

Videokodierung

MPEG

Dirk Theisen

Proseminar Multimediadatenbanken und Retrieval
Institute for Web Science and Technologies
FB4 – Computer Science
Universität Koblenz-Landau,
56070 Koblenz, Germany

dirktheisen@uni-koblenz.de

Abstract. Videokodierung ist in Zeiten von Multimediagesellschaften ein wichtiger Aspekt in der Informatik. Mit dem Aufkommen digitaler Videos stellte sich schnell die Frage nach adäquater Speicherung auf den jeweiligen Medien und damit nach Kompression ohne markanten Qualitätsverlust. Die Anwendungsgebiete reichen dabei von Medien mit großen Bandbreiten wie DVDs oder Blu-rays über digitales Fernsehen bis hin zu mobilen Anwendungen auf PDAs oder Mobiltelefonen. Ein wichtiger Aspekt ist dabei die Standardisierung, in welcher die Expertengruppe MPEG eine bedeutsame Rolle spielt. Weiterhin stellt sich die Frage, welche Anwendungsszenarien es für verschiedene Kompressionsstandards gibt und in welchen Formaten komprimierte Videos gespeichert werden können.

1 Einführung

Der Begriff der Videokodierung bezeichnet allgemein die Darstellung analoger und digitaler Videodaten sowie deren Verarbeitung und Übertragung. Dabei betrifft dies nicht nur natürliche Videos (Filme), sondern auch Computeranimationen.

Ziele sind neben der eigentlichen Darstellung der Videodaten unter anderem das Hinzufügen von weiteren (Meta-)Informationen oder rechtlichem Schutz wie beispielsweise das Digital Rights Management¹. Die größte Herausforderung ist jedoch in allen Bereichen die Videokompression – das Überführen (Enkodieren) von Videodaten in kleinere Datenmengen –, da die Bandbreiten und Ressourcen der meisten Speichermedien und Übertragungswege eine Wiedergabe oder Übertragung unkomprimierter Videodaten nicht zulassen. Aus diesem Grund werden die Begriffe Videokodierung und Videokompression (auch Videokomprimierung) heute oft synonym verwendet.

¹ Das DRM beschreibt Verfahren zur Kontrolle von Nutzung digitaler Medien

Um eine Vorstellung über das Datenaufkommen eines digitalen Videos zu geben und gleichzeitig einige Grundbegriffe der Videotechnik zu erläutern, wird im folgenden Abschnitt zunächst das digitale Fernsehen betrachtet. Davon ausgehend werden Kompressionsmöglichkeiten vorgestellt, welche dann an den MPEG Standards im übernächsten Abschnitt genauer erörtert werden. Weiterhin werden andere Kodierverfahren sowie zuletzt die so genannten Containerformate vorgestellt.

2 Grundlagen

2.1 Digitales Fernsehen

Es gibt heute zwei große Standards in der Fernsehtechnik. Der PAL (Phase-Alternation-Line) - Standard in Europa, Asien, Afrika und Australien bietet eine Bildauflösung von 720 x 576 Pixel bei einer Farbtiefe von 24 Bit mit einer Bildwiederholung von 50 Hz (Bilder pro Sekunde). In Amerika kommt größtenteils der NTSC (National Television Systems Committee) - Standard mit einer Auflösung von 720 x 480 Pixel in 24 Bit Farbtiefe und 60 Hz Bildwiederholung zum Einsatz. Die Entstehung zweier verschiedener Standards ist auf die unterschiedlichen technischen Voraussetzungen der Stromnetze in Europa (50Hz) und Amerika (60Hz) und damit der Röhrenfernseher, welche die Frequenzen des Stromnetzes zur Bildwiederholung nutzen, zurückzuführen. Das Bild liegt in beiden Fernsehsystemen in Halbbildern (*interlaced*) vor, es wird abwechselnd jede zweite Zeile des Bildes gesendet. Das menschliche Auge erkennt durch seine natürliche Trägheit ein Gesamtbild. Effektiv werden also 25 Vollbilder pro Sekunde im PAL Fernsehen und 30 Vollbilder pro Sekunde im NTSC Fernsehen im *Zeilensprungverfahren*² übertragen.

Das Auge nimmt physiologisch bedingt Bildabfolgen ab 16 Bildern pro Sekunde als flüssiges Video wahr. Damit sind die 25 - 30 Vollbilder pro Sekunde für den Menschen als flüssiges Fernsehbild ausreichend.

Standardmäßig liegt Videomaterial im RGB - Farbraum, bei dem jeder Pixel des Bildes additiv aus den drei Farbkomponenten Rot, Grün und Blau zusammengemischt wird, vor. Beim Fernsehen wird jedoch das YCbCr - Farbmodell verwendet. Dieses besteht aus einer Helligkeitskomponente Y und zwei Farbkomponenten Cb (*colour blue*) und Cr (*colour red*). Die wichtigste Information wird – wie in Fig. 1 ersichtlich – bereits als Schwarz-Weiß-Bild in der Y- Komponente gespeichert. Weiterhin erhält die Y-Komponente bereits Grüninformationen.



Fig. 1. YCbCr Bildinformationen

² Zeilensprungverfahren: Synonym für interlaced

Die Gründe für die Verwendung des YCbCr - Farbmodells sind zwei weitere Eigenschaften des menschlichen Auges, welches nämlich zum einen eine geringere Farbauflösung als Helligkeitsauflösung besitzt und weiterhin im grünen Farbbereich besonders empfindlich ist. Im YCbCr-Farbmodell kann nun die Farbauflösung gegenüber der Helligkeitsauflösung verringert werden, ohne dass dies subjektiv für das Auge bemerkbar wird, da Grünanteile und Helligkeitsinformationen weiterhin voll aufgelöst dargestellt werden.

