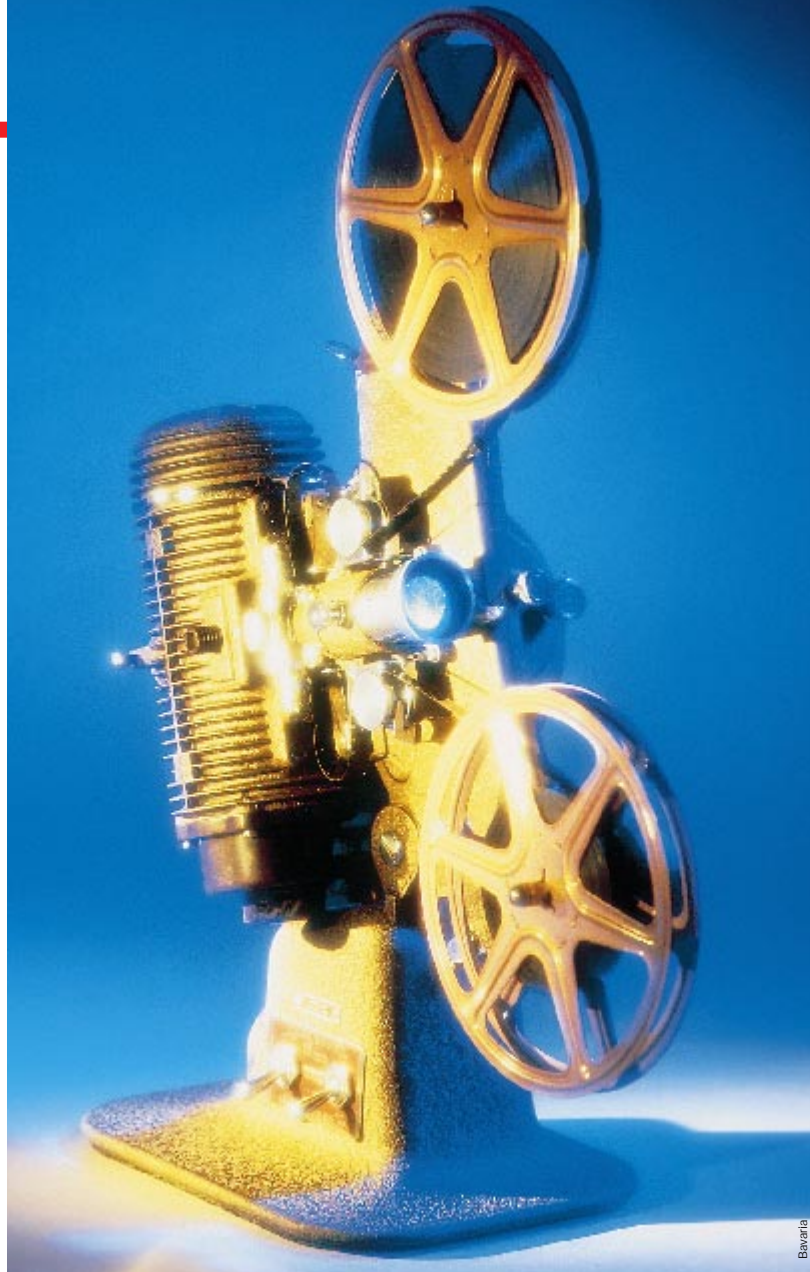


Wenn Sie Ihren PC mit Videomaterial füttern wollen, müssen die bewegten Bilder komprimiert werden. Raffinierte Reduktionsverfahren bändigen die Datenflut. CHIP erläutert Ihnen das Grundprinzip.



So funktioniert MPEG

Seit 1895 die französischen Brüder Lumière ihren ersten Film dem staunenden Publikum vorführten, hat das Bewegtbild nichts an Faszination eingebüßt. Aber erst in jüngster Zeit geht der Wunsch, Videos zur Nachbearbeitung am PC zu digitalisieren, relativ preisgünstig in Erfüllung. Auch vorgefertigte Videos auf CD-ROM (Video-CD, CD-I) lassen sich mit geringem Aufwand auf dem heimischen PC abspielen. Ermöglicht werden solche Anwendungen durch Komprimierungsverfahren nach dem MPEG-Standard.

Die Digitalisierung eines Videos nach der europäischen Fernsehnorm PAL erzeugt den jede Festplatte überwältigenden Datenstrom von rund 23 Megabyte in der Sekunde oder 80 Gigabyte pro Stunde. Damit digitales Video nicht Großrechnern vorbehalten bleibt, muß deshalb der Datenstrom mittels Kompression deutlich verringert werden.

