

Adaption von Videos

Seminar-Arbeit

von

Ralf Diem

aus

Neustadt an der Weinstraße

vorgelegt am

Lehrstuhl für Praktische Informatik IV

Prof. Dr. W. Effelsberg

Fakultät für Mathematik und Informatik

Universität Mannheim

November 2005

Betreuer: Dipl.-Wirtsch.-Inf. Stephan Kopf

Inhaltsverzeichnis:

1	Einleitung	1
2	Grundlagen der Videokompression	2
2.1	Frames	2
2.2	Farbmodell	2
2.3	Farb-Subsampling	2
2.4	Intra-(frame)-Kodierung.....	2
2.4.1	Diskrete Kosinustransformation (DCT).....	2
2.4.2	Quantisierung	3
2.4.3	Entropie-Kodierung.....	4
2.5	Inter-(frame)-Kodierung.....	5
2.5.1	Motion Compensation (MC).....	5
2.6	Variable und konstante Bitrate (VBR und CBR).....	5
2.7	GOP (Group of Pictures).....	6
3	Videostandards	7
3.1	MPEG-1	8
3.2	MPEG-2/H.262	9
3.3	MPEG-4	9
4	Adaption	10
4.1	Transkodierung.....	10
4.1.1	Bitratenanpassung bei fester Auflösung.....	11
4.1.2	Änderung der räumlichen Auflösung.....	11
4.1.3	Änderung der temporären Auflösung.....	12
4.1.4	Erweiterung der Fehlerkorrektur.....	12
4.1.5	Logo- oder Wasserzeicheneinbringung.....	12
4.1.6	Heterogene Transkodierung	13
4.2	Skalierbare Kodierung.....	13
4.2.1	Temporäre Skalierung.....	13
4.2.2	Räumliche Skalierung	13
4.2.3	Datenpartitionierung.....	14
4.2.4	SNR-Skalierung (signal to noise ratio)	14
4.2.5	Fein-granulare-Skalierung (fine granularity scalability, FGS)	14
4.2.6	Objekt-basierte-Skalierbarkeit	15
5	Inhaltsbeschreibung, -suche und -bezug	16
5.1	MPEG-7	16
5.2	MPEG-21	18
5.2.1	Digital-Item-Deklaration.....	18
5.2.2	Digital-Item-Identifikation.....	18
5.2.3	Schutz und Management von geistigem Eigentum	18
5.2.4	Digital-Item-Adaption.....	18
6	Zusammenfassung und Ausblick	19
7	Literaturverzeichnis	II
8	Abbildungsverzeichnis	III
9	Abkürzungsverzeichnis	IV

1 Einleitung

In Zeiten des Internets und der steigenden Vernetzung über die verschiedensten Kanäle ist es immer mehr Menschen möglich ständig auf multimediale Daten zuzugreifen. Das häufig gebrauchte Schlagwort hierfür ist Universal Multimedia Access (UMA). Es obliegt den Datenanbietern die Inhalte so aufzubereiten, dass sie auf einer Vielzahl von Endgeräten in einer akzeptablen Qualität darstellbar sind.

Die vorliegende Seminararbeit beschäftigt sich mit Methoden zur Adaption von Videodaten. Da nichtkomprimiertes Videomaterial sehr hohen Speicherplatzbedarf hat wird es in der Regel komprimiert.

In 2. Kapitel werden einige Grundlagen zur Videokompression dargestellt, damit auch nicht sachkundige Leser in die Materie eingeführt werden, bzw. um evtl. vorhandenes Wissen aufzufrischen.

Kapitel 3 stellt einige wichtige Videokompressionsstandards vor, wobei der wohl am meisten gebräuchliche MPEG-2 ist (z.B. bei DVD, HDTV, DVB-T).

Kapitel 4 stellt dann zwei konkrete Strategien zur Adaption von Videodaten vor.

Im abschließenden Kapitel 5 werden die relativ neuen und sich teilweise noch im Entwicklungsstadium befindlichen Standards MPEG-7 und MPEG-21 beschrieben.

2 Grundlagen der Videokompression

Die im Folgenden vorgestellten Verfahren sind die Grundlage aller blockbasierten Videokompressionsstandards. Dazu gehören z.B. MPEG-1, MPEG-2/H262 und H264¹.

2.1 Frames

Ein Video ist eine Sequenz von schnell aufeinanderfolgenden Einzelbildern, die Frames genannt werden. Typische Frame-Raten sind z.B. 25 Frames pro Sekunde (f/s) im europäischen PAL-System und 29,97 f/s beim nordamerikanischen NTSC.

2.2 Farbmodell

Ausgangslage für die im Folgenden vorgestellten Verfahren ist, dass die Eingabedaten als Bilder im YCbCr Format vorliegen. Bei diesem Farbmodell wird die Farbe eines Bildpunktes (Pixels) nicht durch die Addition der Grundfarben wie im RGB-Farbmodell dargestellt, sondern durch die Grundhelligkeit (Y) und zwei Komponenten Cb und Cr, wobei Cb die Abweichung von Grau in Richtung Blau bzw. Gelb repräsentiert und Cr für die Abweichung in Richtung Rot bzw. Türkis steht.

2.3 Farb-Subsampling

Jedes Bild wird in Makroblöcke (MB's) aufgeteilt, von denen jeder durch eine 16x16 Matrix für den Helligkeitwert (Y) und zwei 8x8 Matrizen für die beiden Chrominanzanteile (Cr und Cb) repräsentiert wird. Das bedeutet, dass die Auflösung für die Chrominanz halb so hoch ist wie für die Helligkeit. Dieses Verfahren berücksichtigt, dass das menschliche Auge Änderungen in der Helligkeit stärker wahrnimmt als Änderungen der Farbe und ist der erste verlustbehaftete Schritt der Kompression. Der 16x16 Block wird noch mal in vier 8x8 Blöcke unterteilt, da alle weiteren Algorithmen auf 8x8 Blöcken basieren.

2.4 Intra-(frame)-Kodierung

Bei der Intra-frame-Kodierung wird ein Frame als ein Einzelbild angesehen, bei dem nur die räumlich redundanten Informationen komprimiert werden.

2.4.1 Diskrete Kosinustransformation (DCT)

Die Diskrete Kosinustransformation ist der zentrale Schritt aller blockbasierten Kompressionsverfahren. Sie „transformiert ein zeitdiskretes Signal vom Orts- in den Frequenzbereich“² und bringt so die Bilddaten in eine gut zu komprimierende Form. Jeder 8x8 Block wird durch folgende Formel³ umgerechnet (s_{xy} ist der Wert an der Stelle (x,y) in der 8x8 Matrix):

¹ Vgl. Ahmad et al: Video Transcoding, S.794

² Wikipedia.org: Die diskrete Kosinustransformation, Stand: September 2005

³ Effelsberg, Steinmetz: Video Compression Techniques, S.30

$$F_{vu} = \frac{1}{4} C_u C_v \sum_{x=0}^7 \sum_{y=0}^7 s_{xy} \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16}$$

wobei

$$C_u, C_v = \begin{cases} \frac{1}{\sqrt{2}} & \text{für } u, v=0 \\ 1 & \text{andernfalls} \end{cases}$$

Das Ergebnis sind 64 Koeffizienten (F_{vu}) im Frequenzbereich (8x8). Der Koeffizient in der oberen linken Ecke wird DC-Koeffizient genannt. Er repräsentiert den Durchschnittswert des zugehörigen 8x8 Blocks. Die restlichen 63 Koeffizienten nennt man AC-Koeffizienten. Die folgende Abbildung zeigt die graphische Darstellung der zugehörigen Basis-Frequenzen zu den berechneten Koeffizienten.

