

Referat:

MPEG-2 - ein ISO-Standard

zum

generischen Kodieren von Bewegtbildern und synchronisiertem Audio

zum Seminar WS 93/94: Workstation-Architekturen für Multimedia-Systeme

von Frank Gadegast

8. Februar 1994

Adresse: Frank Gadegast
Leibnizstraße 30
10625 Berlin
GERMANY

Fon/Fax: ++ 49 30 3128103

E-mail: phade@cs.tu-berlin.de
phade@contrib.de

Inhaltsverzeichnis:

I. Einleitung	1
II. Grundlegende Kodierungsmechanismen.....	2
1. JPEG	2
1.1 YUV-Farbraum	2
1.2 Diskrete Cosinus Transformation	2
2. H.261	3
3. MPEG-I.....	3
3.1 MPEG-Begriffe	3
3.2 MPEG-Parameter	3
3.3 MPEG-Audio	4
3.3.1 Technische Daten	4
3.3.2. Methoden	5
II. MPEG-2-Draft	6
1. ISO - die Organisation	6
2. MPEG-2-Teile	6
III. MPEG-2 Kodierungsmethoden	7
1. Profiles und Levels.....	7
2. Scalability	7
3. Sicherheit.....	10
IV. MPEG-2 Features	11
1. Video.....	11
2. Audio	12
3. System.....	12
V. Ausblick.....	14
1. Andere MPEG-Kodierungen.....	14
2. Die Zukunft für MPEG-2 ?	14
Anhang A: Quellen	15
Anhang B: Abbildungen.....	15

I. Einleitung

Die digitale Darstellung eines Studio-TV-Signals erfordert gemäß der CCIR Recommendation 601 eine Nettodatenrate von 166 Mbit/s. Keine der bereitstehenden digitalen Medien kann eine solche Datenrate bieten. Mit dem hier zu behandelnden Standard MPEG-2 können jedoch auch Video- und Audio-Signale auf bereits bestehende Medien übertragen und gespeichert werden.

MPEG heißt "**M**otion **P**icture **E**xpert **G**roup" und ist die Gruppe der ISO, die sich mit der Standardisierung im Video-Bereich beschäftigt. Im Juli 1993 wurde der "Third Working Draft" des kommenden MPEG-2-Standards mit dem übersetzten Titel "Generische Kodierung von Bewegtbildern und synchronisiertem Audio" vorgelegt.

MPEG-2 ist keine Weiterentwicklung des bereits bestehenden Standard MPEG-I im herkömmlichen Sinne. Eine Entwicklung des Standards MPEG-2 heißt nicht, daß MPEG-I nun veraltet wäre, im Gegenteil. MPEG-2 ist nur eine komplexere, umfassendere Kodierung, geschaffen für ein vollständig anderes Anwendungsprofil. Die verwendeten Mechanismen von MPEG-2 sind jedoch sehr ähnlich zu denen aus MPEG-I.

MPEG-2 wird die Erweiterung von MPEG-I werden, die Rücksicht auf die hardwaretechnischen Neuerungen und Erfordernisse nimmt. Besonders ist hierbei die neu hinzugefügte Funktionalität der "Scalability" zu nennen, die ähnlich dem Standard ODA, ein der Hardware angepaßtes Konsumieren von Videos ermöglichen wird. "The ability of a decoder to ignore some portions of a total bitstream and produce useful audio and video output from the portion which is decoded." [CHI92]. Insbesondere wird dadurch bei einem HDTV (zukünftiges hochauflösendes Fernsehen)-kodierten 16:9-Bild die Abwärtskompatibilität zu einem 'normalem' 4:3 Fernsehbild garantiert. Bis zum November 1993 soll ein 'Comittee Draft' der ISO vorgelegt werden [CT93].

Der Anwendungsbereich für MPEG-2 ist die digitale Bildbearbeitung in der Fernsehtechnik oder Videokommunikation. Folgende einzelne Bereiche lassen sich unterscheiden:

CATV	C Ale T V distribution on optical networks, copper, etc.
CDAD	C able D igital A udio D istribution
DAB	D igital A udio B roadcasting (terrestrial and satellite broadcasting)
DTTB	D igital T errestrial T elevision B roadcast
EC	E lectronic C inema
ENG	E lectronic N ews G athering (including S NG, S atellite N ews G athering)
HTT	H ome T elevision T heatre
IPC	I nter P ersonal C ommunications (videoconferencing, videophone, etc.)
ISM	I nteractive S torage M edia (optical disks, etc.)
NCA	N ews and C urrent A ffairs
NDB	N etworked D ata B ase service (via ATM, etc.)
RVS	R emote V ideo S urveillance
SSM	S erial S torage M edia (digital VTR, etc.)
STV	S atellite T V broadcasting
TTV	T errestrial T V broadcasting

Dabei verfolgt MPEG-2 genau wie MPEG-I das grundlegende Prinzip des ungleichem Encoding-/Decoding-Verhältnis (Studio-/Home-Hardware), um die Kodierung zu optimieren. Folgende grundlegenden Anforderungen müssen durch das Format in den oben genannten Bereichen erfüllt werden:

- innerhalb der Bereiche Satellite Broadcast, Electronic Cinema und Digital Home Television wird eine angepaßte Bandbreite < 3 MBit/s benötigt
- der Videoteil der Kodierung muß eine erhebliche Flexibilität aufweisen. U. a. müssen verschiedenste Bildformate, eine wahlfreie Bildqualität, flexible Bitraten, wahlfreier Zugriff auf verschiedenste Video-Kanäle (Channel hopping), eine nachträgliche und einfache Editierung des kodierten Bitstreams und verschiedenste Trick modes garantiert werden. Eine Wiederholung des Kodier-/Dekodiervorgangs darf nicht zu weiteren Qualitätsverlusten führen.
- der Audioteil der Kodierung muß Mehr-Kanal Audio-Kodierungen, multilinguale Kanäle und niedrigere Sampling-Frequenzen unterstützen.

II. Grundlegende Kodierungsmechanismen

Als Einführung zu MPEG-2 müssen zuerst die grundlegenden Kodierungsmechanismen erleutert werden, dies geschieht am Beispiel der Standards, die die jeweiligen Mechanismen zuerst entwickelt haben.

