

Referat:

# MPEG-Standard ISO 11172

**Kodierung von Bewegtbildern und synchronisiertem Audio zum  
digitalen Speichern bis 1,5 Mbit/s**

zum Seminar: Breitband-Netztechnologien und Multimedia-Systeme

von Metin Çetinkaya und Frank Gadegast

gehalten am 29. Juni 1993

**Adressen: Metin Çetinkaya**  
Allerstraße 10  
12049 Berlin  
GERMANY

**Fon:** ++ 49 30 6223102

**E-mail:** brain@cs.tu-berlin.de

**Frank Gadegast**  
Leibnizstraße 30  
10625 Berlin  
GERMANY

**Fon/Fax:** ++ 49 30 3128103

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

# Inhaltsverzeichnis:

<b>MPEG.....</b>	<b>1</b>
1. JPEG .....	1
1.1 YUV-Farbraum.....	1
1.2 Diskrete Cosinus Transformation .....	1
2. H.261.....	2
3. MPEG-I.....	2
3.1 MPEG-Draft.....	2
3.2 MPEG-Begriffe.....	2
3.3 MPEG-Parameter .....	2
3.4. Resümee .....	3
4. MPEG-Audio.....	3
4.1. Technische Daten.....	3
4.2. Ziele .....	4
4.3. Kodierungsmodi.....	4
4.4. Grundlegendes Kodierungskonzept .....	4
4.5. Die drei Layer .....	5
4.6. Resümee .....	8
5. Sicherheitmechanismen für Multimedia-Daten (Integrität).....	9
5.1. Fehlerarten und Fehlerpropagierung .....	9
5.1.1 Fehler innerhalb von Macro-Blöcken.....	9
5.1.2 Unterschiede von Fehlern in Y- und Cr-/Cb-Planes .....	9
5.1.3 Fehler in P- und B-Frames, Motion Vektoren oder Headern....	10
5.2. Szenarien .....	11
5.2.1 Leichte Fehlerkontrolle - Zusammenfassung.....	11
5.3 Resümee .....	12
6. MPEG-II.....	12
7. DVI .....	13
8. AVI .....	13
9. Beispiele .....	14
9.1 Xing.....	14
9.2 Berkeley und Stanford.....	14
9.3 NVR.....	14
 <b>Anhang A: Quellen.....</b>	 <b>15</b>
 <b>Anhang folgend: Abbildungen.....</b>	 <b>15</b>

# MPEG

Wir wollen hier nur ein Videoformate betrachten, das auch als internationaler Standard verabschiedet wurde. PD-Standards (wie GL, FLI, DL etc.) sollen hier nicht und Industrie-Standards (wie DVI, AVI etc.) sollen hier nur am Rande betrachtet werden.

Die digitale Dastellung 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 den hier zu behandelnden Standards können jedoch auch Video- und Audio-Signale auf bereits bestehende Medien übertragen und gespeichert werden.

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

## 1. JPEG

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

### 1.2 Diskrete Cosinus Transformation

Aus den zur Verfügung stehenden Transformationen hat sich die DCT (Discrete 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\*64 kbit/s
- bereits als Hardware vorhanden
- einfache Umsetzung RGB  $\Rightarrow$  YUV
- Intra-Frame-Kodierung (mit Prädiktion)
- auch DCT und Huffman

## 3. MPEG-I

*"Coded Representation of Picture, Audio and Multimedia/Hypermedia Information"*

MPEG heißt "Motion Picture Expert Group" und ist die Gruppe der ISO, die sich mit der Standardisierung im Video-Bereich beschäftigt. Im Dezember 1992 wurde ein Draft International Standard (DIS 11172) mit dem übersetzten Titel "Kodierung von Bewegungsbildern und assoziierten Audio für digitale Speichermedien mit bis zu 1,5 MBit/s" vorgelegt.

### 3.1 MPEG-Draft

Dieser DIS (ISO/IEC JTC1/SC2/WC11) ist in drei Einzel-Standards unterteilt:

<b>Video</b>	- Kodierungstechniken etc. entsprechend JPEG (s.o.)
<b>Audio</b>	- Psychoakustisches Modell
<b>System</b>	- 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.

### 3.2 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.

### 3.3 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.

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.

### **3.4. Resümee**

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.

## **4. MPEG-Audio**

Der Audio-Teil des MPEG-Draft 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.

### **4.1. 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  $2 \times 768 \text{kbit/s}$ , 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  $44,1 \text{kHz} \times 16 \text{bit/Abtastwert}$ , also ca. 706kbit/s ergibt. Somit errechnet sich für ein stereophones Signal eine Datenrate von  $2 \times 706 \text{kbit/s}$ , 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.

## 4.2. Ziele

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.

## 4.3. Kodierungsmodi

Innerhalb der Kodierung sind verschiedene Modi zu unterscheiden, genauer gesagt sind es vier an der Zahl. Als erstes steht das *Single Channel Coding* zur Verfügung, das zur Kodierung von Monosignalen herangezogen wird, dazu kommt noch das *Dual Channel Coding* als Mittel zur Kodierung von z.B. bilingualen Monosignalen (wie z.B. Zweikanalton im Bereich des TV, wo auf jedem der zwei Kanäle eine andere Sprache ausgestrahlt wird). Weiterhin existiert das *Stereo Coding* zur Kodierung eines Stereosignals, bei diesem Verfahren werden die beiden Kanäle trotzdem separat codiert.

Als letztes ist das *Joint Stereo Coding* zu nennen, das ebenso wie das *Stereo Coding* zur Kodierung eines Stereosignals benutzt werden kann, jedoch mit dem einen Unterschied, daß bei diesem Verfahren die Datenredundanz und -irrelevanz zwischen den beiden Kanälen ausgenutzt und somit eine Datenverminderung erreicht werden soll.

