

# Verbale Beschreibung des Bilddatenkompressionsverfahren nach MPEG

Philipp Spinner, 99-714-784  
Kernvorlesung Multimediale Systeme  
Institut für Informatik der Universität Zürich  
pspinner@access.unizh.ch  
Abgabe: 18. Juni 2002

## Einleitung

Die Moving Pictures Expert Group (MPEG) wurde Ende der 80er Jahre zur Festlegung eines digitalen Standards für Bewegtbilddarstellung ins Leben gerufen. Bis zur Verabschiedung der Norm MPEG-1 standen bereits verschiedene Verfahren zur Verfügung. Zu den bis dato wichtigsten Vertretern gehörten Motion-JPEG (M-JPEG) und die Recommendation H.261 der CCITT<sup>1</sup>.

Das M-JPEG-Verfahren basiert auf der Serialisierung von im JPEG-Format komprimierten Einzelbildern. Der einstellbare Kompressionsgrad liegt für gute Bildqualität bei einem Faktor von 20-25:1. Dieses Einzelbildkompressionsverfahren erwies sich aber bald als sehr ineffizient. Auch die Einbindung von Audiodaten ist nicht im (Quasi-)Standard definiert.

Mit der Recommendation H.261 stand ein audiovisueller Standard der CCITT<sup>1</sup> zur Verfügung. H.261 besitzt durch seine Ausrichtung auf Bewegtbildübertragung bereits einen hybriden Kompressionsalgorithmus, der aus einer Spielart des Deltaverfahrens, einer Restbildkompression sowie einem optional einbindbaren Verfahren zur Bewegungskompensation. Doch der Standard war relativ strikte und so konnte man ihn z.B. auch nicht für TV-Übertragungen oder innerhalb von Multimediaanwendungen benutzen, da entscheidende Modi nicht unterstützt wurden.

Mit den Normen für Bewegtbilddarstellungen, die für die Bereich Heimanwendung (MPEG-1) und Studio (MPEG-2) entwickelt wurden, versuchte man für die Nachteile der oben genannten verfügbaren Standards eine Lösung zu finden; Wobei wesentliche Merkmale von M-JPEG als auch H.261 in die Normung eingingen. MPEG wurde dabei als generische Norm entworfen, wodurch eine Anpassung an verschiedene Anwendungen möglich gemacht werden soll. Diese beiden Standards erkläre ich im nächsten Abschnitt. Auf den MPEG-4 Standard gehe ich im übernächsten ein. Über den MPEG-3 Standard sage ich nichts, da davon das meiste sowieso vom MPEG-2 Standard schon umgesetzt wurde.

## MPEG 1 + 2

Die ältere *MPEG-1* Norm findet Anwendung in der PC-Welt und im Unterhaltungsmarkt (z.B. Filme auf Compact Disks).

*MPEG-2* wurde im Hinblick auf digitales Fernsehen definiert, es unterstützt das Zeilensprungverfahren (Interlace). Die der Kompression zugrunde liegenden Verfahren wie Bewegungskompensation, DCT, Quantisierung und Huffman-Codierung sind für beide MPEG-Normen im Prinzip identisch.

### 1 Schichten

MPEG beschreibt, wie komprimierte Video- und Audio-Daten zu einem Bitstrom zusammengefasst werden. Die Norm umfasst zwei Teile:

*Systemschicht:* definiert, auf welche Weise Bitströme von Video- und Audio-Daten zu einem einzigen Bitstrom zusammengefasst werden. Ausserdem werden Zeitmarken bereitgestellt, die der synchronisierten Wiedergabe von Bild und Ton dienen. Die Systemschicht ist unterteilt in den Pack-Layer und den Packet-Layer. Innerhalb eines Packs werden Packets aneinander gereiht. Packets enthalten neben den Headerinformationen nur Daten einer einzigen Art, also entweder Audio-Daten oder Video-Daten.

*Kompressionsschicht:* Diese Daten werden von den Kompressionsschichten für Video und Audio codiert und zur Übertragung in die Päckchen eingefügt.

### 2 Bewegungskompensation

Das MPEG-Verfahren nutzt die Tatsache, dass in Folgen bewegter Bilder zwischen aufeinander folgenden Bildern grosse Ähnlichkeit besteht. Mit der Ausnahme krasser Szenenwechsel werden sich Bilddetails kontinuierlich von einem Bild zum nächsten fortsetzen, wie zum Beispiel ein sich von links nach rechts bewegendes Fahrzeug oder eine weisse Wolke, die vor dem Hintergrund eines blauen Himmels vorbeizieht. Die Bewegung des Objekts wird einfach durch einen Vektor beschrieben, zum Beispiel durch die Angabe, dass das Objekt sich von einem Bild zum nächsten um 12 Pixel nach rechts und 10 Pixel nach oben bewegt hat. Die Erkennung eines zusammengehörigen Objekts wäre in der Praxis allerdings viel zu aufwendig. Stattdessen werden so genannte