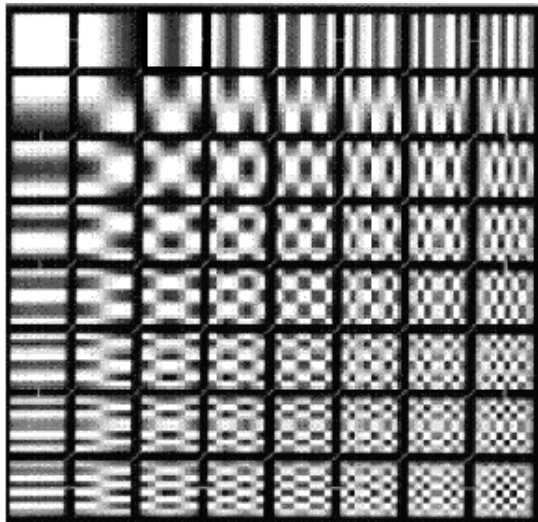


Abbildung 1: Graphische Darstellung der 64 Basis-Frequenzen der 2-Dimensionalen DCT⁴

2.4.2 Quantisierung

Da das menschliche Auge Änderungen in den hohen Frequenzen weniger stark wahrnimmt als in den niedrigen, werden die hochfrequenten AC-Koeffizienten mit einer niedrigeren Genauigkeit gespeichert als die niedrigfrequenten. Dazu wird jeder AC-Koeffizient durch eine positive Ganzzahl geteilt und das Ergebnis auf die nächste Ganzzahl gerundet. Die Quotienten sind in einer exemplarischen Quantisierungsmatrix in Abbildung 2 dargestellt. Man beachte, dass die Quotienten in Richtung der höheren Frequenzen größer werden. Dies bewirkt, dass man bei den hohen Frequenzen viele Null-Einträge hat. Die Quantisierung ist wiederum verlustbehaftet.

⁴ Effelsberg, Steinmetz: A Graduate Course on Multimedia Technology, S.2.2-32

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Abbildung 2: Quantisierungsmatrix⁵

2.4.3 Entropie-Kodierung

Im Anschluss an die Quantisierung werden die Daten noch zweimal komprimiert. Zunächst durch eine Lauflängenkodierung⁶, bei der die AC-Koeffizienten in einer vorgegebenen Reihenfolge abgearbeitet werden. Abbildung 3a zeigt die Zick-Zack-Reihenfolge, die bei MPEG-1 Verwendung findet. MPEG-2 nutzt ebenfalls diese Reihenfolge, allerdings ist in diesem Standard noch eine Alternativ-Reihenfolge definiert, die in Abbildung 3b dargestellt ist.

Der zweite Komprimierungsschritt ist eine Variable-Length-Kodierung die bei MPEG-1/2 auf einer Huffman-Kodierung⁷ basiert. Es wird eine standardisierte (De-)Kodiertabelle verwendet, die aus Versuchen gewonnen wurde.

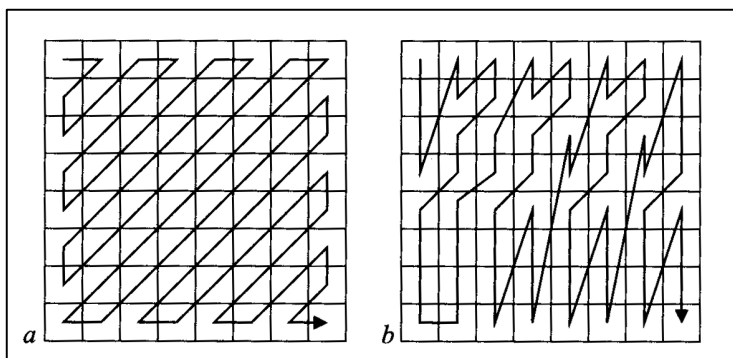


Abbildung 3: Mögliche Scanreihenfolgen der AC-Koeffizienten bei der Entropiekodierung⁸

Die wichtigeren DC-Koeffizienten (die ja die Basisfarbe des Blocks darstellen) werden durch eine Differenzial-Kodierung komprimiert.⁹ D.h. der aktuelle DC-Koeffizient wird als die Differenz zum DC-Koeffizienten des vorhergehenden Blocks kodiert.

Die Entropie-Kodierung ist wiederum verlustfrei.

⁵ Ghanbari: Standard Codecs, S.162

⁶ nachzuschlagen in: Effelsberg, Steinmetz: Video Compression Techniques, S.9

⁷ nachzuschlagen in: Effelsberg, Steinmetz: Video Compression Techniques, S.14

⁸ Ghanbari: Standard Codecs, S.188

⁹ vgl. Effelsberg, Steinmetz: A Graduate Course on Multimedia Technology, S.2.2-37

2.5 Inter-(frame)-Kodierung

Bei einem Video gibt es aber auch die Möglichkeit die zeitlich redundanten Informationen zu reduzieren. Diese Möglichkeit wird Inter-(frame)-Kodierung oder auch Non-Intra-(frame)-Kodierung genannt. Sie beruht darauf, dass sich ein Frame von seinem Vorgängerframe nur in relativ wenigen Pixeln unterscheidet, die dadurch entstehen, dass sich z.B. Objekte bewegen, der Bildausschnitt durch Zoomen oder Schwenken der Kamera verändert wird etc. Das heißt, anstatt den ganzen Frame zu kodieren wird nur der Unterschied zum Vorgängerframe kodiert.

2.5.1 Motion Compensation (MC)¹⁰

Die Differenz zwischen den Frames wird folgendermaßen bestimmt:

Zunächst muss die Bewegung der Objekte bestimmt werden (Motion Estimation, ME). Dies geschieht meist durch einen Block-Matching-Algorithmus (BMA), bei dem ein Frame in Blöcke zu $M \times N$ oder meist N^2 Pixeln unterteilt wird. Jeder Block wird daraufhin im nachfolgenden oder vorhergehenden Frame in einem begrenzten Gebiet um die Ursprungsposition gesucht. Wird er gefunden wird nur die Verschiebung des Blockes durch einen Bewegungsvektor (Motion Vector, MV) kodiert. Dies ist in Abbildung 4 illustriert.

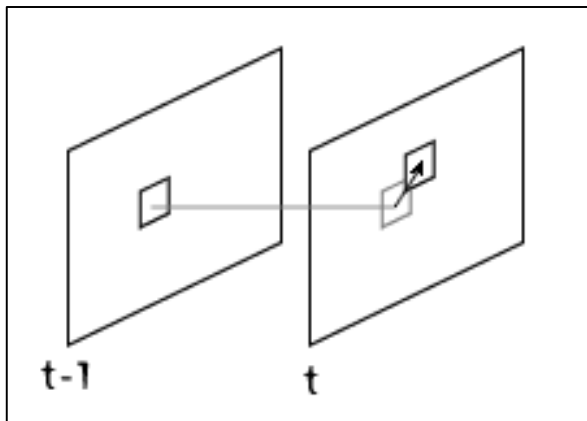


Abbildung 4: Bestimmung eines Bewegungsvektors

Anschließend wird das Bild, das sich aus den verschobenen Blöcken zusammensetzt, mit dem Originalbild verglichen. Die Abweichung der beiden Bilder, die dadurch entsteht, dass nicht alle Blöcke zugeordnet werden können, wird intra-frame-kodiert. Die Motion Compensation ist der rechenintensivste Schritt bei der Kodierung und benötigt bis zu 70% der Gesamtzeit.

2.6 Variable und konstante Bitrate (VBR und CBR)

Außerdem gibt es die Möglichkeit, Videosequenzen mit sich schnell änderndem Inhalt weniger stark zu komprimieren als Sequenzen, die nur wenig inhaltliche Bewegung haben. Dies resultiert dann in einer variablen Bitrate bei konstanter Bildqualität. Bei konstanter Bitrate hingegen muss diese so hoch gewählt werden, dass auch schnelle Passagen qualitativ gut dargestellt werden, was im Vergleich zur VBR zu einem höheren Speicherplatzbedarf führt.

¹⁰ vgl. Ghabari: Standard Codecs, S.35ff

2.7 GOP (Group of Pictures)

Bei MPEG-1/2 werden die Frames in die folgenden 4 Gruppen unterschieden:¹¹

1. I-Pictures (intra coded pictures) werden unabhängig von anderen Bildern intra-frame-kodiert.
2. P-Pictures (predictive coded pictures) werden inter-frame-kodiert in Bezug auf das letzte vorhergehende I- oder P-Picture.
3. B-Pictures (bidirectionally predictive coded pictures) werden ebenfalls inter-frame-kodiert, allerdings können sie Bezug auf das direkt vorhergehende oder das direkt nachfolgende Bild haben.
4. D-Pictures (DC coded picture) enthalten nur die DC-Koeffizienten der einzelnen Blöcke und sind für Vorschaubilder o.ä. gedacht. Diese Gruppe wird in der Praxis nicht oder kaum angewendet.

Eine Abfolge von Bildern von einem I-Picture zum nächsten heißt Group of Pictures. Abbildung 5 zeigt eine beispielhafte GOP in Anzeigereihenfolge.