Das Verringern der Farbauflösung bezeichnet man auch als Farbrunterabtastung (Chroma Subsampling). Es gibt verschiedene Varianten des Subsampling, die unterschiedliche Datenmengen in Anspruch nehmen. Bezeichnet werden die Varianten durch die drei Werte A:B:C.

Der erste Wert A beschreibt den Faktor der Abtastung der Helligkeit und ist standardmäßig 4. Er ergibt sich aus der NTSC - Videobandbreite im analogen Fernsehen und wurde für das digitale Fernsehen der Einfachheit halber übernommen, obwohl er teilweise nicht mehr korrekt ist. Der zweite Wert B steht für den Faktor der Abtastung der beiden Farbkomponenten Cb und Cr. Der dritte Wert C sagt aus, wie die Farbkomponenten in der zweiten Dimension abgetastet werden.

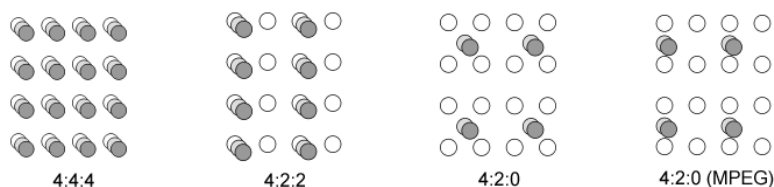


Fig. 2. YCbCr Subsampling

Die in Fig. 2 dargestellten Grafiken beschreiben das Verhältnis der Abtastung der Y - Komponente (weiße Kreise), und der beiden Farbkomponenten (graue Kreise). Es ergeben sich folgende Abtastraten:

4:4:4 – Hier erfolgt keine Farbrunterabtastung, da alle 3 Komponenten gleichhäufig abgetastet werden. Die Informationsmenge entspricht somit dem original vorliegenden RGB-Raum. Anwendung findet dies in der hochwertigen Bildverarbeitung, in der die vollen Farbinformationen benötigt werden.

$$Y + Cb + Cr = 3 YCbCr$$

4:2:2 – Diese Farbrunterabtastung entstand im NTSC - Standard. In horizontaler Richtung wird im Gegensatz zur vertikalen Richtung nur jede zweite Farbinformation abgetastet. Dies gleicht sich durch das Zeilensprungverfahren in NTSC und PAL aus, da somit in beiden Dimensionen jedes zweite Pixel farbabgetastet wird. Die Einsparung an der Datenmenge liegt beim 4:2:2 Subsampling bei 33% im Vergleich zu einem RGB- oder 4:4:4 gesampelten Bild.

$$Y + 1/2 Cb + 1/2 Cr = 2 YCbCr$$

4:2:0 – Diese Art der Abtastung wird bei Bildern im JPEG Format und beim MPEG-1 – mit leicht versetzter Farbabtastung – verwendet. Es ergibt sich eine Einsparung von 50%.

$$Y + 1/4 Cb + 1/4 Cr = 1,5 YCbCr$$

Mittels Chroma Subsampling lassen sich also die Größen der Einzelbilder eines Videos um 33% oder 50% verringern, ohne dass dies subjektiv für den Betrachter bemerkbar wird. Die folgende Beispielrechnung soll jedoch das Datenaufkommen eines Fernsehbildes im PAL - Standard verdeutlichen.

Spezifikation: 25 Vollbilder in der Sekunde, Auflösung 720 x 576, 24 Bit Farbtiefe:

$$25/s \times 720 \times 576 \times 24\text{Bit} \cong 248 \text{ Mbit/s}$$

Durch das 4:2:2 Subsampling lässt sich diese Rate um ein Drittel auf ≈ 166 Mbit/s verringern. Jedoch beträgt die maximale Übertragungsrate beim digitalen Fernsehen (DVB³) 5-8 Mbit/s, selbst aktuelle DSL Bandbreiten von bis zu 50 Mbit/s würden eine Übertragung nicht ermöglichen können, wodurch weitere Kompressionsverfahren angewendet werden müssen.

2.2 Videokompression

2.2.1 Voraussetzungen

Ein im YCbCr-Farbmodell vorliegendes Video hat wie bereits gesehen eine deutlich zu hohe Datenmenge für die meisten Anwendungsfälle. Daher gibt es weitere Kompressionstechniken, die die Datenmenge verringern.

Man unterscheidet zunächst zwischen

- verlustfreier Kompression und
- verlustbehafteter Kompression

Mit der verlustfreien Kompression werden Redundanzen in Bild / Video und kodiertem Bitstrom entfernt, ohne dass sich dabei die Qualität verringert. Man spricht hierbei auch von einer Redundanzreduktion. Bei der verlustbehafteten Kompression werden möglichst für den Menschen unwichtige Bildinformationen eingespart (Irrelevanzreduktion). Je nach Kompressionsgrad ist es jedoch nicht mehr möglich, nur irrelevante Informationen zu entfernen, und es muss ein Kompromiss zwischen Qualität und Datenmenge geschlossen werden. Je höher die Kompression, desto niedriger die Datenmenge und Qualität des Videos. Mit der verlustbehafteten

³ Digital Video Broadcasting: Digitales Fernsehen in den Varianten - S (Sat), - S2 (HD-Sat), - T (Terrestrisch), - C (Kabel)

Kompression lassen sich im Gegensatz zu der nur begrenzt möglichen verlustfreien Kompression beliebig hohe Kompressionen – bis hin zur Unkenntlichkeit eines Bildes oder Videos – erzielen.