Verlustbehaftete Kompressionsverfahren wie JPEG sind beim Packen deutlich effizienter als verlustlose Verfahren (beispielsweise GIF), aber naturgemäß auf einige Anwendungsgebiete (Bilder, Videos,

Audio) beschränkt. Sie machen es sich zunutze, daß die menschlichen Sinnesorgane nur einen Teil der aufgenommenen Informationen an das Gehirn weitergeben. Der Rest der ursprünglichen Bild- und Audioinformation ist schlicht überflüssig und wird durch den Kompressionsalgorithmus „weggerechnet“. So reagiert bei der Bildbetrachtung das Auge auf Farbveränderungen schwächer als auf Helligkeitsveränderungen. Infolge dessen lassen sich die Farbinformationen ohne sichtbare Einbußen reduzieren, solange die Helligkeitsverteilung erhalten bleibt.

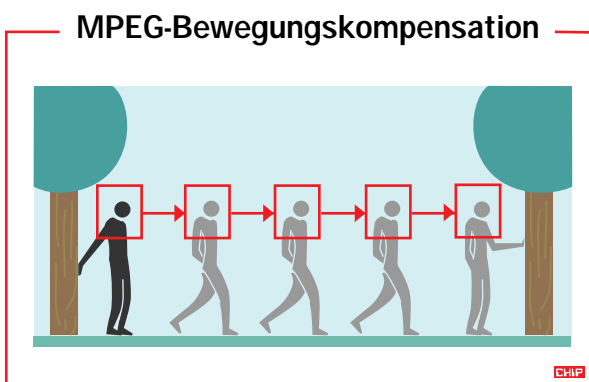
Bei der Kompression von Musik macht man sich die physiologische Tatsache zunutze, daß das menschliche Ohr leise Töne in lauten Passagen nicht mehr wahrnimmt. Gleiches gilt für zwei Töne, deren Frequenzen fast gleich sind. Der Mensch hört nur den lauterer der beiden Töne.

Es gibt eine beachtliche Zahl auf verschiedene Anwendungsbereiche zugeschnittene Kodierungslösungen. MPEG zählt in der Gruppe der Audio- und Video-Komprimierer zu den zukunftssträftigsten Verfahren.

○ Die Anfänge von MPEG

Die „Motion Picture Expert Group“ (MPEG) begann Ende der 80er Jahre mit der Entwicklung eines Standards für die digitale Speicherung von Bewegtbildern. Als erste Norm wurde von diesem Standardisierungsgremium MPEG-1 festgelegt. Die Vorläufer dieses Verfahrens sind M-JPEG (Motion - Joint Photographic Expert Group) und die H.261-Empfehlung des CCITT (Comité Consultatif International Télégraphique et Téléphonique).

Das relativ simple Motion-JPEG-Verfahren setzt im JPEG-Format komprimierte Einzelbilder aneinander. Die Kompression der Einzelbilder hat den Vorteil, daß M-JPEG-Sequenzen mit geeigneten Programmen problemlos nachbearbeitet werden können. Von Nachteil



Vektorisiert:
Ein Objekt wird in 16×16 Blöcke geteilt. Ändert das Objekt von einem Frame zum nächsten seine Position, wird nur noch sein Bewegungsvektor berechnet.

ist die geringe Kompressionsrate. M-JPEG ist nie ganz vereinheitlicht worden, und es existiert kein Integrationsstandard für Audioinformationen in M-JPEG-Files.

Aufgrund der Erkenntnisse, die mit M-JPEG und H.261 gesammelt wurden, entwickelte das MPEG-Gremium die Normen MPEG-1 und MPEG-2. MPEG ist im Vergleich zu seinen Vorläufern flexibel. Eine Anpassung an verschiedene Anwendungen bereitet deshalb keine größeren Probleme.

Die Kodierung von MPEG-Videos ist sehr rechenintensiv. Herkömmliche PC sind für Kodierung in Echtzeit völlig ungeeignet. Sogenannte Software-Codecs, also Programme, die MPEG-Files aus AVI-Videos produzieren, benötigen für die Umsetzung eines einminütigen Videos 30 Minuten und mehr.

Die MPEG-Dekodierung ist glücklicherweise nicht so aufwendig. Es gibt preiswerte Zusatzkarten für den PC (zum Beispiel die Motion von Elsa) und manche Grafikkartenhersteller bieten MPEG-Dekoder integriert oder als Aufsteck-Option an. Auf einem Pentium-Rechner läßt sich unter Zuhilfenahme eines Software-Dekoders (beispielsweise des Xing-Player) mit ein paar Geschwindigkeitseinbußen MPEG wiedergeben (dazu auch CHIP 2/96).