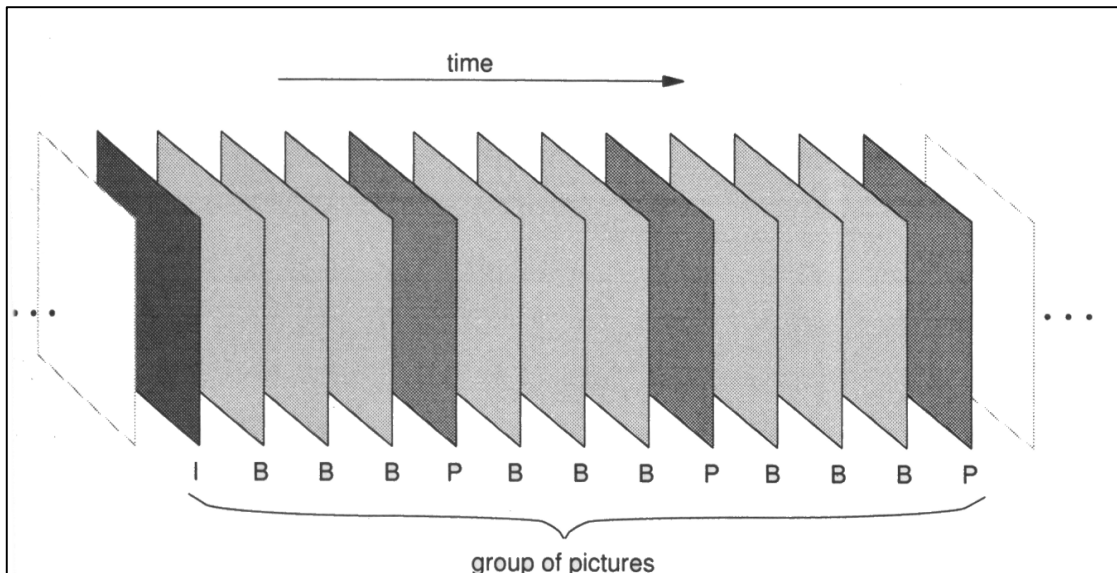


Abbildung 5: Eine typische GOP in Anzeigereihenfolge¹²

¹¹ Effelsberg, Steinmetz: Video Compression Techniques, S.40f

¹² Mitchel et al.: MPEG video compression standard, S. 23

Die Dekodierreihenfolge muss natürlich wegen der B-Pictures etwas anders sein. Dies ist in Abbildung 6 dargestellt.

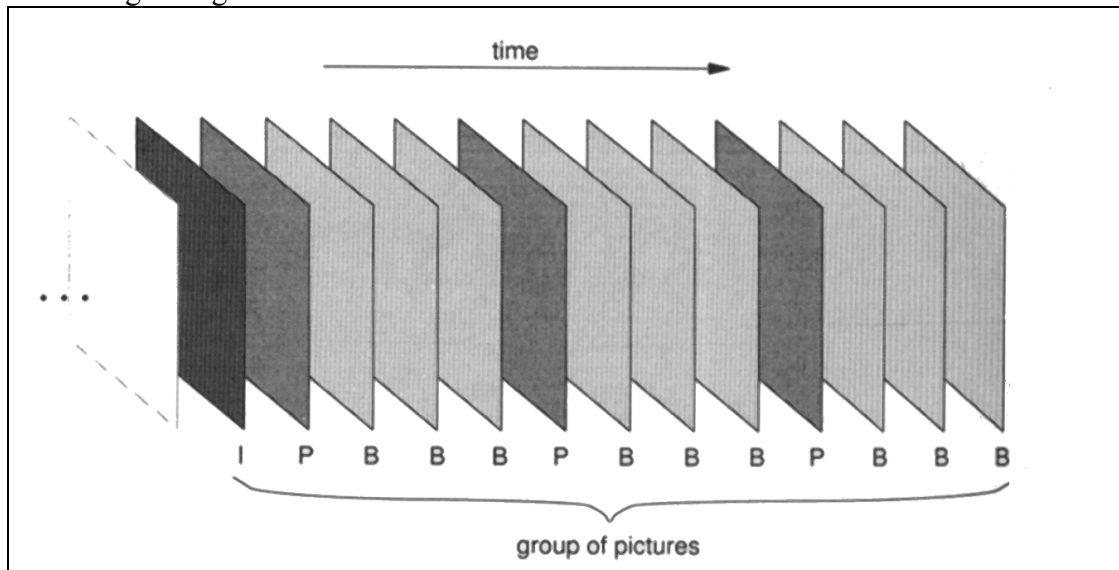


Abbildung 6: Eine typische GOP in Dekodierreihenfolge¹³

3 Videostandards

Es gibt weltweit zwei allgemein anerkannte Standardisierungsorganisationen. Das ist zum einen die International Telecommunication Union (ITU-T) und zum anderen die International Standardisation Organisation (ISO), zu der die Moving Pictures Experts Group (MPEG) gehört. Abbildung 7 zeigt die zeitliche Entwicklung der einzelnen Video/Bildkompressionsstandards beider Gruppen. In der Regel wird nur die Syntax der Daten und der Dekodierervorgang genormt, die Kodierung und woher man die Daten für die Kodierung erhält, wird offen gelassen.

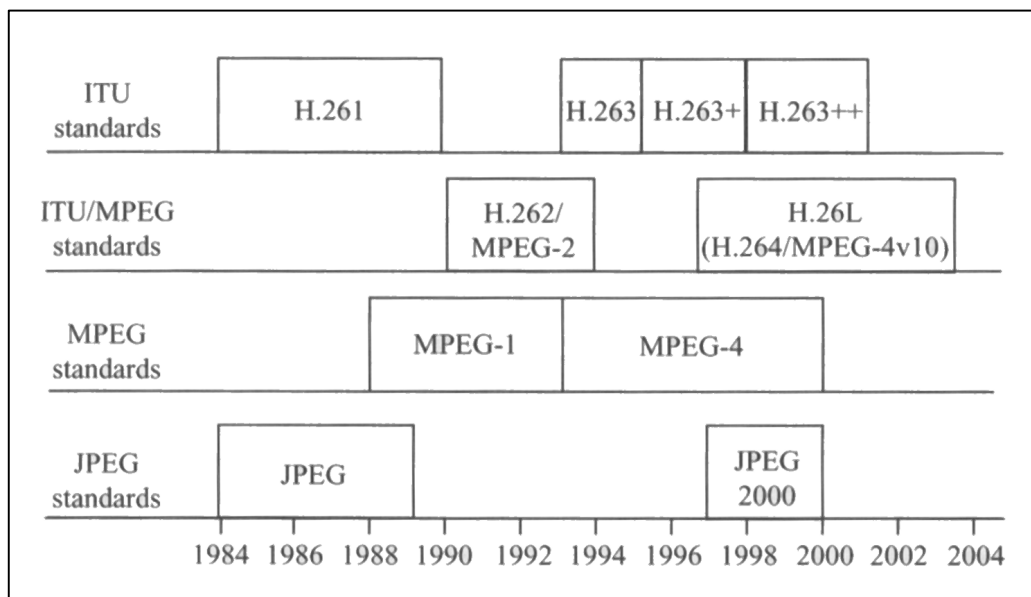


Abbildung 7: Evolution der Video/Bildkompressionsverfahren¹⁴

¹³ Mitchel et al.: MPEG video compression standard, S.24

Dass MPEG-2/H.262 eine so weite Verbreitung gefunden hat (z.B. DVD, HDTV, DVB-T) ist sicherlich damit begründet, dass der Standard von beiden Organisationen in Gemeinschaftsarbeit entwickelt wurde.

3.1 MPEG-1

Die Ziele, die die Moving Pictures Experts Group bei der Entwicklung hatte und auch umsetzen konnte, waren:¹⁵

1. Eine Datenrate von 1,5 Mbit/s zu erreichen, was der Datenrate einer T1 Verbindung in den USA und der Streamingrate der Nutzdaten einer CD-Rom entspricht.
2. Die vorhandenen Standards mit 525 Zeilen (entspricht 480 Zeilen digital) und 625 Zeilen (entspricht 576 Zeilen digital) zu unterstützen. Daher wurde beschlossen, dass MPEG-1 progressiv arbeiten soll und interlaced Material vorher konvertiert werden muss.
3. Die Bildqualität sollte mindestens der einer analogen Aufzeichnung mit einem Videokassettenrekorder entsprechen.
4. Außerdem sollten beim Abspielen Funktionen wie schnelles Vor- und Zurückspulen, Standbild, Zeitlupe und zufälliger Zugriff vorhanden sein.

Abbildung 8 zeigt einen einfachen Kodierer.

