über Breitbandnetze, das Internet, DVD, PDAs, Mobiltelefone und Set-Top-Boxen. Als Anwendungen kommen in Frage: Unterhaltung, Kommunikation zwischen einzelnen Personen oder Gruppen (Bildtelefon, Videokonferenzen), Produktmarketing, Fernunterricht, interaktive Videospiele, Multimedia-Mailing, Fernüberwachung durch Webcams usw.

MPEG-4 entspricht den Eigenheiten der menschlichen visuellen Wahrnehmung (Human Visual System HVS) besser als MPEG-2. Der Mensch nimmt eine Szene durch Umrisse (shapes), Bewegung, Texturen und Farbe wahr. MPEG-4 trägt dem durch die Verwendung visueller Objekte Rechnung. Die Realität wird dadurch HVS-freundlicher abgebildet, als es Makroblöcke und Pixel vermögen. Und nicht zuletzt ist das Schneiden und Nachbearbeiten von MPEG-2-komprimiertem Material nur in geringem Umfang möglich.

MPEG-4 wurde 2000 formell als Standard ISO/IEC 14496 verabschiedet. Die Arbeiten daran halten an und berücksichti-

gen die Entwicklungen auf dem Gebiet der Codierungsforschung und der verfügbaren Hard- und Software. Neben einer höheren Datenverdichtung bietet MPEG-4 einen Baukasten von Technologien, die zu anwendungsspezifischen „Profiles“ zusammengefasst werden und gleichermaßen nützlich sind für Autoren von Inhalten, Diensteanbieter und Endanwender:

- Autoren können mit MPEG-4 Inhalte flexibler und mit höherer Wiederverwendbarkeit erstellen bei besseren Möglichkeiten zum Verwalten ihrer Urheberrechte.
- Für Diensteanbieter bietet MPEG-4 transparente Informationen über die Servicequalität (QoS – Quality of Service), die auf die verwendete Übertragungsnetzart umgesetzt werden müssen. So wird der Transport über heterogene Netze optimiert. Die verbesserte Interaktion mit dem Endanwender erschließt neue Geschäftsmodelle.
- Für Endanwender steigert ein höherer Grad an Interaktion mit den Diensten (nach Maßgabe des Inhalteautors) deren Attraktivität. Der Zuwachs an Multimedialität erstreckt sich auch auf neue Übertragungsstrukturen, insbesondere solche mit relativ niedrigen Bitraten wie Mobiltelefone, Wireless LANs usw., und lässt völlig neue Anwendungen zu.

Teile und herrsche!

Schon die alten Römer hatten die Macht dieses Prinzips erkannt. Auch bei MPEG-4 sorgt es für eindrucksvolle Resultate. Ein Wesensmerkmal von MPEG-4 ist die Zerlegung des Bildinhalts in statische und bewegte Bild- und Tonobjekte (natürliche und synthetische!) sowie räumliche Ebenen und deren getrennte Bearbeitung und zeitgleiche Übertragung. Die resultierende Bildqualität verblüfft in Anbetracht des geringen Speicherbedarfs bzw. der niedrigen Übertragungsrate (s. Bild auf erster Seite dieses Artikels und Abbildung 1).

Alles nur Theater!

Gewiss: Vergleiche hinken! Aber ein MPEG-4-Empfangsgerät, mit dem sich der Zuschauer audiovisuelles Programmmaterial zeigen lässt, hat Ähnlichkeit mit einem Puppentheater. Der Programmanbieter überträgt zunächst das Bühneninventar (Hintergründe, Mobiliar, Schauspieler usw.) und dann das Drehbuch, welches vorschreibt, wie damit zu verfahren ist. Im MPEG-4-Empfänger kommt nun das Schauspiel zur Aufführung. Dabei hat der Zuschauer Eingriffsmöglichkeiten in das Geschehen, indem er das Drehbuch oder die Regieanweisungen entsprechend abändert.