1. JPEG

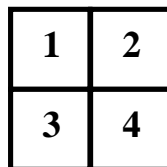
Der internationale Standard Joint Picture Experts Group" (JPEG) (ISO/IEC SC29/WG10) ist ein Standard zur Kodierung von photographischen Standbildern. Der Basisalgorithmus beschreibt eine Transformationskodierung auf der "Diskreten Cosinus Transformation" (DCT).

1.1 YUV-Farbraum

Da sich gezeigt hat, daß das menschliche Auge Farbwerte in einer geringeren Auflösung als Helligkeitswerte wahrnimmt, wird der gängige RGB-Farbraum (Rot-Grün-Blau) bei den digitalen Bildstandards in drei Planes umgewandelt, eine Luminanz-Plane (Grauwerte, Y) und zwei Chrominanz-Planes (Farbwerte U/V oder auch Cr/Cb). Die Farbwertplanes können dann von ihrer Kantenlänge halbiert werden (das ist dann ein Viertel der Fläche !). Dies reduziert die Datenmenge enorm, ist jedoch ein Datenverlust. Diese Art der Kodierung nennt man den 4:1:1 YUV-Farbraum.

Abbildung I: Farbraum

4:1:1



Y



Cb



Cr

1.2 Diskrete Cosinus Transformation

Aus den zur Verfügung stehenden Transformationen hat sich die DCT (Diskrete Cosine Transform, abgeleitet aus der Diskreten Fourier Transformation) als besonders effizient erwiesen. Eine auf einen 8x8 Pixelblock angewendete DCT ergibt wiederum einen 8x8 Pixelblock. Die Koeffizienten der DCT lassen sich als Spektrum des 8x8 Eingabeblock interpretieren.

Der Koeffizient mit den Frequenzen Null in horizontaler und vertikaler Richtung wird als DC-Koeffizient bezeichnet. Die restlichen 63 Koeffizienten werden als AC-Koeffizienten bezeichnet. "Während die Energie des Bildsignals zufällig verteilt sein kann, konzentriert sich die Energie des korrespondierenden DCT-Blocks vorzugsweise auf Koeffizienten mit niedrigen Frequenzen" [MUS93]. Werden die Koeffizienten im Zick-Zack durchnummeriert, ergeben sich als zu speichernde Werte ein DC-Koeffizient, dann wenige niedrige AC-Koeffizienten und viele AC-Koeffizienten nahe Null. Die DCT ist ein verlustfreies Verfahren, da die Kodierung komplett umkehrbar ist.

Nun eignen sich diese Werte jedoch optimal um sie mit dem Huffman-Verfahren (eine Art Top-Down Shannon-Fano-Kodierung; häufige Bytewerte werden durch kurze Bitfolgen ersetzt) und anschließend nach dem Lauflängen-Verfahren zu kodieren (sollten sieben Nullen im Strom nacheinander folgen, braucht man nur die 7 und die 0 zu kodieren).

Da man aber beim JPEG-Verfahren auch die Kompressionsrate angeben kann (die höherfrequenten Signale des Blocks werden dann ignoriert), fassen wir die Kodierung daher wie folgt zusammen:

- verlustbehaftete Kodierung
- 4:1:1 YUV-Farbraum
- DCT - Discrete Cosine Transform
- Entropie und Huffman-Kodierung

2. H.261

"Der H.261-Standard kann zur Kompression von Bildsequenzen für Videokonferenzen oder für Bildfernsehen verwendet werden." [MUS93]. Zusammenfassend hat H.261 folgende Charakteristika:

- Standard zur Übertragung von digitalen Video-Sequenzen der CCITT
- bei $p \approx 64$ kbit/s
- bereits als Hardware vorhanden
- einfache Umsetzung RGB \Rightarrow YUV
- Intra-Frame-Kodierung (mit Prädiktion)
- auch DCT und Huffman

3. MPEG-I

Dieser "Draft International Standard" (DIS) ISO 11172 (ISO/IEC JTC1/SC2/WC11) ist in drei Einzel-Standards unterteilt:

Video	- Kodierungstechniken etc. entsprechend JPEG (s.o.)
Audio	- Psychoakustisches Modell
System	- behandelt Synchronisation und Multiplexing (auf diesen Teil wird hier durch fehlende Informationen nicht näher eingegangen)

Ein MPEG-Stream läßt sich also in mindestens 32 Video- und Audio-Spuren und 2 Systemspuren zur Synchronisation zerlegen.

3.1 MPEG-Begriffe

Ein MPEG-Video-Stream wird durch Aneinanderreihung von Intra- (I), Predicted- (P), und Bidirectional- (B) beschrieben. I-Frames sind komplett im JPEG-Format abgelegt, in P-Frames werden nur die Differenzen zu einem I-Frame kodiert, ein B-Frame kodiert die mittleren Differenzen eines I- und eines P-Frames. Jeder Frame wird wiederum in drei Planes, eine Luminanz-Plane (Grauwerte, Y) und zwei Chrominanz-Planes (Farbwerte, Cr und Cb), zerlegt. Alle Planes werden in sogenannte Macroblöcke unterteilt. Alle Planes werden in 8x8 Pixel-Blöcke aufgeteilt. Jeder dieser Blöcke wird mittels einer Discrete-Cosine-Transform (DCT) kodiert, dabei wird der erste Wert jedes Blocks als DC-Koeffizient (DC) bezeichnet; die restlichen Differenzwerte als AC-Koeffizienten (AC). DC's werden untereinander auch als Differenzen gespeichert. "Echte" DC's bezeichne ich als Master-DC's (MDC); einen als Differenz zu einem MDC gespeicherten DC als Differenz-DC (DDC).

Zusätzlich können Frames auch als Verweise der Macroblöcke auf gleiche, in vorherigen Frames gespeicherten Blöcken kodiert werden (hier bezieht man sich jedoch auf 16x16 Pixel-Blöcke). Dieses Verfahren nennt man Motion Compensation.

3.2 MPEG-Parameter

Im Gegensatz zu JPEG werden in MPEG feste Quantisier- und Huffman-Tabellen verwendet, diese sind zwar nicht für jede Video-Übertragung optimal, können dann jedoch auch in Hardware kodiert werden. Dies ist unerlässlich für die Fernsehtechnik.

Die Bildgröße ist variabel (in 16-Pixel-Schritten) und kann maximal horizontal 720 pels und vertikal 576 pels betragen, dies ist die Auflösung eines normalen S-VHS Signals. Dies wird jedoch nur durch die gewünschte Datenrate begrenzt und nicht durch die Kodierungsvorschriften (hier kann man bis zu $4096 * 4096$ pixel bei 60 Hz kodieren).