Jedoch muss bei allen Kompressionsverfahren auch der Zeitaufwand für Kodierung und Dekodierung – also die zeitliche Effizienz – beachtet werden. So bedeuten bei Echtzeitanwendungen wie der Videotelefonie Kodierzeiten von wenigen Sekunden bereits ein unbrauchbares Ergebnis.

2.2.2 Kompressionsverfahren

Die Kompressionsverfahren lassen sich in zwei weitere Gruppen eingliedern:

1. Bildkompression (Kompression der Einzelbilder)
2. Kompressionen im Bildablauf (Bewegungskompensation)

Bei der Bildkompression werden ähnlich zum JPG-Format die Informationen und damit die Größe der einzelnen Bilder eines Videos verringert. Hierdurch lassen sich bereits gute Einsparungen erzielen. Beispielsweise kann die Farbtiefe oder mit der Farbunterabtastung die Auflösung der Farbinformationen verringert werden. Die Idee, jedes Bild einzeln zu komprimieren, wird im Motion-JPEG Format umgesetzt. Immerhin kann auf diese Art die Datenrate für PAL-Fernsehen bei guter Qualität auf ca. 30 Mbit/s verringert werden.

In vielen Videos ändern sich im Bildverlauf jedoch meist nur kleine Bildbereiche, beispielsweise bleibt der Hintergrund in vielen Szenen über längere Zeit gleich. Hier macht es Sinn, Kompressionen im Bildablauf zu verwenden, in dem nur die Unterschiede aufeinanderfolgender Bilder erfasst werden.

Bildkompressionen wie das Subsampling werden heute in fast allen Kompressionsstandards eingesetzt. Es gibt verschiedene Möglichkeiten, Kompressionen im Bildablauf zu erzielen. Im folgenden Kapitel werden die Videokodierungs- und Kompressionsstandards der MPEG (Moving Pictures Expert Group) vorgestellt.

3 Standards und Formate

3.1 MPEG – Moving Pictures Expert Group

Das wichtigste Gremium in der Standardisierung der Videokodierung der letzten 20 Jahre ist die 1988 gegründete Moving Pictures Experts Group. Diese besteht aus rund 350 Experten aus internationaler Wirtschaft und Forschung, welche drei bis vier Mal jährlich zu fünftägigen Treffen zusammenkommen und sich mit Video- und Audiostandards beschäftigen. Dabei wird der Name MPEG heute meist eher für die von ihr entwickelten ISO-Standards als für die Expertengruppe selbst verwendet.

Eine weitere wichtige Institution ist die Internationaler Fernmelde Union (ITU), die die Standards mit der Bezeichnung H.26x entworfen hat. ITU und MPEG arbeiten mittlerweile jedoch eng zusammen und entwickeln ihre Standards in Gremien gemeinsam.

3.2 MPEG Standards

Es gibt mehrere MPEG Standards, die im Folgenden vorgestellt werden. Für alle Standards gilt, dass die MPEG nur den Bitstrom und den Dekodiervorgang standardisiert. Für den Einkodierer werden lediglich Musterimplementierungen vorgegeben, die nicht zwangsläufig effizient sind. Effiziente Einkodierer werden meist durch kommerzielle Drittanbieter entwickelt.

Da alle späteren Standards der MPEG auf dem MPEG-1 Standard aufbauen, wird dieser zunächst näher erläutert und an ihm die Kompressionsverfahren erklärt.

3.2.1 MPEG-1

MPEG-1 wurde als verlustbehafteter Kompressionsstandard für Video und Audio 1993 von der MPEG mit dem Ziel, Video in VHS-Qualität auf CD mit einer maximalen Datenrate von 1,5Mbit/s darzustellen, verabschiedet. Er verwendet viele Techniken des von der ITU bereits drei Jahre zuvor entwickelten Standards H.261, welcher jedoch Bildtelefonie und Videokonferenzen über ISDN standardisierte und somit geringere Bandbreiten zum Ziel hatte.

Verwendung fand der MPEG-1 Standard auf der Video-CD, dem Vorläufer der DVD. Der wohl bekannteste Teil des MPEG-1 Standards ist das mp3 - Audioformat, welches im MPEG-1 Audio Layer 3 standardisiert wurde.

Der MPEG-1 Standard besteht aus 5 Teilen, in denen verschiedene Punkte spezifiziert werden:

1. System: Speicherung und Synchronisation von Video und Audio
2. Video: Videokompression
3. Audio: Audiokompression
4. Testen: Testverfahren zur Korrektheit der Implementierung
5. Referenzsoftware: Beispielsoftware zur Demonstration der Kodierung

In dieser Arbeit werden Schwerpunkte auf die Video- und Audiokompression gelegt, da diese den komplexesten und wichtigsten Teil der Videokodierung ausmachen.