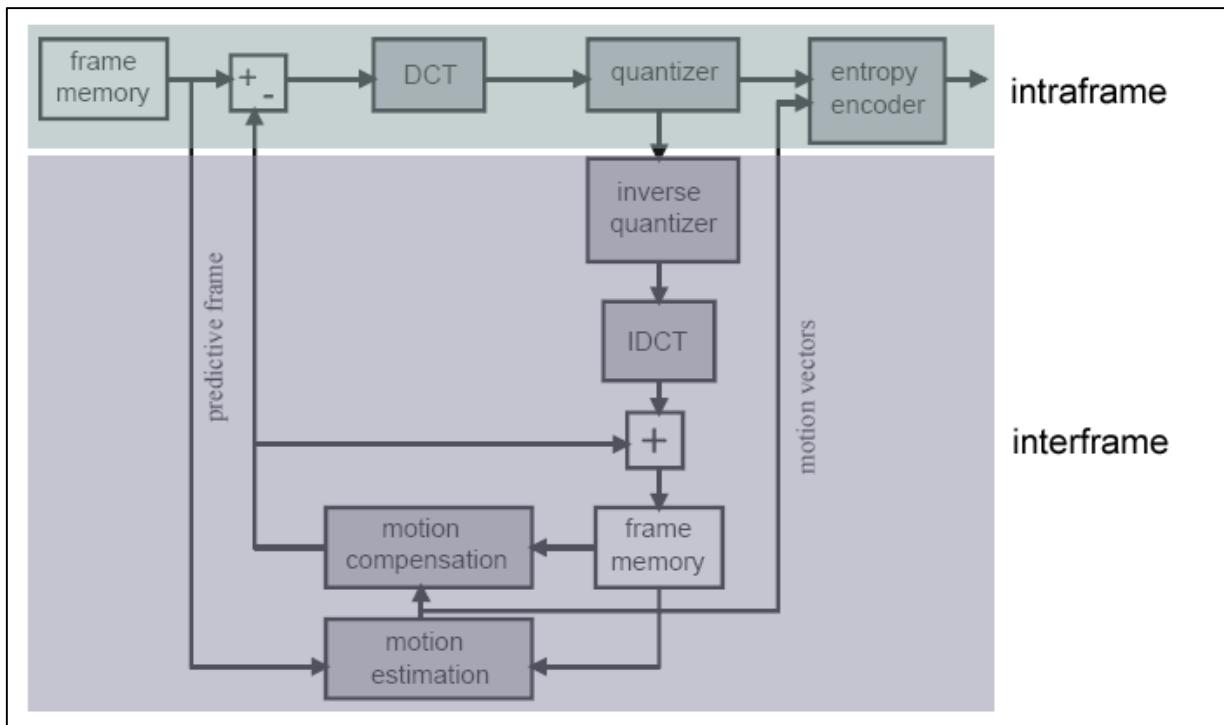


Abbildung 8: Einfacher MPEG-Kodierer¹⁶

¹⁴ Ghanbari: Standard Codecs, S.5

¹⁵ Effelsberg, Steinmetz: A Graduate Course on Multimedia Technology, S.2.3-1f

¹⁶ Effelsberg, Steinmetz: A Graduate Course on Multimedia Technology, S.2.3-5

3.2 MPEG-2/H.262

MPEG-2 sollte flexibler als MPEG-1 werden, um in einer Vielzahl von Anwendungen eingesetzt werden zu können. Dazu mussten z.B. Mechanismen zur Fehlerkorrektur und Skalierbarkeit¹⁷ eingeführt werden. Daher wurden verschiedene Profile definiert, die je nach Anwendung zum Einsatz kommen können (siehe Abbildung 9).

	Simple profile no B frames not scalable	Main profile B frames not scalable	SNR scalable profile B frames SNR scaling	Spatially scalable pro- file B frames spatial scal- ing	High profile B frames spatial or SNR scaling
High level 1920x1152x60		<=80 Mb/s			<=100 Mb/s
High-1440 level 1440x1152x60		<=60 Mb/s		<=60 Mb/s	<=80 Mb/s
Main level 720x576x30	<=15 Mb/s	<=15 Mb/s	<=15 Mb/s		<=20 Mb/s
Low level 352x288x30		<=4 Mb/s	<=4 Mb/s		

Abbildung 9: Profile in MPEG-2 aufgetragen in Bezug zur Auflösung¹⁸

MPEG-3 war als Standard für sehr hohe Datenraten geplant, allerdings wurde er letztendlich in MPEG-2 integriert.

3.3 MPEG-4

MPEG-4 verfolgt einen anderen Ansatz als seine Vorgänger. Die Grundidee ist eine objektbasierte Kodierung des Videos, bei der eine Szene in voneinander unabhängige Layer von Objekten eingeteilt wird (Video Object Planes, VOP's). Diese Objekte (z.B. Hintergrund, Object of Interest) können dann auch unabhängig voneinander mit unterschiedlichen Methoden kodiert werden. Außerdem sind die verwendeten Kodierungstechniken für mobile und paketbasierte Netzwerke optimiert¹⁹. Der Standard beschreibt wie die einzelnen Objekte kodiert werden, allerdings lässt er die Objekterkennung offen.

¹⁷ vgl. S.13ff

¹⁸ Effelsberg, Steinmetz: A Graduate Course on Multimedia Technology, S.2.3-11

¹⁹ vgl. Effelsberg, Steinmetz: A Graduate Course on Multimedia Technology, S.2.3-13

4 Adaption

Audiovisuelle Daten sind inzwischen durch die Anbindung mobiler Endgeräte ans World Wide Web (WWW) von überall her abrufbar. Dazu ist es aber notwendig, die Daten an die Gegebenheiten wie z.B. Beschaffenheit des Netzwerkes (LAN, WLAN, UMTS etc.), Kapazitäten des Endgerätes (Smartphone, Handy, PDA, Notebook etc.) und die Nutzerwünsche anzupassen. Hinsichtlich der Netzwerkbeschaffenheit müssen unter Umständen die Bitrate an die Bandbreite angepasst werden sowie Mechanismen zur Fehlerkorrektur bei störungsanfälligen Netzwerken eingefügt werden. In Bezug auf die Endgeräte muss evtl. die Framerate reduziert werden, weil entweder nicht genug Rechenleistung zur Dekodierung bei hoher Framerate zur Verfügung steht, oder aber die Auflösung des Endgerätes zu gering ist und deshalb angeglichen werden muss. Außerdem sollte der Nutzer selbst auch noch die Möglichkeit haben, die Qualität des gelieferten Materials zu beeinflussen, denn er muss ja letztendlich dafür bezahlen.

Es ist wünschenswert, dass diese Adaption möglichst transparent für den Endnutzer vollzogen wird. Frameworks für Universal Multimedia Access beschäftigen sich mit genau dieser Problematik. Dabei gibt es zwei Strategien, die nachfolgend vorgestellt werden: Zum einen die Transkodierung und zum anderen die skalierbare Kodierung.

4.1 Transkodierung

Transkodierung wird benötigt, um bereits komprimiertes Videomaterial von einem Standard in einen anderen zu konvertieren, oder, um innerhalb eines Standards einzelne Parameter (Bitrate, Auflösung etc.) anzupassen. Bei diesem Vorgang ist es auch möglich, Logos oder Wasserzeichen einzufügen.

Man kann so von dem Original mehrere Kopien in unterschiedlichen Qualitätsstufen oder Standards erzeugen und bei einer Anfrage die passendste Version wählen. Es ist aber auch möglich, bei einer Anfrage in Echtzeit zu transkodieren und das erhaltene Material zu streamen. Abbildung 10 gibt eine Übersicht über die Transkodieroperationen und ihre Klassifikation.