Bei einem zu erzielenden Datendurchsatz von 1,5 Mbit/s werden maximal 386 Kbit/s für den Audio-Bereich verwandt.

Durch das Kodierungsverfahren, sind Einzelbildzugriff, schnelles suchen (Positionierung auf I-Frames), rückwärts spielen und die Editierbarkeit des Datenstroms gegeben.

MPEG wurde durch die Wahl der Kodierung auf asymmetrische Anwendungen (Fernseh-Technik, Multimedia-Mail, etc.) zugeschnitten, mit vermehrtem Hardware-Einsatz können jedoch auch symmetrische Verfahren implementiert werden. Insbesondere benötigt die Dekodierung keine Hardwareunterstützung.

MPEG ist nicht die optimalste Videokodierung, allerdings ist sie der einzig internationale Konsens in diesem Bereich und wird, auch dank seiner Verwandtschaft mit den Telekommunikations-Standards (H.261, S-VHS) weite Verbreitung finden.

3.3 MPEG-Audio

Der Audio-Teil des MPEG-Drafts beschreibt Mechanismen und Algorithmen, mit denen die digitale Speicherung von Audiosignalen auf kostengünstigen Speichermedien auf der einen Seite und die digitale Übertragung von Audiosignalen auf Kanälen mit begrenzter Kapazität auf der anderen Seite ermöglicht werden sollen. Bei all diesem Zielstreben steht jedoch die Erhaltung der Qualität in einem bestimmten Bereich im Vordergrund. Hier im Audio-Bereich wird einer der Compact Disc nahekommende Qualität angestrebt.

3.3.1 Technische Daten

Nachfolgend werden zum Vergleich der Datenverhältnisse einige verschiedene Beispiele aus der praktischen Anwendungswelt und der hier behandelte Standard aufgeführt, woraus schon die Leistungsfähigkeit dieses Standards ersichtlich wird.

Studioformat

Die Darstellung eines stereophonen Audiosignals im Studioformat erfordert eine Abtastfrequenz von 48kHz und eine gleichförmige Quantisierung von 16bit pro Abtastwert. Daraus ergibt sich eine Datenrate von 768kbit/s für ein Monosignal, als Produkt der Multiplikation der 48kHz mit den 16bit/Abtastwert. Daraus resultierend ergibt sich für ein Stereosignal eine Datenrate von 2×768 kbit/s, also ca. 1,5Mbit/s.

Compact Disc

Als Vergleich dazu wird auf einer Compact Disc mit einer Abtastfrequenz von 44,1kHz bei der gleichen Quantisierung von 16bit/Abtastwert gearbeitet, wodurch sich für ein Monosignal eine Datenrate von $44,1\text{kHz} * 16\text{bit/Abtastwert}$, also ca. 706kbit/s ergibt. Somit errechnet sich für ein stereophones Signal eine Datenrate von 2×706 kbit/s, also ca. 1,4Mbit/s.

MPEG-Audio-Standard

Im MPEG-Audio-Standard werden zwei Abtastfrequenzen verwendet, nämlich zum einen 32,441kHz und zum anderen 48kHz. Aber im Gegensatz zu den oben beschriebenen Fällen ergeben sich hier im Endeffekt Datenraten zwischen 32kbit/s und 192kbit/s für ein Monosignal. Für ein Stereosignal ergeben sich Datenraten zwischen 128kbit/s und 384kbit/s. Mit einer Datenrate unter 128kbit/s (bis 64kbit/s) können leider noch keine zufriedenstellende Ergebnisse erzielt werden, näheres hierzu sind weiter unten aufgeführt.

3.3.2. Methoden

Das Ziel des Standards ist mit einer von 1,5Mbit/s im Studioformat auf 256kbit/s reduzierten Datenrate eine der Qualität einer Compact Disc ebenbürtige Qualität zu erreichen, wobei auch bei niedrigeren Datenraten wie 192kbit/s bis hinunter zu 128kbit/s noch akzeptable Qualitäten erzielt werden sollen. Hierbei ist jedoch die Einschränkung zu machen, daß in diesen unteren Datenraten-Bereichen bei kritischen Testsignalen Qualitätsminderungen durch das menschliche Gehör wahrnehmbar sind, d.h. hier treten bei der Kodierung Fehler auf, die sich nicht mehr in einem akzeptablen Rahmen bewegen (nähere Erläuterungen sind weiter unten aufgeführt).

Das menschliche Gehör ist im Allgemeinen ja sowieso bei Störungen im Audio-Bereich empfindsamer als im visuellen Bereich, d.h. kurzzeitiges "Rauschen" und "Knacken" ist störender als "Flimmern" oder sonstiges im visuellen Bereich, daher wird eine der CD vergleichbare Qualität angestrebt.

In Layer I wird mit einer Datenrate von 384kbit/s, in Layer II und Layer III mit einer Datenrate von 256kbit/s eine der Compact Disc vergleichbare Qualität erreicht. Bei Datenraten von 192 bzw. 96 kbit/s kommt man sowohl mit dem Layer II Codec als auch mit dem Layer III Codec der gewünschten Qualität sehr nahe, jedoch sind bei einigen kritischen Testsignalen Kodierungsfehler deutlich wahrnehmbar.

Als sehr nützlichen Nebeneffekt erhält man bei dieser Art der Datenkomprimierung und Kodierung in diesem Standard die Rauschfreiheit, d.h. es sind keine zusätzlichen Verfahren zur Geräuschkürzung wie z.B. das Dolby System nötig.

Die in Layer I und Layer II beschriebenen und verwendeten Coder sind bereits vollständig als integrierte Schaltungen realisiert und stehen als Endgeräte bzw. Computerkarten zur Verfügung. Die Layer I Kodierung wird bereits in DCC-Recordern (**D**igital-**C**ompact-**C**assette) und einigen Multimedia-Systemen verwendet.

Die Layer II Kodierung wird vom europäischen DAB (**D**igital **A**udio **B**roadcasting) System verwendet. Zukünftiges Ziel dürfte es sein mit dem Layer III die CD-Qualität schon bei einer Datenrate von 2x64kbit/s zu erreichen, hierzu wird eine ideale Realisierung des *Joint Stereo Coding* notwendig sein. Dies ist in naher Zukunft zu erwarten.