---

<sup>1</sup> CCITT: The International Telegraph and Telephone Consultative Committee

Makroblöcke mit einer Pixelgröße von 16x16 untersucht. Diese Makroblöcke entsprechen 4 Blöcken, wie sie bei JPEG codiert werden. Im nächsten Schritt wird die Differenz aus dem realen Makroblock im ersten Bild und dem verschobenen im zweiten Bild gebildet. Dieses Fehlerbild muss neben dem Verschiebungsvektor zur Beobachtung der Fehlerfortpflanzung codiert und gespeichert werden. Der geringste Speicheraufwand entsteht natürlich, wenn der Unterschied zwischen den verschobenen Makroblöcken und den tatsächlich dargestellten Blöcken so klein ist, dass auf die Codierung der Differenz ganz verzichtet werden kann.

MPEG steuert die Darstellung von komprimiertem Video, durch die Festlegung einer Syntax [1].

### 3 Bildtypen

MPEG verwendet drei Bildtypen: I-(*Intra*), P-(*Predicted*) und B-(*Bidirectional predicted*) Bilder. *Intracodierte Bilder* verwenden nur Informationen aus einem einzigen digitalisierten Vollbild. Bei I Bildern findet die Bewegungskompensation somit keine Anwendung. Sie dienen aber als Ausgangspunkt für die Ermittlung von P- und B-Bildern. P-Bilder werden durch Bewegungskompensation aus zeitlich zurückliegenden Bildern vom Typ I oder P abgeleitet. Auch P-Bilder dienen demnach als Referenzbilder für die Ermittlung von Verschiebungsvektoren. B-Bilder werden entweder aus früheren oder späteren P- bzw. I-Bildern abgeleitet oder auf Basis von Nachbarbildern interpoliert. Sie dienen niemals zur Berechnung von Bewegungs-Vektoren.

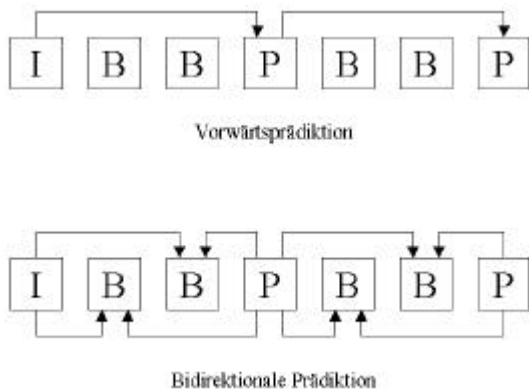


Abbildung 1: Zusammenspiel der Bildtypen [2]

### 4 Bildkomprimierung

Nebst der Bewegungskompensation bedient man sich folgender Komprimierungstechniken:

- *Bildformat YUV*: Das normalerweise verwendete RGB-Format werden die Farbanteile zu einer Mischfarbe kombiniert. Das Bildformat enthält nicht drei Farbanteile, sondern einen Anteil Helligkeit und zwei Differenzfarbanteile. Da nun das menschliche Auge für Helligkeitsunterschiede deutlich empfindlicher ist als

für Farbunterschiede, wird bei diesem Format nur jeder zweite oder vierte Farbwert gespeichert. Dadurch entsteht der kleinere Speicheraufwand.

- *DCT (Diskrete Kosinus Transformation)*: Eine weitere wesentliche Funktion die Überführung der Pixeldaten in den Frequenzraum. Hier wird entsprechend der zweidimensionalen Ausdehnung eines Einzelbildes bzw. der bereits zusammengefassten Blöcke von 8\*8 Pixeln eine zweidimensionale Diskrete Kosinustransformation angewendet. Sie ist ähnlich der Fouriertransformation. Im Frequenzraum lassen sich dann auch die
- *Quantisierungen*: anwenden. Dies sind Filterfunktionen und Wertungskriterien, welche auch bei JPEG-Bildern angewandt werden, um die für das Auge irrelevanten Frequenzen wegfällen zu lassen.
- *Laufängencodierung*: Der letzte Komprimierungsschritt. Hier werden die quantisierten DCT-Werte aufgrund ihrer statistischen Häufigkeit codiert. Komprimierungsmethoden sind, z.B.: Huffman-Codierung oder Arithmetische Codierung.

### 5 Übertragungsraten

Für eine unkomprimierte digitale Videodarstellung sind, z.B. für das Fernsehen eine Datenrate von 290 Mbits/s bis 370 Mbits/s nötig. Mit einer verlustfreien Komprimierung wird zirka ein Komprimierungsfaktor von 2 bis 4 zu 1 realisiert. Durch die vorher genannten verlustbehafteten Komprimierungstechniken erreicht der MPEG-1+2 Standard mit der Bewegungskompensation meist 0.5 bis 15 Mbit/s (192 KByte/s) je nach Auflösung, Bildwiederholrate (bis 30pps möglich) und Detaileinstellungen der Komprimierungstechniken.