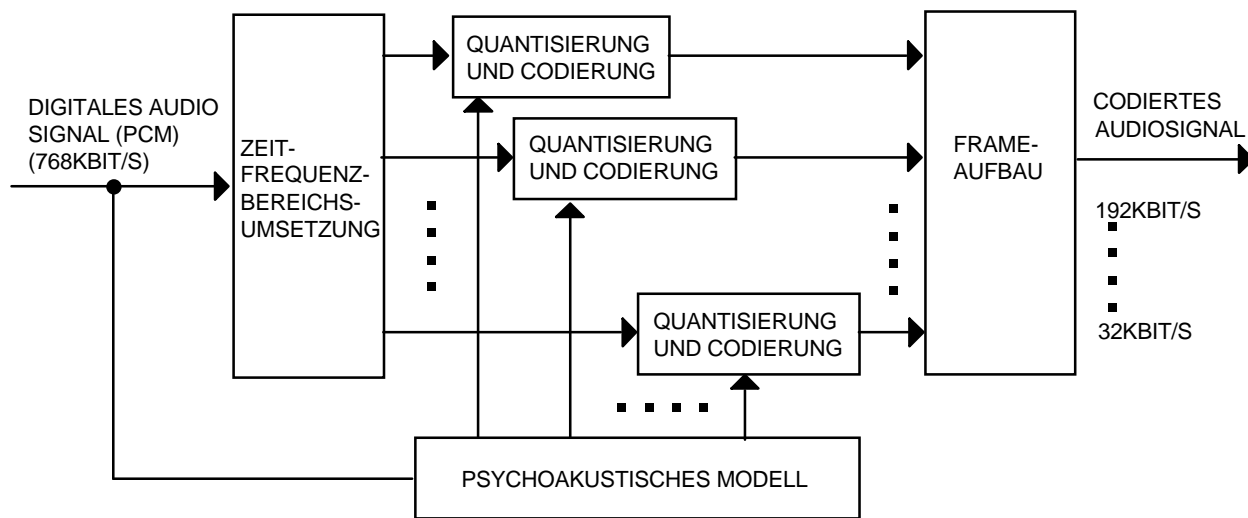
Dieses zuletzt genannte Verfahren ist jedoch noch nicht realisiert, das wird eine Aufgabe für die zukünftige Erweiterung bzw. Verbesserung des Standards sein.

## 4.4. Grundlegendes Kodierungskonzept

Das digitale Eingangssignal wird in 32 gleichförmige Spektralkomponenten (Frequenzgruppen, Teilbänder) zerlegt, dieses Grundprinzip entspricht dem Vorgang im menschlichen Gehör (Psychoakustik). Der Vorgang wird als Zeit-Frequenzbereichs-Umsetzung bezeichnet. Die Spektralkomponenten werden dann in Abstimmung auf die Wahrnehmungseigenschaften des menschlichen Gehörs codiert. Diese Kodierung wird von einem der drei Layer durchgeführt. Die Quantisierung und Kodierung wird unter Einbeziehung einer Maskierungsschwelle realisiert.

Diese Maskierungsschwelle wird vom *Psychoakustischen Modell* für jede Komponente individuell durch eine sogenannte *Discrete Fourier Transform* berechnet und gibt die maximal erlaubte Quantisierungsfehlerleistung an, mit der noch codiert werden darf, ohne daß eine Wahrnehmung dieses Fehlers durch das menschliche Gehör befürchtet werden muß. Es ist ein sehr sensibles Verfahren, da bei dem kleinsten Fehler, sprich Abweichung von der Maskierungsschwelle, eine Störung durch das Gehör wahrnehmbar würde.

Nach der Kodierung der einzelnen Spektralkomponenten wird der Frame zusammengesetzt und als codiertes digitales Audiosignal ausgegeben, dessen Datenrate sich zwischen 32kbit/s und 192kbit/s bewegt.



**Bild 1:** Generelles Blockschaltbild des Audio Coders

## 4.5. Die drei Layer

Die oben erwähnten drei Layer des MPEG-Audio-Standard arbeiten alle nach dem beschriebenen Grundprinzip, jedoch gibt es zwischen ihnen natürlich Unterschiede im Bereich der Zerlegung des Eingangssignals und der Kodierung, die im folgenden angeschnitten werden sollen.

### Layer I

Der Layer I Code wurde unter dem Aspekt der leichten Implementierbarkeit entwickelt, woraus sich auch nur eine moderate Datenreduktion ergibt.

In diesem Layer werden Frames zu je 384 PCM-Abtastwerten verarbeitet, was bei einer Abtastfrequenz von 48kHz einem Zeitintervall von 8ms entspricht (385 PCM-Abtastwerte / 48kHz). Dabei wird mit einer Polyphasenfilterbank mit Aliasing-Kompensation die Zeit-Frequenzbereichs-Umsetzung realisiert.