Das Visual-Decoder-Modell

Die Umsetzung des MPEG-4-Pro-

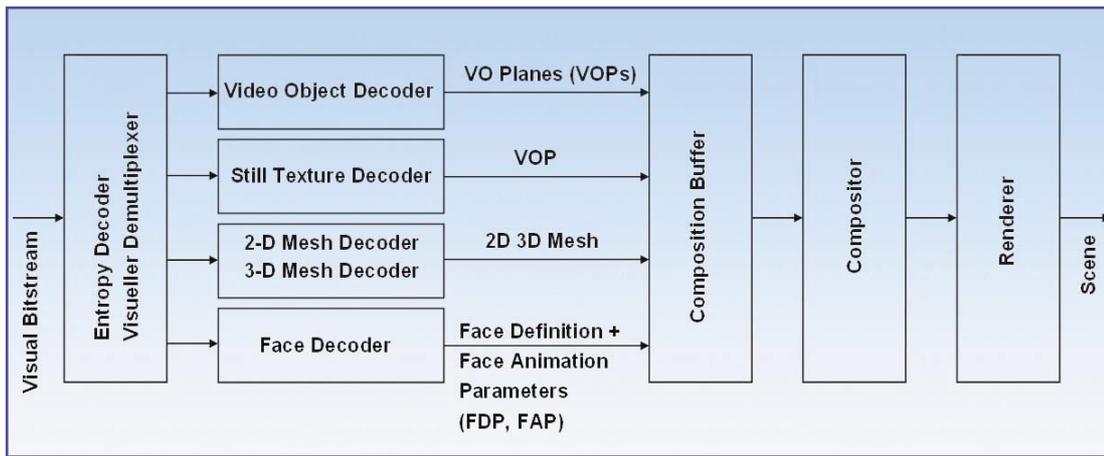


Bild 2: MPEG-4 Visual Decoder (Quelle: Kazem Najafi)

grammstroms in Szenenbilder geschieht im MPEG-4-Visual-Decoder-Modell (Abbildung 2). Es sind die zum Coder inversen Baugruppen, die aus den übermittelten Bildbestandteilen wieder das Ausgangsbild herstellen.

Die Werkzeuge zur Darstellung von Video-Aufnahmen im MPEG-4-Standard erlauben das effiziente Speichern, Übertragen und Manipulieren kleinster Einheiten von Bild- und Video-Inhalten, den Video-Objekten (VO). Ein Video-Objekt kann z. B. eine sprechende Person ohne Hintergrund sein, die dann mit anderen audiovisuellen Objekten zu einer Szene zusammengeführt wird. Die Szenenkomposition erfolgt durch die MPEG-4-Szenenbeschreibungssprache BIFS (Binary Format for Scenes).

Abbildung 3 erläutert dies. Aus dem MPEG-4-Datenstrom werden durch Demultiplexen die Elementarströme gewonnen, welche die primitiven AV-Objekte (Hintergrund, Golfspielerin, Ball, Flagge

und Ton) repräsentieren. Ebenfalls im MPEG-4-Datenstrom eingebettet ist die Szenenbeschreibung in Gestalt eines hierarchischen Szenenbaums. Damit kann der Decoder die Szene wiederherstellen. Als Interaktion ist z. B. die Zahlung einer Gebühr für das Sehen der Sendung denkbar, die das Video-Objekt Golfball sichtbar macht und so dem Betrachten des Matches erst einen Sinn gibt. Weil die Objekte einzeln codiert sind, können sie unabhängig voneinander manipuliert werden. So ist eine objektabhängige Auflösung denkbar, der Austausch, das Einbinden oder Löschen von Objekten, die Wahl der Perspektive, des Ausschnitts und vieles mehr.

Mesh, Morph, Render und Co.