Wie der bereits 1990 vorgestellte H.261 Standard nutzt auch MPEG-1 diverse Methoden zur Redundanz- und Irrelevanzreduktion. Als maximale Auflösung unterstützt MPEG-1 Auflösungen von bis zu 4095 x 4095 Pixel, effektiv werden aber relativ geringe Auflösungen im SIF (Source Input Format) von 352 x 240, 352 x 288 oder 320 x 240 Pixel eingesetzt. Als Farbmodell wird YCbCr verwendet und nach der Umwandlung aus dem RGB-Raum wird das bereits vorgestellte 4:2:0 Subsampling durchgeführt, wodurch bereits 50% der Daten ohne markante Einschränkungen eingespart werden.

Als eine sehr effektive Möglichkeit, weiter zu komprimieren, bietet sich die Kompression im Bildablauf an. Dazu unterscheidet MPEG-1 vier Bildtypen, auch Frames genannt:

- I-Frames (intraframe)
- P-Frames (predicted frame)
- B-Frames (bidirectional frame)
- D-Frames (direct frame)

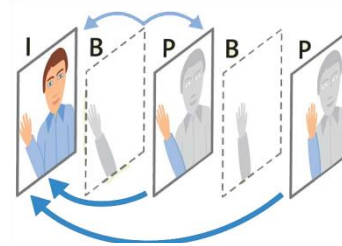


Fig. 3. Frametypen in MPEG-1

I-Frames werden als so genannte Schlüsselbilder unabhängig von vorausgehenden oder nachfolgenden Bildern kodiert. Sie benötigen am meisten Speicherplatz, sind aber notwendig, um im Video (fast) beliebig hin- und herspringen zu können. Die I-Framekodierung ist zwar schnell, jedoch ist die Datenrate recht hoch, da das Bild Pixel für Pixel kodiert werden muss – vergleichbar mit einem JPEG Bild – man spricht daher auch von Standbildern. Die eigentliche Bewegungskompression beginnt nun mit den weiteren Frametypen: Die P-Frames werden abhängig von vorausgehenden P- oder I-Frames kodiert, indem die Unterschiede zum vorausgehenden Bild und nicht die komplette Bildinformation gespeichert wird. Davon ausgehend, dass sich bis auf Szenenwechseln meist nur kleine Teile des Bildinhalts ändern, lässt sich hier viel Speicherplatz einsparen. Die höchste Kompression bieten die *B-Frames*, welche Informationen von dem vorhergehenden und nachfolgenden I- oder P-Frame benötigen. Da beim Zugriff auf B-Frames auch bereits zeitlich zurückliegende Frames aufgegriffen werden müssen, werden so genannte Bildpuffer eingesetzt, um die vergangenen Frames zwischen zu speichern. Weiterhin lässt sich mit B-Frames künstlich die Framerate eines Videos erhöhen, in dem man vereinzelt weitere B-Frames als Zwischenbilder einfügt.

Als vierter Frametyp gibt es die D-Frames, welche beim Vor- oder Rücklauf eines Videos als Vorschau angezeigt werden. Sie sind für den Bildaufbau bei normaler Wiedergabe nicht von Bedeutung und werden von keinem der anderen Frametypen referenziert oder benutzt. Es wird nur eine Farb- und Helligkeitsinformation pro 8x8 Pixelblock hinterlegt, wodurch die Qualität gering ist.

Die Abfolge mehrerer Frames von I-Frame zu nächstem I-Frame wird auch als Group of Pictures (GOP) bezeichnet. Ein Videostrom im MPEG Format besteht daher aus vielen GOPs. Die Größe der GOPs ist somit ein Indikator für die Videoqualität eines MPEG-1 Videos. Je kleiner die GOP, das heißt je mehr I-Frames, desto besser die Bildqualität, desto weniger Kompression kann aber auch nur erzielt werden. MPEG-1 benutzt GOPs meist in der Größe von 15-18 Frames (1 I-Frame und 14 bis 17 B- oder P-Frames), was also einer Dauer von etwa einer halben Sekunde pro GOP entspricht.

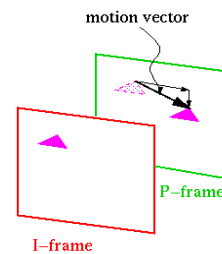


Fig.4: Bewegungsvektor

Um die Differenzinformationen in P- oder B- Frames zu vorherigen I-Frames zu kodieren werden die in Bezug stehenden Bilder in so genannte Makroblöcke von 8 x 8 Pixel zerlegt, die auf Unterschiede zwischen 2 Bildern überprüft werden und diese Unterschiede dann als Bewegungsvektoren in Referenz zum vorherigen Frame speichern.

Eine weitere Komprimierung der Daten wird dann durch diverse weitere Kompressionsalgorithmen durchgeführt. Diese werden in vielen Kompressionsanwendungen verwendet und daher hier nur kurz angesprochen.

Zunächst wird jeder Makroblock mittels vorwärts gerichteter Diskreter Kosinustransformation (FDCT) in einen Block von wiederum 8x8 Werten überführt, dessen Werte keine absoluten Farb- oder Helligkeitsinformationen, sondern eine Abweichung (Offset) zum Durchschnittswert des Blocks, enthalten. Die Genauigkeit dieser Werte wird anschließend mit der Quantisierung der Informationen auf gröbere Werte herabgesetzt. Weiterhin werden Lauflängenkodierung (RLE) zur vereinfachten Speicherung von gleichen, aufeinanderfolgenden Informationen sowie die Huffman-Kodierung zur vereinfachten Speicherung von Wiederholungen und Zyklen in den Bitfolgen eingesetzt. Als Ergebnis dieser Schritte erhält man schließlich den bekannten MPEG-Bitstrom.