### MPEG-4

Für die in der letzten Zeit entstandenen Multimediaanwendungen, die über Kanäle wie z.B. ISDN, aber auch Mobilfunk gesendet werden, und dabei sehr viel kleinere Datenübertragungsraten benötigen, musste ein einheitlicher Standard gebildet werden, der sehr flexibel gehandhabt werden kann.

### 1 Szene

Der MPEG-4 Standard ist mit aufeinander folgenden Szenen aufgebaut. Im Gegensatz zu allen bisherigen Komprimierungsstandards steht dabei nicht mehr die Pixelmatrizenkomprimierung im Vordergrund, sondern vielmehr der eigentliche Szeneninhalte. Die Szene wird in ihre Bestandteile zerlegt, in primitive Audio- und Videoobjekte. Diese Objekte können synthetisch (computergeneriert) hergestellt sein oder aber natürlich aufgenommen. Diese Mischform wird das

erste Mal im MPEG-4 Standard aufgenommen. Daneben hat jede Szene ihren Szenengraph, der für die räumliche und zeitliche Positionierung der einzelnen Objekte sorgt, und auch den Kamerablickwinkel bestimmt. Um nun in die Szene eingreifen zu können, um z.B. Manipulationen vorzunehmen, muss nur der Szenengraph verändert werden.

## 2 Videoobjekte und Umgebung

Die Informationen eines Videoobjektes werden in drei Bereiche aufgeteilt:

- *Form (Shape coding)*: Die Form kann entweder durch eine binäre Matrix (welche so gross ist wie die Form) dargestellt werden, bei der die Bereiche innerhalb der Matrix, welche zur Form gehören mit 1 die anderen mit 0 gekennzeichnet werden. Oder aber die Bereiche haben jeweils einen Wert zwischen 0 und 255, welcher die Durchsichtigkeit des Bereiches anzeigt.
- *Bewegungsinformationen (Motion estimation und compensation)*: Übernommene Idee von MPEG-1+2, die mit I-, P- und B-Bilder arbeiten. Hier wird dies übertragen auf die Videoobjekte angewendet.
- *Texturen (Texture coding)*: Für die Komprimierung der auf die Form aufgesetzten Textur werden altbewährte Algorithmen eingesetzt. Sie werden wieder blockweise mit der DCT transformiert und danach quantisiert.

Daneben kann auch die *Umgebung (Sprite coding)* über den Bildausschnitt hinaus gespeichert werden. Es wird dazu nur noch die relevante Kameraposition gespeichert, damit der sichtbare Bereich innerhalb der Umgebung bestimmt ist.

### Ausblick MPEG-7

Mit dem MPEG-7 Standard wird in der Zukunft die Suche nach Multimediadaten sehr erleichtern. MPEG-7 erweitert die MPEG-2 und 4 codierten Daten um so genannte „Header-Bits“. In denen wird der Inhalt der codierten Daten beschrieben. Diese Daten können dann auch von Suchmaschinen benutzt werden.

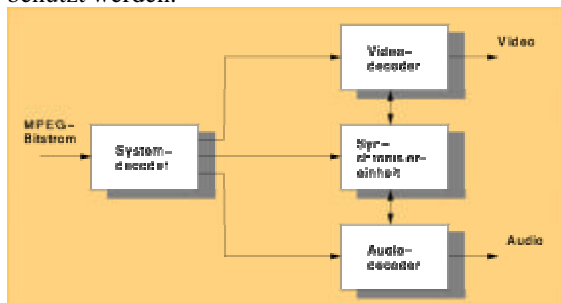


Abbildung 2: Bild eines MPEG Decoders [1]

## Literaturangaben

- [1] Klaus Schroiff, Multimedia-Datenformate; Kapitel 6: MPEG I+II, 1995. Internetreferenz

[http://i31www.ira.uka.de/docs/semin94/06\\_MPEG/main\\_html.html](http://i31www.ira.uka.de/docs/semin94/06_MPEG/main_html.html) (letzter Zugriff: 18.Juni 2002)

- [2] Björn Eisert, „Was ist MPEG?“ – Ein Erklärungsversuch, 1995. Internetreferenz: <http://www.cybersite.de/german/service/Tutorial/mpeg/> (letzter Zugriff: 18.Juni 2002)
- [3] Thomas Angermayer, „MPEG Videokompression“, Seminararbeit am Technikum Wien, 2000. Internetreferenz: [http://www.angermayer.com/referate/pdfs/referat\\_mpeg-videokompression.pdf](http://www.angermayer.com/referate/pdfs/referat_mpeg-videokompression.pdf)