Zur Codierung synthetischer, d. h. im Computer erstellter Objekte wird die Mesh-Technologie eingesetzt. Dabei bildet man das Objekt durch eine Hülle aus Polygonen (mesh = Maschennetz) ab, die mit Textu-

ren ausgefüllt werden (Rendern) (Abbildung 4). Aber auch reale Video-Objekte lassen sich durch ein gerendertes Maschennetzmodell darstellen. Die erforderliche Auflösung bestimmt die lokale Maschengröße (progressive mesh). Bewegungen und Geometrieänderungen werden nachgebildet durch Morphing-Befehle. Je nach Komplexität der Oberfläche des Objekts und seiner Aktionen sind extreme Einsparungen gegenüber den herkömmlichen Verfahren zur Digitalisierung von Bildausschnitten möglich.

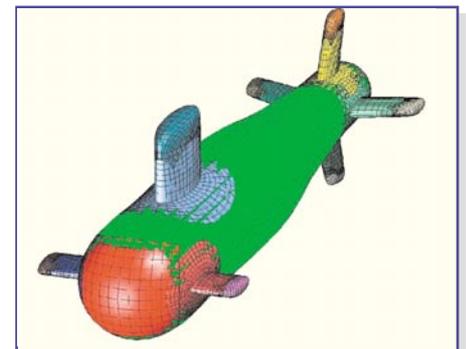


Bild 4: Ein Objekt mit Maschennetzoberfläche (Quelle: Bill Henshaw, Lawrence Livermore National Laboratory, http://www.llnl.gov/CASC/Overture/henshaw/publications/using_mappings.pdf)

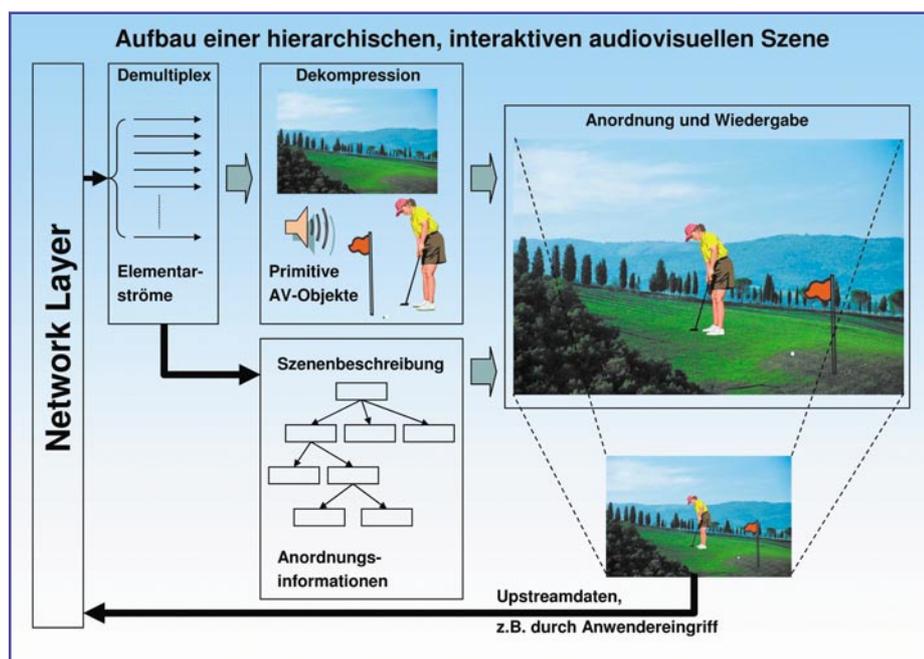


Bild 3: Aus Objekten und Kompositionsanweisungen wird das Szenenbild „zusammgebaut“.