Aufgrund dieser Entwicklungen in diesem Bereich darf man im Hinblick auf Multimedia-/Hypermedia-Systeme auch Fortschritte erwarten, da durch die Reduzierung der Datenmengen die Einbindung von Video- und Audiodaten einen realistischeren Bezug erhält. Zu weiteren Informationen betreffs MPEG-I verweise ich auf [**GAD93.4**].

II. MPEG-2-Draft

1. ISO - die Organisation

Der MPEG-2 "Third Working Draft" mit dem Titel "*Generic Coding of Moving Pictures and Associated Audio*" wurde von der "Motion Picture Expert Group" der ISO erarbeitet. Diese setzt sich aus den folgenden Untergruppen zusammen.

- ISO/IEC JTC1/SC29/WG11 xxxxx
 - ISO = International Organization for Standardization
 - IEC = International Electrotechnical Commission
 - JTC1 = Joint Technical Committee 1
 - WG11 = Work Group 11

Da dies bis jetzt noch ein "Working Draft" (WD) ist, wurde auch noch keine ISO-, bzw CCITT-Nummer vergeben, der Draft nennt sich selber noch "ITU-TS Rec. H.26x". Der Arbeitsplan für die Weiterentwicklung des Standards sieht wie folgt aus:

- Working Draft 1 November 1992
- Working Draft 2 März 1993
- Working Draft 3 July 1993
- Committee Draft November 1993

- Draft International Standard März 1994
- International Standard November 1994

2. MPEG-2-Teile

Auch der MPEG-2-Draft besteht wiederum aus drei Hauptteilen, nämlich dem Video-, dem Audio- und dem Systemteil. Im Draft selber wird zwar ein weiterer, vierter Conformance-Teil (der die Charakteristiken von MPEG-2-Bitstreams und Testergebnisse enthalten soll) erwähnt, er wurde jedoch bis jetzt noch nicht publiziert.

III. MPEG-2 Kodierungsmethoden

Drei grundlegende neue Kodierungsmethoden sollen in MPEG-2 verwirklicht werden.

1. Profiles und Levels

Erstens werden sog. Profiles und Levels definiert. Diese schränken die zur Verfügung stehenden Parameter der Kodierung ein, um dies dann in den Kompressionsalgorithmen ausnutzen zu können. Gleichzeitig standardisieren diese Profiles und Level auch die Kodierungsparameter für die speziellen, oben genannten Anwendungsbereiche. Die Kombination von Profiles und Levels läßt dabei aber immer genügend Spielraum für die unterschiedlichsten Anwendungen.

- Profiles = complexity of compression
- Level parameters = sample rate, frame dimension, coded bitrates

Folgende Profiles und Levels sind im Standard bis jetzt festgeschrieben:

Abbildung II: Profiles und Levels

Profile	B pictures?	Syntax Comments
Simple	No	Intended for software applications, perhaps CATV.
Main	Yes	Most decoder chips, CATV, satellite. 95% of users.
Main+	Yes	Main with Spatial scalability.
Next	Yes	Main+ with 4:2:2 macroblocks.

Level	Max. frame sample size	Pixels/sec	Max. bitrate	DC bits	Significance
Low	352 x 288	3.05 M	4 Mb/s	8,9,10	CIF, consumer tape equiv.
Main	720 x 480	10.40 M	15 Mb/s	8,9,10	CCIR 601, studio TV
High 1440	1440 x 1152	47.00 M	60 Mb/s		4x 601, consumer HDTV
High	1920 x 1080	62.70 M	80 Mb/s		production SMPTE 240M std

Abbildung III: Sampling dimensionen and bitraten für MPEG-2

Dimensions	Coded rate	Comments
352x480x24 Hz (progressive)	2 Mbit/sec	Half horizontal 601. Looks almost NTSC broadcast quality, and is a good (better) substitute for VHS. Intended for film src.
544x480x30 Hz (interlaced)	4 Mbit/sec	PAL broadcast quality (nearly full capture of 5.4 MHz luminance carrier). Also 4:3 image dimensions windowed within 720 sample/line 16:9 aspect ratio via pan&scan.
704x480x30 Hz	6 Mbit/sec	Full CCIR 601 sampling dimensions.

2. Scalability

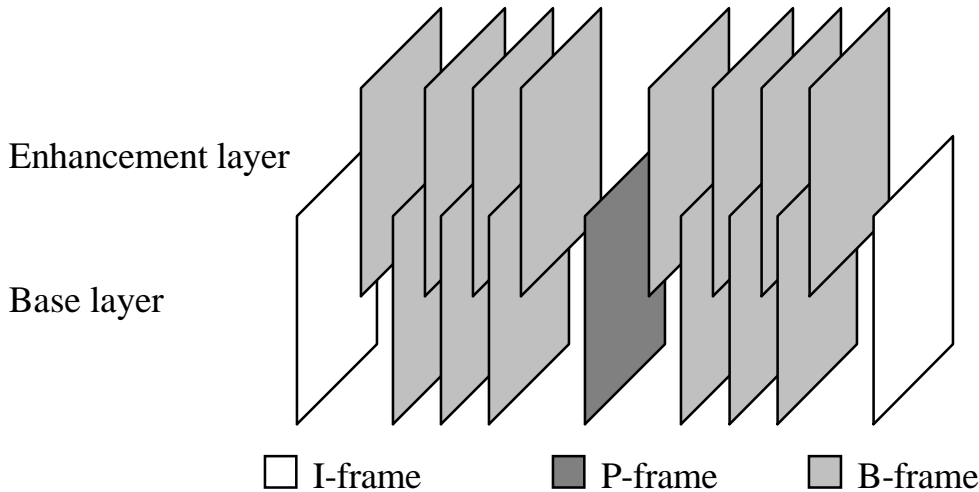
Zweitens wird die Scalability von Video und Audio verwirklicht. Sie liefert die Möglichkeit von unterschiedlicher Endnutzung bei ein- und derselben feststehenden Kodierung. So kann der Endnutzer entscheiden, welche Teile einer Übertragung er konsumieren möchte.

"Scalability: the ability of a decoder to ignore some portions of a total bitstream and produce useful audio and video output from the portion which is decoded."