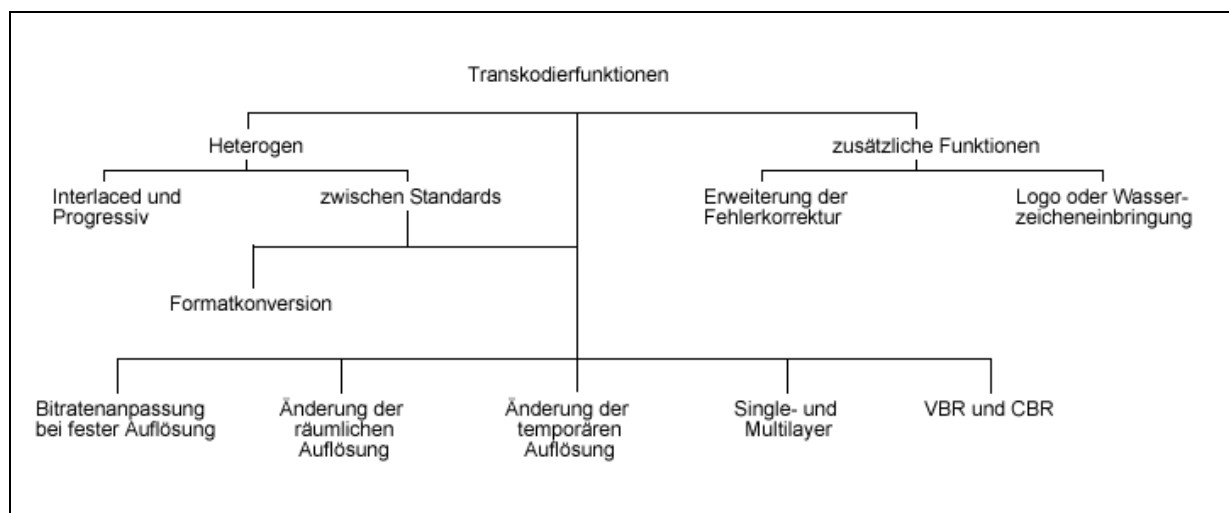


Abbildung 10: Klassifikation der Transkodierfunktionen²⁰

²⁰ nach Ahmad et al: Video Transcoding, S.794

Die einfachste Architektur für einen Transkoder ist die Hintereinanderschaltung eines Decoders und eines Kodierers. Dies ist aber relativ ineffizient, da recht wenige der Eingangsinformationen (z.B. Bewegungsvektoren, evtl. vorhandene Statistiken, etc.) wiederverwendet werden. Es existieren auch komplexere Kodierer, die die Eingangsinformationen besser ausnutzen (vergleiche hierzu [1]).

Nachfolgend werden ausgewählte Änderungsoperationen und die Probleme, die durch sie entstehen beschrieben.

4.1.1 Bitratenanpassung bei fester Auflösung

Die Anpassung der Bitrate dient hauptsächlich dazu, das Videomaterial an eine gegebene Bandbreite eines Netzwerkes zu adaptieren (z.B. für Streaming) und die Dateigröße zu reduzieren. Sie wird erzielt, indem man die Einträge der Quantisierungsmatrix erhöht,²¹ damit mehr AC-Koeffizienten den Wert Null annehmen. Da die Quantisierung aber ein verlustbehafteter Schritt ist, nimmt dadurch die Bildqualität ab. Dies kann sogar so weit führen, dass im Bild sogenannte Blockartefakte sichtbar werden. Abbildung 11 zeigt ein Beispiel. Allerdings ist dieser Transkodierschritt effizient auszuführen, da die Bewegungsvektoren nicht neu berechnet werden müssen.



Abbildung 11: Bild links: unkomprimiert, Bild rechts: 200-fach komprimiert

4.1.2 Änderung der räumlichen Auflösung

Die Änderung der räumlichen Auflösung wird nötig, wenn das Endgerät ein Display mit einer geringeren Auflösung besitzt. Sie kann zum Beispiel durch Cropping zustande kommen, d.h. man schneidet aus dem Bild die Region, die von Interesse ist (Region of Interest), aus. Oder man nutzt entsprechende Filter, die die Auflösung des gesamten Bildes reduzieren.

Problematisch ist es im ersten Fall die Region of Interest zu bestimmen. Dies muss, um eine gewisse Performance zu erhalten (noch) über Metadaten geschehen.

Im zweiten Fall stimmen die Bewegungsvektoren unter Umständen nicht mehr überein und sie müssen entweder neu berechnet werden oder die Ausgangsvektoren müssen auf neue Vektoren abgebildet werden (motion vector mapping). Abbildung 12 zeigt ein Beispiel.

²¹ nach Ahmad et al: Video Transcoding, S.794.

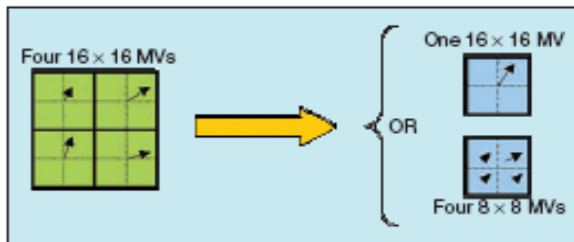


Abbildung 12: Illustration von 4:1 und 1:1 motion vector mapping²²

4.1.3 Änderung der temporären Auflösung

Falls ein Endgerät zu wenig Rechenleistung hat, um ein Video mit hoher Framerate zu dekodieren, müssen Frames gelöscht werden. Dazu nimmt man vorzugsweise inter-frame-kodierte Bilder. Hier entsteht wieder das Problem, dass die Bewegungsvektoren nicht mehr stimmen. Allerdings müssen die Vektoren nicht über verschiedene Blöcke, sondern über mehrere Frames hinweg neu berechnet werden (motion vector reestimation). Veranschaulicht ist dies in Abbildung 13.

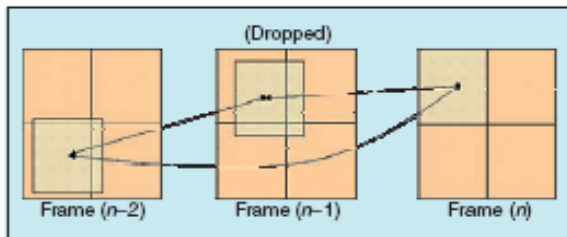


Abbildung 13: Motion vector reestimation²³

4.1.4 Erweiterung der Fehlerkorrektur

Bei der Versendung über kabellose Netzwerke (z.B. WLAN, UMTS) entstehen in der Regel mehr Übertragungsfehler als bei kabelgebundenen Netzwerken. Um diese zu kompensieren, muss der Anteil an intra-frame-kodierten Bildern oder Makroblöcken erhöht werden. Denn angenommen, in einem I-Picture ist ein Fehler, so setzt sich dieser über die gesamte Group of Pictures fort bis zum nächsten I-Picture. Durch das Einfügen redundanter Daten vergrößert sich natürlich die Datenmenge/Datenrate, was man gerade in diesen Netzwerken vermeiden will und es entsteht ein gewisser Trade-off.

4.1.5 Logo- oder Wasserzeicheneinbringung

In manchen Fällen ist es notwendig, sichtbare oder auch transparente Logos oder Wasserzeichen einzubringen (z.B. Senderlogo). Dabei kann man einen Großteil des ursprünglichen Materials nutzen, man muss nur in den Regionen, in denen das Logo/Wasserzeichen eingefügt wird, die Bewegungsvektoren neu berechnen.

²² Vetro et al: Video Transcoding Architectures and Techniques, S.22

²³ Vetro et al: Video Transcoding Architectures and Techniques, S.22

4.1.6 Heterogene Transkodierung

Ein heterogener Transkoder muss nicht nur Basisfunktionalitäten wie motion vector mapping, Änderung der räumlichen und temporären Auflösung und Bitratenanpassung beherrschen, sondern zusätzlich auch die Syntax des Eingangsmaterials in die des Ausgangsmaterials konvertieren können.²⁴ Dies verkompliziert natürlich den gesamten Aufbau des Transkoders erheblich. Besonders problematisch ist es, wenn der Eingabe- und Ausgabestandard unterschiedliche Ansätze haben, wie es z.B. bei MPEG-2 und MPEG-4 der Fall ist (blockbasiert kontra objektbasiert).

4.2 Skalierbare Kodierung

Ein anderer Ansatz zur Adaption von Videos ist die skalierbare Kodierung. Dabei wird das Video in einen Basis-Layer und einen oder mehrere Verbesserungs-Layer (enhancement layer) zerlegt. Hat man z.B. nur eine geringe Bandbreite zur Verfügung, sendet man nur den Basis-Layer, der für sich alleine dekodierbar ist. Dieser liefert dann das Video, allerdings in einer geringen Qualität. Bei steigender Bandbreite kann man dann auch die Verbesserungs-Layer verschicken und die Bildqualität nimmt zu. Zwei hervorstechende Nachteile hat dieses Verfahren allerdings. Zum einen ist die Anzahl der Verbesserungs-Layer ist begrenzt und zum anderen man kann mit der Bitrate höchstens so niedrig gehen, wie die Bitrate des Basis-Layers ist. Nachfolgend werden sechs Methoden für skalierbare Kodierung vorgestellt.