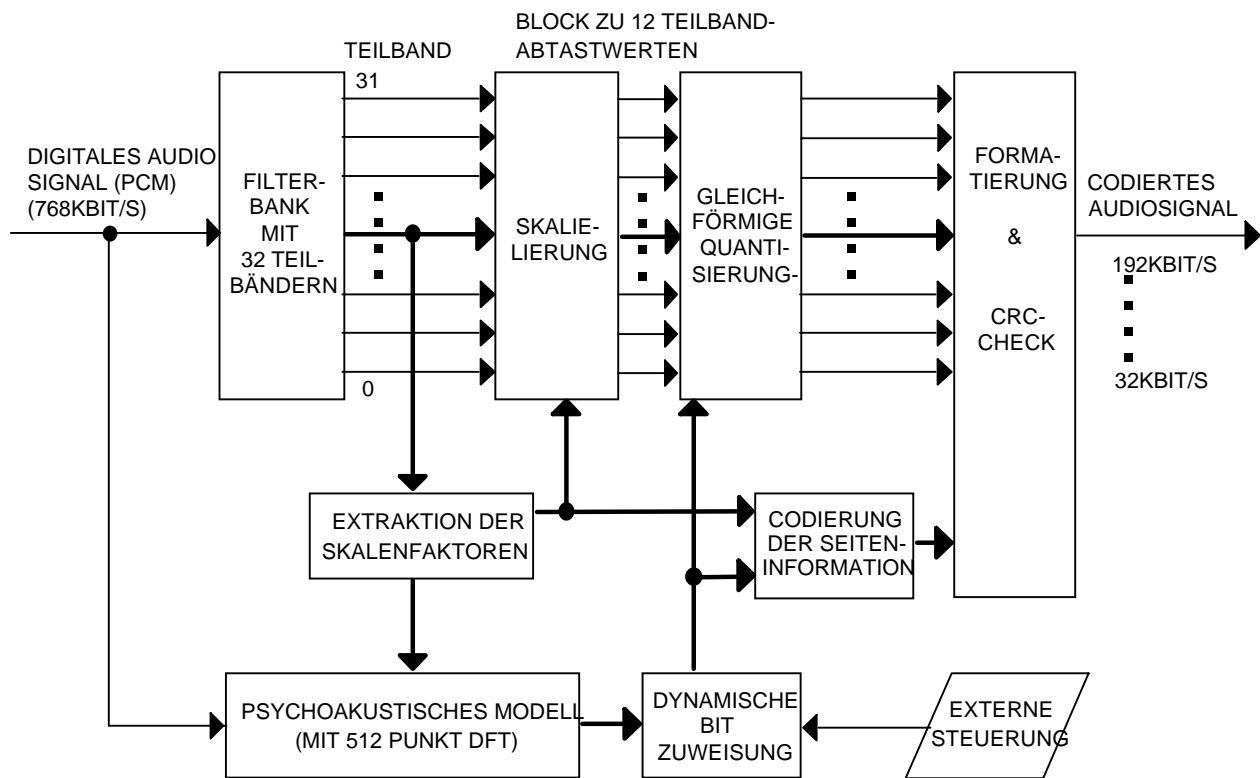
Das Eingangssignal wird in 32 gleichbreite Teilbänder zerlegt, das ergibt eine Bandbreite von 750Hz pro Teilband (48kHz / 32 Teilbänder).

Bei einer Zerlegung in 32 Spektralkomponenten oder auch Teilbänder ergibt das 12 Abtastwerte pro Teilband (384 PCM-Abtastwerte / 32 Spektralkomponenten), die Teilbandabtastwerte. Aus diesen wird der maximale Absolutwert bestimmt und mit 6 Bit kodiert, dieser Wert wird Skalenfaktor genannt. So werden die Teilbandabtastwerte skaliert und entsprechend der vom Psychoakustischen Modell gesteuerten dynamischen Bitzuweisung quantisiert. Die Bitzuweisungsinformation wird mit 4bit codiert.

Bei der Quantisierung wird je nach Anzahl der zugewiesenen Bits einer von 14 verschiedenen gleichförmigen Quantisierern mit  $2^{(n-1)}$  Quantisierungsstufen benutzt, wobei  $n$  zwischen 2 und 15 liegt.

Nach der Quantisierung findet die Formatierung statt, parallel dazu kann optional ein CRC-Check der Seiteninformation durchgeführt werden. Die Seiteninformation ist im Bezug auf Übertragungsfehler der sensibelste Teil der codierten Daten und setzt sich hier aus der Bitzuweisungsinformation aller Teilbänder (d.h. 32 mit 4bit codierte Werte) und den Skalenfaktoren für alle Teilbänder mit von Null verschiedener Bitzuweisung (d.h. bis zu 32 mit je 6bit codierte Werte) zusammen.

Das codierte Audiosignal bewegt sich zwischen 32kbit/s und 192kbit/s, somit wird in diesem Layer mit einer Datenrate von 384kbit/s für ein Stereosignal eine vergleichbare CD-Qualität erreicht.



**Bild II:** Blockschaltbild des MPEG-Audio Layer I Coders

## Layer II

Die Struktur dieses Layer-Coders ist der des Layer I sehr ähnlich, die in Layer I noch nicht benutzten Blöcke sind in der unteren Abbildung grau hinterlegt, wogegen die modifizierten Blöcke lediglich schraffiert sind.

Im Gegensatz zum Layer I werden hier Frames zu je 1152 PCM-Abtastwerten verarbeitet, was bei einer Abtastfrequenz von 48kHz einem Zeitintervall von 24ms entspricht (1152 PCM-Abtastwerte / 48kHz), durch die Verdreifung der Framelänge ergeben sich 36 statt 12 Teilbandabtastwerte. Diese werden in drei Blöcke mit je 12 Abtastwerten gruppiert, für jeden Block wird entsprechend Layer I ein Skalenfaktor bestimmt, somit ergeben sich in jedem Frame drei Skalenfaktoren pro Teilband.

Die Kodierungsstrategie sieht zur Datenratenreduktion eine Variierung der Anzahl der Skalenfaktoren vor. So werden z.B. im Fall von stationären tonalen Signalen alle 36 Teilband-Abtastwerte unter Verwendung eines einzigen Skalenfaktors skaliert, da hier die Unterschiede zwischen aufeinanderfolgenden Skalenfaktoren sehr gering sind.

Es kann aber genausogut sein, daß alle drei Skalenfaktoren verwendet werden müssen, z.B. im Fall von Anschlägen. Durch diese Technik erreicht man bereits eine Reduzierung der Datenraten für die Skalenfaktoren um den Faktor zwei. Diese Auswahl findet in dem grau hinterlegten Block "Skalenfaktor-Auswahl-Information" statt.

Als Unterschied zum Layer I lassen sich noch die Quantisierer aufführen, denn in diesem Layer wird im unteren Frequenzbereich ein Satz mit 15, im mittleren mit 7 und im oberen mit 3 verschiedenen gleichförmigen Quantisierern verwendet. Je nach Satzgröße werden 4, 3 oder 2 Bit zur Kodierung der Bitzuweisungsinformation benötigt.

