

Multimedia – Daten und Formate

Zusammenfassung der Folien + Zusatzbemerkungen aus dem Web.

by MarvinTheRobot

Die folgenden Seiten beinhalten einen Überblick über den Stoff der in der Vorlesung Multimedia – Daten und Formate durchgemacht wurde. (Image Retrieval wurde nicht in die Zusammenfassung aufgenommen)

Rechtschreib- / Übersetzungs- und Verständnisfehler vorbehalten

Teil 0:
Multimedia Definitionen

ein Medium lässt sich auf zwei Arten definieren:

- eine Substanz, über die etwas übertragen werden kann; ein Mittel zur Verbreitung oder Präsentation von Information

- ein Mittel zur „Massenkommunikation“ (wie Zeitungen, Magazine und Fernsehen)
daher: Multimedia = „eine Kombination von mehreren Medien“

Aber nicht jede beliebige Kombination von verschiedenen Medien kann man mit dem Begriff „Multimedia“ bezeichnen. Dazu müssen folgende Kriterien erfüllt sein:

- die Erzeugung, Manipulation, Präsentation und Speicherung der Informationen ist computergesteuert
- die Informationen sind voneinander unabhängig
- die Information wird decodiert durch ein kontinuierliches (zeitabhängiges) und ein diskretes (zeitunabhängiges) Medium

Trotzdem wird der Begriff Multimedia sehr oft für Systeme verwendet, die nicht alle oben genannten Bedingungen erfüllen.

Teil 1:
Medientypen:

Folgende Medientypen werden behandelt:

- Text
- Graphik (Bilder)
- Audio
- Video

Text Operationen:

String und character Operationen

Editieren

Formatieren – Interaktiv, nicht Interaktiv

Sortieren

Komprimieren (Huffman, Lempel – Ziv)

Verschlüsselung

Sprachspezifische Operationen (Rechtschreibung, Grammatik, statistische Analyse des Schreibstils)

Medientyp – Text – Attribute:

- ASCII / ISO charset
- Markup Text – HTML, SGML, XML, etc.
- Hypertext – Hypercard, World Wide Web

Medientyp – Graphik:

- Achromatisches und farbiges Licht
- Intensität
- Farbmodelle
- Bild Attribute
- Bild Formate
- Bild Operationen

LICHT

Chromatisch (Farbe)

Achromatisch (schwarz-weiss)

Die Farbe eines Objekts hängt von

- Der Objektoberfläche: Reflexion, physische Eigenschaften, etc
- Lichtquelle (Intensität, Art, Anzahl der Quellen)
- Farbe der Umgebung
- Visuelles System das verwendet wird um „wahrzunehmen“ (Auge, Hirn, Photozelle)

ab.

ACHROMATISCHES LICHT

Schwarz, weiss, normalerweise Grautöne (Schwarz und weiss kommen fast nie in reiner Form vor)
Gemessen wird dieses Licht nur in der Menge oder in der Energie mit der es ausgestrahlt wird.

Das Verhältnis zwischen minimaler und maximaler Intensität nennt man: Dynamic Range.

HALBTON APPROXIMATION

Dithering:

Da für den monochromen Druck nur eine - nämlich schwarze - Druckfarbe eingesetzt wird, können die verschiedenen Grautöne nicht wie beim Malen durch Hinzufügen eines weißen Farbstoffs gewonnen werden. Man bedient sich deshalb eines Tricks: Auf die weiße Grundfläche des Papiers werden mehr oder weniger große Punkte gedruckt. Dazu verlässt man sich auf die Tatsache, dass das menschliche Auge die schwarzen und weißen Farbanteile zu dem gewünschten grauen Farbeindruck vermischt. Außerdem setzt man die einzelnen Bildpunkte eher diagonal, da diagonale Linien vom Menschen eher als zusammengehörig gesehen werden als senkrechte oder waagrechte.

FARBE:

Es gibt mehrer Farbmodelle die heute noch verwendet werden.

Print und Graphik Industrie verwendet Farbbücher, Tabellen und Codes

Künstler verwenden „tint“, „shade“, und „tone“

CRTs verwenden Hardware Modelle (RGB)

Physiker verwenden optische Modelle (Wellenlänge, Energie-Intensität)

Die menschliche Wahrnehmung von Farbe hängt davon ab wie das Gehirn auf die komplette visuelle Umgebung reagiert.

Wir empfangen elektromagnetische Strahlen die auf unsere „Retina“ fallen, dadurch entsteht Farbsehen.

Wir sehen Licht zwischen 400 und 700 Nanometern Wellenlänge.

Im Auge auf der Netzhaut befinden sich Stäbchen und Zäpfchen. Die Stäbchen sind für das Farbsehen verantwortlich, es gibt jeweils Zäpfchen für Rote, Grüne und Blaue Wellenlängen. Die Stäbchen liegen alle konzentriert im Zentrum der Netzhaut. Rund um diese schließen die Zäpfchen an, die für das Hell-Dunkel-Sehen in der Nacht verantwortlich sind.

Die oben beschriebene Tristimulus-Theorie hat allerdings einen Haken – mehrere Farben, die wir sehen, können nämlich nicht als Kombination von roten, grünen und blauen Wellenlängen dargestellt werden.

Betrachtet man die Farbkurven im Tristimulus-Modell, so erkennt man, dass die Kurve der blauen Wellenlänge kurz in den negativen Bereich absinkt. Da aber eine negative Reizung nicht möglich ist, muss dieses Modell transformiert werden. Durch mathematische Umformungen erhält man eine Tabelle, in der alle drei Wellenlängen immer im positiven Bereich liegen.

Dieses neu entstandene Modell wird das CIE-Farbmodell genannt. Durch einen Schnitt in der Ebene erhält man eine Hufeisenkurve, auf deren Rand alle Spektralfarben vorhanden sind. Im Mittelpunkt der Kurve liegt weiß, und auf den anderen Punkten der Fläche zwischen der Kurve liegen alle möglichen Farben. Sie entstehen durch Mischung der Randfarben.

FARBMODELLE für RASTER GRAPHIK

RGB:

Hardware orientiertes Modell, das die Farben Rot, Grün und Blau in unterschiedlichen Intensitäten mischt.

Diese Farbmodell ist additiv, das heisst Rot, Grün und Blau werden zusammengemischt um „alle“ Farben zu erhalten.

Nicht alle Farben können exakt erreicht werden, die Exaktheit ist abhängig von der „performance“ des CRT Displays.

Die Darstellung des RGB Modells lässt sich mit einem Würfel am besten veranschaulichen.

CMYK:

Cyan, Magenta und Gelb mit Schwarz sind die Grundfarben des CMYK Modells. Dieses Modell wird häufig für den Druck verwendet. CMYK ist ein subtraktives Modell, da Papier kein Licht abstrahlt.

HSV:

„hue saturation value“: basierend auf „tint“, „shade“ und „tone“.

Die Darstellung des HSV Modells lässt sich am ehesten mit einer am Kopf stehenden sechsseitigen Pyramide veranschaulichen. Das HSV Modell entspricht dem menschlichen Farbsehen am ehesten. Im Zentrum der Grundfläche der Pyramide befindet sich die Farbe Weiss, an den Ecken der Grundfläche findet man die Farben Grün, Gelb, Rot, Magenta, Blau und Cyan, an der Spitze befindet sich die Farbe Schwarz.

GRAPHIK Attribute:

- Anzahl der Kanäle (Art des Farbmodells + die Alpha Kanäle)
- Kanaltiefe (Anzahl der bits pro Pixel)
- Pixel Seitenverhältnis
- Interlacing
- Kompression (verlustfrei vs. Verlust)
- Indexing (Farben werden anstatt direkt, in eigene Tabellen gespeichert – „color lookup table (CLUT). Im Bild werden dann in den einzelnen Bereichen auf die Farben in der Tabelle verwiesen. z.B. gif Format)

Interlacing ist eine Bildeigenschaft, die man in Grafikprogrammen beim Abspeichern von Bildern festlegen kann. Interlacing bewirkt, dass sich das Bild in mehreren Durchgängen aufbaut: Erst sieht man es nur grob, aber nach und nach wird es deutlicher, bis es schließlich fertig geladen ist. (z.B.: beim gif Format)

Einige Bildformate:

- EPS – Encapsulated Post Script
- GIF – Graphics Interchange Format. (CompuServer Information Service, ursprünglich für Einwahlverbindungen)
- Group 3 und Group 4 Facsimile
- PICT – kann geometrische Daten beinhalten (Linien, Polygone, etc.)
- JPEG – Joint Photographic Experts Group, lossy (verlustig)

GRAPHIK Operationen:

- Editieren „paintbrush operations“, selektion
- Punkt Operationen
- Filtern
- Compositing
- Geometrische Transformationen
- Konvertierung

VIDEO:

Video ist eine Abfolge von Einzelbildern

Charakteristik von analogem Video

- Frame rate (25, ~30)
- Scan Lines (525, 625, >1000)
- Seitenverhältnis (4:3, 16:9)
- Interlacing (2:1, non-interlaced = progressive)
- Signal Qualität (Endverbraucher, Professionell, Übertragung)
- Standards für Fernsehen:
PAL: Österreich, Deutschland, Westeuropa, ...
NTSC: USA, Japan, ...
SECAM: Osteuropa, Frankreich

Bei Component Video gibt es getrennte Kanäle für Helligkeit und Farbe, was eine bessere Qualität ergibt, allerdings mehr Bandbreite und eine gute Synchronisation verlangt. Composite Video hat nur einen Kanal und daher eine schlechtere Qualität, wird aber wegen seiner einfacheren Handhabung zum Broadcast verwendet.

S-Video ist eine Kombination aus beiden.

NTSC verwendet:

- 525 Scan Lines pro Bild
- 29,97 Bilder pro Sekunde
- Interlaced Scan Lines teilen das Bild in 2 Teile, jeder teil hat 262,5 Lines
- YIQ Farbmodell

PAL verwendet:

- 625 Scan Lines pro Bild
- 25 Bilder pro Sekunde
- Interlaced Scan Lines teilen das Bild in 2 Teile, jeder teil hat 312,5 Lines
- YUV Farbmodell

Videoband Formate für Analoges Video

- VHS
- Betamax
- Video 8
- S-VHS
- Hi 8
- Umatic
- Betacam
- 1 inch

Video Ausrüstung:

- Routing switcher, (matrix switch) zum hin und her schalten zwischen mehreren Signalen; ein Input kann auf mehrere Outputs umgelegt werden.
- Timebase Correctors – Ausbessern von Schwankungen von VTR und VCR
- Sync Generator – Master Clock um sicher zu gehen dass alle Geräte mittels Impuls zugleich beginnen/stoppen, etc...
- Frame (delay) Buffer: Video wird auf externem Gerät gespeichert und passend zum Takt (Master Clock) ausgegeben.

Post Production – Video Production Switcher

Special Effects, Untertitel, n input Streams werden zu 1 Stream zusammengefasst.
Heute wird für die Post Production immer mehr Software eingesetzt.

Digitales Video:

Digital Video „Größen“:

- „rohes“ – „raw“ digital Video:
- 20 MByte / sek
- 70 GByte / h
- 1 GByte = 50 sekunden

- „low data rate“ digital Video:
- 200 kbit / sek – 5 Mbit / sek
- 1 GByte = 30min – 10h

Während analoges video ein kontinuierliches Signal ist, verwendet digitales Video diskrete numerische Werte.

Den Übergang vom analogem zum digitalen Video bezeichnet man als Sampling. Dabei wird das analoge Signal zu bestimmten Zeitpunkten abgetastet und für die jeweiligen Werte eine digitale Repräsentation gesucht. Die Frames werden dann durch ein Pixel-Array dargestellt.

Der CCIR 601 - Standard bezeichnet eine Norm zur Digitalisierung analoger Videosignale. Mithilfe von CCIR kann man sowohl PAL, als auch NTSC und Secam mit der gleichen Rate abtasten. CCIR 601 basieren auf dem im Fernsehen verwendeten YCC-Farbmodell, das dadurch charakterisiert ist, dass Y den Helligkeitskanal und die beiden Cs zwei Farbkanäle bezeichnen. Die einzelnen „Familien“ innerhalb des CCIR 601 - Standards unterscheiden sich durch unterschiedliche Werte für m:n:l. m ist die Basisrate, mit der Y abgetastet wird. n und l bestimmen, mit welcher Rate die beiden C-Kanäle abgetastet werden.