4.2.1 Temporäre Skalierung

Bei der temporären Skalierung enthält der Basis-Layer nur wenige Frames pro Sekunde (vorzugsweise I- und P-Pictures) und die Verbesserungs-Layer ergänzende Frames (B-Frames), so dass die temporäre Auflösung mit jedem Verbesserungs-Layer steigt (siehe Abbildung 14).

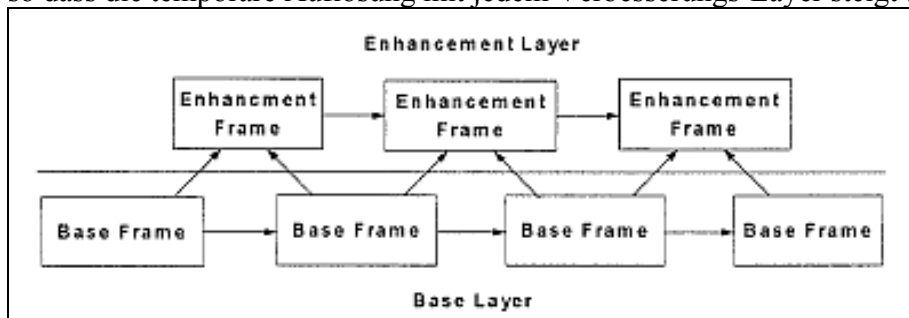


Abbildung 14: Darstellung der temporären Skalierung²⁵

4.2.2 Räumliche Skalierung

Bei dieser Methode wird der Basis-Layer mit einer niedrigen räumlichen Auflösung kodiert und die Verbesserungs-Layer enthalten Daten zur Interpolation des Basis-Layers auf die nächst höhere Auflösung²⁶.

²⁴ vgl. Ahmad et al: Video Transcoding, S.800

²⁵ Li: Overview of Fine Granularity Scalability in MPEG-4 Video Standard, S.303

²⁶ Dörries: Videostreaming, Teil 2, S.27

4.2.3 Datenpartitionierung

Bei der Datenpartitionierung werden im Basis-Layer priorisierte Daten wie die niedrigfrequenten DCT-Koeffizienten und die Bewegungsvektoren kodiert und der Verbesserungs-Layer enthält ergänzende höherfrequente DCT-Koeffizienten.

4.2.4 SNR-Skalierung (signal to noise ratio)

Die SNR Skalierung ist der Datenpartitionierung ähnlich, allerdings enthält hier der Basis-Layer alle DCT-Koeffizienten mit einem hohen Quantisierungsschritt und der Verbesserungs-Layer Daten zur Verbesserung der Quantisierung des Basis-Layers. Der Name der Methode kommt daher, dass die Verbesserungs-Layer das Signal-Rausch-Verhältnis (signal to noise ratio) verbessern.

4.2.5 Fein-granulare-Skalierung (fine granularity scalability, FGS)²⁷

Bei der FGS werden die DCT-Koeffizienten mit der sogenannten Bit-Plane Methode kodiert. Der Unterschied zur Lauflängenkodierung ist, dass die Koeffizienten nicht als absolute Ganzzahlwerte angesehen werden, sondern als Binärzahl mit einer bestimmten Anzahl Bits. Von jedem DCT-Block werden die 64 Koeffizienten in ihrer Binärdarstellung in sogenannte Bit-Planes (MSB = most significant bit) eingeteilt, die aus jeweils 64 Bits bestehen. Das folgende Beispiel zeigt diese Einteilung:

Koeffizienten	DC	AC ₁	AC ₂	AC ₃	AC ₄	AC ₅	AC ₆	AC ₇	AC ₈	AC ₉	AC ₁₀AC ₆₃
Absolutwerte	10	0	6	0	0	3	0	2	2	0	0.....0
MSB-Plane	1	0	0	0	0	0	0	0	0	0	0.....0
MSB-1-Plane	0	0	1	0	0	0	0	0	0	0	0.....0
MSB-2-Plane	1	0	1	0	0	1	0	1	1	0	0.....0
MSB-3-Plane	0	0	0	0	0	1	0	0	0	0	0.....0

Binärdarstellungsrichtung

Die einzelnen MSB-Planes werden dann durch Variable-Längen-Kodierung folgendermaßen komprimiert:

Die Anzahl der aufeinanderfolgenden Nullen wird mit dem EOP (end of plane = nur noch Nullen folgen, wobei 1 dem EOP entspricht) kombiniert.

Aus dem obigen Beispiel ergibt sich:

- MSB-Plane: (0,1)
- MSB-Plane-1: (2,1)
- MSB-Plane-2: (0,0),(1,0),(2,0),(1,0),(0,1)
- MSB-Plane-3: (5,1)

Dieses Ergebnis wird dann nochmals mit einer vorgegebenen Tabelle variable-length-kodiert. Der Vorteil dieser Technik ist, dass die Bildqualität proportional zur Anzahl der erhaltenen Bits steigt und dass der Bitstrom des Verbesserungs-Layers jederzeit unterbrochen werden kann. Außerdem sind die Informationen direkt nach ihrer Wichtigkeit gegliedert. Abbildung 15 zeigt die Ergebnisse von der Darstellung eines Bildes, dessen Intensität in 8 Bit-Planes unterteilt wurde.

²⁷ Li: Overview of Fine Granularity Scalability in MPEG-4 Video Standard, S.304

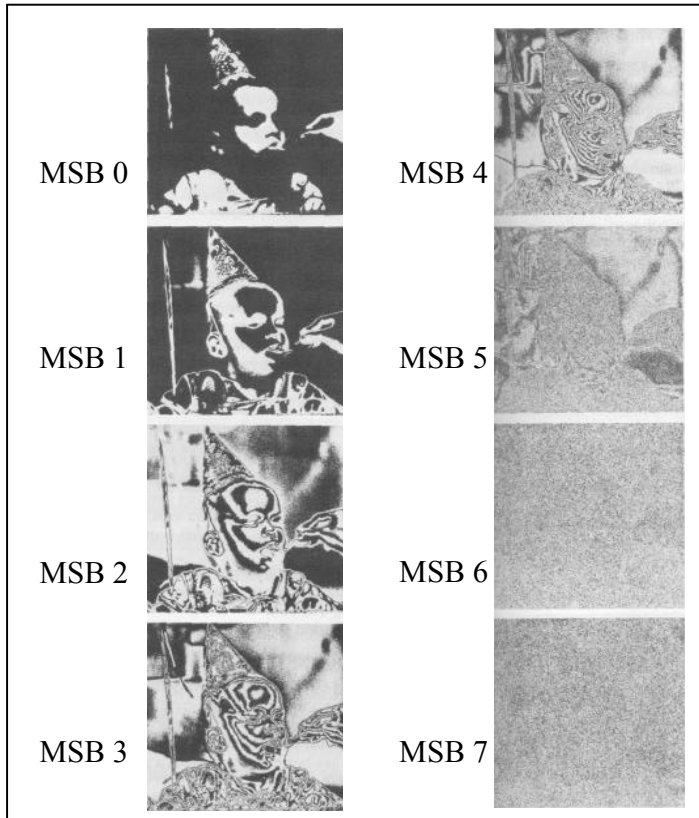


Abbildung 15: Darstellung der einzelnen MBR-Planes der Helligkeitswerte eines Bildes²⁸

4.2.6 Objekt-basierte-Skalierbarkeit

Die bisher beschriebenen Methoden zur Skalierbarkeit beziehen sich alle auf ganze Frames. MPEG-4 definiert eine weitere Skalierbarkeit, bei der das Object of Interest mit einer höheren Framerate kodiert wird als der Hintergrund (object-based temporal scalability, OTS). Dabei werden zwei Typen unterschieden: Bei Typ 1 wird der Hintergrund zusammen mit dem Object of Interest im Basis-Layer kodiert und der Verbesserungs-Layer enthält weitere P-Pictures mit dem Object of Interest. Typ 2 hingegen enthält im Verbesserungs-Layer B-Pictures des Object of Interest. Abbildung 16 veranschaulicht dies.