Der MPEG-1 Standard beschreibt ebenfalls Standards zur Audiokodierung. Es gibt drei in der Komplexität ihrer Kodierung ansteigende Audiolayer. Für das Audiosignal sind Abtastraten von 32kHz, 44,1kHz und 48kHz vorgesehen. Ziel ist auch hier die Kompression der Audiodaten. Dabei werden analog zur Videokompression möglichst viele vom Menschen nicht hörbare Informationen eliminiert.

Der Audio Layer 1 ist eine vereinfachte Variante des Audio Layer 2 und wurde von Philips in der digital compact cassette, die Nachfolger der klassischen analogen Audiokassette (compact cassette) werden sollte, verwendet.

Der Audio Layer 2 – auch bekannt unter dem Namen MP2 oder Musicam – war lange Zeit Standardformat im Radiowesen, da er sich mit einer Kompression von 256kbit/s gut über die vorhandene ISDN-Infrastruktur in Europa verbreiten lies.

Das beispielsweise mit 48kHz abgetastete Audiosignal wird hierbei in 32 Teilbänder aufgeteilt, um in diesen jeweils Maximalwerte der Lautstärke zu bestimmen und kurz vor- oder nachher auftretende leise Töne zu löschen. Das menschliche Ohr bemerkt leise Töne kurz vor oder nach lauten Pegeln psychoakustisch bedingt nicht, wodurch sich mit dem Löschen dieser Töne Speicherplatz sparen lässt.

Der bekannte Audio Layer 3 – die MP3 – besitzt den höchsten Komplexitätsgrad der Kodierung und ist damit das effektivste Verfahren der drei Layer. Besonders durch Online-Tauschbörsen und den Einsatz auf tragbaren Flashspeichern ist dieses Format heute de-facto Standard bei komprimierten Audiodateien.

3.2.2 MPEG-2

MPEG-2 wurde 1994 als Nachfolger von MPEG-1 eingeführt. Er wurde durch die DVD und das digitale Fernsehen (DVB) einer der wichtigsten Videostandards.

Die wesentlichen Videokodierungsmerkmale in MPEG-2 wurden aus MPEG-1 beibehalten, wodurch er auch kompatibel zu MPEG-1 ist. Jedoch zielt der Standard auf höhere Bildqualität und damit auf höhere Datenraten von bis zu 15 Mbit/s, die jetzt durch die DVD als Speichermedium übermittelt werden konnten.

MPEG-1 hatte trotz des Erfolges einige bekannte Schwächen, die nun behoben werden sollten. So konnten nur progressive Bilder (Vollbilder) dargestellt werden, eine Unterstützung des Zeilensprungverfahrens fehlte. MPEG-2 wurde aber im Gegensatz zu seinem Vorgänger mit Hauptaugenmerk auf das digitale Fernsehen entwickelt und unterstützte das Zeilensprungverfahren von Anfang an.

Weiterhin wurde MPEG-2 sehr skalierbar und flexibel gestaltet, wodurch der Standard eine große Bandbreite von Anwendungen von mobilen Anwendungen bis hin zu HDTV abdecken kann. Eine einzige Spezifikation der Bildeigenschaften für alle Anwendungen wie in MPEG-1 schien nicht sinnvoll, daher wurden in MPEG-2 mehrere Profile und Levels definiert. In den Profilen werden kodiertechnische Eigenschaften wie Bildkodierung oder Subsampling festgelegt, in den Levels darstellungstechnische Eigenschaften wie Bildwiederholungsfrequenzen, Auflösungen oder Übertragungsraten.

Ein Anwendungsbereich, der den MPEG-2 Standard nutzt, kann nun entscheiden, welches Profil und Level er verwendet. Beispielsweise arbeitet die DVD im Main Profile und Main Level.

| Abk. | Name | Bilder | Chroma |
|------|-----------------|---------|--------|
| SP | Simple Profile | I, P | 4:2:0 |
| MP | Main Profile | I, P, B | 4:2:0 |
| 422P | 4:2:2 Profile | I, P, B | 4:2:2 |
| SNR | SNR Profile | I, P, B | 4:2:2 |
| SP | Spatial Profile | I, P, B | 4:2:2 |
| HP | High Profile | I, P, B | 4:2:2 |

Table 1. MPEG-2 Profile

| Abk. | Name | Format | Mbit/s |
|------|------------|-----------|--------|
| LL | Low Level | 352x288 | 4 |
| ML | Main Level | 720x576 | 15 |
| H14 | High 1440 | 1440x1152 | 60 |
| HL | High Level | 1920x1080 | 80 |

Table 2. MPEG-2 Levels

MPEG-1 Audio bot maximal zwei Audiokanäle (Stereo). In MPEG-2 wurde auch der Mehrkanalton standardisiert. Dabei wurde zunächst der MPEG-1 Layer 2 zum MPEG-2 BC (Backward compatible) – auch MPEG-2 Multichannel genannt – weiterentwickelt. Dieses Format sollte in Europa auf DVDs als Ersatz für *Dolby Digital*⁴, welches ursprünglich nicht für Europa vorgesehen war, eingesetzt werden. Jedoch funktionierten die Kodierung des Mehrkanaltons und die Rückwärtskompatibilität zum Stereoton nicht zuverlässig, wodurch sich der Standard bei den Herstellern nicht durchsetzen konnte. Als dann kurze Zeit später Dolby Digital auch in Europa eingesetzt werden durfte, verlor MPEG-2 Multichannel endgültig an Bedeutung.