Einige Beispiele:

4:4:4 pro Pixel werden alle Werte abgetastet, ergibt höchste Qualität

4:2:2 nur von jedem 2. Pixel pro werden Farbwerte angetastet (für Broadcast)

4:1:1 von jedem vierten Pixel pro Zeile werden Farbwerte angetastet (ergibt VHS-Qualität)

Faktoren, die die Qualität von digitalem Video beeinflussen:

- Größe der Frames
- Framerate
- Art der Kompression

Digitale Video Formate:

- Digital Betacam
- Sony DV Format
- Panasonic / JVC DV Format
- Sony Betacam SX – MPEG Codierung
- DVI (Intel)
- Quick Time (Apple)
- Px64

Kompression:

- Verlustig vs. verlustfrei
- Echtzeit (symmetrisch)
- Spatial vs. temporal
- Skalierbar
- Art des Quellmaterials

Faktoren die die Video Qualität beeinflussen

- Framegröße und tiefe
- Framerate, Schlüsselbild Rate (Key Frame Rate)
- Algorithmus Parameter, Decodierungszeit
- Rate der Komprimierten Daten
- Größe des komprimierten Files

Möglichkeiten mit digitalem Video:

- Lagerung (HD, CD-ROM, RAID, DVTR)
- Editieren (non-linear)
- Digital video Effekte (Schlüsselbilder setzen, transitions,.....)

Digitales Video übertrifft in Zugriffszeit, Kapazität (nur limitiert durch Speichermedium), Flexibilität, Portabilität und Integration. Der einzige Nachteil den digitales Video gegenüber analogem Video hat sind die Kosten für die Lagerung.

AUDIO:

Was ist Schall? Ein kurzer Ausflug in die Akustik:

Physikalische Definition: Schall ist die wellenförmige Ausbreitung von Druckschwankungen in elastischen Medien, Periodisches Komprimieren und Dekomprimieren des Mediums.

Eigenschaften:
Amplitude, Frequenz

Unterschiede:

- Ton (eine einzige Sinusschwingung)
- Klang (Überlagerung von Grundton und Obertönen)
- Geräusch (kein Grundton mehr erkennbar, Ansammlung von verschiedenen Frequenzen)

Schalldruckpegel, Lautstärkepegel

Schall – Mathematische Beschreibung:

Schall beruht auf Schwingungen, diese lassen sich durch ihren zeitlichen Schwingungsverlauf beschreiben. -> Sinusschwingung ist am einfachsten.

Wird beschrieben durch:

- Frequenz (Anzahl der Schwingungen / sek)
- Periodendauer (Dauer der Schwingung)
- Amplitude (Druckwert bei maximaler Kompression)
- Wellenlänge (während der Periodendauer zurückgelegter Weg)
- Schallgeschwindigkeit (abhängig vom Ausbreitungsmedium)

Aus der Sinusschwingung lassen sich andere Schwingungen wie z.B. Rechteckschwingung, Dreieckschwingung ableiten.

Die Wellenlänge ist nicht direkt hörbar, spielt aber eine wichtige Rolle bei Hörphänomenen: Beugung: wenn Wellenlänge größer als das Hindernis: „um die Ecke hören“

Frequenz ist ein Maß für die Tonhöhe. Je höher die Frequenz desto höher der Ton.

Amplitude der Druckwelle = Lautstärke

Eigenschaften von Schall

Brechung: Schallwellen ändern Ihre Richtung sobald die Gebiete mit unterschiedlicher Schallgeschwindigkeit durchlaufen.

Dispersion: Für verschiedene Tonhöhen ist die Stärke der Brechung verschieden.

Der Schalldruckpegel wird in Dezibel gemessen. Er wurde eingeführt, um den großen Hörbereich der Lautstärke (der Unterschied zwischen Hörschwelle und Schmerzgrenze ist 1 : 2.000.000) sinnvoll beschreiben und darstellen zu können. Durch die Umrechnung in einen kleineren Wertebereich liegt die Hörschwelle dann bei 0dB und die Schmerzgrenze bei 120dB. Außerdem steigt der Schalldruckpegel dadurch logarithmisch an, was mehr unserem Lautstärkeempfinden entspricht.

Es gibt gewisse Frequenzen die das menschliche Ohr als besonders „laut“ empfindet obwohl die Lautstärke gleich laut ist wie bei den nicht so „lauten“ Frequenzen – deswegen auch: Einführung der Lautstärkepegel

Digitales Audio:

Ein Mikro wandelt eine Schallwelle in ein elektrisches Signal um.
Spannungsverlauf analog dem Druckschwankungsverlauf an der Membran des Mikros.

Analoge Audiotechnik:

Analoges Signal gespeichert auf einem Magnetband, Schallplatte,.....
Editieren: mechanisches Schneiden, überspielen

Digitale Audiotechnik:

Analoges Signal wird mittels Analog / Digital Konverters in binäres Zahlensystem umgewandelt, gespeichert, am Computer editiert, mittels Digital/Analog Konverter wieder als Analogsignal an Lautsprecher ausgegeben.

Vergleich Digital / Analog:

Vorteil Digital:

- Weniger Rauschen
- Höhere Dynamik
- Verlustloses Kopieren
- Höhere Linearität (Frequenzgang, Übertragungskurve, -> Klirrfaktor)
- Temperatur unempfindlich
- Keine Gleichlaufschwankungen

Nachteile Digital:

- Anfällig für Datenverlust
- Höhere Übertragungsbandbreiten
- Aufwendigere Hardware

Abtasten:

Prozedur: in periodischen Zeitintervallen wird dem analogen Eingangssignal eine Probe (=sample) entnommen.

Abtastfrequenz = 2 x die höchste im Signal vorkommende Frequenz
(Nyquist – Shannon Theorem)

Quantisierung:

Prozedur: Abgetasteten Spannungswerten werden diskrete Zahlenwerte zugeordnet:
Gesamtspannungsbereich in Quantisierungsintervalle Q unterteilt -> kontinuierlicher Wert nächstgelegenen Zahlen zugeordnet.
Qualität Q soll möglichst klein sein, ergibt sich aus der Länge des Datenworts.

Beim Quantisieren können auch Fehler auftreten. Wenn der Zahlenwert vom abgetasteten wert $\leq 0,5$ Q abweicht, wird der Fehler als Quantisierungsrauschen hörbar.

Schallwiedergabe:

Quantisierungsrauschen erzeugt einen Schalldruckpegel L(noise).
Maximales Nutzsignal Schalldruckpegel L(max)

Dynamikbereich:

$$\text{SNR} = L(\text{max}) - L(\text{noise})$$

SNR = signal to noise Ratio. (in dB)

Lineare und nicht lineare Quantisierung:

Lin. Quantisierung: Q konstant; übliche Methode in der Audiotechnik

Nicht Lin. Quantisierung: je größer der Wert umso größer Q.

Kleine Werte -> kleiner Quantisierungsfehler

Große Wert -> großer Quantisierungsfehler, aber Rauschen von großem Nutzsignal maskiert;

Auch mit kleiner Auflösung passable Qualität. Verwendung: Für Systeme mit geringer Übertragungsbandbreite.

Codierung:

Signale von Audio sind bipolar (pos, neg.)

Zweierkomplementcode: 0 und 1 wobei 0 ein neg. wert und 1 ein positiver wert ist.
Bei der Addition zweier Signale gibt es kein Offset (=Nullpunktverschiebung)

Pulscodemodulation:

Prozedur: Binärwerte werden seriell in Form von modulierten Spannungspulsen über eine einzige elektrische Leitung übertragen.

Datenrate: Abtastfrequenz x Bitauflösung = Datenrate

Bsp: mono – 44,1 kHz – 16 Bit → 705600 bit/s

Bandbreite: nötiger Frequenzbereich, um PCM verlustlos zu übertragen.

Bandbreite = Datenrate / 2

CD – mono: 352,8 kHz

Bei der Quantisierung werden den abgetasteten Werten diskrete Zahlenwerte zugeordnet. Dazu müssen die ermittelten ungenauen Werte auf einen diskreten Zahlenbereich abgebildet werden. Am besten erreicht man das, wenn man den Gesamtspannungsbereich in Quantisierungsintervalle unterteilt. Den analogen Werten werden dann die diskreten Werte des nächstgelegenen Intervalls zugeteilt. Natürlich kommt es dabei zu Rundungsfehlern, die sich dann beim Rekonstruieren als Quantisierungsrauschen bemerkbar machen. Wichtiger als wie laut dieses Rauschen ist, ist die Frage, wie laut es im Verhältnis zum eigentlichen Ton ist. Dieser Wert heißt SNR (signal-to-noise ratio); je größer er ist desto besser. Die lineare Quantisierung, bei der die Quantisierungsintervalle konstant gehalten werden, ist die übliche Methode in der Audiotechnik. Die nichtlineare Quantisierung ist aufwändiger, liefert aber bessere Ergebnisse. Hier wird die Größe der Intervalle variiert: je größer die Frequenz, desto mehr Intervalle gibt es.

Bei der anschließenden Codierung mittels PCM (Pulscodemodulation) werden die Binärwerte mittels elektrischer Impulse über eine einzige elektrische Leitung übertragen.

Attribute von digitalem Audio:

- Samplingrate
- Quantisierung
- Anzahl der Kanäle
- Interleaving
- Komprimierung

MIDI

Midi ist die Abkürzung für Musical Instrument Digital Interface, also eine digitale Schnittstelle für Musikinstrumente. Diese können sein ein Computer, Synthesizer, MIDI-Gitarren, usw..., die über diese Schnittstelle miteinander kommunizieren können. Eine MIDI-Leitung kann die Daten nur seriell übertragen und immer nur in eine Richtung. Ein Synthesizer ist ein Klangerzeuger, der oft über Prozessor, Keyboard und Bedienfeld verfügt. Ein Sequenzer wird benötigt, entweder als Hardware oder als Softwareimplementation, um MIDI-Daten zu speichern.

Der Sequenzer hat mehrere MIDI-Eingänge und -Ausgänge. Eine MIDI-Schnittstelle verfügt über 16 Kanäle. Jedes MIDI-Gerät sendet oder empfängt nur auf bestimmten Kanälen Nachrichten, so dass man die verschiedenen Klangerzeuger gezielt ansprechen kann.

Achtung! MIDI selbst erzeugt keine Klänge, MIDI ist nur ein Steuersignal!

MIDI wird zu Synchronisation von mehreren Synthies, Keyboards oder anderen Instrumenten verwendet. Genauso können auch Video mit Audio synchronisiert werden.

Teil 2:

Multimedia Environments:

Die Videodisc:

Die Videodisc steht am Scheitelpunkt der digitalen und analogen Medien, ist aber selber ein digitales Medium. Alle Videodiscs sind flache Scheiben und vom Durchmesser mit einer Langspielplatte zu vergleichen. In den siebziger Jahren konkurrierten verschiedene Videodisc-Systeme miteinander. Das Problem aller Systeme bestand darin, gewaltige Informationsmengen kodiert auf relativ begrenzten Flächen unterzubringen. Das System, das am meisten verbreitet war, ist LV (Laser Vision). Dabei wurde ein Laser verwendet, um kleine pits (Vertiefungen) in die Oberfläche der einer Master-Disc zu brennen. Die pits waren in einer langen Spirale, die in der Mitte startete, angeordnet. Von der Master-Disc wurde ein Abdruck gemacht, und aus diesem wurden die Discs für den Verkauf gegossen. Gelesen wurden die Videodiscs von einem Laser, der die pits abtastete und dadurch in binäre Signale umwandelte. Logischerweise waren diese Discs read-only.

LV – discs werden in 2 Formaten angeboten:

CLV – constant linear velocity

CAV – constant angular velocity

CAV Discs haben ein flexibles Playback: Pause (Still, Bild einfrieren), Play, Forward, Reverse in verschiedenen Geschwindigkeiten.

CLV Discs haben die doppelte Kapazität

Die CD

CD Audio wurde von Philips und Sony entwickelt, in Anlehnung an die LV Videodisc

Unterschiede:

Echtes digitales Medium (LV benutzt nur analoge Informationen)

Die CD verwendet NUR eine konstante lineare Geschwindigkeit (CLV)

Wie bei der Videodisc wird die Information in pits gespeichert, die bei der Herstellung auf eine Aluminiumschicht gepresst werden. Die Aluminiumschicht wird mit einem Schutzlack versiegelt und ist damit vor Beschädigungen geschützt. Gelesen werden die Daten ebenfalls durch ein Laser-System von innen nach außen gelesen. Der erste Track befindet sich also fast im Zentrum der CD ganz im Gegensatz zur Schallplatte.