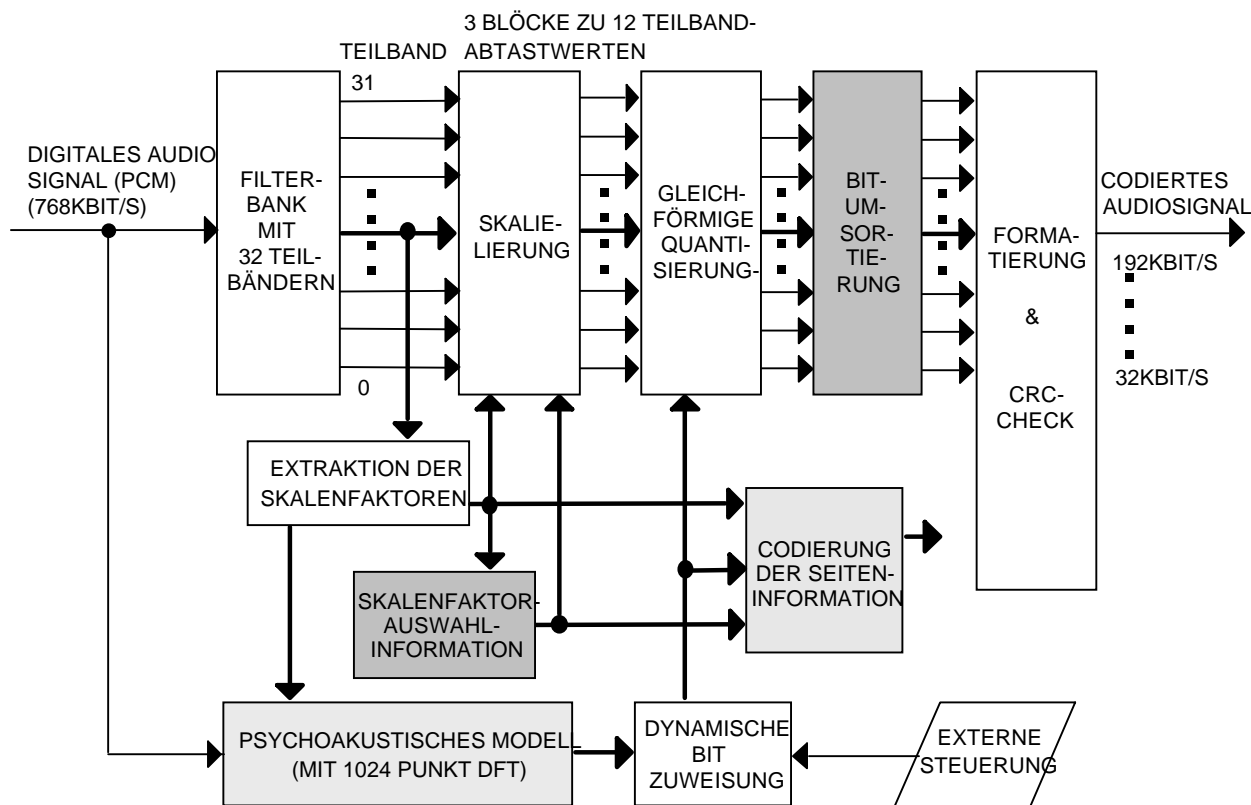
Da für bestimmte Quantisierungsstufen ( {3, 5, 7, 9, 15, 31, 63, ..., 65535} ) keine effiziente Kodierung mit 4, 3 oder 2 Bit möglich ist, werden in diesen Fällen drei aufeinanderfolgende Teilband-Abtastwerte zu einer Gruppe zusammengefaßt, dadurch ist auch eine zusätzliche Bitumsortierung von Nöten.

Die Seiteninformation in diesem Layer setzt sich aus der Bitzuweisungsinformation für alle 32 Teilbänder (d.h. 32 mit je 4, 3 oder 2 Bit codierte Werte), der Skalenauswahl und den Skalenfaktoren für alle Teilbänder mit von Null verschiedener Bitzuweisung (d.h. bis zu 32 mit je 2 Bit codierte Werte bzw. 3x32 mit je 6 Bit codierte Werte). Ansonsten gilt für die Seiteninformation das Gleiche wie in Layer I.



Durch die entsprechende Quantisierung und Kodierung wird in diesem Layer mit einer Datenrate von 256kbit/s eine vergleichbare CD-Qualität erreicht.

Dieser Layer ist im Verhältnis zum Layer I komplexer aber im Bezug auf die Kodierung auch effizienter.



**Bild III:** Blockschaltbild des MPEG-Audio Layer II Coders

### **Layer III**

Der Layer III besitzt die größte Komplexität und zugleich die höchste Effizienz. In der unteren Abbildung sind die neuen Blöcke grau hinterlegt, die modifizierten Blöcke schraffiert.

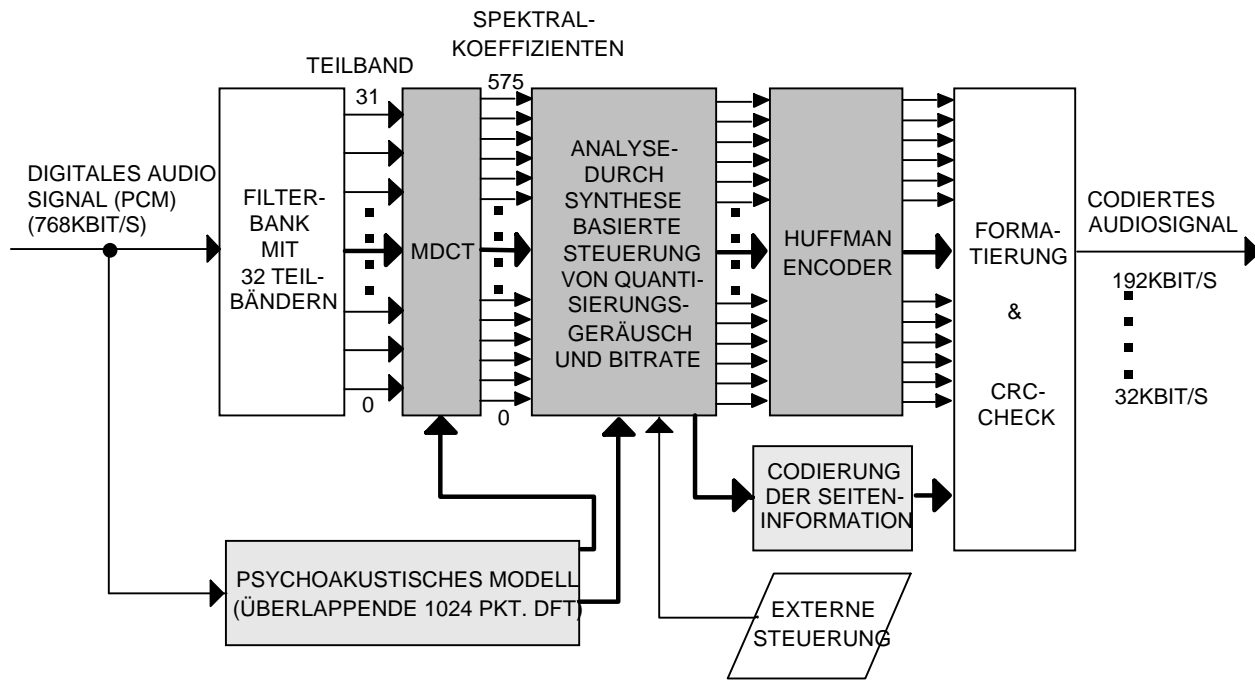
Im Unterschied zu den vorherigen Layern verwendet dieser Layer eine Hybrid-Filterbank, bestehend aus einer Kaskadierung der Polyphasenfilterbank mit einer Modified Discrete Cosine Transform (MDCT), mit adaptiver Fensterumschaltung (s.u.). Diese Filterbank bietet einerseits Kompatibilität mit den anderen Layern, andererseits ermöglicht es eine verbesserte Spektralzerlegung.