○ Farb- und Auflösungsreduktion

Üblicherweise liegt Video-Ausgangsmaterial mit der Farbauflösung der 4:2:2-Form vor. Diese Angabe betrifft das Abtastverhältnis der Videosignalkomponenten Y (Helligkeit) sowie Cr und Cb (Farb-

MPEG - Normen

MPEG-1

Video-CD, CD-I, Video on Demand, CD-Karaoke und interaktive Videospiele werden nach der MPEG-1-Norm kodiert. Die Übertragungsrate liegt bei den dort üblichen Auflösungen bei 192 Kilobyte pro Sekunde. Da in der Mehrzahl der Fälle die MPEG-Daten auf einer CD gespeichert sind, wurde der Datenstrom für Single-Speed-CD-ROM-Laufwerke optimiert. Außer der CD-ROM eignen sich auch lokale Netzwerke und ISDN als Medium für MPEG-1.

MPEG-2

Die MPEG-2-Norm wurde für das digitale Fernsehen definiert. MPEG-2 kommt beim Satellitenfernsehen sowie bei Video-on-Demand via Breitbandkabel zum Einsatz. Die der Kompression zugrunde liegenden Verfahren sind für MPEG-1 und -2 im Prinzip identisch. Beide Verfahren erzielen in der Regel identische Datenverdichtungen.

Folgende Funktionen werden von der MPEG-2-Norm zusätzlich unterstützt:

- Wahlfreier Zugriff innerhalb des Videos
- Wiedergabe rückwärts
- Schnelles Vor- und Rückwärts-suchen
- Audiovisuelle Synchronisation
- Fehler toleranz
- Toleranz bei Kapazitätsengpässen
- Editierbarkeit
- Variable Bildgröße und Bildwiederhol-frequenz
- 16:9-HDTV-Modus
- Interlace-Unterstützung (Halbbild-Modi)
- Surround-Sound

MPEG-3

Dieser Standard war für die Komprimierung von HDTV-Filmen mit bis zu 1920×1080 Pixeln entworfen worden. Es stellte sich aber schnell heraus, daß für diese Belange der

MPEG-2-Standard genauso gut geeignet ist. Darum wurde die Entwicklung an MPEG-3 eingestellt.

MPEG-4

Die Arbeit an MPEG-4 – Beginn 1993 – ist noch nicht abgeschlossen, das Standardisierungsgremium rechnet mit Ende 1998. MPEG-4 soll für Videos mit geringen Bitraten (4,8 bis 74 Kilobit pro Sekunde) eingesetzt werden. Geeignet ist diese Norm dann für Bildtelefone einschließlich mobiler Bildtelefonie, Spiele, Multimedia Electronic Mail und Multimedia Videotext. Angestrebt werden zehn Bilder pro Sekunde in guter Qualität einschließlich Ton. Die Auflösung wird voraussichtlich 176×144 Bildpunkte betragen. Die dafür benötigten extrem hohen Kompressionsraten sind mit herkömmlichen Mitteln nicht mehr erreichbar. Deshalb sind neue Verfahren wie fraktale und morphologische Kompression im Gespräch.

FACHBEGRIFFE

DCT (Diskrete Cosinus-Transformation) ist ein mathematisches Verfahren, zum Beispiel für die Zerlegung eines Videobildes. MPEG und JPEG benutzen die zweidimensionale, orthonormale 8×8 -DCT. Als Resultat entsteht eine Matrix von Transformations-Koeffizienten, die als Ausgangsprodukt für weitere Transformationen Verwendung findet.

Interpolation ist eine mathematische Methode, um zusätzliche Werte in eine numerische Sequenz einzufügen. Die zusätzlichen Werte werden unter Zuhilfenahme der Differenzenrechnung durch Polynome beschrieben.

Die Interpolation wird bei der Veränderung von Bilddaten, beispielsweise bei der Änderung der Bildgröße, angewendet. So werden bei einer Bildgrößenverdoppelung neue Pixel aus den nebenliegenden Punkten kalkuliert, was vom Betrachter als deutliche Qualitätsverbesserung des Bildes empfunden wird.

Huffman-Kodierung ist ein mathematisch-statistisches Verfahren für die Kodierung numerischer Werte. Dabei werden statistisch häufig auftretenden Werten kurze Codes zugeordnet (zum Beispiel 011), selten auftretende Werte erhalten dagegen lange Codes (zum Beispiel 100110110).