CD Familie:

- CD-DA – Compact Disc – Digitales Audio
- CD+G, CD+MIDI
frühes CD Format, welches Text, einfache Graphiken und MIDI in den 6 nicht verwendeten subcode bits (R-W) beinhaltet, playback ist auf jedem CD Player möglich.
- CD – ROM
- Photo CD
- Photo CD master – 35mm, bis zu 100 Bilder
- Pro Photo CD Master – 35mm, 70 mm, 120mm, 6-100 Bilder
- Photo CD Portfolio, audio, text, graphik, und Bilder in schlechterer Qualität, bis zu 800 Bilder
- Photo CD Catalog, bis zu 600 Bilder in schlechter Qualität
- CD-R, CD-WO – Compact Disc Recordable / Write Once
- CD-ROM XA – Extended Architecture (offene Form von CD-i)
- CD-I – Compact Disc – Interactive
- Video CD
- CD-V – Compact Disc – Video (hybrid digital audio / analog video)

CD - ROM

Unterschied CD-Rom und CD-DA – Es gibt eine zusätzliche Schicht für Fehlererkennung und Korrektur

Data Block – Die disc ist in consecutive data blocks unterteilt (98 frames, 2352 bytes)

2 Block Formate

- Mode 1 – 2048 bytes application Data
- Mode 2 – 2333 bytes application Data

DVD

Als man mit der CD speziell bei Video-Wiedergabe sehr schnell an ihre Grenzen stieß, entwickelten Sony und Philips zusammen mit JVC aus dem SD/MMCD-Standard die DVD (Digital Versatile Disc). Auf der DVD können die Daten nicht nur viel dichter als auf einer CD gespeichert werden, sondern auch in mehreren Layern und sie kann beidseitig beschrieben werden. Sie hat daher den 7fachen Speicherplatz einer CD.

Eine DVD kann ein oder zweiseitig beschrieben sein. Sofern man keinen Player besitzt der die ober und unterseite der DVD nacheinander lesen kann muss man die DVD am Ende einer Seite natürlich im Laufwerk umdrehen. Nachteil an der beidseitig beschriebenen DVDs ist auch dass die „Schutzschicht“ wegfällt die bei einer einseitig beschriebenen DVD oder CD das CD – Label bot.

DVD Formate:

- DVD-5
- DVD-9
- DVD-10
- DVD-18
- DVD-R
- DVD-RAM (CAV, 30 Jahre lebensdauer, kann bis zu 100000 beschrieben werden)
- DVD-RW (CLV, PC-Backup, kann das Videotape ersetzen, bis zu 1000 mal wieder beschreibbar)
- DVD+RW (CAV für daten, CLV für Video, Kompatibel mit DVD-Roms)

DVD+RW wird vom DVD Forum nicht als Standard anerkannt!

4 Arten von Daten:

- Video – 1 Stream; Encoding in MPEG-1 & MPEG-2
- Audio – bis zu 8 streams: Inkludiert Dolby Digital, MPEG-1 & MPEG-2, LPCM und DTS
- Subpictures – Graphiken die über den Film gelegt werden, werden oft für Untertitel verwendet (bis zu 32 Sprachen) und natürlich für Menüs
- Navigation – Program chains, Player Kontrolle, Menüs
- Bit Rate (Maximum: 9,8 MB/s (video+audio+subpictures. Mittelwert: 4MB/s)

CD-i

Die Compact Disc Interactive ist ein von Philips entwickelter Standard. Auf CD-i-Playern lassen sich Audio-, Photo- und VideoCDs auch interaktive CDs mit Spielen oder Lernsoftware abspielen. Dazu benötigt man spezielle Hard- und Software. Videos werden im MPEG-1-Standard gespeichert.

DVI

DVI (Digital Video Interactive) besteht aus Algorithmen zur Kompression und Dekompression von Videos in Echtzeit, einer Benutzerschnittstelle und einem festgelegtem Datenformat. Entwickelt wurde es von IBM und Intel.

Quick Time

Allgemeines:

QuickTime ist eine Software-Erweiterung des Macintosh-Systems. Es verfügt über die Fähigkeit, kontinuierliche Mediendaten aufzunehmen, diese zu speichern und zu verwalten, sie zu synchronisieren und anzuzeigen.

Medientypen:

Video und Bilder – basierend auf Apples PICT Format, beinhaltet viele verschiedene kompressions Algorithmen

MPEG-1 Compression, compact video compression, video compression, animation compression, photo compression, graphics compression, raw compression

Audio:

Bandbreite der Sample Rate maximal 65 kHz

Mono oder stereo

PCM mit sample grÖÙe von 8, 16 oder 32 bits.

Macintosh Audio Compression und Expansion (MACE), erlaubt 3:1 oder 6:1 kompression, verlustig

Text und Graphik:

Wird inkludiert in PICT daten oder Filmen, grundsätzlich wird dies von QuickTime ignoriert.

QuickTime verwendet digitalisiertes Video als Standarddatentyp im System und erlaubt eine einfachere Handhabung anderer kontinuierlicher Medien, wie Audio und Animation. Alle Arten von kontinuierlichen Medien werden in einem *Filmdokument* gespeichert. Hierin werden weiterhin Daten, wie bspw. die Information, wann ein Film erzeugt und modifiziert worden ist, oder seine Dauer verwaltet. Mit jedem Film ist ein Poster-Frame assoziiert, der in der Dialogbox von QuickTime erscheint.

Eine Spur repräsentiert einen Informationsstrom (von Audio- oder Videodaten), der parallel zu jeder anderen Spur läuft. Mit jeder Spur werden Informationen gespeichert, wie das Erzeugungs- und Modifikationsdatum, die Dauer, die Spurnummer, räumliche Eigenschaften (Transformationsmatrix, Anzeigefenster, Clipping-Region), eine Liste von verbundenen Spuren, die Lautstärke und Startzeit, die Dauer, die Abspielrate und eine Datenreferenz für jedes *Mediensegment*.

Ein Mediensegment ist eine Menge von Referenzen auf Audio- und Videodaten, die zeitliche Informationen (Erzeugung, Modifikation, Dauer), die Sprache, die Anzeige- und Klangqualität, den Mediendatentyp und Datenzeiger beinhalten. Zukünftige Weiterentwicklungen werden neben Audio und Video sog. „*Custom Tracks*“ beinhalten, bspw. eine Spur für Untertitel. Alle Spuren können parallel betrachtet oder angehört werden. Spuren eines Films sind dabei immer synchronisiert. Da Filme Dokumente darstellen, können sie nicht nur abgespielt (inklusive Pausen oder Sprüngen innerhalb des Films), sondern auch editiert werden.

Editieroperationen sind bspw. das Ausschneiden, Kopieren und Einfügen von Daten. Filmdokumente können auch Teil anderer Dokumente sein. QuickTime ist weiterhin skalierbar und erlaubt die Verwendung von Hardware-Komponenten, wie Beschleuniger- oder Kompressor/Dekompressorkarten.

QuickTime Media Organisation

Zeitgebung ist explizit

Werte werden von „unsigned integers“ repräsentiert.

Time Scale – Units pro Sekunde

Dauer – Maximaler Zeit Wert

Timebase – Playback rate

Zeitkoordinaten System (TCS, scale und dauer)

Konzeptionelles Level:

- Data Entity – Repräsentation der aktuellen Lagerung die für a/v Daten verwendet wird.
- Media Entity – temporäre Sequenz gemessen in media time
- Track Entity
- Movie Entity

Physische Organisation der Daten:

Atom – Grundeinheit, besteht aus:

- Größe (in bytes)
- Code aus 4 characters (dient zur identifizierung des Typs)
- Inhalts Section (Felder oder andere Atome)

Die QuickTime-Architektur besteht aus drei wichtigen Komponenten (siehe Abb. 11-8): der *Film-Toolbox*, die eine Menge an Diensten anbietet, die es dem Benutzer erlauben, Filme in Anwendungen zu verwenden. Dieser kann dann Charakteristika von Audio- und Videodaten eines Films direkt manipulieren. Ein Film wird so in die Desktop-Umgebung des Benutzers integriert. Filmdateien können mit dem System-Clipboard importiert und exportiert werden; das Editieren eines Films erfolgt in der Film-Toolbox.

Die zweite Komponente, der *Kompressionsmanager*, bietet eine allgemeine Schnittstelle für die Datenkompression und -dekompression zur und von der Festplatte, CD-ROM und Floppy an, die unabhängig von der jeweiligen Implementierung arbeitet. Hier wird auch ein Verzeichnisdienst angeboten, der bei der Auswahl der richtigen Kompressionskomponente hilft. Der Kompressionsmanager unterstützt verschiedene Schnittstellenebenen für unterschiedliche Anwendungserfordernisse. Die Kompressionstechniken bestehen in einer proprietären Bildkompression, einem proprietären Kompressor für digitalisierte Videodaten (wodurch man einen Kompressionsfaktor von 8:1 erhält bzw. 25:1, wenn auch die zeitlichen Abhängigkeiten beseitigt werden) und einer JPEG-Implementierung. Ein Animationskompressor kann digitale Daten im verlustbehafteten und im verlustfreien (fehlerfreien) Modus komprimieren.

Weiterhin ist in QuickTime auch ein Grafikkompressor verfügbar. Die Umwandlung der Pixeltiefe in Bits pro Pixel kann hierzu als Filter zusätzlich zu anderen Kompressoren angewendet werden.

QuickTime Software Architektur:

Movie Toolbox

- Ein high Level Interface für playback, editieren, aufnehmen und Modifizierung des Mediums, Tracks oder Films.
- Beinhaltet ein Movie Controller Fenster
- Komponenten werden für die jeweilige Vorgänge konfiguriert (e.g. decompression)
- Bild Compression Manager
- Component Manager (Applications oder die Toolbox können instanzen von Komponenten Typen erstellen. Der CM ist zuständig um die Einzelkomponenten zusammen zu halten und den Clients zu erlauben untereinander verbunden zu bleiben und zu kommunizieren)

QuickTime Component Examples

- Movie Controller Components (Filme können unter der Verwendung eines Standard User Interfaces abgespielt werden)
- Standard Bildkomprimierung (Dialog Komponenten)
- Sequence Ausleser
- Mediendaten Austausch
- Clock (Taktgeber)
- Vorschau
- Etc etc

Das Komponentenkonzept in QuickTime erlaubt eine einfache Erweiterung, ohne daß Anwendungen betroffen werden. Hierbei wird versucht, eine hierarchische Funktionalitätsstruktur durch Komponenten zu erreichen. Die Komponente des Film-Controllers vereinfacht die Schnittstellenprogrammierung durch den Benutzer. Ein Nachteil von QuickTime ist, daß keine klare Schichtenbildung von Abstraktionen für Programmierer besteht und daß die Funktionalität der Manager und Komponenten manchmal überlappt.

Teil 3:

Kompression:

Es gibt drei Arten der Komprimierung:

1. entropy-coding-Verfahren sind verlustfrei (lossless) Codierverfahren, d.h. nach der Dekompression sind die dekomprimierten Daten identisch zu den Ursprungsdaten. Die entropy-coding-Verfahren erzielen bei Bild- und Tonkomprimierung nicht besonders große Kompressionsraten, sind daher mehr für Textdateien oder ähnlich strukturierte Daten mit häufig wiederholten Bitfolgen geeignet.
2. source-coding-Verfahren sind verlustbehaftet (lossy) und erzeugen ein den unkomprimierten Daten ähnliches Abbild. Aus den komprimierten Daten lässt sich das Original nicht mehr (ohne Zusatzinformationen) wiederherstellen. Bei steigender Komprimierungsrate verlieren die komprimierten Daten immer mehr an Ähnlichkeit gegenüber den Originaldaten.
3. hybrid-coding-Verfahren kombinieren in geeigneter Weise verlustreiche und verlustfreie Komprimierungen.

1. Entropy coding / Run Length Encoding

Inhaltsabhängiges Coding

RL – Encoding ersetzt eine Sequenz von gleichen aufeinanderfolgenden Bytes mit der Anzahl der vorkommenden Häufigkeit.

Die Anzahl der vorkommenden Häufigkeit wird durch ein spezielles Zeichen (!) angezeigt.

Vorgangsweise:

Wenn dasselbe Byte mindestens 4 mal hintereinander vorkommt zähle die Häufigkeit der selben Bytes

Schreibe die komprimierten Daten in das Format, bei den Bytes die eben öfter vorkommen schreibe das Byte einmal, danach ein ! und danach die Anzahl wie oft das Byte hintereinander in der unkomprimierten Zeichenkette vorkommt:

Beispiel: (unkomprimiert)

ABCCCCCCCCDEEEFFFGGGGG

(komprimiert)

ABC!9DE!4FFFG!6

Es existieren auch Variationen dieser Methode speziell für z.B. Text

Diese Kompression verwendet z.B. das tif oder pcx Format

Statistisches Encoding

Bei der statistischen Codierung werden häufig zu erwartende Symbole durch kurze Codewörter und seltenere Symbole durch längere Codewörter beschrieben.