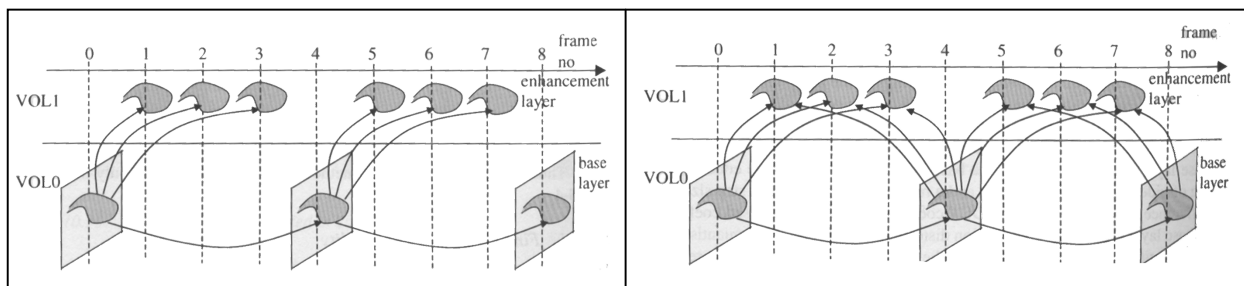


Abbildung 16: OTS bei MPEG-4; links Typ 1, rechts Typ 2²⁹

²⁸ Bit plane Coding – An explanation, S.3

²⁹ Ghanbari: Standard Codecs S.340

5 Inhaltsbeschreibung, -suche und -bezug

Bei den bisher vorgestellten Verfahren werden große Datenmengen erzeugt, da bei der Transkodierung von einem Video mehrere Versionen und bei der Skalierung verschiedene Layer erzeugt werden. Um diese Datenmengen effizient zu verwalten und auf ihnen effektive Suchalgorithmen zu implementieren (die nicht textbasiert sein werden) benötigt man eine standardisierte Datenbeschreibungssprache. Diese Datenbeschreibungssprache ist unter dem Namen MPEG-7 standardisiert worden. Außerdem wird durch den Wunsch nach Universal Multimedia Access ein Standard nötig, der den Austausch von multimedialen Daten über Netzwerkgrenzen hinweg ermöglicht, aber zugleich die Rechte der Urheber berücksichtigt. UMA bedeutet hierbei „universeller oder nahtloser Zugang zu multimedialem Inhalt durch automatische Inhaltsauswahl und –adaption basierend auf der Umgebung des Endnutzers“³⁰. Das Projekt MPEG-21 hat sich zum Ziel gesetzt ein Framework für UMA zu entwickeln.

Nachfolgend werden die grundlegenden Entwicklungsziele der beiden Standards vorgestellt.

5.1 MPEG-7

MPEG-7 definiert ein Interface zur Beschreibung multimedialen Inhalts. Es ist also ein Interface für Metadaten, die „bits about the bits“³¹. Es basiert auf XML und ist wie dieses durch das Konzept von Namespaces und Schemadefinitionen erweiterbar und an vielfältige Gegebenheiten anpassbar. Wie bei den anderen Standards wird nur die Darstellung und Dekodierung der Beschreibung genormt, wie man zum Inhalt der Beschreibung kommen soll wird nicht beschrieben. Abbildung 17 illustriert diesen Sachverhalt.

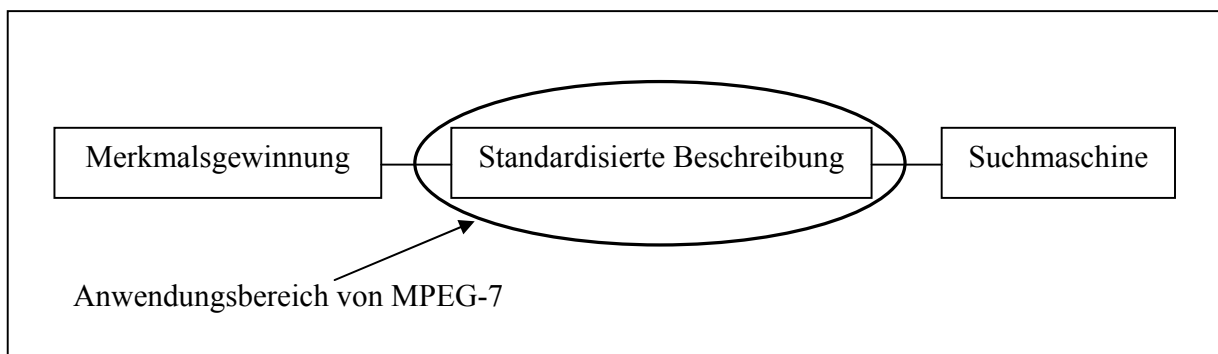


Abbildung 17: Der Anwendungsbereich von MPEG-7³²

Ein Beispiel für die Anwendung von MPEG-7 ist: Jemand sucht eine Konzertaufzeichnung eines Sängers, er hat aber nur ein Portrait und ein Lied dieses Sängers. Er gibt das Bild und die Musikaufzeichnung in eine Suchmaschine ein und diese durchsucht dann die ihr zur Verfügung stehenden MPEG-7 Daten. Wird eine Übereinstimmung gefunden, gibt die Suchmaschine das assoziierte Video aus.

MPEG-7 unterstützt also UMA, indem es Werkzeuge für die Beschreibung von Objekten, Transkodierhinweisen, Variationen und Zusammenfassungen multimedialer Inhalte bereitstellt. Außerdem werden Werkzeuge zur Beschreibung von Nutzerpräferenzen und Nutzungshistorie

³⁰ van Beek et al: Metadata-Driven Multimedia-Access, S.40

³¹ Manjunath et al: MPEG-7 S.8

³² nach Ghanbari: Standard Codecs S.349

definiert.³³ Dies geschieht über Audio- und Video-Deskriptoren auf unterschiedlichen Abstraktionsebenen. Die Informationen auf niedrigem Niveau können meist automatisch generiert werden, wohingegen die Informationen auf höheren Leveln (noch) manuell extrahiert werden müssen. Im Folgenden sind einige Videodeskriptoren aufgeführt:

Farbdeskriptoren³⁴ für

- Farbraum
- Dominante Farbe(n)
- Skalierbare Farbe
- GOP Farbe
- Farbstruktur
- Farblayout

Strukturdeskriptoren für

- Homogene Struktur
- Strukturbrowsing
- Kantenhistogramm

Formdeskriptoren für

- Bereichsbasierte Form
- Konturbasierte Form
- 3-D Form

Gesichtsdeskriptor

Bewegungsdeskriptoren für

- Bewegungsaktivität
- Kamerabewegung
- Bewegungsablauf
- Parametrische Bewegung

Deskriptoren für die Nutzungshistorie

Deskriptoren für Transkodierhinweise

Deskriptoren für das Medienformat

Vor allem die letztgenannten Deskriptoren sind für die Adaption von großem Nutzen. Zum Beispiel kann der Videoinhalt (objekt-, segment- oder regionenbasiert) durch die Deskriptoren für Transkodierhinweise gewichtet werden und so kann bei Cropping³⁵ oder der Erzeugung von Layern für Objekt-basierte-Skalierung³⁶ die Region bzw. das Object of Interest effizienter bestimmt werden. Aber auch die Deskriptoren auf niedrigerem Level wie z.B. der Deskriptor für

³³ vgl. van Beek et al: Metadata-Driven Multimedia-Access, S.41

³⁴ vgl. hierzu und zum Folgenden: Manjunath et al: MPEG-7, S.187ff

³⁵ vgl. 4.1.2., S.11

³⁶ vgl. 4.2.6., S.15

die Bewegungsaktivität können die Transkodierung verbessern, da man ihn nutzen kann, um den Suchbereich für die Bewegungsvektoren besser festlegen zu können.

5.2 MPEG-21

MPEG-21 ist ein erweiterbares Framework für Universal Multimedia Access, das sich teilweise noch in der Entwicklung befindet. Es definiert „Metadaten, die vor allem Relevanz für die Steuerung der Kommunikation besitzen“³⁷ sowie Interfaces zur Einbindung verschiedenster Services. Nachfolgend sind die wichtigsten Zielsetzungen und Konzepte erläutert:

5.2.1 Digital-Item-Deklaration³⁸

Heutzutage herrscht eine große Vielfalt unterschiedlicher digitaler Elemente (digital items, DI's), beispielweise HTML-Dokumente, MP3-Musikstücke, JPEG-Bilder, MPEG-Videos und viele mehr. Doch was macht diese Elemente aus? MPEG-21 hat sich zum Ziel gesetzt einen Standard zu entwickeln um beliebige digitale Elemente zu definieren und deklarieren.

5.2.2 Digital-Item-Identifikation

Digitale Elemente werden mit einem digitale-Elemente-Identifikator (digital item identifier, DII) versehen, der nach Möglichkeit auch auf globaler Ebene eindeutig sein sollte.³⁹

5.2.3 Schutz und Management von geistigem Eigentum

Bei transparentem Zugang für den Nutzen zu multimedialen Daten muss der Schutz von geistigem Eigentum gewährleistet sein. MPEG-21 soll ein Framework bieten, das diesen Schutz über Netzwerk – und Gerätegrenzen hinweg garantiert und verwaltet. Dazu bedient sich MPEG-21 der Rights Expression Language und dem Rights Data Dictionary.