Die Anzahl der Abtastwerte und Zerlegung in Spektralkomponenten entspricht Layer II, ergänzend werden im Layer III die Abtastwerte am Ausgang eines jeden der 32 Teilbänder einer MDCT (Modified Discrete Cosine Transform) zugeführt, die dann 18 Spektralkoeffizienten pro Teilbandsignal erzeugt. Hierbei kann je nach Anwendung die spektrale Auflösung der MDCT für ausgewählte oder alle Teilbänder auf 6 Spektralkoeffizienten reduziert werden. Dieses Verfahren nennt sich adaptive Fensterumschaltung und sorgt für eine höhere zeitliche Auflösung bei Wahl einer geringeren Frequenzauflösung (sprich 6 Spektralkoeffizienten pro Teilband statt 18). D.h. die Anzahl der Spektralkoeffizienten kann von 576 (32x18) auf 192 (32x6) reduziert werden.

Die Spektralkoeffizienten werden mit einer vom Psychoakustischen Modell gesteuerten Analyse-durch-Synthese Technik codiert, in der der aktuelle Quantisierungsfehler mit der berechneten Maskierungsschwelle verglichen und die Quantisierung entsprechend angepaßt werden (ungleichförmige Quantisierung).

Auf die quantisierten Spektralkoeffizienten wird zuerst eine Huffman Kodierung und zuletzt eine Lauflängenkodierung vorgenommen. Ansonsten entspricht das Verfahren in diesem Layer dem des Layer II.

Die CD-Qualität wird hierbei mit der selben Datenrate wie in Layer II erreicht, bei niedrigeren Datenraten erreicht man noch akzeptable Ergebnisse, mit den oben bereits erwähnten Einschränkungen.



**Bild IV:** Blockschaltbild des MPEG-Audio Layer III Coders

## 4.6. Resümee

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 Geräuschminderung wie z.B. das Dolby System nötig.

Die in Layer I und Layer II beschriebenen und verwendeten Coder sind bereits fast vollständig als integrierte Schaltungen realisiert und werden wohl bald zur Verfügung stehen, falls sie denn nicht schon zur Verfügung stehen, da die Layer I Kodierung bereits in DCC-Recordern (**D**igital-**C**ompact-**C**assette) und einigen Multimedia-Systemen verwendet werden.

Die Layer II Kodierung wird vom europäischen DAB (**D**igital **A**udio **B**roadcasting) System verwendet werden. 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.

## 5. Sicherheitmechanismen für Multimedia-Daten (Integrität)

### 5.1. Fehlerarten und Fehlerpropagierung

Je nachdem wo in dem MPEG-Video-Stream Fehler bei der Übertragung bzw. Speicherung auftreten, sind auch die zu beobachtenden Fehler des Videobildes unterschiedlich. Da es sich bei MPEG um ein hochkomprimierendes Datenformat (fast redundanzfrei) handelt, sind Fortpflanzungsfehler (Fehlerpropagierung) nicht zu vermeiden und bedürfen einer besonderen Analyse.

#### 5.1.1 Fehler innerhalb von Macro-Blöcken

Sollte ein AC-Koeffizient fehlerhaft sein, hat dies zur Folge, daß der Rest des Blockes auch fehlerhaft wird. Da jedoch AC's untereinander als Differenzen gespeichert werden, ergibt sich ab dem fehlerhaften Wert eine Helligkeitsverschiebung. Ist der Fehler zwischen Original-Wert und Fehler-Wert nicht zu gross ( $< 32$ , d.h. ein Bitkipper in den unteren 5 bits eines bytes), kann diese Verschiebung toleriert werden, dies wären 5/8 aller Bitkipper, sofern sie nur diesen Makroblock (siehe unten) betreffen. Allerdings machen AC's nur ca. 25-40 % der Videodaten aus.

Sollte ein DDC-Koeffizient fehlerhaft sein, ist nicht nur der gesamte Macroblock (im oberen Sinne) betroffen, sondern auch auch alle nachfolgenden Blöcke, deren DC's als Differenz zu diesem DDC gespeichert werden. Dies ist im günstigsten Falle kein weiterer Block. Im MPEG-Stream ist jedoch nur ca. jeder 60ste Macroblocke mit einem MDC gespeichert, d.h. das im ungünstigsten Fall ein fehlerhafter DDC ca 60 weitere Blöcke beeinflusst.

Diese Beeinflußung geschieht in der oben beschriebenen Verschiebung im Helligkeitsbereich.