Huffman Encoding:

Eine Form der statistischen Codierung ist die Huffman-Codierung. Es ist ein Kompressionsverfahren zur verlustfreien Kompression von Daten, wobei es sich die Häufigkeit von Informationen zunutze macht. Beim Codieren wird ein Binärbaum aufgebaut. Dieses Verfahren wird von einer Vielzahl von Kompressionsalgorithmen oft unterstützend verwendet, z.B. gif oder zip. Unter bestimmten Umständen (z.B. Gleichverteilung der Worthäufigkeiten) kann die Huffman-Codierung allerdings keine Kompression bewirken.

JPEG – Encodierung

Die JPEG-Codierung (Joint Photographic Experts Group) ist ein Komprimierungsverfahren auf Grundlage der Diskreten Cosinus Transformation (DCT).

Kurz und bündig gesagt macht die JPEG Codierung folgendes:

1. Blockeinteilung in 8x8 Pixel-Blöcke die dann als Vektor der Dimension 64 interpretiert und mit DCT (Diskreter Cosinus Transformation) in den Frequenzbereich transformiert werden.
2. Quantisierung des DCT-Ergebnisvektors
3. Codierung des quantisierten Vektors mittels Huffman- oder arithmetischer Codierung.

Die Quantisierung (Schritt 2), der eigentliche Komprimierungsschritt, sieht im einzelnen folgendermaßen aus:

Das zu komprimierende Bild wird in 8x8 Pixel große Blöcke unterteilt. Diese 64 Pixel werden als Vektor der Dimension 64 aufgefasst und mit DCT in den Frequenzbereich transformiert. Die 64 Einträge des DCT-Ergebnisvektors werden in eine Zick-Zack Linie gebracht. Jeder "Frequenzwert" entlang dieser Zick-Zack-Linie wird ganzzahlig durch den entsprechenden Wert des Quantisierungsvektors dividiert, wodurch sich eine Gewichtung ergibt.

Bemerkungen:

Applikationen müssen nicht Decoder und Encoder implementiert haben wenn der Komprimierungsprozess einer gewöhnlichen (standardisierten) Tabelle entspricht.

Der Encodierte Stream hat ein fixes Interchange Format

- Encodierte Bild Daten
- Gewählte Parameter
- Tabellen des Codierungs Prozesses

Normalerweise inkludiert das Interchange Format alle Informationen die für das Decodieren notwendig sind ohne vorher etwas über den Coding Vorgang zu wissen.

JPEG – Längere Erklärung:

Anwendung auf farbige und grauskalierte Bilder

Durch schnelle Kodierung und Dekodierung von Einzelbildern können auch bewegte Sequenzen bearbeitet werden. → Motion JPEG

Teile von JPEG sind heute schon als reines Software Paket (in jedem Graphikprogramm, z.B. Photoshop) oder mit spezieller Hardware Unterstützung verfügbar. Achtung! Nur die unbedingt notwendigen Algorithmen werden in den gängigen Produkten unterstützt. Daher ist im weitesten Sinne nur der „Basis-Modus“ (mit einer bestimmten Vorgabe der Bildaufbereitung – Ebenen und Farbkodierung sind begrenzt) von JPEG erhältlich.

Was soll JPEG können?

- Die Unabhängigkeit der Bildgröße soll gewährleistet werden.
- Es muß für jedes Verhältnis Höhe zu Breite eines Bildes und Höhe zu Breite eines Pixels verwendbar sein.
- Der verwendete Farbraum und die Farbvielfalt müssen unabhängig voneinander sein.
- Der Bildinhalt muß eine beliebige Komplexität und beliebige statistische Eigenschaften des Bildes aufweisen können.
- Es muß dem Stand der Technik in bezug auf Kompressionsgrad und erreichbarer Bildqualität entsprechen (oder nahezu entsprechen).
- Die Komplexität der Verarbeitung muß eine Software-Lösung auf einer großen Menge verfügbarer Standardprozessoren ermöglichen und soll mit Spezial-Hardware drastisch reduziert werden können.
- Bilder sollen beim Dekodieren sequentiell (Zeile für Zeile) und progressiv (ein Bild wird als Ganzes immer weiter verfeinert) bearbeitet werden können. Auch eine verlustfreie und eine hierarchische Kodierung mit unterschiedlichen Auflösungen desselben Bildes sollen möglich sein.

Der Anwender kann aufgrund der Parametrisierbarkeit eine individuelle Abwägung zwischen der Qualität des reproduzierten Bildes, der Dauer der Kompression und der Größe des komprimierten Bildes vornehmen.

JPEG unterscheidet vier Modi, die selbst noch verschiedene Varianten enthalten:

- Der *verlustbehaftete, sequentielle DCT-basierte Modus (Baseline Process, Basis-Modus)* muß von jedem JPEG-Dekoder unterstützt werden.
- Der *erweiterte verlustbehaftete DCT-basierte Modus* stellt eine Menge weiterer Alternativen zum Basis-Modus zur Verfügung.
- Der *verlustfreie Modus* zeichnet sich durch einen wesentlich geringeren Kompressionsfaktor aus und ermöglicht eine fehlerfreie Rekonstruktion des Originalbildes.
- Der *hierarchische Modus* beinhaltet Bilder verschiedener Auflösungen und verwendet dazu Algorithmen, die in den anderen drei Modi definiert wurden.

Techniken die der Baseline Process verwendet:

- Block
- MinimumCodedUnit (MCU)
- FDCT
- Run-Length
- Huffman

Zur Bildaufbereitung stellt JPEG ein Modell vor, das möglichst viele allgemein verwendete Arten von Standbildern beschreibt. Hierbei geht man bspw. nicht explizit von drei Bildebenen mit einer 9-bit-YUV-Kodierung und einer festen Anzahl von Zeilen und Spalten aus. Auch die Abbildung der kodierten Farbwerte auf die tatsächlich dargestellten Farben wird nicht mitkodiert. Damit wird man der für JPEG aufgestellten Forderung der Unabhängigkeit von Bildparametern wie der Bildgröße oder des Seitenverhältnisses des Bildes und der Pixel gerecht.

Die Auflösung der einzelnen Ebenen kann unterschiedlich sein. Ein grau skaliertes Bild wird demnach meist nur aus einer Komponente bestehen, ein RGB-Farbbild hingegen aus drei Komponenten mit jeweils derselben Auflösung (gleiche Anzahl von Zeilen $Y1 = Y2 = Y3$ und derselben Anzahl von Spalten $X1 = X2 = X3$). YUV-Farbbilder mit Unterabtastung der Chrominanzkomponenten verwenden, bezogen auf die Bildaufbereitung von JPEG, drei Ebenen mit $Y1 = 4Y2 = 4Y3$ und $X1 = 4X2 = 4X3$. Jedes Pixel wird mit p bit dargestellt und kann folglich die Werte von 0 bis 2^{p-1} annehmen. Sämtliche Pixel aller Komponenten eines Bildes müssen die gleiche Anzahl Bits aufweisen. Die verlustbehafteten Modi von JPEG gehen von einer Auflösung von entweder 8 oder 12 bit pro Pixel aus. Der verlustfreie Modus läßt eine Anzahl zwischen minimal 2 und maximal 12 bit pro Pixel zu. Eine JPEG-Anwendung, die Bilder mit davon abweichender Anzahl an Bits verwendet, muß nach einem ihr geeignet erscheinenden Verfahren eine Transformation des Bildes auf eine der hier festgelegten Anzahl von Bit pro Pixel vornehmen.

In die komprimierten Daten werden nicht die Werte X_i und Y_i aufgenommen, sondern die Werte X (Maximum aller X_i) und Y (Maximum aller Y_i), sowie die Faktoren H_i und V_i . Für jede Ebene stellen H_i und V_i die relative Auflösung in bezug auf die minimalen Auflösungen dar.

H_i und V_i müssen ganzzahlig sein und dürfen nur die Werte 1 bis 4 annehmen. Diese zunächst umständlich erscheinende Festlegung wird bei der Verzahnung verschiedener Komponenten (*Interleaving*) benötigt.

Zur Kompression erfolgt bei der Bildaufbereitung eine Zerlegung des Bildes in Dateneinheiten. Der verlustfreie Modus versteht unter einer Dateneinheit genau ein Pixel. Die verlustbehafteten Modi verwenden Blöcke zu je 8×8 Pixeln als Dateneinheiten. Dies beruht auf der hier verwendeten DCT, die immer zusammenhängende Blöcke transformiert

Bisher werden die Dateneinheiten meistens Ebene für Ebene aufbereitet und in dieser Reihenfolge der folgenden Bildverarbeitung angeboten. Pro Ebene werden die Dateneinheiten pro Zeile jeweils von links nach rechts bearbeitet. Dies ist eine nicht-verzahnnte Bildaufbereitung (*non-interleaved data ordering*).

Ein über mehrere Ebenen verzahntes Bearbeiten der Dateneinheiten erfolgt auf Grund der endlichen Verarbeitungsgeschwindigkeit des Dekodierers. Bei einem RGB-kodierten Bild mit extrem hoher Auflösung und nicht-verzahnnter Bildaufbereitung könnte die Wiedergabe auf dem Bildschirm erst nur rot, anschließend rot-grün und dann erst in den korrekten Farben erfolgen.

Es werden die *Minimum Coded Units (MCUs)* in der folgenden Reihenfolge gebildet:

Im verzahnten Modus werden mehrere Dateneinheiten verschiedener Ebenen zu Minimum Coded Units zusammengefaßt. Haben alle Komponenten die gleiche Auflösung ($X_i \times Y_i$), so enthält eine MCU von jeder Komponente genau eine Dateneinheit. Beim Dekodieren wird das Bild MCU-weise zur Anzeige gebracht. Hiermit erfolgt selbst bei nur teilweise dekodierten Bildern eine korrekte Farbdarstellung.

Gemäß JPEG können maximal vier Komponenten verzahnt kodiert werden. Dies ist keine Einschränkung, weil für Farbbilder im allgemeinen drei Komponenten für die Darstellung verwendet werden. Jede MCU darf maximal zehn Dateneinheiten beinhalten. Es können jedoch innerhalb eines Bildes einige Komponenten verzahnt und andere nicht-verzahnt kodiert werden.

Bildverarbeitung

Als erster Schritt der Bildverarbeitung erfolgt in diesem grundlegenden Modus (*Baseline Process*) eine Transformationskodierung mit Hilfe der Diskreten Kosinus-Transformation DCT. Dafür wird der Wertebereich der Pixel in das um den Nullpunkt symmetrische Intervall $(-128, 127)$ verschoben. Diese Dateneinheiten mit je 8×8 verschobenen Pixelwerten werden als Syx bezeichnet, wobei x und y im Bereich zwischen null und sieben liegen.

Quantisierung

Die sich an die Bildverarbeitung anschließende *Quantisierung* aller DCT-Koeffizienten arbeitet ebenfalls verlustbehaftet. Die JPEG-Anwendung stellt für den folgenden Schritt eine Liste mit 64 Einträgen bereit. Jeder Eintrag wird zur Quantisierung eines der 64 DCT-Koeffizienten verwendet. Damit kann jeder der 64 Koeffizienten separat eingestellt werden; die Anwendung kann damit die relative Bedeutung der verschiedenen Koeffizienten beeinflussen. Bestimmte Frequenzen lassen sich so höher bewerten als andere; dies sollte je nach Beschaffenheit des zu komprimierenden Bildmaterials erfolgen. An dieser Stelle wird der mögliche Kompressionsgrad auf Kosten der erreichbaren Bildqualität beeinflusst.

Entropiekodierung

JPEG bietet für die eigentliche Entropiekodierung die *Huffman-Kodierung* und die *arithmetische Kodierung* an. Für den in diesem Abschnitt betrachteten verlustbehafteten sequentiellen DCT-basierten Basismodus ist nur die Huffman-Kodierung definiert. In beiden Arten der Entropiekodierung findet vorher für die quantisierten AC-Koeffizienten eine Lauflängen-Kodierung von Nullwerten statt. Zusätzlich werden die von Null verschiedenen AC-Koeffizienten, ebenso wie die DC-Koeffizienten in eine spektrale Repräsentation umgewandelt, um die Daten weiter zu komprimieren. Die hierzu erforderliche Anzahl von Bits hängt vom Koeffizientenwert ab. Ein von Null verschiedener AC-Koeffizient wird mit 1 bis 10 bits repräsentiert. Zur Repräsentation der DC-Koeffizienten wird eine höhere Auflösung von 1 bis maximal 11 bit verwendet.