Ein schönes Beispiel für diese Technik sind dreidimensionale Humanoiden (auch Avatare genannt), die den natürlichen Menschen durch intelligente Geometriedefinitionen (BDP: Body Definition Parameter Set) und Parametersets zur Änderung der Körperstellung (BAP: Body Animation Parameter Set) oder des Gesichtsausdrucks (FAP: Face Animation Parameter Set) verblüffend echt nachbilden. Auf der Homepage der Humanoid Animation Working Group, <http://www.h-anim.org>, kann man sich einen Eindruck verschaffen, wie die mathematische Beschreibung eines Humanoiden funktioniert. Wer selber einmal ein Gesicht animieren will: Auf <http://mrl.nyu.edu/~perlin/demox/Face.html> kann er mit wenigen Tas-



Bild 5: Mit wenigen Animationspunkten lässt sich der Gesichtsausdruck verändern. (Quelle: Ken Perlin)

tenanschlägen traurige, wütende, lachende, erstaunte usw. Gesichtsausdrücke zaubern (Abbildung 5). 68 in Gruppen eingeteilte FAPs, die in unterschiedlicher Intensität miteinander kombiniert werden können, erlauben die Belebung eines künstlichen Gesichts bis in die Feinmotorik. Wie man solche Animationen in die MPEG-4-Szenenbeschreibungssprache BIFS umsetzt, kann man auf <http://www.research.att.com/projects/AnimatedHead/joern2.html> nachlesen.

Mehr Kompressionseffizienz durch objektorientierte Bewegungskompensation

MPEG-4 erlaubt die Anwendung unterschiedlicher Verfahren zur Bewegungskompensation:

Standard 8x8- oder 16x16-Pixel-Block (wie bei MPEG-2)

Global Motion Compensation (GMC) Mit Hilfe weniger zu übertragender Parameter kann die Bewegung eines ganzen Objekts beschrieben werden. Durch Techniken wie Image Warping (schrittweise

Veränderung von Bildern), Motion Trajectory Coding (Codierung einer Bewegungspur) und Texturcodierung sind erhebliche Codierungsgewinne möglich.

Global Motion Compensation für statische „Sprites“

Ein Sprite ist ein grafisches Bild, welches bewegt oder animiert werden kann. Ein Beispiel: Wird ein Panorama-Hintergrund (Zuschauer auf Tribüne) als Sprite hinter einem Objekt (Tennisspieler) verschoben, hat der Betrachter den Eindruck, der Tennisspieler bewegt sich vor einem Ausschnitt der Zuschauerkulisse (Abbildung 6).

Formangepasste DCT (Shape-adaptive DCT: SA-DCT)

Im Bereich der Texturcodierung von Objekten mit beliebigen Umrissen verbessert die SA-DCT die Codierungseffizienz erheblich. Durch eine 4x4-DCT mit einer Integer-Arithmetik (anstelle der bei MPEG-2 üblichen 8x8-DCT mit Floating-Point-Arithmetik) werden unerwünschte Effekte wie Blockbildung und Überschwingen an Kontrastgrenzen (ringing) (Abbildung 7) sowie Driftartefakte infolge von