**Bild V:**

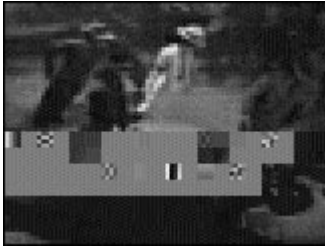
*Fehlerhafter Grauwert-Macroblock mit folgender Helligkeitsverschiebung des restlichen Frames.*

Sollte ein MDC-Koeffizient fehlerhaft sein, beeinflusst er nicht nur den kodierten Macroblock, sondern ALLE nachfolgenden, mit einem DDC-Koeffizienten kodierten Blöcke (ca. 60). Diese Fehlerart kann nur tolerierbar sein, wenn es sich um einen fortlaufenden und nur einmal verwendeten Video-Stream handelt.

#### 5.1.2 Unterschiede von Fehlern in Y- und Cr-/Cb-Planes

Da das menschliche Auge unempfindlicher für Farben als für Grauwerte ist, wurden die Chrominanz-Planes bei der MPEG-Kodierung halbiert. Fehler innerhalb von Macroblöcke (wie oben skizziert), sollten nun größere Auswirkungen auf Farb-Macroblöcke haben als auf Grauwert-Macroblöcke, da die Farb-Macroblöcke größer sind (16x16 Pixel).

Dies trifft in der Praxis um so eher zu, da Fehler in den Chrominanz-Planes eher zu einer totalen Zerstörung der folgende Farb-Werte führen (bis zum nächsten MDC-kodierten Macroblock). Dies kann größere Ausmaße annehmen, und wird vom Auge eher wahrgenommen, als eine gleichgroßflächige Helligkeitverschiebung.



**Bild VI:**

*Fehlerhafter Farbwert-Macroblock mit folgender Zerstörung der nächsten Macroblöcke. Dieser Frame "repariert" sich durch den nächsten MDC-kodierten Macroblock selber.*

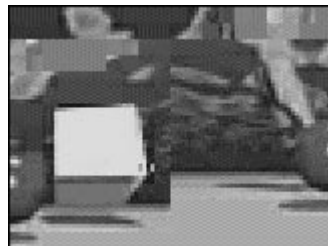
Besonders sind Farbverschiebungen in Computeranimationen als sehr störend anzusehen, denn hier werden Flächen eher durch die Farbgebung als durch unterschiedlich Grauwerte voneinander abgegrenzt. Ausserdem werden zumeist nur begrenzt viele Farben des Farbraums benutzt. Eine Verschiebung in großen gleichfarbigen Flächen ist als sehr störend zu empfinden. Solche Animationen zeichnen sich auch durch eine verstärkte Motion Compensation (siehe unten) im Farbbereich aus, was wiederum zu einer verstärkten Fehlerpropagierung führt.

Fehler in den Chrominanz-Planes von Computer-Animationen sind nicht tolerierbar.

### **5.1.3 Fehler in P- und B-Frames, Motion Vektoren oder Headern**

Bisher hatten wir nur Fehler in Intra-Frames betrachtet. Da P-Frames als Differenz zu I-Frames (B-Frames sogar als mittlere Differenz von I- und P-Frames) gespeichert werden, sind diese extrem anfällig für Fehler in "ihren" I-Frames (bzw. B-Frames auch "ihrer" P-Frames). Deshalb sollte die (relative, doch dazu später) Integrität von I-Frames immer gegeben sein.

Durch Motion-Compensation werden Blöcke nicht als Blöcke sondern als Vektoren zu bereits kodierten (und gleichen) Blöcken gespeichert. Motion-Compensation bedeutet somit auch immer Fehlerpropagierung. Fehler in den Referenzframes ziehen hier räumliche Verschiebungen des Inhalt des mit Motion-Compensation kodierten Frames nach sich.

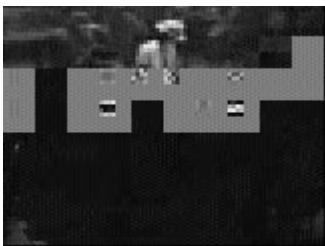


**Bild VII+VIII:**

*Motion - Compensation - Fehler.  
Referenzframe.  
Inhaltsverschiebung.*

Fehler im Header (Vorspann) sind im MPEG-Format nur zu Verzeihen, wenn es sich um einen verbindungsorientierten MPEG-Stream (Videokonferenz etc.) handelt. Ansonsten kann ein Fehler im Header (z.B. der Framegröße) den ganzen Stream zerstören. Der zusätzliche Rechenaufwand zum Schutz des Header holt sich in Grenzen; die durchschnittliche Länge des Header ist kleiner 1 K.

Bisher hatten wir nur Fehler auf bit- bzw. Byte-Ebene (Bitkipper) betrachtet. Diese sind, wie gesehen, in bestimmten Bereichen tolerierbar. Sind jedoch ganze Bytefolgen, Planes oder Frames beschädigt, ist dies nur zu tolerieren, wenn es sich um einen fortlaufenden, einmalig zu "sehenden" MPEG-Stream handelt.



**Bild IX:**

*Kombinierter Fehler. Der ganze Frame ist zerstört.*

Lücken (meistens durch Übertragung entstanden) im MPEG-Video-Stream führen zu noch nicht genau untersuchten Fehlern. Fehlt z.B. ein sogenannter 'sequence ending code' verweigern die bekannten MPEG-Viewer die Interpretation des MPEG-Streams. Leider konnte bis jetzt noch kein lückenhafter MPEG-Stream konstruiert werden, dessen betroffener Frame auch anzeigbar wäre.

Die (jetzige) Schlußfolgerung ist, daß ein Integritäts-Check bezogen auf die Länge des Video-Streams bei Speicherung unabdingbar wird.

## **5.2. Szenarien**

Es können verschiedene Szenarien bei der Behandlung von Fehlern (oder deren Vermeidung) in Bezug auf die Integrität eines MPEG-Stream unterschieden werden.

- **Keine Fehlerkontrolle.** Wie oben gezeigt sind hochkomprimierte Multimedia-Daten anfällig für Fehlerpropagierung. Keine Fehlerkontrolle ist nur ratsam, wenn es sich um einen "einmaligen" MPEG-Stream handelt, z.B bei Bildtelefonie oder Video-Konferenzschaltungen (ohne Aufzeichnung). Gefahrlos sind hier Fehler im Header des MPEG-Streams, da diese hier meist nicht benutzt werden.