Als Weiterentwicklung des MPEG-2 BC entstand MPEG-2 NBC (non backward compatible), welcher auch unter dem Namen AAC (advanced audio coding) bekannt ist. Hierbei wurden die Schwächen in der Kodierung von MPEG-2 BC und MPEG-1 Layer 3 (MP3) verringert, so das bei gleicher Qualität höhere Kompressionsraten erzielt werden konnten. Obwohl AAC eine Weiterentwicklung von MPEG-2

⁴ Dolby Digital: Mehrkanal-Tonsystem der Fa. Dolby

Multichannel ist, hat es sich als Konkurrenzformat zu MP3 im Stereobereich etabliert und soll nach Wunsch der MPEG dort längerfristig MP3 als Standard ablösen.

Durch Onlineshops und Musikverwaltungsprogramme wie Apples iTunes sowie in Audiostreams im Internet findet es bereits heute große Anwendung.

Obwohl MPEG-2 ein sehr erfolgreicher Standard war und ist, löste er MPEG-1 nicht vollständig ab. MPEG-2 wurde nie für niedrige Bandbreiten entwickelt, wodurch dort MPEG-1 effektiver und im Vorteil ist.

3.2.3 MPEG-3

MPEG-3 sollte als Standard für hochauflösende Signale wie HDTV eingesetzt werden. Es wurde jedoch deutlich, dass die Neuerungen im bereits vorhandenen MPEG-2 eingepflegt werden konnten und daher kein neuer Standard von Nöten war. Somit wurde MPEG-3 nie veröffentlicht.

3.2.4 MPEG-4

Das ursprüngliche Ziel vom 1998 veröffentlichten Standard MPEG-4 war, auf Systemen mit niedrigen Bandbreiten und Ressourcen wie Mobiltelefonen, PDAs und vor allem dem immer mehr aufkommenden Internet (Videostreaming und Übertragung) eine möglichst hohe Qualität zu bieten. Die ITU hatte mit dem H.263 Standard bereits viele der gewünschten Punkte umgesetzt, so dass der H.263 Standard direkt in den MPEG-4 Standard mit einfluss. MPEG-4 nimmt viele Funktionen und Techniken aus den beiden vorherigen MPEG Standards auf und fügt weiterhin neue Funktionen wie das Digital Rights Management oder die Unterstützung für 3D-Rendering mit auf.

Der größte Unterschied zwischen MPEG-4 und seinen Vorgängern ist jedoch seine Objektorientierung. Es werden nicht mehr nur Audio- und Videodaten, sondern auch weitere Inhalte wie Animationen, computergenerierte Objekte oder Szenenbeschreibungen mit kodiert. Man spricht auch von der objektorientierten Videokodierung. Jeder Teil einer Multimediaszene wird als Medienobjekt erfasst, dessen Eigenschaften wie Aussehen oder Bewegungen in Attributen deklariert werden. Dadurch können verschiedene Objekten im Bild differenziert behandelt und auch komprimiert werden – so beispielsweise ein Hintergrund höher als eine im Vordergrund agierende Person.

Der MPEG-4 Standard besteht aus vielen einzelnen Unterpunkten, die teilweise heute noch weiterentwickelt werden. So wurde unter anderem auch das in MPEG-2 standardisierte AAC weiterentwickelt und spezifiziert. Ein weiterer wichtiger Teil – besonders für die Zukunft des hochauflösenden Fernsehens – ist der MPEG-3 Part 10.

3.2.5 MPEG-4 AVC

Der MPEG-4 AVC (Advanced Video Coding) Standard ist kein eigenständiger Standard, sondern in MPEG-3 Part 10 enthalten. Er wurde zusammen mit der ITU entwickelt und wird deshalb auch als H.264 bezeichnet. Ziel war es, eine deutlich effizientere Kodierung gegenüber MPEG-2 zu ermöglichen und die Kodierung auch für hochauflösende Videos einsetzen zu können.

MPEG-4 AVC erreicht heute eine etwa dreimal höhere Kompression als MPEG-2, das heißt, bei gleicher Qualität ist die Datenmenge um zwei Drittel geringer. Jedoch benötigen die effizienteren Kodierverfahren auch das Doppelte bis Dreifache an Zeit. Als Anwendungsgebiet findet sich der Videostandard heute auf Blu-ray Discs sowie beim hochauflösenden digitalen Fernsehen über Satellit (DVB-S2) wieder. Hier wird der bis dato Standard MPEG-2 immer weiter abgelöst. Bei der Kodierung von Blu-ray Discs spielt der höhere Rechenaufwand eine geringere Rolle, weiterhin arbeiten die Decoder in Abspielgeräten oder auch Satellitenreceivern heute deutlich schneller, so dass das Mehr an Rechenaufwand kompensiert werden kann.

Die bessere Komprimierung wird durch viele kleine Verbesserungen gegenüber den vorherigen MPEG Standards erzielt. Unter anderem wird nun auch in den speicherintensiven I-Frames eine Form der Bildvorhersage eingesetzt, indem Makroblöcke aus umliegenden Blöcken abgeschätzt und nur die Differenzen abgespeichert werden. P- und B- Frames können im AVC mehrere vorliegende und nachfolgende Frames referenzieren, wodurch Bewegungszyklen erkannt und besser kodiert werden können. Referenzierte Frames können beliebig gewichtet in andere Frames eingehen, wodurch sich Überblendungen sehr effizient kodieren lassen.