Damit wird eine Darstellung gemäß dem *ISO-Intermediate-Symbol-Sequence-Format* erzeugt, das im wesentlichen alternierend drei Angaben vorsieht:

1. die Anzahl der folgenden Koeffizienten mit dem Wert Null,
2. die für die Darstellung des danach folgenden Koeffizienten benutzte Anzahl an Bits,
3. den Wert des Koeffizienten, dargestellt mit der angegebenen Anzahl an Bits.

Von Vorteil bei der anschließenden Huffman-Kodierung ist die kostenfreie Verwendbarkeit, weil sie nicht durch Patente geschützt ist. Nachteilig ist die Forderung, daß die Anwendung die Kodierungstabellen bereitzustellen hat, weil JPEG keine Tabellen vorschreibt. Es können in diesem Modus je zwei unterschiedliche Huffman-Tabellen für die AC-Koeffizienten und für die DC-Koeffizienten verwendet werden.

In dieser sequentiellen Kodierung wird in einem einzigen Durchlauf das gesamte Bild kodiert und dekodiert. Bei der Dekodierung mit sofortiger Darstellung auf dem Bildschirm ergibt sich der Ablauf: Das Bild wird komplett von oben nach unten aufgebaut.

(btw: wer das versteht, ist wirklich gut. ;-)

MPEG – Motion Picture Experts Group

Mpeg Standards

MPEG-1 (Zielgruppe: niedrige Datenraten – VHS bei 1.5 Mbit/s [1992])

MPEG-2 (Zielgruppe: hohe Datenraten – hohe Qualität – Studio Qualität bis zu 15Mbit/s [1994])

MPEG-4 (Zielgruppe: sehr niedrige Datenraten - <64 kbs – kleine Bilder [1998])
(MPEG-7 und MPEG-21)

Was will man mit MPEG erreichen?

- Videodaten in akzeptabler Qualität liefern, wobei die Komprimierung zwischen 1.0 und 1.5 Mbit/s liegen sollte (MPEG-1)
- Unterstützung von symmetrischer/asymmetrischer Komprimierungs- und Dekomprimierungsapplikationen
- Audio und Video können synchronisiert wiedergegeben werden.
- Schneller Zugriff auf zufällig gewählte Teile des Videos (random access) ist genauso wie schnell vorwärts / rückwärts, Pause etc gewährleistet.
- Extrem schlechtes Verhalten bei Datenfehlern wird vermieden.
- Sofern benötigt, soll die Komprimierungs- / Dekomprimierungs- Verzögerung kontrollierbar sein.
- Editierbarkeit durch andere Applikationen (wenn benötigt) soll gewährleistet sein.
- Genügend Flexibilität in Hinsicht auf das Format ist gewährleistet, um das Video in Fenstern abzuspielen.
- Die Rechenzeit sollte keine „low-cost chipsets“ ausgrenzen die bereits fähig wären das Video in „real time“ zu encodieren.

MPEG – Grundsätzliche Informationen

- MPEG definiert Audio und Video Codierung und System Daten Streams mit Synchronisation.
- MPEG bedient sich explizit anderer Standards z.B. JPEG.

Ein MPEG Stream beinhaltet Informationen über

- Das Seitenverhältnis des Videos (1:1, 4:3, 16:9)
- Die Refresh Frequenz (8 Frequenzen definiert: 23. 976 Hz, 24 Hz, 25 Hz, 29. 97 Hz, 30 Hz, 50 Hz, 59. 94 Hz and 60 Hz)

Audio und Video Komprimierung wurde so designed dass sie problemlos bei der Geschwindigkeit eines CD-Roms (Video – ca. 1.150 Mbit/s - Audio – ca. 0.256 Mbit/s) funktioniert.

Die Audiokomprimierung arbeitet mit einem „subband coder“ mit einem psychoakustischen Modell um die bitrate zu reduzieren.

Die Videokomprimierung verwendet einen „block transform coder“ mit bewegungskompensierendem „inter- frame coding“.

MPEG Video Standard

Ein Bild muß aus drei Ebenen (Komponenten) bestehen: Neben der Luminanz Y sind dies die beiden Farbdifferenzkomponenten Cr und Cb (ähnlich dem YUV-Format). Die Luminanzkomponente hat in horizontaler und vertikaler Richtung die doppelte Auflösung wie die anderen beiden Ebenen (Color-Subsampling), dabei sollte die Luminanzkomponente die Anzahl von 768 x 576 Pixel nicht überschreiten. Es wird mit einer Tiefe von 8 bit pro Pixel in jeder Ebene gearbeitet.

Eine Prädiktion von Einzelbildern wirkt sich meistens in einer erheblichen Reduktion der Datenmenge aus. Bildbereiche mit unregelmäßig starken Bewegungsabläufen werden hier jedoch nur zu einer der Intraframe-Kodierung ähnlich großen Datenmenge reduziert. Die Verwendung von zeitlichen Prädiktoren erfordert die Speicherung einer Vielzahl von vorher bestimmten Informationen und Bilddaten. Deshalb kann in Abhängigkeit des erforderlichen Speicherplatzes und des erreichbaren Kompressionsgrades abgewogen werden, in welchem Umfang die Prädiktion eingesetzt wird.

Die Prädiktion ist jedoch meistens nur für Bildbereiche (und nicht für das gesamte Bild) sinnvoll. Das Bild wird deshalb aus solchen Bildbereichen (*Makroblöcke*) zusammengesetzt. Ein MPEG-Makroblock wird zu je 16×16 Pixel in der Luminanz-Komponente, inklusive je 8×8 Pixel in den beiden Chrominanzebenen, aufgeteilt. Diese Größen eignen sich zur Kompression auf der Basis einer Bewegungsschätzung besonders gut. Sie sind ein Kompromiß zwischen dem Rechenaufwand, der zur Schätzung betrieben werden muß, und der sich aus der Schätzung ergebenden Datenreduktionsrate. Innerhalb eines Makroblocks werden die 6 Blöcke zu je 8×8 Pixel sequenzialisiert; zuerst werden die vier Luminanzblöcke, anschließend die zwei Chrominanzblöcke gebildet. Es existieren keine vom Anwender zu definierenden MCUs wie in JPEG, weil der Bildaufbau über die vorgeschriebene Bildwechselfrequenz innerhalb von maximal 41,7 ms erfolgt. Die drei Komponenten werden gemeinsam komprimiert und dekomprimiert. In MPEG kann aus Benutzersicht ein progressiver Bildaufbau keine wesentlichen Vorteile gegenüber einem sequentiellen Aufbau besitzen.

Zur Bildverarbeitung müssen in MPEG vier Arten der Kodierung eines Bildes unterschieden werden. Der Grund hierfür liegt in der widersprüchlichen Anforderung einer effizienten Kodierung bei gleichzeitigem wahlfreiem Zugriff. Um einen hohen Kompressionsgrad zu erreichen, müssen zeitliche Redundanzen aufeinanderfolgender Bilder ausgenutzt werden. Der schnelle wahlfreie Zugriff bedingt eine Einzelbildkodierung. Im einzelnen werden folgende Bilder unterschieden, wobei der Begriff „Bild“ hier als Synonym zu „Einzelbild“ bzw. „Frame“ verwendet wird:

- *I-Bilder (Intra Coded Pictures)* werden ohne zusätzliche Informationen bzgl. anderer Einzelbilder kodiert (Intraframe-Kodierung). Ein I-Bild wird wie ein Standbild behandelt; hier wurde in MPEG auf die Ergebnisse von JPEG zurückgegriffen. Die Kompression muß jedoch im Gegensatz zu JPEG auch in Echtzeit möglich sein. Die Kompressionsrate ist deshalb am geringsten. I-Bilder bilden die Anker für den wahlfreien Zugriff.
- *P-Bilder (Predictive Coded Pictures)* benötigen zur Kodierung und Dekodierung Informationen bzgl. vorangegangener I- oder P-Bilder. Zur Dekodierung ist hier die Dekompression des vorherigen I-Bildes und aller eventuell dazwischenliegenden P-Bilder nötig. Dafür ist die Kompressionsrate gegenüber I-Bildern erheblich höher. Ein P-Bild läßt den Zugriff auf das folgende P-Bild zu, wenn dazwischen kein I-Bild liegt.
- *B-Bilder (Bidirectionally Predictive Coded Pictures)* benötigen zur Kodierung und Dekodierung Informationen vorangegangener und später auftretender I- oder P- Bilder. Die höchste Kompressionsrate ist durch die Kodierung als B-Bild erreichbar. Es entsteht als Differenzbild zur Prädiktion aus einem vorherigen und einem nachfolgenden I- oder P-Bild, kann aber niemals als Referenz auf andere Bilder dienen.
- *D-Bilder (DC Coded Pictures)* sind Intraframe-kodiert und werden für einen schnellen Vorlauf verwendet. Hier werden bei der DCT nur die DC-Parameter kodiert, die AC-Parameter werden vernachlässigt.

Die Gesetzmäßigkeit einer Folge von I-, P- und B-Bildern wird von der MPEG-Anwendung bestimmt. Eine extreme Auflösung in bezug auf die Forderung nach wahlfreiem Zugriff würde am besten durch eine Kodierung des gesamten Stroms als I-Bilder erreicht. Die höchste Kompressionsrate ist über eine Kodierung mit möglichst vielen B-Bildern zu erzielen. Für praktische Anwendungen hat sich die Folge von IBBPBBPBB IBBPBBPBB... als sinnvoll herausgestellt: Ein wahlfreier Zugriff hat dann eine Auflösung von neun Einzelbildern (d. h. ca. 330 ms) bei einer sehr guten Kompressionsrate. Alle 15 Bilder sollte ein I-Bild vorkommen

MPEG Performance

Decodierung ist sehr einfach

- Mit den heutigen Computern ist ein real-time software encoding kein Problem.
- Ein breites Spektrum an billigen Decoderchips sind vorhanden – verwendet in kiosk, VoD und anderen Spiele Applikationen.

Encodierung ist teuer

- Wenn man gute Qualität erreichen will, dauert das encodieren auf heutigen Computern bis zu 20mal länger als „real-time“ (DVD 2 MPEG Ripper kennen dieses Problem ;-))

- Real-time Encoding ist möglich, die Maschinen dafür sind aber für den Normalverbraucher nicht leistbar. (damn...)

Man geht heute davon aus, daß sich die Qualität einer nach MPEG komprimierten Videosequenz bei der vorgegebenen maximalen Datenrate von ca. 1,5 Mbit/s nicht mehr wesentlich verbessern läßt. Hierbei soll nur das Ergebnis (der Kompressionsfaktor und die Qualität) zählen, der dazu nötige Aufwand allerdings nicht. Im Bereich der Bewegtbild-Kodierung wird deshalb an Kompressionsverfahren für einen Bereich von bis zu 100 Mbit/s gearbeitet. Dies wird mit „MPEG-2“ bezeichnet. Während MPEG-1 die bisher etablierten Verfahren kennzeichnet, zielt MPEG-2 auch auf eine höhere Bildauflösung, ähnlich der digitalen Videostudionorm CCIR 601, ab. Die MPEG-Gruppe entwickelte den MPEG-2-Videostandard, der den kodierten Bitstrom für digitales Video hoher Qualität spezifiziert. Als kompatible Erweiterung setzt MPEG-2-Video auf dem fertiggestellten MPEG-1-Standard auf, indem verzahnte Videoformate (interlaced Video) und eine Reihe weiterer Eigenschaften, unter denen die Unterstützung von hochqualitativem digitalen Fernsehen zu nennen sind, integriert werden.

Als generischer internationaler Standard wurde MPEG-2-Video in Form von erweiterbaren Profilen definiert, von denen jedes die Möglichkeiten unterstützt, die eine Anwendungsklasse benötigt. Das MPEG-2 Main Profile wurde festgelegt, um die Übertragung digitalen Videos im Bereich von ca. 2 bis 80 Mbit/s über Kabel, Satellit und andere Broadcast-Kanäle zu unterstützen. Eine weitere wichtige Aufgabe ist hierbei die Unterstützung der digitalen Speicherung und anderer Kommunikationsanwendungen. Die Parameter des Main Profile und die des High Profile können auch für hochqualitatives Fernsehen verwendet werden.