Zusätzlich wird MPEG-2 somit weitgehends Speichermedienunabhängig. Innerhalb der Einzelbild- (oder Frame-) scalability lassen sich z.B. die folgenden beiden Modi denken (andere Kombinationen sind natürlich möglich):

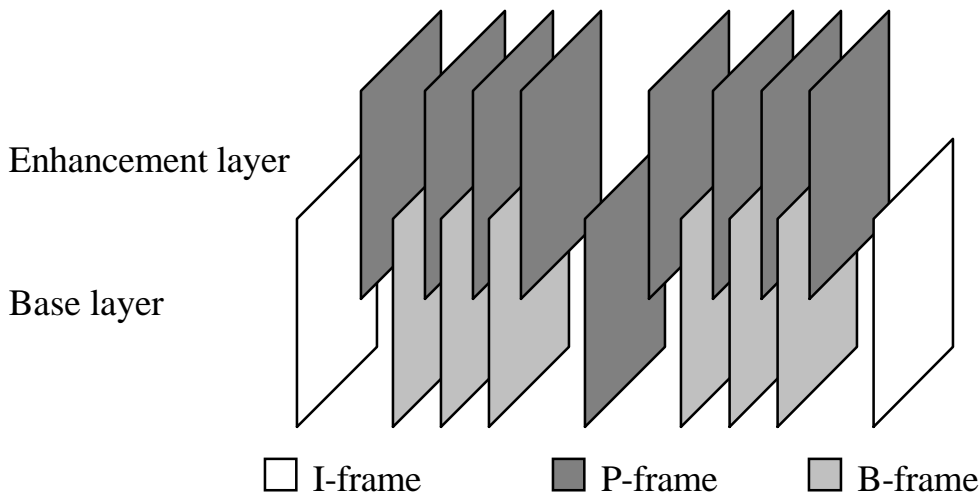
Erstens die zeitliche Scalability, d.h. sollten die Ressourcen des konsumierenden Systems ausreichen, kann die Framerate erhöht werden, indem in den normalen Ablauf des "Base Layers" zusätzlich B-frames des "Enhancement Layers" dekodiert werden.

Abbildung IV: Temporal scalability



Zweitens die Qualitäts-Scalability, d.h. anstatt der ungenaueren B-frames des "Base Layers" werden P-frames des "Enhancement Layers" dekodiert und angezeigt. Dies erfordert natürlich mehr Rechenzeit.

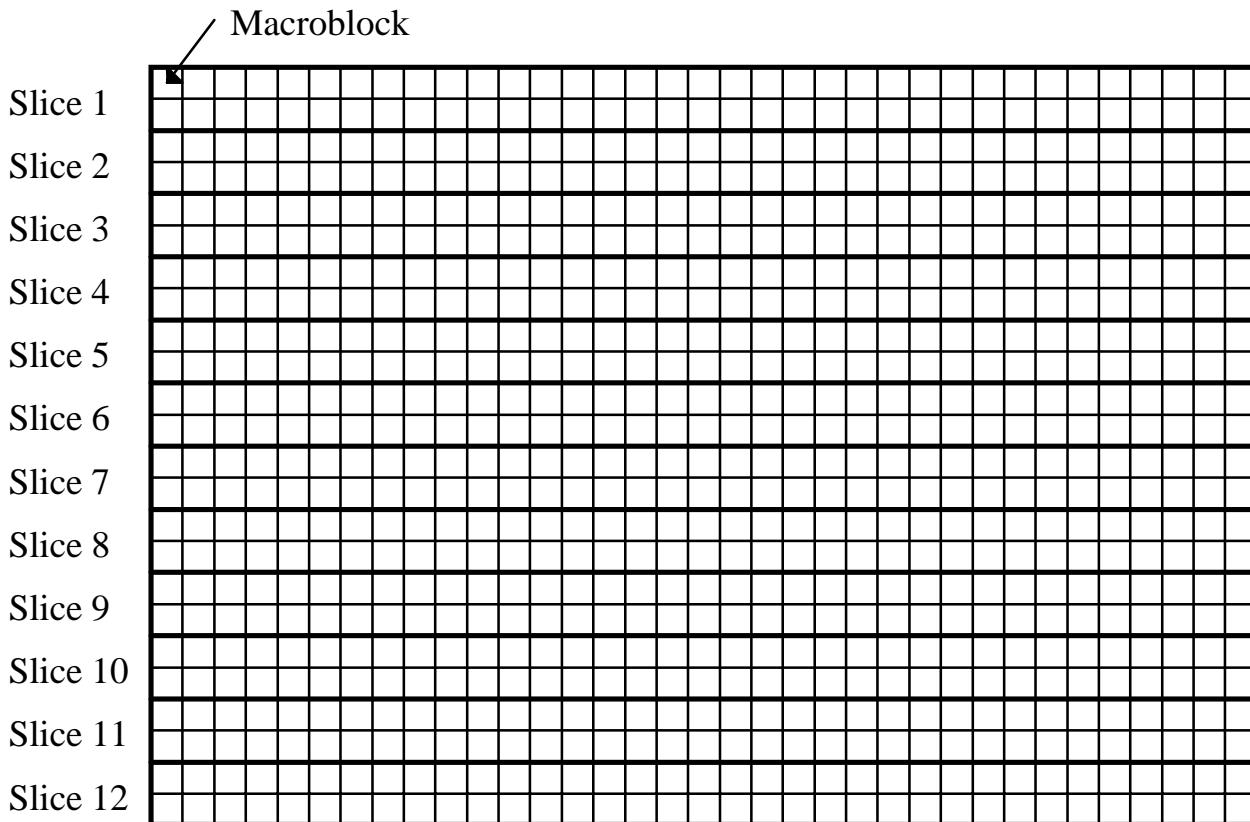
Abbildung V: Quality scalability



Zusätzlich sind noch die Kodierung von 3-dimensionalen Bildern (stereoscopic view) und die sichere oder geheime Übertragung von Teilstreams mit der Scalability-Funktion möglich.