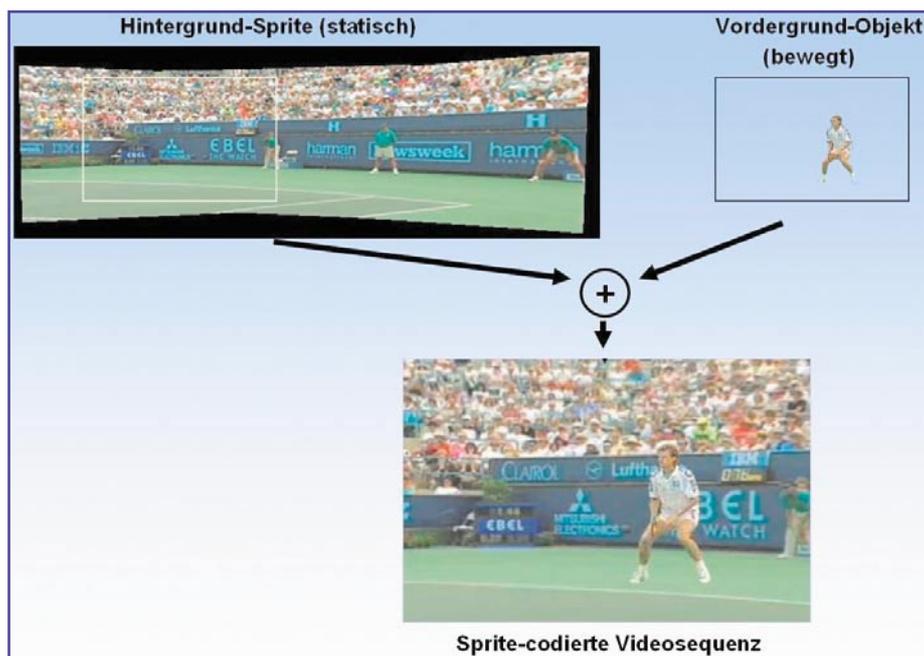


Bild 6: Spritecodierung einer Videosequenz (Quelle: Overview of the MPEG-4 Standard, http://www.chiariglione.org/ride/inside_MPEG-4_part_B/inside_MPEG-4_part_B.htm)

Rundungsfehlern verringert. Das „Ringing“ ist übrigens das sichtbar gewordene Gibb'sche Phänomen, welches besagt, dass die harmonische Synthese an Unstetigkeitsstellen der nachzubildenden Funktion stets Überschwinger produziert.

Weitere Maßnahmen zur Effizienzsteigerung

CABAC (Context-adaptive Binary arithmetic Coding) als effizienterer Ersatz für die variable Lauflängencodierung (VLC) passt sich an die vom Bildinhalt abhängige Wahrscheinlichkeitsverteilung der Symbole an, indem es die Korrelationen zwischen den Symbolen auswertet. Damit ist eine erhebliche Steigerung der Kompressionsrate möglich. Die Auffälligkeit von Blockartefakten verringert eine so genannte Inloopfilterung. Dabei wird in



Bild 7: MPEG-typische Bildfehler: Blockbildung (blocking) und Überschwingen an Kontrastgrenzen (ringing). (Quelle: Mitsubishi Electric Research Institute)

den Randbereichen der Blöcke eine Angleichung an die Nachbarn vorgenommen. Eine adaptive Quantisierung macht sich die geringere Detailauflösung des menschlichen Auges in hellen und dunklen Bildbereichen zunutze. In solchen Bereichen wird durch Wahl einer größeren Quantisierungskennlinie stärker komprimiert.

Resistenz gegen Übertragungsfehler

Durch ausgefeilte Methoden zur schnellen Resynchronisation, Datenwiederherstellung und Fehlerverdeckung wird in MPEG-4 eine hohe Robustheit gegen Übertragungsfehler sichergestellt. Diese ist natürlich auch wieder abhängig vom Bildmaterial.

Bei sich schnell und stark verändernden Bildfolgen wirken sich Paketverluste stärker aus als in ruhigen Szenen (Abbildungen 8 bis 11). (Quelle: Peter Parnes, www.cdt.luth.se/projects/siram/mp4)

Extreme Skalierbarkeit

Vom Sub-QCIF mit 172 x 144 Pixel (Quar-



Bild 8: Fehlerfreie Übertragung einer Bildfolge mit relativ geringen Inhaltsänderungen



Bild 9: Übertragung einer Bildfolge mit relativ geringen Inhaltsänderungen bei 5 % Paketverlusten => geringe Auswirkungen



Bild 10: Fehlerfreie Übertragung einer Bildfolge mit starken Inhaltsänderungen



Bild 11: Übertragung einer Bildfolge mit starken Inhaltsänderungen bei 5 % Paketverlusten => starke Auswirkungen

ter Common Intermediate Format für Videoconferencing und -telefonie) bis HDTV mit 1440 x 1152 Pixel (High-Definition TV: hochauflösendes Fernsehen), entsprechend Bitraten von 5 KBit/s bis 10 MBit/s, ist MPEG-4 für alle Qualitätsanforderungen geeignet. Aus der Vielzahl der Parameter, welche die Eigenschaften der MPEG-4-Codierung und Übertragung (Videoauflösung, Datenraten ...) beeinflussen, sind 19 Profile zusammengestellt worden, die wiederum in zahlreiche „Level“ gegliedert sind.