Viele der Änderungen oder Neuerungen in MPEG-4 AVC waren bereits in den vorherigen Standards realisierbar, wurden aber durch den erhöhten Rechenaufwand nicht umgesetzt. Durch die Steigerung der Rechenleistung der De- und Enkodierer ist dies nun möglich und auch effektiv einsetzbar.

3.2.6 Weitere MPEG Standards

Die MPEG hat noch weitere Standards verabschiedet, die sich jedoch nicht direkt auf die Video- und Audiokompression beziehen. So wurde MPEG-7 im Jahre 2002 verabschiedet, um multimediale Inhalte mit Hilfe von Metainformationen zu beschreiben. Ein weiterer Standard – MPEG-21 – bietet einen standardisierten Rahmen für Erzeugung, Produktion, Freigabe und Handel mit multimedialen Inhalten. Er stellt diese Inhalte als Prozess (Workflow) dar, in den alle Beteiligten eingebunden werden, um standardisierte Zusammenarbeit zu ermöglichen.

Bei beiden Standards handelt es sich also nicht um Audio- oder Videokodierstandards.

3.3 Zusammenfassung MPEG

Die MPEG hat in den letzten 20 Jahren durch die veröffentlichten Standards maßgeblich zum einheitlichen Nutzen von Video- und Audiodaten weltweit

beitragen. War die Nutzung von MPEG-1 noch auf die Videodarstellung am Computer oder auf Video-CDs beschränkt, so wurde das skalierbarere MPEG-2 durch seine Anwendung im digitalen Fernsehen und auf der DVD allgegenwärtig.

Beide Standards spielen jedoch durch ihre hohen Datenraten im Internet eine geringere Rolle. Selbst die Bandbreite von MPEG-1 mit 1,5 Mbit/s kann im Internet erst seit wenigen Jahren übertragen werden. Hier greift MPEG-4 mit seiner Skalierbarkeit auf eine Breite von Anwendungen ein. Ziel der MPEG war es, mit MPEG-4 die vielen proprietären Formate im Netz abzulösen. Dies ist, wie sich im Folgenden zeigen wird, nur teilweise gelungen, da es im Internet viele etablierte Standards und Formate, die nicht von der MPEG spezifiziert wurden, gibt. Mit MPEG-4 AVC ist jedoch ein Standard entstanden, der gerade durch das Aufkommen von HDTV und Blu-ray Discs in dem letzten Jahren eine große Bedeutung für die Zukunft spielen wird.

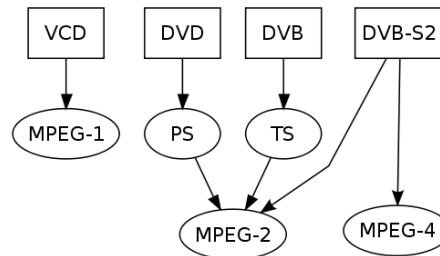


Fig. 5. Anwendung der MPEG Standards

3.4 Codecs, Container und Videokompression

Bislang wurden ausschließlich Kompressions- und Kodierstandards der MPEG betrachtet. Die MPEG spezifiziert jedoch auch Containerformate, welche Daten verschiedener Codecs und damit Kodiervorgängen aufnehmen können.

Als Codec bezeichnet man Verfahren oder auch Software, die Daten codieren und decodieren. Beim Erstellen der Videocodecs orientieren sich die meisten Unternehmen an den MPEG (bzw. ITU) Kompressionsstandards. So beruht beispielsweise der bekannte DivX Codec der Firma DivX Inc. auf MPEG-4.

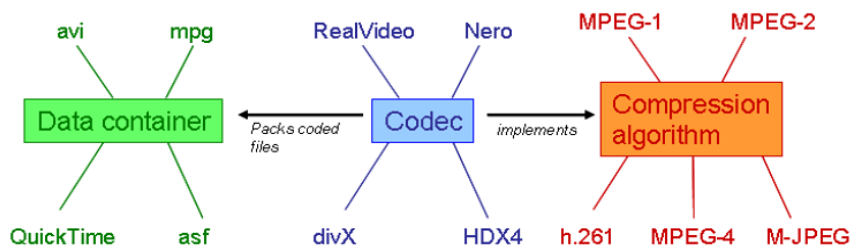


Fig. 6. Beziehung: Container – Codec – Kompression

Neben den auf MPEG beruhenden Codecs gibt es auch eigene, proprietäre Entwicklungen, wie beispielsweise Realvideo oder das Microsoft Media Video. Meist

weisen diese vom Aufbau Ähnlichkeiten zu den MPEG-Standards auf, sind jedoch oft nicht kompatibel.

3.5 Containerformate

Die Videokodierstandards der MPEG besitzen eigene Standardcontainer mit Dateiformaten wie beispielsweise *.mpeg*, *.mpg* oder *.mp4*. Es gibt jedoch auch Container, die unterschiedliche Datenformate und Codecs sowie weitere Daten wie Untertitel oder Menüstrukturen beinhalten können. Im Folgenden werden einige bekannte und wichtige Containerformate vorgestellt.

3.5.1 3GGP (.3gp)

Beim *3rd Generation Partnership Project* handelt es sich um ein Standardisierungsgremium für Mobilfunk, das unter anderem auch das für Mobiltelefone optimierte *.3gp* - Format spezifiziert hat, welches einen MPEG-4 basierten Videocodec verwendet.