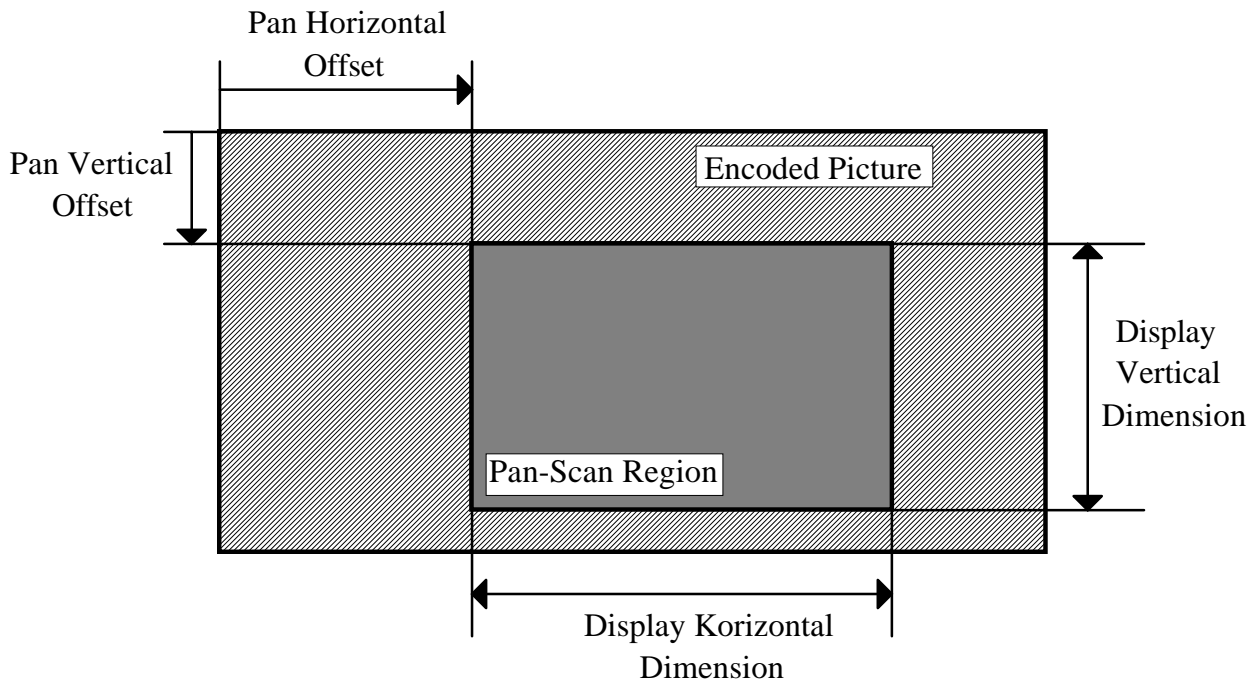
Weiterhin wird noch die Macroblock-scalability definiert. Dabei handelt es sich um einen zusätzlichen Eintrag in jedem Macroblock, der auf den nächsten wichtigen (mit vielen Änderungen gegenüber dem letzten Frame) Macroblock verweist, sollte also die Dekodierzeit nicht ausreichen, können "unwichtige" Macroblöcke übersprungen werden.

Abbildung VI: Macroblock scalability



Schlußendlich gibt es noch die Pan-Scan-Scalability. Diese ermöglicht die Definition von Ausschnitten im aktuellen Bild. Diese Ausschnitte können zwar von Frame zu Frame unterschiedlich eingeteilt werden, die Hauptanwendung ist jedoch die Definierung eines 4:3 Fernseh-Bildes innerhalb eines 16:9 HDTV-Bildes.

Abbildung VII: Pan-Scan



3. Sicherheit

MPEG-2 ist auch der erste der vielen Multimedia-Video-Definitionen, die auch an den Bereich der Sicherheit (Vertraulichkeit und Integrität) denkt; MPEG-I hatte schon im Audio-Streams einen rudimentären CRC-check eingeführt. MPEG-2 soll jedoch robust gegenüber Bit-Error und Zellen-Verlusten werden, die Möglichkeit sog. "Stream Identification Signals" bieten und einen speziell auf MPEG-2 optimierten, schnellen Cypher-Algorithmus enthalten, der ein "Scrambling" des kompletten Streams ermöglicht.

IV. MPEG-2 Features

Da es sich bei dem aktuellen MPEG-2-Dokument noch um einen "Working Draft" handelt und bis jetzt auch noch kaum andere Publikationen zu diesem Thema erschienen sind, kann man an dieser Stelle nur die beabsichtigten Features von MPEG-2 erwähnen, die zugrundeliegenden Algorithmen und Methoden sind noch zu unklar und im Draft nicht erwähnt.

1. Video

MPEG-2 soll laut Anwendungsprofil auf alle Eventualitäten in der digitalen Video-, Übertragungs- und Fernsehtechnik vorbereitet sein. Ein wichtiges Feature ist daher die Rückwärtskompatibilität zu existierenden digitalen Kodierungen wie MPEG-I oder H.261. Diese Rückwärtskompatibilität ist keine Bit-Übereinstimmung der resultierenden Streams, sondern sorgt über die Benutzung und Variierung der Profiles und Levels nur dafür, daß Hard- und Software die z.B. zum Einsatz von MPEG-I bestimmt war, auch mit MPEG-2 (mit beschnittenen Features dann natürlich, denn MPEG-2 ist von der benötigten Rechenleistung zur Kodierung viel aufwendiger) zu betreiben. Allerdings sind auch sog. Sub-Koder innerhalb eines MPEG-2-Koders angedacht die exakt die rückwärtskompatiblen Streams erzeugen.

Z.B. ergibt die Kombination von Video Main Profile (wie MPEG-I) und Video Main Level (CCIR 601, Studio TV) eine MPEG-I-ähnliche Kodierung, sie wird im Kurzen MP@ML genannt.

Eine erweiterte Komprimierungsmethode mit dem Namen "Block-based Motion Compression Prediction" (MCP) soll in MPEG-2 benutzt werden. Sie ermöglicht "Frame Motion Prediction" (das Auffinden von ganzen, gleichen Frames innerhalb des Streams), "Field Motion Prediction" (das Auffinden von gleichen, wie auch immer geformten Teilen in verschiedenen Frames und auch innerhalb des aktuellen Frames) und das zeitliche Vorhersehen von Kodierungsfehlern. Zusätzlich werden sog. "Backward Motion Vectors" eingeführt, in MPEG-I waren Motion Vektoren immer nur auf nachfolgende Frames gerichtet.

Motion Vektoren in MPEG-2 können auch genau bis auf ein halbes Pixel innerhalb des Frames bestimmt werden, dies ermöglicht das Auffinden von langsamen Verschiebung.

Diese Methoden ermöglichen zwar eine wesentliche genauere Analyse der Videodaten, erhöhen den benötigten Rechenaufwand jedoch enorm.