5.2.4 Digital-Item-Adaption

Natürlich muss ein UMA-Framework auch die Aufbereitung und Adaption der verwalteten digitalen Daten beherrschen. Abbildung 18 zeigt schematisch eine solche Adaption.

³⁷ Kosch, Heuer: MPEG-7

³⁸ vgl. hierzu und zum Folgenden: Bormans et al: MPEG-21, S.56f

³⁹ vgl. Ghanbari: Standard Codecs S.365

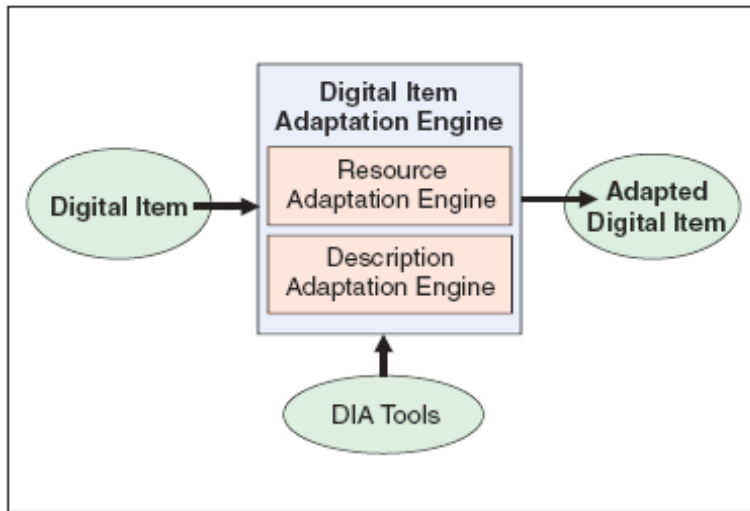


Abbildung 18: Illustration der Digital-Item-Adaption⁴⁰

Gerade für die Digital-Item-Adaption sind die in Kapitel 4 vorgestellten Methoden⁴¹ von großer Bedeutung, da die Digital-Item-Adaption-Engine natürlich auch einen Transkoder für Videomaterial, bzw. Mechanismen zur Ausgabe von skalierbaren Videos implementieren muss.

6 Zusammenfassung und Ausblick

Abschließend ist zu sagen, dass die Menge von digitalen multimedialen Inhalten in den nächsten Jahren stark zunehmen wird. Die hier vorgestellten Methoden beschreiben Ansätze, diese Inhalte zu strukturieren und für jedermann immer und überall zugänglich zu machen. Allerdings ist eine konsequente Umsetzung der vorgestellten Konzepte für Universal Multimedia Access und die Weiterentwicklung von allen Seiten (Forschung, Industrie, Anwender) her nötig, damit der Endnutzer auch wirklich davon profitieren kann.

⁴⁰ Bormans et al: MPEG-21, S.58

⁴¹ vgl. Kapitel 4, S.10ff

7 Literaturverzeichnis

- [1] Ahmad, I. und Wei, X. und Sun, Y. und Zhang, Y.: Video Transcoding: An Overview of Various Techniques and Research Issues. IEEE TRANSACTIONS ON MULTIMEDIA, Vol. 7, No. 5, Seiten 793-804, Oktober 2005
- [2] Bit plane Coding – An explanation.
<http://www.wmin.ac.uk/itrg/IS/DPI/Digital%20Image%20Processing/Bit%20Plane%20Coding.pdf>, University of Westminster
- [3] Dörris, G.: Videostraming, Teil 2. Fraunhofer Institut für Medienkommunikation, <http://www2.inf.fh-bonn-rhein-sieg.de/~kjonas2m/lv/WS2002/mbtv2/vorlesung/streaming2.pdf>, 2002
- [4] Effelsberg, W. und Steinmetz, R.: A Graduate Course on Multimedia Technology. <http://www.informatik.uni-mannheim.de/pi4/lectures/ws0506/mm/>, Mannheim, 2006
- [5] Effelsberg, W. und Steinmetz, R.: Video Compression Techniques. dpunkt-Verlag, Heidelberg, 1998
- [6] Ghanbari, M.: Standard Codecs: Image Compression to Advanced Video Coding. The Institution of Electrical Engineers, London, 2003
- [7] Kosch, H und Heuer, J.: MPEG-7. <http://www.gi-ev.de/service/informatiklexikon/informatiklexikon-detailansicht/meldung/52/>, Gesellschaft für Informatik e.V., Informatiklexikon, 2003
- [8] Li, W: Overview of Fine Granularity Scalability in MPEG-4 Video Standard. IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, Vol. 11, No. 3, Seiten 301-317, März 2001
- [9] Manjunath, B.S. und Salembier, P. und Sikora, T.: Introduction to MPEG-7: Multimedia Content Description Interface. John Wiley & Sons Ltd., West Sussex, 2002
- [10] Mitchell, J.L. und Pennebaker, W.B. und Fogg, C.E. und LeGall, D.J.: MPEG Video Compression Standard. Digital Multimedia Series, Chapman & Hall, New York, 1996
- [11] van Beek, P. und Smith, J.R. und Ebrahimi, T. und Suzuki, T. und Askelof J.: Metadata-Driven Multimedia-Access. IEEE SIGNAL PROCESSING MAGAZINE, Seiten 40-52, März 2003
- [12] Vetro, A. und Christopoulos, C. und Sun, H.: Video Transcoding Architectures and Techniques. IEEE SIGNAL PROCESSING MAGAZINE, Seiten 18-29, März 2003
- [13] Wikipedia.org: http://de.wikipedia.org/wiki/Diskrete_Kosinustransformation, Stand September 2005

8 **Abbildungsverzeichnis**

Abbildung 1: Graphische Darstellung der 64 Basis-Frequenzen der 2-Dimensionalen DCT.....	3
Abbildung 2: Quantisierungsmatrix.....	4
Abbildung 3: Mögliche Scanreihenfolgen der AC-Koeffizienten bei der Entropiekodierung....	4
Abbildung 4: Bestimmung eines Bewegungsvektors.....	5
Abbildung 5: Eine typische GOP in Anzeigereihenfolge.....	6
Abbildung 6: Eine typische GOP in Dekodierereihenfolge.....	7
Abbildung 7: Evolution der Video/Bildkompressionsverfahren.....	7
Abbildung 8: Einfacher MPEG-Kodierer.....	8
Abbildung 9: Profile in MPEG-2 aufgetragen in Bezug zur Auflösung.....	9
Abbildung 10: Klassifikation der Transkodierfunktionen.....	10
Abbildung 11: Bild links: unkomprimiert, Bild rechts: 200-fach komprimiert.....	11
Abbildung 12: Illustration von 4:1 und 1.1 motion vector mapping.....	12
Abbildung 13: Motion vector reestimation.....	12
Abbildung 14: Darstellung der temporären Skalierung.....	13
Abbildung 15: Darstellung der einzelnen MBR-Planes der Helligkeitswerte eines Bildes.....	15
Abbildung 16: OTS bei MPEG-4; links Typ 1, rechts Typ 2.....	15
Abbildung 17: Der Anwendungsbereich von MPEG-7.....	16
Abbildung 18: Illustration der Digital-Item-Adaption.....	19

9 Abkürzungsverzeichnis

AC	Alternate Current
BMA	Block Matching Algorithm
CBR	Constant Bit Rate
DC	Direct Current
DCT	Diskrete Kosinustransformation
DI	Digital Item
DII	Digital Item Identifier
DVB-T	Digital Video Broadcasting-Terrestrial
DVD	Digital Versatile Disc
HDTV	High Definition Television
HTML	Hypertext Markup Language
ISO	International Standardisation Organisation
JPEG	Joint Photographic Experts Group
LAN	Local Area Network
MB	Makroblock
MC	Motion Compensation
ME	Motion Estimation
MP3	MPEG Audio Layer 3
MPEG	Moving Pictures Experts Group
MSB	Most Significant Bit
MV	Motion Vector
NTSC	National Television Systems Committee
OTS	Object-based Temporal Scalability
PAL	Phase Alternating Line
PDA	Personal Digital Assistant
SNR	Signal-to-Noise-Ratio
UMA	Universal Multimedia Access
UMTS	Universal Mobile Telecommunications System
VBR	Variable Bit Rate
VOP	Video Object Plane
WLAN	Wireless Local Area Network
XML	Extensible Markup Language