3.5.2 Apple Quicktime (.mov)

Apple Quicktime wurde zunächst als Containerformat für den *Mac* entworfen, ist heute aber plattformübergreifend verfügbar. Es beruht grundsätzlich auf MPEG-4, kann aber auch viele andere Videocodecs einbinden und findet neben dem Internet besonders in der Apple Produktfamilie (*Mac*, *iPhone*, *iPod*) Anwendung. Es handelt sich nicht nur um ein reines Audio und Videoformat, sondern es kann auch multimediale Inhalte wie Animationen, Hyperlinks oder 3D Objekte enthalten, was der objektorientierten Videokodierung in MPEG-4 entspricht.

3.5.3 Audio Video Interleave (.avi)

Das AVI Format wurde von Microsoft 1992 eingeführt und beruht auf der *Video for Windows* Kodierschnittstelle. AVI speichert Audio- und Videodaten in einander verzahnt ab. Es kann verschiedene Codecs enthalten und ist durch die Bindung an Windows heute sehr weit verbreitet, obwohl es oftmals nicht das effizienteste Format ist.

3.5.4 Videoobject (.vob)

Beim *Videoobject* – Format handelt es sich um das Containerformat für DVD-Videos. Als Kompressionsalgorithmus kommt wie bereits erwähnt hauptsächlich MPEG-2 zum Einsatz.

3.5.5 Advanced Streaming Format (.asf)

ASF wurde von Microsoft für Videostreaming im Internet entwickelt. Es verwendet meist Windows Media Audio (*wma*) und Windows Media Video (*wmv*) als Inhalte, es kann aber auch beliebige andere Videocodes enthalten.

4 Zusammenfassung

Ziel dieser Arbeit war es, einen Überblick über die digitale Videokodierung zu bieten. Dazu wurden zunächst die Grundlagen digitaler Videos am Beispiel des digitalen Fernsehens erläutert. Es stellte sich heraus, dass bei digitalen Videos enorme Datenmengen auftreten, die über die vorhandenen Wege nicht übertragen werden können. Daher wurde die Videokompression an Hand der wichtigsten Videostandards – den MPEG Standards – dargestellt.

Die Anwendung und Verbreitung der auf MPEG-Kodierung beruhenden Videoformate zeigt die Bedeutung und die Wichtigkeit der Standardisierung. Diese gelang den Bereichen Digitales Fernsehen sowie DVD. Bemerkenswert ist die Tatsache, dass MPEG-2 hier seit über 15 Jahren eingesetzt wird und bestehen bleibt, obwohl sich in der schnelllebigen Computerwelt, deren Rechenkapazitäten sich in dieser Zeit exorbitant vergrößert haben, viele Dinge meist nur wenige Jahre halten können.

Im Internet hat sich jedoch eine Vielzahl an Kodieralgorithmen und Containerformaten standardisierter und proprietärer Art etabliert. Das Ziel der MPEG, hier mit dem Standard MPEG-4 ebenfalls eine Vereinheitlichung zu schaffen, ist – wie gezeigt wurde – noch nicht gelungen, und die Zukunft wird zeigen, ob dies vergleichbar zum MPEG-2 Standard gelingen kann.

Jedoch bleibt auch zu bemerken, dass Standardisierung nicht gleichzeitig Kostenfreiheit bedeutet. Viele Teile der MPEG Standards wurden von Unternehmen und Instituten entwickelt (*mp3*) und sind heute noch zur Codierung (vor allem im professionellen Einsatz) lizenzpflichtig.

Quellen

1. Adobe, A Digital Video Primer,
<http://bscw.uni-koblenz.de/bscw/bscw.cgi/d1315421/VideoCoding02.pdf>
2. Optibase, A Guide to MPEG-4,
<http://bscw.uni-koblenz.de/bscw/bscw.cgi/d1315417/VideoCoding01.pdf>
3. Djordje Mitrovic, Video Compression,
<http://bscw.uni-koblenz.de/bscw/bscw.cgi/d1315425/VideoCoding03.pdf>
4. Digital Video for the next Millenium,
<http://bscw.uni-koblenz.de/bscw/bscw.cgi/d1315429/VideoCoding04.pdf>
5. Oliver R. Ahlemann, Methoden der digitalen Audiobearbeitung
<http://www.fh-wedel.de/~si/seminare/ss02/Ausarbeitung/9.digitalaudio/audio2.htm>
6. Fernsehnormen – Standards und Technik
<http://www.paradiso-design.net/videostandards.html>
7. MPEG Website
<http://www.chiariglione.org/mpeg/>

Abbildungen

- Fig.1. YCbCr Bildinformationen,
http://de.academic.ru/pictures/dewiki/66/Barns_grand_tetons_YCbCr_separation.jpg
- Fig.2. YCbCr Subsampling,
http://de.wikipedia.org/wiki/Chroma_Subsampling
- Fig.3. Frametypen in MPEG-1,
http://www.chip.de/ii/11176688_f004c7bfab.jpg
- Fig.4. Bewegungsvektor
<http://www.cmlab.csie.ntu.edu.tw/cml/dsp/training/coding/mpeg1/motionvector.gif>
- Fig.5. Anwendung der MPEG Standards
<http://upload.wikimedia.org/wikipedia/commons/thumb/5/5b/Mpeg.svg/500px-Mpeg.svg.png>
- Fig.6. Beziehung Container – Codec – Kompression,
<http://bscw.uni-koblenz.de/bscw/bscw.cgi/d1315425/VideoCoding03.pdf>