Die MPEG-Experten erweiterten die Eigenschaften des Main Profile zusätzlich um die Definition eines hierarchischen bzw. skalierbaren Profils. Dieses soll Anwendungen, wie etwa das kompatible terrestrische Fernsehen, Videosysteme auf der Basis von Paketnetzwerken, die Abwärtskompatibilität mit existierenden Standards (MPEG-1 und H.261) und andere Anwendungen, für die eine mehrschichtige Kodierung erforderlich ist, unterstützen. Ein derartiges System könnte dem Klienten bspw. die Option anbieten, entweder einen kleinen portablen Empfänger zur Dekodierung von Standard-Fernsehen einzusetzen oder einen größeren fest installierten Empfänger, der aus demselben Broadcast-Signal hochauflösendes Fernsehen dekodieren kann.

MPEG-2 4:2:2 Profil

- MPEG 4:2:2 ergibt eine bessere Bildqualität als Motion JPEG.
- Die Anzahl der Bits in einem Makroblock sind frei.
- Unterstützt bis zu 8 Generationen an Standards

MPEG Audio

Mp3 – der meistgesuchte Begriff im Jahre 1999

MPEG 1/2 Layer-3 = mp3

- MPEG 1/2 Layer-3 ist ein offener Standard
- Die Spezifikation von mp3 ist gegen einen geringen Betrag offen für jedermann.
- Mp3 gehört niemandem bestimmten
- Das Format ist eindeutig und gut definiert
- Verfügbarkeit von Encodern und Decodern
- Viele Technologien unterstützen mp3: Soundkarten und Computer sind schnell genug um Audio zu decoden und zu encoden, kleine Datenmengen haben den Vorteil für das Senden im Internet, CD-ROM und CD-Writer unterstützen mp3 (lesen,schreiben,mp3 to Audio CD on the fly), Viele portable CD-Player haben mp3 implementiert.

Seit MPEG-1 Layer 3 im Jahre 1991 entwickelt wurde, wurde immer mehr versucht

- Die Qualität der Daten in möglichst gleich bleiben zu lassen die Komprimierung aber immer besser zu machen.
- Die encodierten Files sollten so (oder fast genauso) wie das Original klingen
- Software und Hardware für diesen Vorgang soll möglichst einfach und für jeden leistbar sein

Audio Encoder

Es erfolgt eine Zerlegung des Tonsignals in 32 Frequenzbänder. Die Quantisierung der verschiedenen Spektralanteile kann unterschiedlich erfolgen. Hierfür wird parallel zur eigentlichen FFT über ein psychoakustisches Modell der Rauschpegel in jedem dieser Frequenzbänder bestimmt. Bei einem starken Rauschen kann auch eine gröbere Quantisierung erfolgen. Ein geringer Rauschpegel führt zu einer hohen Auflösung. Die so quantisierten Spektralanteile werden in der ersten und zweiten Stufe einfach PCM-kodiert und in der dritten Stufe anschließend nach Huffman kodiert.

Die Audiokodierung kann sich hier auf einen Kanal, zwei unabhängige Kanäle oder den Stereoton beziehen. Der Stereoton kann nach der MPEG-Spezifikation in zwei Arten vorliegen: Zum einen lassen sich beide Kanäle völlig unabhängig voneinander verarbeiten, zum anderen gibt es den Modus des Joint Stereo. In diesem Fall berücksichtigt MPEG die zwischen beiden Kanälen vorhandenen Redundanzen und erreicht somit einen höheren Kompressionsgrad.

Jede Ebene definiert 14 feste Bitraten für den kodierten Audiodatenstrom, die in MPEG durch einen Bitratenindex adressiert werden. Der Minimalwert ist hierbei stets 32 Kbit/s. Diese Ebenen unterstützen verschiedene maximale Bitraten: Ebene 1 erlaubt eine maximale Bitrate von 448 Kbit/s, Ebene 2 von 384 Kbit/s und Ebene 3 von 320 Kbit/s. Ein Dekoder muß für die Ebenen 1 und 2 keine variable Bitrate unterstützen. In Ebene 3 wird eine variable Bitrate derart spezifiziert, daß der Bitratenindex umgeschaltet werden kann. In Ebene 2 sind nicht alle vorstellbaren Kombinationen von Bitrate und Modus erlaubt:

- 32 Kbit/s, 48 Kbit/s, 56 Kbit/s und 80 Kbit/s sind nur für einen einzigen Kanal erlaubt.
- 64 Kbit/s, 96 Kbit/s, 112 Kbit/s, 128 Kbit/s, 160 Kbit/s und 192 Kbit/s sind für alle Modi erlaubt.
- 224 Kbit/s, 256 Kbit/s, 320 Kbit/s, 384 Kbit/s sind für die Modi *Stereo*, *Joint Stereo* und *Dual Channel* erlaubt.

Überlegungen bezüglich Qualität

Artefakte

- Verlust der Bandbreite
- pre-echoes (Geräusche setzen ein bevor die Musik eigentlich ein „Geräusch“ erzeugen sollte).
- Roughness, double-speak (wird die Bitrate geringer und die sampling frequenz niedriger gibt es größere unterschiede zwischen der Zeitauflösung von Coder und der Zeitstruktur der Signale)

Messung der Codec Qualität

Hörtests:

Einzig wirklich zielführende Methode ist der Hörtest mit Versuchspersonen.

Einfache sachliche Messungstechniken:

Beispielmessungen – Signal-zu-Geräusch-Rate, Bandbreite des Decodierten Signals

perceptual measurement techniques (pmt):

Diese Techniken können Hörtests nicht komplett ersetzen sind in manchen Bereichen aber schon „state of the art“.

MPEG – 4

MPEG-4 beinhaltet neue algorithmische Techniken wie die modellbasierte Bildkodierung der menschlichen Interaktion mit multimedialen Umgebungen, sowie die Sprachkodierung mittels niedriger Bitraten, die in Umgebungen wie dem *Europäischen Mobilien Telefonsystem* (GSM) eingesetzt wird. Die wichtigste Neuerung von MPEG-4 ist die *Flexibilität*. Ein Benutzer kann das Kompressionsverfahren auf verschiedene Art und Weise einsetzen, um Systeme zu entwickeln und diese für eine Vielzahl von Anwendungen zu konfigurieren.

MPEG-4 ist kein fester Standard, der nur für einige wenige Anwendungen zu gebrauchen ist.

Gegenwärtig vorhandene audiovisuelle Dienste erlauben meist nur eine Abspielfunktionalität. Im Gegensatz dazu liegt in MPEG-4 ein deutlicher Schwerpunkt auf Interaktionsmöglichkeiten. Deshalb sind unter den Hauptzielen der Standardisierung von MPEG-4 der wahlfreie Zugriff auf Video- und Audioszenen und die Möglichkeit der Inhaltsbearbeitung. MPEG-4 soll daher eine universelle effiziente Kodierung verschiedener audiovisueller Datenformen, die man auch als *audiovisuelle Objekte* bezeichnet, erlauben. Dies bedeutet im Grunde genommen, daß die Intention von MPEG-4 eine Repräsentation der realen Welt ist, die als Komposition audiovisueller Objekte verstanden wird.

Eine große Schwäche anderer audiovisueller Kodierungsstandards ist die Einschränkung der Zahl der verwendeten Audio- und Videodatentypen. MPEG-4 versucht, natürliche und synthetische audiovisuelle Objekte harmonisch zu integrieren, bspw. Mono-, Stereo- und Mehrkanalaudio. Weiterhin kann in MPEG-4 entweder 2-D- oder 3-D- bzw. Mono-, Stereovideo oder solches aus weiteren Kameraperspektiven (sog. *Multiview Video*) eingesetzt werden. Diese Integration sollte auch hinsichtlich audiovisueller Beziehungen erweitert werden, wodurch der gegenseitige Einfluß und die Abhängigkeit zwischen diesen Informationstypen mit in die Verarbeitung einbezogen werden könnte. Weiterhin werden neue und bereits verfügbare Analyse- und Kodierungswerkzeuge in MPEG-4 integriert. Insgesamt wird sich MPEG-4 auch in Zukunft weiterentwickeln, wenn neue oder bessere Werkzeuge, Datentypen oder Funktionalitäten verfügbar sind.

Der schnelle technologische Fortschritt der letzten Jahre zeigt deutlich, daß Standards, die die kontinuierliche Weiterentwicklung von Hardware und Methodik nicht berücksichtigen und die auf diese Art und Weise nur eine spezielle Lösung realisieren, nach kurzer Zeit nicht mehr dem Stand der Technik entsprechen. Ein wichtiges Ziel von MPEG-4 ist daher *Flexibilität und Erweiterbarkeit*. Das in MPEG-4 verwendete Konzept zur Umsetzung der Flexibilität bzw. der Erweiterbarkeit ist die syntaktische Beschreibungssprache „MPEG-4 Syntactic Description Language (MSDL)“. Der MSDL-Ansatz ist im Bereich der Standards zur audiovisuellen Kodierung revolutionär, da die Erweiterbarkeit nicht nur durch die Integration neuer Algorithmen durch eine Auswahl und durch ein Einbinden vordefinierter Werkzeuge (Ebene 1) erfolgt, sondern auch dadurch, daß neue Werkzeuge, die vom Kodierer heruntergeladen wurden, „gelernt“ werden können. Gleichzeitig unterliegt auch die MSDL der Weiterentwicklung, da der Standard jederzeit neue Werkzeuge, Techniken und Konzepte in MSDL aufnehmen kann, die bessere oder neue Funktionalitäten unterstützen.

Die Vision des MPEG-4-Standard kann am besten mit den acht neuen oder verbesserten Funktionalitäten erklärt werden, die in der Paketbeschreibung des MPEG-4-Vorschlags enthalten sind.

- MPEG-4 muß auch noch einige andere wichtige sog. Standardfunktionalitäten unterstützen, die in den bisher verfügbaren Standards enthalten sind. Beispiele hierfür sind die Synchronisation von Audio und Video, Modi kurzer Verzögerungen und die Zusammenarbeit über Netzwerke.
- MPEG-4 ermöglicht eine Skalierbarkeit des Inhalts, der räumlichen und zeitlichen Auflösung, der Qualität und der Komplexität in feiner Granularität. Eine *inhaltsbasierte Skalierbarkeit* kann hierbei die Existenz eines Prioritätsmechanismus für Objekte einer Szene implizieren. Die Kombination verschiedener Skalierungsarten kann zu interessanten Szenenrepräsentationen führen, in denen die wichtigeren Objekte in einer höheren räumlich-zeitlichen Auflösung repräsentiert sind.
- MPEG-4 beinhaltet eine Syntax und verschiedene Kodierungsverfahren zur Unterstützung der *inhaltsbasierten Manipulation* und zur *Bearbeitung des Bitstroms*, bei denen ein sog. *Transcoding* (Umwandlung von einer Kodierungsinstanz in eine andere) nicht erforderlich ist. Dies bedeutet, daß der Benutzer in der Lage sein sollte, auf ein spezifisches Objekt einer Szene bzw. des Bitstroms zuzugreifen. Hierdurch besteht die Möglichkeit, die Eigenschaften eines Objekts zu ändern.
- MPEG-4 bietet *effiziente Werkzeuge zum inhaltsbasierten Zugriff auf multimediale Daten* und auf deren Organisation. Diese Zugriffsmöglichkeiten können in einer Indizierung, dem Anlegen von Hyperlinks, Abfragen, Anschauen der Daten, dem Herauf- oder Herunterladen und dem Löschen der Daten bestehen.
- MPEG-4 unterstützt effiziente Methoden zur Kombination synthetischer mit natürlichen Szenen (z. B. Text- und Grafiküberlagerungen), die Fähigkeit, natürliche und synthetische

Audio- und Videodaten zu kodieren sowie vom Dekoder kontrollierbare Methoden zum Mischen synthetischer Daten mit gewöhnlichem Video oder Audio. (Da fallen Dinge wie „branden“ von Videos mit Firmenlogos während der ganzen Spieldauer oder das einblenden von Untertiteln rein,nehm ich an...)

- MPEG-4 bietet die Möglichkeit, verschiedene Ansichten/Tonspuren einer Szene, bzw. eine ausreichende Synchronisation zwischen den resultierenden elementaren Strömen, effizient zu kodieren. Videoanwendungen mit Stereobildern oder mehrfachen Ansichten werden von MPEG-4 dahingehend unterstützt, daß die Redundanz in den mehrfachen Ansichten derselben Szene ausgenutzt werden kann. Dies gestattet weiterhin Lösungen, die mit normalem (Mono-)Video kompatibel sind. Diese *Kodierung mehrfacher gleichzeitiger Datenströme* sollte eine effiziente Repräsentation von natürlichen 3D-Objekten beinhalten, wenn für diese eine ausreichende Anzahl von Ansichten verfügbar ist.
- Universelle Zugriffsmöglichkeiten implizieren den Zugriff auf Anwendungen über verschiedenste (drahtlose oder kabelgestützte) Netzwerke und Speichermedien. MPEG-4 muß daher in *fehleranfälligen Umgebungen* robust arbeiten. Insbesondere für Anwendungen mit niedriger Bitrate muß in schweren Fehlerfällen eine ausreichende Fehlerrobustheit vorgesehen sein. Die hierbei verwendete Idee ist es, nicht die Fehlerkontrolltechniken auszutauschen, die im Netzwerk implementiert sind, sondern eine Elastizität gegenüber Restfehlern anzubieten. Hierzu kann bspw. eine selektive Vorwärtsfehlerkorrektur, eine Fehlereindämmung oder eine Fehlermaskierung erfolgen.