Einfacher Datenzugriff

MPEG hat das „Delivery Multimedia Integration Framework“ (DMIF) als Rahmenwerk formuliert, um die Lieferung multimedialer Inhalte unabhängig von ihrem Speicherort und der Zugriffs- und Bereitstellungsmethode zu ermöglichen. Damit sind die Aspekte der multimedialen Verarbeitung von denen der Übermittlung getrennt. So kann z. B. eine Multimedia-Anwendung wie ein MPEG-4-Abspielgerät auf Inhalte zugreifen, ohne dabei das Lieferszenario und die Technologien der daran beteiligten Netze berücksichtigen zu müssen.

Transport

In MPEG-4 sind keine Transportlayer definiert. Dadurch lassen sich mit entspre-

chenden Anpassungen MPEG-4-codierte AV-Inhalte vielfältig transportieren, z. B. als Inhalt von Datencontainern in MPEG-2-Transportströmen (MPEG-4 over MPEG-2) oder über IP-basierte Netze.

Ausblick

Obwohl wir noch lange nicht am Ende der Entwicklungs- und Normierungsarbeiten stehen, hat MPEG-4 heute schon das Potential, MPEG-2 als allgemeinen Standard für die digitale TV-Übertragung abzulösen. Insbesondere da, wo Speicher- und Übertragungskapazitäten knapp sind und ein hohes Maß an Interaktivität und Skalierbarkeit gewünscht wird, kann MPEG-4 durch den objektbasierten Ansatz seine Vorteile ausspielen. Die folgenden Erweiterungen sind bereits im fortgeschrittenen Stadium.

MPEG-7 als „Multimedia Content Description Interface“ ist ein Standard zur Beschreibung von MPEG-4-Daten. Die dabei eingesetzte Description Definition Language DDL ist eine Erweiterung der Seitenbeschreibungssprache XML (Extensible Markup Language), die vornehmlich im Webseitendesign eingesetzt wird. Sie erlaubt die Beschreibung, Verwaltung und Organisation von audio-visuellen Datenbeständen und wird für Navigation, Zugriff, Präsentation und Interaktion einge-

setzt. Auch Mechanismen zur Wahrung von Urheberrechten (DRM: Digital Rights Management) sind in MPEG-7 eingebaut. Eine genaue Beschreibung findet sich in http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm#2.1_MPEG-7_Systems.

MPEG-21 Multimedia Framework ist ein MPEG-Standard, der eine Infrastruktur für Erzeugung, Produktion, Freigabe und Handel von multimedialen Inhalten festlegt. Er fasst alle am Arbeitsfluss Beteiligten – vom Produzenten bis zum Konsumenten – zusammen und sorgt somit für eine standardisierte und flexible Zusammenarbeit (Def. Wikipedia). MPEG-21 beschreibt die vollständige Infrastruktur zur Übertragung angeforderter und bereitgestellter multimedialer Inhalte. Deren Nutzungsbedingungen sind in der MPEG-21-Architekturkomponente Intellectual Property Management & Protection (IPMP) geregelt. Details sind unter http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm#_Toc23297968 zu finden.

Fazit: Mit MPEG-4 und seinen Erweiterungen MPEG-7 und MPEG-21 ist interaktives Multimedia aus allen denkbaren Quellen über alle heute bekannten elektronischen Verbreitungs Kanäle unter Beachtung urheberrechtlicher und kommerzieller Aspekte wirtschaftlich realisierbar. **ELV**