Zusätzliche Features des MPEG-2-Videoteils sind (neben den oben und vorher genannten):

- variable Macroblockgrößen
- downloadable length coding tables
- Huffman oder JPEG-Methoden zur Variablen-Kodierung
- Quant matrix download extension
- Interlace prediction and compression extension
- alternative Scanmethoden für "Interlaced Video"
- Trick modi durch Ausnutzung der Scalability

Um die unterschiedlichsten Anwendungsprofile zu unterstützen wird auch ein variabler Farbraum eingeführt. 4:1:1 wird wegen MPEG-I weiterhin unterstützt, für das Broadcasting (Übertragung von Fernsehbildern zwischen verschiedenen Sendern) und dem professionellen Einsatz von digitaler Studioteknik ist jedoch ein Farbverlust nur in geringem Maße oder gar nicht zu akzeptieren.

Abbildung VIII: Variabler Farbraum

4:2:2

1	2
3	4

Y

5
6

Cb

7
8

Cr

4:4:4

1	2
3	4

Y

5	6
7	8

Cb

9	10
11	12

Cr

2. Audio

Der Audioteil von MPEG-2 ist wiederum genau wie bei MPEG-I auf die Kodierung von Audiosignalen hoher Qualität (CD, Studio) zugeschnitten, weist aber eine viel höhere Flexibilität auf. Zusätzlich zu den festgelegten 'normalen' Sampling-Frequenzen von 22,05, 44.1 oder 48 kHz sind auch solche von 8, 11, 16 oder 24 kHz möglich, dies insbesondere zur Kodierung von Signalen aus digitalen Quellen wie ISDN, G.722 oder Studiosamples. Ein variables Bit/Sample-Verhältnis wird erlaubt.

Um die Rückwärtskompatibilität zu MPEG-I und insbesondere die Unterstützung von HDTV und Dolby-Surround zu ermöglichen, kann man in MPEG-2 bis zu 5 Kanäle kodieren (plus einen Kanal mit tiefstfrequenten Anteilen bis zu 100 Hz für Spezialeffekte). Diese Kanäle können in 3 aufgeteilt werden (links, mitte, rechts = HDTV), die anderen beiden werden als Raumkanäle genutzt (Dolby-Surround). Speziell über die 3 Kanäle lassen sich vielfältige Kombinationen zur multilingualen Übertragung wählen.

Durch die Vielfältigkeit der Kodierung und Kanäle werden auch neue Möglichkeiten zur Kompression der Audiosignale erschlossen (zusätzlich zu den von MPEG-I bekannten Mechanismen). Zu nennen sind

- Intensity Stereo Coding
- Phantom Coding of Center
- Dynamic Transmission Channel Switching
- Dynammic Cross Talk
- Adaptive Multi-Channel Prediction

Eine erweiterte Funktionalität im Bereich "Motion Compensation" und "Motion Estimation" werden definiert. Alle Samples (bzw. Offsettabelle und Subbandinformationen) können nun durch Integritätschecks geschützt werden.

3. System

Dem Systemteil bei MPEG-2 kommt nun, da mehrere Video- (Base und Enhancementlayer) und Audio-Streams innerhalb eines MPEG-2-Stream definierbar sind, eine immer größere Bedeutung zu. Ein neues, generelles Feature, nämlich die Unterteilung in Transport- und Program-Stream wird nötig.

Der Transportstream ist eine Kodierung, die mehrere einzelne Program-Streams zusammenfaßt und in Pakete aufteilt (splicing). Außerdem ermöglicht er

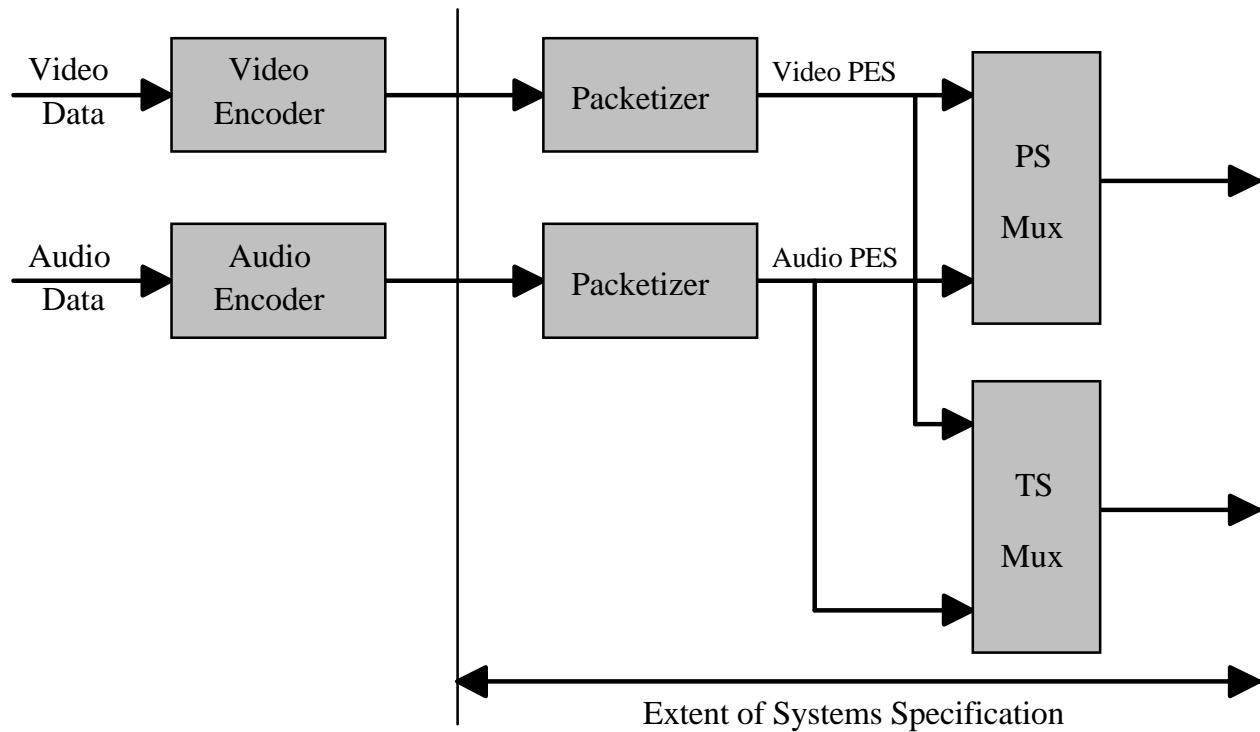
- Error protection and error correction (durch CRC-checks)
- Padding von Audio-Streams

- Entropy coding
- Scrambling

Ein einzelner Programstream wiederum ermöglicht das Multiplexen (MUX) von einzelnen Video- bzw. Audio-Streams, sorgt für die zeitgenaue Synchronisation derselben (dies war eigentlich die einzige Aufgabe des MPEG-I-Systemteils !) und speichert eine Art Verzeichnis mit Informationen und Einsprungsadressen der Streams.