MPEG-4 beinhaltet effiziente Methoden zum verbesserten zeitlich wahlfreien Zugriff von Teilen einer audiovisuellen Sequenz in einem begrenzten Zeitintervall mit einer feinen Auflösung. Hierin sind auch konventionelle Möglichkeiten des wahlfreien Zugriffs bei sehr niedrigen Bitraten enthalten. Audiovisuelle Szenen bestehen in MPEG-4 aus audiovisuellen Objekten (AVO), die in einer hierarchischen Art und Weise organisiert sind.

Auf der untersten Hierarchiestufe sind primitive AVOs zu finden, wie:

- Ein zweidimensionaler fester Hintergrund,
- Das Bild einer sprechenden Person (ohne Hintergrund) oder
- Die mit der Person assoziierte Sprache.

MPEG-4 standardisiert eine Reihe dieser primitiven AVOs, mit denen sowohl natürliche als auch synthetische Inhaltstypen repräsentiert werden können, die entweder 2- oder 3-dimensional sind. MPEG-4 definiert die kodierte Repräsentation derartiger Objekte, bspw.:

- Text und Grafik,
- Köpfe sprechender Darsteller und der damit assoziierte Text, der vom Empfänger zur Synthetisierung der Sprache und zur Animation des Kopfs verwendet wird,
- Animierte menschliche Körper.

MPEG-4 verwendet verschiedene Möglichkeiten zur Kodierung visueller Objekte. Diese sind für natürliche Bilder und Videos als eine Menge von Werkzeugen und Algorithmen zur effizienten Kompression von Bildern, Videos, Texturen und 2-D- und 3-D-Netzen (Meshes) sowie für geometrische Ströme, die diese Netze in einer zeitlich variierenden Art animieren, definiert. Es stehen Werkzeuge zum wahlfreien Zugriff und zur Manipulation aller Typen von visuellen Objekten zur Verfügung. Weiterhin wird eine inhaltsbasierte Kodierung sowie eine inhaltsbasierte räumliche, zeitliche und qualitative Skalierung unterstützt. Für natürliche Inhalte können Mechanismen, die eine Fehlerrobustheit und -elastizität in fehleranfälligen Umgebungen realisieren, verwendet werden.

Ebenso wie MPEG-1 und MPEG-2 beschreibt auch MPEG-4 Ströme. Da MPEG-4 Inhalte in mehrere Objekte unterteilt, betreffen die Stromeigenschaften das Multiplexing, das Demultiplexing und die Synchronisation mehrfacher Ströme.

Die AVO-Daten werden zu einem oder zu mehrere elementare Ströme (im folgenden mit dem Fachbegriff Elementary Streams bezeichnet) gebündelt. Diese sind durch die Dienstgüte (Quality of Service, QoS) charakterisiert, die sie für die Übertragung anfordern (z. B. maximale Bitrate oder Bitfehlerrate) sowie durch andere Parameter, wie z. B. die Stromtypinformation, mit deren Hilfe die erforderlichen Dekoderressourcen und die Präzision der Zeitinformation beim Kodierer bestimmt werden. Die Art und Weise, wie eine derartige Information über Stromeigenschaften in einer synchronisierten Art unter Ausnutzung verschiedener vom Netzwerk zur Verfügung gestellter QoS von

der Quelle zur Senke transportiert wird, wird im Sinne einer Zugriffseinheitsschicht (Access Unit Layer) und eines konzeptionellen zweischichtigen Multiplexers festgelegt.

Die Access Unit Layer erlaubt die Identifikation von Zugriffseinheiten (Access Units), z. B. Video- oder Audio-Frames, Szenenbeschreibungskommandos in Elementary Streams, die Wiederherstellung der Zeitbasis des AV-Objekts oder der Szenenbeschreibung sowie die Synchronisation zwischen diesen. Der Header einer Zugriffseinheit kann in einer Vielzahl von Wegen konfiguriert werden, wodurch ein breites Spektrum von Systemen erlaubt wird.

Die FlexMux-Schicht (Flexibles Multiplexing) wird von MPEG vollständig spezifiziert. Sie enthält ein Multiplexing-Werkzeug, das eine Gruppierung von Elementary Streams (ESs) mit einem geringen Multiplexing-Verwaltungsaufwand erlaubt. Hierdurch können bspw. Elementary Streams mit ähnlichen Dienstgüteanforderungen gruppiert werden.

Die TransMux-Schicht (Transport Multiplexing) modelliert die Schicht, die zu den geforderten Dienstgüte passende Transportdienste anbietet. MPEG-4 spezifiziert hierfür lediglich die Schnittstelle zu dieser Schicht. Dadurch kann jede geeignete Transportprotokollarchitektur, wie z. B. (RTP)/UDP/IP, (AAL5)/ATM oder der Transportstrom von MPEG-2, über eine geeignete Sicherungsschicht zu einer spezifischen TransMux-Instanz werden. Die jeweilige Auswahl wird dem Endbenutzer/Dienstanbieter überlassen, wodurch MPEG-4 in einer Vielzahl von Arbeitsumgebungen verwendet werden kann.

In derselben Art und Weise, in der MPEG-1 und MPEG-2 das Verhalten eines ideal arbeitenden Dekodierungsgeräts zusammen mit der Syntax und Semantik des Bitstroms beschreiben, definiert MPEG-4 ein Modell eines Systemdekoders. Hierdurch wird eine präzise Definition einer Terminaloperation möglich, ohne unnötige Annahmen über Implementierungsdetails treffen zu müssen. Gerade dies ist essentiell, um Entwicklern die Freiheit zu geben, MPEG-4-Terminals und Dekodierungsgeräte in einer Vielzahl von Wegen zu implementieren. Diese Geräte reichen von Fernsehempfängern, die keine Möglichkeit haben, mit einem Sender zu kommunizieren, zu Rechnern, die vollständig in der Lage sind, bidirektional Daten auszutauschen. Einige Geräte werden MPEG-4-Ströme über isochrone Netzwerke empfangen, während andere nicht-isochrone Mittel (bspw. das Internet) einsetzen werden, um MPEG-4-Informationen auszutauschen. Das Modell des Systemdekoders stellt derart ein gemeinsames Modell dar, auf dem alle Implementierungen von MPEG-4-Terminals basieren können.

Der Schritt des MPEG-4-Demultiplexing ist im Sinne eines konzeptionellen zweischichtigen Multiplexers spezifiziert, der aus einer TransMux-Schicht und einer FlexMux-Schicht sowie aus der Access Unit Layer besteht, die Synchronisationsinformationen übermittelt.

Der allgemeine Begriff TransMux-Schicht wird dazu verwendet, um von der (existierenden oder zukünftig) zugrunde liegenden Multiplexing-Funktionalität zu abstrahieren, die zum Transport von MPEG-4-Strömen geeignet ist. Es ist zu beachten, daß diese Schicht nicht im Kontext von MPEG-4 definiert ist. Beispiele hierfür sind der MPEG-2-Transportstrom, H.223, ATM AAL 2 und IP/UDP. Die TransMux-Schicht wird mit Hilfe einer Schutz-Subschicht und einer Multiplexing-Subschicht modelliert, die verdeutlichen, daß diese Schicht für das Anbieten spezifischer QoS-Parameter verantwortlich ist. Die Funktionalität der Schutz-Subschicht beinhaltet einen Fehlerschutz und Werkzeuge zur Fehlerentdeckung, die für das gegebene Netzwerk oder das Speichermedium geeignet sind. In einigen TransMux-Instanzen ist es eventuell nicht möglich, diese Subschichten separat zu identifizieren. In jedem konkreten Anwendungsszenario werden eine oder mehrere spezifische TransMux-Instanzen verwendet. Jeder TransMux-Demultiplexer realisiert den Zugriff auf die TransMux-Kanäle. Die Anforderungen an die Datenschnittstelle in bezug auf den Zugriff auf den TransMux-Kanal sind hierbei für alle TransMux-Instanzen dieselben. Sie beinhalten eine zuverlässige Fehlerentdeckung, eine Auslieferung fehlerhafter Daten mit einer geeigneten Fehleranzeige (falls möglich) und eine Einteilung der Nutzlast in Frames. Diese Einteilung besteht entweder aus Strömen, die von der AL in Pakete eingeteilt werden, oder aus FlexMux-Strömen. Die Anforderungen werden in einer informellen Art und Weise in der TransMux-Schnittstelle im Systemteil des MPEG-4-Standards zusammengefaßt.

Die FlexMux-Schicht wird auf der anderen Seite vollständig von MPEG beschrieben. Sie stellt ein flexibles Werkzeug mit geringem Verwaltungsaufwand und geringer Verzögerung zur optionalen Verzahnung der Daten dar, die insbesondere dann nützlich ist, wenn die Paketgröße oder der Verwaltungsaufwand der darunter liegenden TransMux-Instanz groß ist. Das FlexMux ist selber nicht robust gegenüber Fehlern. Es kann entweder auf TransMux-Kanälen mit hohen QoS verwendet werden oder zur Bündelung von Elementary Streams. Beide sind in bezug auf die Fehlertoleranz gleich. Das FlexMux erfordert eine zuverlässige Fehlerentdeckung und eine ausreichende Einteilung der FlexMux-Pakete in Frames (zum wahlfreien Zugriff und zur Fehlerbehebung), die von der darunter

liegenden Schicht zur Verfügung gestellt werden muß. Diese Anforderungen werden in der Strom-Multiplex-Schnittstelle zusammengefaßt, die den Datenzugriff auf individuelle Transportkanäle definiert. Der FlexMux-Demultiplexer liest aus den FlexMux-Strömen von der AL in Pakete eingeteilte Ströme.

Die Access Unit Layer verfügt über eine minimale Menge an Werkzeugen zur Konsistenzüberprüfung, zur Auffüllung von Headern (Padding), zur Übermittlung von Informationen der Zeitbasis und zur Beförderung von Zugriffseinheiten mit Zeitmarken eines Elementary Streams. Jedes Paket besteht aus einer Zugriffseinheit oder aus einem Fragment einer solchen. Diese mit Zeitmarken versehenen Einheiten stellen die einzige semantische Struktur von Elementary Streams dar, die in dieser Schicht sichtbar ist. Die AU Layer erfordert eine zuverlässige Fehlerentdeckung und eine Einteilung in Frames von individuellen Paketen der darunter liegenden Schicht, die z. B. vom FlexMux vorgenommen werden kann. Die Art und Weise, in der durch die Kompressionsschicht auf Daten zugegriffen werden kann, ist in der informellen Elementary Stream-Schnittstelle zusammengefaßt, die ebenfalls im Systemteil des MPEG-4-Standards zu finden ist. Die AU Layer liest Elementary Streams aus den Strömen, die von der AL in Pakete unterteilt werden.

Zur Vorhersage des Verhaltens eines Dekoders bei der Dekompression der verschiedenen elementaren Datenströme, die eine MPEG-4-Session bilden, versetzt das Dekodermodell des Systems den Kodierer in die Lage, die minimalen Pufferressourcen zu spezifizieren und zu überwachen, die zur Dekodierung einer Session erforderlich sind. Die benötigten Ressourcen werden dem Dekoder mit Objektdeskriptoren während der Einrichtung einer MPEG-4-Session übermittelt, so daß dieser entscheiden kann, ob er in der Lage ist, diese Session durchzuführen.

Durch die Verwaltung der endlichen Menge an Pufferplatz erlaubt bspw. das Modell dem Sender, Nicht-Echtzeit-Daten zu früh zu übertragen, wenn beim Empfänger genügend Platz zu deren Speicherung vorhanden ist. Auf die vorgeschichteten Daten kann dann im geeigneten Augenblick zugegriffen werden, wodurch in diesem Moment Echtzeitinformationen einen größeren Anteil der Kanalkapazität verwenden können, falls dies erforderlich ist.