Eine Fehlerkontrolle bei der Speicherung durch das jeweilige Betriebssystem ist nur begrenzt gegeben. Sollte die eingesetzte Hardware jedoch auch eine leichte Fehlerkontrolle zulassen (dies sollte nicht "zu" zeitaufwendig sein), ist diese zu bevorzugen.

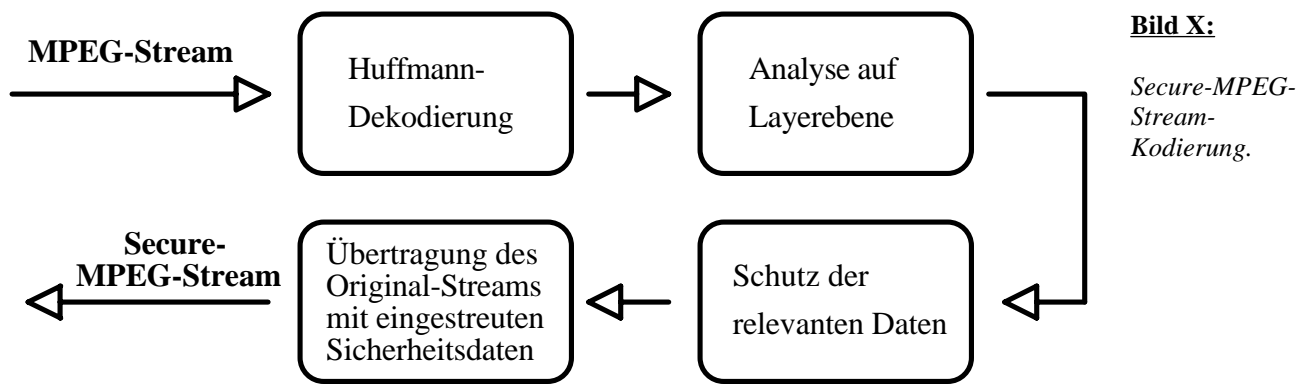
- **Leichte Fehlerkontrolle.** Sollen MPEG-Video-Daten für längere Zeit gespeichert oder über fehleranfällige Leitungen übertragen werden, empfiehlt sich eine teilweise Fehlerkontrolle. Die zu schützenden Teilinformation wurden durch die Fehlerbeschreibung (oben) skizziert und werden in den Lösungsansätzen weiter erläutert. Der Header sollte auch hier schon geschützt werden.
- **Totale Fehlerkontrolle.** Sollen MPEG-Video-Daten archiviert oder in professioneller Art und Weise bearbeitet (geschnitten, editiert, Tricktechnik, etc) muss eine Fehlerkontrolle auf den gesamten MPEG-Stream angewandt werden. Besonders bei der Weiterverarbeitung von Einzelbildern sind auch noch so kleine Fehler nicht zu tolerieren. Zur Fehlerkontrolle können hier jedoch bereits bekannte Verfahren eingesetzt werden (CRC-Check).

### **5.2.1 Leichte Fehlerkontrolle - Zusammenfassung**

Folgende Grundsätze können betreffs der Integrität von MPEG festgehalten werden, wenn ein leichte Fehlerkontrolle eingesetzt werden soll:

- Es ist notwendig den MPEG-Stream zu analysieren. Eine zufällige teilweise Kontrolle des Streams bietet keinen Schutz, denn es werden nicht nur einfache Inhalte zerstört, sondern deren Folgefehler führen mit hoher Wahrscheinlichkeit zur Zerstörung des gesamten Streams. Insbesondere sind die bekannten MPEG-Player nicht gegen verkürzte MPEG-Streams geschützt; größere Systemfehler sind die Folge.
- Um nach der Analyse des MPEG-Streams nicht ein weiteres Mal Huffman dekodieren zu müssen, sollte der Original MPEG-Stream mit eingestreuten Sicherheits- oder Schutz-Informationen übertragen bzw. gespeichert werden.

Dies bedeutet nun, daß ein Sicherheitsalgorithmus folgenden allgemeinen Ablauf haben sollte:



Zusammenfassend sind folgende Daten des MPEG-Streams zu schützen:

- der gesamte Header
- alle DC's (MDC's sowie DDC's) beider Farb-Planes aller I-Frames
- alle MDC's der Grauwert-Plane aller I-Frames, sowie alle anderen MDC's in B- oder P-Frames, egal ob Grauwert-Plane oder Farb-Plane
- alle Motion-Vektoren
- der gesamte Trailer (falls vorhanden)
- Zusätzlich ist eine Check-Summe betreffs der Anzahl der Bytes jedes Frames zu erstellen, um eine Verkürzung des MPEG-Streams überwachen zu können. Verkürzte Frames sind beim dekodieren zu entfernen.

Es ist noch fraglich inwieweit besondere Steuerinformationen zu schützen sind. Sollten z.B fehlende 'sequence ending codes' zum Versagen der Player-Software (sowie weiterer Systemfehler) führen, ist das Vorhandensein dieser Steuerinformationen zu gewährleisten.

### 5.3 Resümee

Mit der beschriebenen Mechanismen wird eine Methode zur Verfügung gestellt, die bei minimalen Rechen- und Kodierungsaufwand maximalen Schutz (in Beziehung auf die Feststellung einer Veränderung der Daten) gewährleistet. Eine Fehlerkorrektur kann in Hinsicht auf das Volumen nicht zur Verfügung gestellt werden.

Verschieden Stufen des Integritätsschutzes sind denkbar. Schutz vor Veränderung der relevanten Daten in Bezug auf Manipulation, Schutz vor Unbrauchbarkeit (Versagen der Software !) oder Schutz der Abspielqualität können durch Auswahl der oben beschriebenen schutzrelevanten Daten lassen sich implementieren.

## 6. MPEG-II

MPEG-II 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-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].