Der Systemteil muß auch 'Wissen' über die kontrollierenden Medien haben (Digital Storage Media Control / Time Stamp Coding). Weiterhin definiert er den "Packetized Elementary Stream" (PES) und konvertiert zwischen Transport und Programm Datenstrom.

Abbildung IX: System Encoder



V. Ausblick

"Computer: the including, combining AND substituting media ?"

1. Andere MPEG-Kodierungen

Nicht unerwähnt sollen die weiteren Bestrebungen der MPEG-Gruppe sein. MPEG-3 war speziell für die Kodierung für HDTV gedacht, hat aber durch die Flexibilität von MPEG-2 an Gewicht verloren. HDTV überträgt mit bis zu 1920 * 1080 pixel bei 30 Hz, das ergibt eine Nettodatenrate von 20-40 Mbit / sec. Durch eine "fine-getune"te Kodierung mit MPEG-2 (High Level und Main+ Profile) läßt sich eine genügend gute Kodierung für HDTV finden, außerdem werden damit gleichzeitig die Standardisierungsstreitigkeiten zwischen USA/Japan und Europa wegen HDTV entschärft. MPEG-3 wird es wahrscheinlich nicht geben.

MPEG-4 ist noch nicht konkretrisiert. Es soll ebenso wie MPEG-2 generische Methoden zur Kodierung von Audio-Visuellen Systemen bieten, allerdings für kleinste Datenraten (bis zum 10fachen von kbit/s, also maximal 100.000 kbit/s). Diese Übertragungsraten werden heutzutage von digitalen Telefonleitungen erbracht und würden dann eine Video- und Audioübertragung bei ca. 176 * 144 pixel bei 10 Hz gestatten, z.B. über ISDN. Um diese Kompression zu erreichen werden folgende Kompressionmethoden angedacht:

- Fraktale Kompression
- Morphing
- Pyramid Image Coding

2. Die Zukunft für MPEG-2 ?

MPEG-2 ist im Gegensatz zu MPEG-I nicht für Softwarelösungen auf heutigen Hardwareplattformen geeignet. MPEG-2 setzt dafür eine viel zu hohe Rechenleistung für die Analyse der gegebenen Video- und Audiodaten voraus. Da MPEG-2 auch für hochauflösende Video- und Audioübertragungen konzipiert wird, kann auch keine Dekodierung des Datenstroms ohne Hardwareunterstützung erfolgen.

MPEG-2-Hardware ist bis heute allerdings noch nicht vorhanden, Entwicklungen befinden sich in der Testphase und werden in Software getestet.

Nach dem durchschlagenden Erfolg von MPEG-I in der digitalen Videokodierung ist (mit dem notwendigen Druck durch die Bildverarbeiter wie z.B. der Satelitentechnik und dem Fernsehen) damit zu rechnen, daß MPEG-2 ein integraler Bestandteil in der Fernsehtechnik auch beim Endbenutzer werden wird. In vielleicht 5 Jahren fragt man beim Kauf eines Fernsehgeräts nicht mehr, ob es mit PAL oder NTSC arbeitet, sondern ob es HDTV oder MPEG-2 dekodieren kann.

Anhang A: Quellen

- [BAD93] Badura, Rolf-Stefan, Meyer-Zajontz, Jörg: Quicktime, DVI und MPEG, Referat, Jan. 93
- [BOR91.3] Bormann, Ute: ISIS 1 Internationale Standards for Informationstechnik - Systemarchitektur, 1991
- [CHI92] Chiariglione, Leonardo: Multimedia Communication (MPEG-2), Brussels 1992
- [CT93] c't (Kürzel ku): MPEG-2-Standard festgeklopft, Heft 6, 1993
- [GAD93] Gadegast, Frank: Offene Dokumentverarbeitung mit multimedialen Standards, Referat, Jan. 1993
- [GAD93.2] Gadegast, Frank, Jürgen Meyer: Integrität von Multimedialen Daten am Beispiel MPEG, Semesterarbeit, Jun. 1993
- [GAD93.3] Gadegast, Frank, : Graphische Benutzeroberflächen und die Multimediafähigkeit, Referat, Jun. 1993
- [GAD93.4] Gadegast, Frank, Metin Çetinkaya : MPEG-I ISO 11172, Referat, Jun. 1993
- [HUN93] Hung, Andy C.: PVRG-MPEG-CODEC 1.1, Manual, Stanford, 1993
- [HUN93.2] Hung, Andy C.: PVRG-JPEG-CODEC 1.1, Manual, Stanford, 1993
- [ISO92] ISO/IEC Draft International Standard (DIS) 11172: Information technology - Coding of moving pictures and associated audio for digital storage media up to about 1,5 Mbit/s (MPEG), International Organization for Standardization, Geneva, 1992
- [ISO93] ISO/IEC Third Working Draft xxxxx, H.26x: Information technology - Generic Coding of moving pictures and associated audio (MPEG-2), International Organization for Standardization, Geneva, 1993
- [MUS93] Musman, Hans-Georg; Werner, Oliver; Fuchs, Hendrik: Kompressionsalgorithmen für interaktive Multimedia-Systeme, IT+TI 2/93, Hannover 1993
- [PAT92] Patel, Ketan; Smith, Brian C.; Rowe, Lawrence A.: Performance of a Software MPEG Video Decoder, Berkeley, 1992

Anhang B: Abbildungen

Abbildung I: Farbraum

Abbildung II: Profiles und Levels

Abbildung III: Sampling dimensionen and bitraten für MPEG-2

Abbildung IV: Temporal scalability

Abbildung V: Quality scalability

Abbildung VI: Macroblock scalability

Abbildung VII: Pan-Scan

Abbildung VIII: Variabler Farbraum

Abbildung IX: System Encoder