Für den Echtzeitbetrieb wird ein Zeitmodell angenommen, in dem die Ende-zu-Ende-Verzögerung von der Signalausgabe eines Kodierers zur Signaleingabe am Dekoder konstant ist. Weiterhin müssen die übertragenen Datenströme implizite oder explizite Zeitinformationen tragen. Es existieren zwei derartige Informationsarten: Die erste wird zur Übermittlung der Geschwindigkeit des Zeitgebers, oder der Zeitbasis des Kodierers an den Dekoder verwendet. Die zweite, die aus Zeitmarken besteht, die an Teile der kodierten AV-Daten angehängt werden, beinhaltet die gewünschte Dekodierungszeit für Zugriffseinheiten oder die Kompositions- und Ablaufzeit für Kompositionseinheiten. Diese Information wird in den AL-PDU-Headern übermittelt, die in der Access Unit Layer generiert werden. Mittels dieser Zeitinformation kann das Intervall zwischen den Bildern und die Audioabtastrate an den Dekoder angepaßt werden, um mit den Werten des Kodierers zum Zweck eines synchronisierten Betriebs übereinzustimmen.

Verschiedene AV-Objekte können von Kodierern mit verschiedenen Zeitbasen zusammen mit der auftretenden leicht unterschiedlichen Geschwindigkeit kodiert werden. Es ist jederzeit möglich, diese Zeitbasen auf die des empfangenden Terminals abzubilden. In diesem Fall kann aber keine reale Implementierung eines empfangenden Terminals die von Zeit zu Zeit auftretende Wiederholung oder den Verlust von AV-Daten vermeiden, was durch das zeitliche Aliasing (relative Reduktion oder Verlängerung der Zeitscala) bedingt ist.

Obwohl Systemoperationen ohne jegliche Zeitinformation erlaubt sind, ist die Definition eines Puffermodells nicht möglich.

MPEG 7

MPEG-7 ist weniger als Kompressionsformat gedacht: Intention der MPEG-7-Arbeitsgruppe der ISO war die Etablierung eines Metadaten-Standards, der den jeweiligen in einem anderen Format (wie bspw. MPEG-4) kodierten Inhalt auf einer weiteren Spur ergänzt. Die Verwendung von Metadaten zielt auf eine neuartige Erweiterung der Kodierung multimedialer Daten ab: hauptsächlich auf die Verbesserung von Suchverfahren und Anzeigestrategien, aber auch auf Konsistenzüberprüfungen und bspw. eine Skalierung, die Konsistenz und Priorität mit einbezieht. Derart beabsichtigt MPEG-7 die Integration anderer Medieninhaltsformate (speziell MPEG-4) und deren Erweiterung. Neben der Kodierung von Metadaten wird MPEG-7 ausschließlich Schnittstellen zur Zusammenarbeit mit Werkzeugen der automatischen Inhaltsanalyse und Suchmaschinen definieren, aber nicht diese Dienste selber.

Ein gänzlich unterschiedliches Kodierungsverfahren ist die Technik der fraktalen Bild- und Videokompression. Hierbei wird keinerlei aktuelle Pixelinformation übertragen, sondern lediglich eine Transformationsfunktion, die ein zum Zielbild ähnliches Bild als Fixpunkt beinhaltet. Somit besteht der Dekodierungsprozeß aus einer iterativen Anwendung dieser Funktion auf ein beliebiges Ursprungsbild.

Hieraus resultieren verschiedene spezifische Attribute dieser Kodierungstechnik: zunächst arbeitet der Dekodierungsprozeß progressiv und die Dekodierungseffizienz ist skalierbar, da die Qualität des dekodierten Bildes mit jedem Iterationsschritt zunimmt. Außerdem ist der Prozeß unabhängig von der Auflösung. Die Abbildungsfunktion kann wiederholt angewendet werden, um mehr und mehr Details zu erhalten. Das dritte Attribut ist eine Asymmetrie zwischen Kodierungs- und Dekodierungsprozeß. Es stellt sich daher die Frage, wie ein Bild als eine Transformationsfunktion kodiert werden kann. Der Algorithmus nutzt hierbei ein fraktales Attribut von Bildern aus: die Selbstähnlichkeit. Bilder bestehen aus Bereichen, die einander ähnlich sind. Eine Transformationsfunktion besteht dann aus einer Abbildung von Bildbereichen auf die jeweils ähnlichsten Bildteile. Diese Abbildung beinhaltet ein Stauchen, Strecken, Rotieren, Verschieben der Form und Kontrast- und Helligkeitsänderungen der Bildbereiche. Dies ist eine Art der Vektorquantisierung, wobei keine feste Menge von Quantisierungsvektoren verwendet wird. Um die Existenz eines Fixpunkts der Abbildung zu gewährleisten, also ein resultierendes Bild, muß als einzige Bedingung, die die Theorie von der Abbildung verlangt, eine Kontraktionseigenschaft gegeben sein. Der absolute Wert des Kontraktionsfaktors muß im Intervall $[0,1)$ liegen.

Die Abbildung wird in heutigen Implementierungen erreicht, indem das Originalbild in Blöcke (von 8×8 Pixeln) unterteilt wird und für jeden Block der ähnlichste Block von Bildbereichen gefunden wird (also überlappende 16×16 Pixelblöcke). Die Menge der mit jedem (8×8 -) Originalblock zu vergleichenden Blöcke vergrößert sich durch die Möglichkeiten, die durch die geometrischen Transformationen gegeben sind. Der ähnlichste Block kann gefunden werden, indem ein Distanzmaß minimiert wird. Hierzu minimiert man gewöhnlicherweise die Summe der quadrierten Pixeldifferenzen.

Für natürliche Bilder kann die fraktale Kompression hohe Kompressionsraten (bis zu 1000) mit einer sehr guten Bildqualität erreichen. Der größte Nachteil dieser Kodierungsmethode ist die Berechnungskomplexität und die wenig effektive Anwendbarkeit auf grafische Bilder. Um diese Komplexität in praktikablen Grenzen zu halten, wird nur eine Teilmenge aller Transformationen betrachtet, z. B. nur die Rotationswinkel 0 Grad, 90 Grad, 180 Grad und 270 Grad. Es muß aber immer noch jeder Originalblock mit einer sehr großen Anzahl von Blöcken verglichen werden, um die Abbildung auf den ähnlichsten Block zu finden. Neben der Berechnungskomplexität erzeugt diese Kodierungstechnik Verluste, da lediglich eine Ähnlichkeit der Blöcke, nicht aber deren Identität, verwendet wird.

Zusammenfassung der wichtigsten Dinge

JPEG ist der Standard für die Einzelbildkodierung, der in Zukunft die größte Bedeutung haben wird. Er läßt über seine sehr weitgefaßte Definition eine Vielzahl von Freiheitsgraden zu. Hier können bspw. bis zu 255 Bildkomponenten, d. h. Ebenen, existieren. Ein Bild kann aus bis zu 65535 Zeilen bestehen, von denen jede bis zu 65525 Pixel beinhalten kann. Als Leistungsfähigkeits-Koeffizient läßt sich hier ein Maßstab in *Bit pro Pixel* angeben. Er ist ein Mittelwert, der sich als Quotient aus der Anzahl der in kodierter Form enthaltenen Bits und der Anzahl der im Bild enthaltenen Pixel ergibt. Hiermit lassen sich folgende Aussagen für DCT-kodierte Einzelbilder treffen.

- 0,25 bis 0,50 bit/Pixel: Mäßige bis gute Qualität, für einige Anwendungen ausreichend.
- 0,50 bis 0,75 bit/Pixel: Gute bis sehr gute Qualität, für viele Anwendungen ausreichend.
- 0,75 bis 1,50 bit/Pixel: Ausgezeichnete Qualität, für die meisten Anwendungen ausreichend.
- 1,50 bis 2,00 bit/Pixel: Oft vom Original kaum mehr zu unterscheiden. Genügt fast allen Anwendungen, selbst bei höchsten Qualitätsansprüchen.

Im verlustfreien Modus wird trotz der erstaunlichen Einfachheit des Verfahrens im Mittel ein Kompressionsgrad von ca. 2:1 erreicht. JPEG ist als Produkt heute sowohl in Software als auch in Hardware verfügbar und wird oft in Multimedia-Anwendungen eingesetzt, die eine hohe Qualität erfordern. Das primäre Ziel ist hierbei die Kompression eines Einzelbildes. In der Form des *Motion JPEG* kann JPEG aber auch zur Videokompression in Anwendungen, wie z. B. der medizinischen Bildverarbeitung, eingesetzt werden.

H.261 und H.263 sind bereits etablierte Standards, die vor allem von den Betreibergesellschaften der Telefon- und Weitverkehrsnetze gefördert werden. Über eine sehr beschränkte Auflösung im QCIF-Format und mit reduzierter Bildwechselfrequenz bereitet die Realisierung von H.261- und H.263-Kodierern und Dekodierern heute keine größeren technische Probleme. Dies gilt besonders dann, wenn die Bewegungskompensation und der optische Tiefpaß nicht Bestandteil der Implementierung sind. Die Qualität ist dann jedoch nicht immer zufriedenstellend. Wenn man das Bild im CIF-Format bei 25 Bildern/s mit einer Bewegungskompensation komprimiert, ist die Qualität jedoch durchaus akzeptabel. H.263 wird meist in Netzwerkumgebungen für Anwendungen im Dialogmodus, wie z. B. der Videotelefonie und Konferenzsystemen, eingesetzt. Die resultierenden kontinuierlichen Bitraten sind hervorragend für heutige Weitverkehrsnetzwerke geeignet, die mit ISDN, gemieteten Leitungen oder sogar GSM-Verbindungen arbeiten.

MPEG ist der vielversprechendste Standard für zukünftig zu verwendendes komprimiertes Video und Audio. Während die JPEG-Gruppe ein System beinhaltet, das auch immer für Video verwendet werden kann, ist diese zu sehr auf eine Animation von Einzelbildern fokussiert, anstatt die Eigenschaften von Bewegtbildern zu verwenden. Die Qualität von MPEG-Video bei etwa 1,2 Mbit/s (ohne Ton) ist vergleichbar mit VHS-Aufzeichnungen, auch für CD-ROM-Geräte geeignet. Der Kompressionsalgorithmus arbeitet für eine Auflösung von ca. 360 × 240 Pixeln sehr gut. Selbstverständlich sind auch höhere Auflösungen dekodierbar, bei bspw. 625 Zeilen geht dies jedoch auf Kosten der Qualität. Die Zukunft von MPEG zeigt in Richtung MPEG-2, das eine Datenstromkompatibilität mit MPEG-1 definiert, aber Datenraten bis zu 100 Mbit/s zuläßt. Dies verbessert die heute verfügbare Qualität von MPEG-kodierten Daten signifikant.

MPEG definiert weiterhin einen Audiostrom mit verschiedenen Abtastraten, die bis zur DAT-Qualität von 16 bit/Abtastwert reichen. Ein weiterer wichtiger Teil der Arbeit der MPEG-Gruppe ist die Definition einer Datenstromsyntax.

MPEG wurde weiterhin optimiert, indem ein Abfragemodell für Anwendungsbereiche, wie Tutorensysteme auf der Basis von CD-ROMs, und interaktives Fernsehen verwendet wurde. Diese in MPEG-2 eingebettete Optimierung wird eine Fernsehqualität auf Kosten einer höheren Datenrate erlauben. MPEG-4 wird hohe Kompressionsraten bei Video und damit assoziiertem Audio erreichen und ist weiterhin ein geeignetes Werkzeug, um ganze Klassen neuer Multimedia-Anwendungen zu

schaffen. Allerdings wird der komplizierte und noch sehr junge MPEG-4-Standard in weitverbreiteten kommerziellen Anwendungen derzeit noch wenig eingesetzt.

JPEG, H.263, MPEG und andere Techniken sind nicht als Alternativen bezüglich der Datenkompression zu sehen. Ihre Ziele sind unterschiedlich und teilweise komplementär. Die meisten Algorithmen sind sich sehr ähnlich, wenn auch nicht gleich. Die technische Qualität sowie die Verfügbarkeit am Markt bestimmen, welche dieser Techniken in zukünftigen Multimediasystemen eingesetzt werden. Hieraus wird sich eine *Kooperation* und eine *Konvergenz* der Techniken ergeben. Ein zukünftiger Multimedia-Computer könnte bspw. Einzelbilder in JPEG generieren, H.263 oder MPEG-4 für eine Videokonferenz anwenden und zum Lesen abgespeicherter Multimedia-Informationen MPEG-2 benötigen. Dies ist allerdings eine rein hypothetische Feststellung, die keine Art der zukünftigen Entwicklung oder Strategie für diese Systeme präjudiziert.

Fehler und Irrtümer vorbehalten.