Weitere 'Features' von MPEG-II werden sein:

- angepaßte Bandbreite < 3 MBit/s wegen Bedarf von
  - Satellite Broadcast
  - Electronic Cinema
  - Digital Home Television, etc.
- Scalability von Audio und Video
- MPEG-1 and H.261 abwärtskompatibel

- Multi-Channel-Mode, multi-lingual (5 Kanäle mit voller Bandbreite, 1 Kanal für niedrige Frequenzen)
- intelligente Schachtelung von Video und Audio zur Vermeidung von Synchronisationsverlusten (auto-sync.)
- Encryption, Security
- MPEG-4 ? - Bitrates up to tens of kbit/s - fraktale Kompression

## 7. DVI

"Digital Video Interactive (DVI) ist ein früher Versuch der Firma Intel (in Zusammenarbeit mit der Firma IBM) einen Standard in der Kompression von kontinuierlichen Medien durchzusetzen. Intel kaufte 1988 die Rechte an DVI vom David Sarnoff Research Center. Ein Jahr später stellte man das erste Produkt vor. Es war ein 386er PC mit sieben Steckkarten." [BAD93].

Zwei Kompressionsformate werden von DVI unterstützt: RTV (**R**eal **T**ime **V**ideo) wurde vom Benutzer mittels seiner lokalen Hardware gebraucht; PLV (**P**roduction **L**evel **V**ideo) erreichte Kompressionsraten von 160:1, konnte jedoch nur mittels von Intel autorisierten Betrieben benutzt werden.

Bis jetzt ist DVI weder auf anderen Rechnerwelten noch ohne extreme Hardwareunterstützung denkbar. Außerdem scheint sich Intel mit dem neuen Verfahren "Indeo" mittlerweile gegen das eigene DVI gestellt zu haben. Trotzdem hat die Firma Fast nun ein DVI-Board produziert (eigentlich die sog. Screen Machine von Fast mit einer neuen Software unter Windows 3.1 und einem DVI-Kompressionschip).

## 8. AVI

AVI (Audio-Video-Interlaced) scheint Microsofts Antwort auf Apple's Quicktime zu sein, kann dies jedoch nicht leisten. AVI stellt unter der Benutzeroberfläche MS-Windows 3.1 mehrere Utilities und Driver zu Verfügung, die es erlauben, kleinste Video-Sequenzen (20 Sekunden 160x120 Pixel bei 8 bit Farbe nehmen immerhin 2 MB Daten ein) per Hardware zu digitalisieren, zu editieren, komplett in die Windows-Umgebung zu integrieren, mit von Hardware gesammeltem Audio zu synchronisieren und natürlich auch alles zusammen abzuspielen.



**Bild XI:**

*AVI-Editor von Microsoft's Video for Windows.*

Insbesondere wird hier ein Kompressionsverfahren der Firma Intel (Intel Indeo) eingesetzt, das ein 2-dimensionales 'scaling' des Video-Filmes erlaubt, jedoch ist mit diesem 'scaling' nicht 'scalability' gemeint.

Das große Manko von AVI ist die noch ungenügende Kompression, da AVI bis jetzt auch nur auf einer Hardware-Plattform implementiert ist (und AVI kein internationaler Standard ist) wird sich AVI zwar einer großen Verbreitung in der Intel/Microsoft-Welt erfreuen, jedoch wohl kaum in den Bereich der Telekommunikation vordringen können.

## 9. Beispiele

### 9.1 Xing

Die MPEG-Player der Firma Xing Technologies waren die ersten PD-Programme die MPEG zu jedermann brachten. Die Software zeichnet sich durch ein sehr gute (Windows-typische) 'Usability' aus. Xing hat auch die ersten Versuche zur Audio-Integration unternommen (Synchronisation über die Zeit mit Windows-WAV-Files).



**Bild XII:**

*Xing's MPEG-Player-Interface.*

Allerdings setzt Xing immer noch nicht auf dem MPEG-I-Standard auf. Es werden nur mit ausschließlich I-Frames kodierte Videos und keine variablen Framegrößen unterstützt.

### 9.2 Berkeley und Stanford

Die Gruppe um Lawrence A. Rowe, Ketan Patel und Brian Smith an der Computer Science Division-EECS, Univ. of Calif. at Berkeley stellte den ersten PD-Player auf Unix-basis vor. Dieser Code, der den kompletten MPEG-I-Standard implementiert war so einfach und 'sauber' programmiert, daß er bis heute auf dutzende verschiedene Plattformen und System portiert wurde.

Ebenso das erste PD-Codec von Andy C. Hung, University of Stanford. Es erlaubt das parametrisierte Erzeugen von MPEG-Videos. Beide Tools kann man mittlerweile als den MPEG-Standard-Code bezeichnen.

### 9.3 NVR

NVR (North Valey Research, Inc.) haben mit ihrem "Digital Media Development Kit, Version 1.0" das erste integrierte, komplett interaktiv zu steuernde MPEG-System vorgestellt. Es ist für Sun-Computer unter der Benutzeroberfläche Open-Look (bez. X11) mit Hilfe von Motif erstellt worden und unterstützt den kompletten MPEG-Standard incl. 'frame-grabbing', Kodierung, Player etc.. Als kommerzielle Lösung sind jedoch die preislichen (bzw. Hardware-) Voraussetzungen (Sun IPC mit Parallax-Board plus Software = DM 20.000.-) noch als zu hoch anzusehen.



## 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-II), 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
- [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
- [ISO90.4] ISO/IEC International Standard : Standard Music Description Language (SMDL), Hypermedia/Time-base Subset (HyTime), International Organization for Standardization, Geneva, 1990
- [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
- [MUS93] Musman, Hans-Georg; Werner, Oliver; Fuchs, Hendrik: Kompressionsalgorithmen für interaktive Multimedia-Systeme, IT+TI 2/93, Hannover 1993
- [NVR93] North Valey Reserach: Digital Media Development Kit, Reference Manual, Version 1.0, Jan. 1993
- [PAT92] Patel, Ketan; Smith, Brian C.; Rowe, Lawrence A.: Performance of a Software MPEG Video Decoder, Berkeley, 1992

## Anhang folgend: Abbildungen