Das hat – abhängig vom Ausgangsmaterial – eine deutliche Reduktion der Datenmenge ohne Verlust zur Folge. Die Zuordnung der Daten zu den Codes erfolgt nach einer Tabelle.

MPEG benutzt eine Variante der Huffman-Kodierung. Auch Faxgeräte der Gruppe 3 nutzen die Huffman-Kodierung.

differenzen). Zur Datenreduktion erfolgt in sogenannten Dezimierungsfiltren eine Auflösungsreduzierung: Die horizontale Auflösung wird durch die gewichtete Mittelung eines Pixels mit seinen später nicht mehr verwendeten Nachbarpixeln um die Hälfte verringert. In vertikaler Richtung erfolgt eine ähnliche Filterung, oder es wird einfach jede zweite Zeile weggelassen. So entsteht ein 4:2:0-Abtastungsmuster.

Bei der Darstellung komprimierten Videomaterials muß umgekehrt die anfängliche Auflösung wiederhergestellt werden. Dazu wird eine gewichtete Mittelung (Interpolation) durchgeführt. Der Effekt ist, daß zum Beispiel aus der Wertefolge 4-8-12 eine Folge 4-6-8-10-12 erzeugt wird. Die Gewichte sind die Filterkoeffizienten des sogenannten Interpolationsfilters.

Die Bewegungskompensation

In Videosequenzen ähneln sich aufeinanderfolgende Einzelbilder relativ stark. Selbst bei sehr bewegten Szenen, beispielsweise einem Autorennen, sind die Unterschiede zwischen den einzelnen Bildern (Frames) nicht groß. Diese Tatsache macht sich das MPEG-Verfahren zunutze. Die Bewegung eines Gegenstandes oder einer Person zwischen zwei Bildern wird zu einem Vektor (Motion Vector) vereinfacht.

Die Erkennung eines komplexen Gegenstandes ist in der Praxis jedoch zu

MPEG-Bildtypen


Zur Realisierung der Motion Compensation verwendet MPEG drei Bildtypen: I- (Intra), P- (Predictive-coded) und B- (Bi-directional predictiv-coded) Bilder.

Ein I-Bild speichert die komprimierte Information eines einzigen Vollbilds. Es ist mit den Frames von Motion-JPEG vergleichbar. Intra-Bilder werden etwa alle 0,5 Sekunden in den Datenstrom eingelagert. Das ermöglicht das Vor- und Zurückspringen auf dem komprimierten Videomaterial. I-Bilder dienen als Ausgangsbilder für die Ermittlung von P- und B-Bildern.

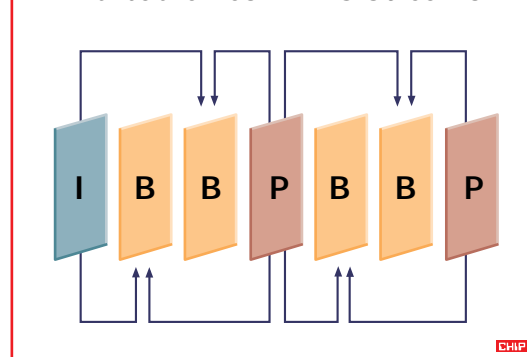
P-Bilder werden durch Bewegungsabschätzung (Motion Estimation) aus zeitlich zurückliegenden Bildern vom Typ I oder P abgeleitet. P- wie I-Bilder dienen auch als Referenzbilder zur Ermittlung der Verschiebungsvektoren.

Ein B-Bild wird aus einem früheren und einem späteren – daher bidirektional – P- oder I-Bild abgeleitet oder auf Basis der Nachbarbilder durch Interpolation berechnet.

Bei der Wiedergabe eines MPEG-Videos werden die originalen Frames wiederhergestellt. Wenn als Beispiel die Frames in der Reihenfolge ...IBBPBBP... angeordnet sind, erfolgt die Entschlüsselung folgendermaßen: Zuerst wird das I-Bild dekodiert. Das ist einfach, da I-Bilder die volle Information enthalten. Dann erfolgt die Dekodierung des ersten P-Frames unter Zuhilfenahme des I-Frames. Beide werden im Speicher gehalten. Danach werden die beiden B-Bilder aus dem I- und dem P-Bild rekonstruiert.

Die Kompressionsmethode für die 16×16 -Blöcke in den I- und P-Bildern basiert auf der Huffman-kodierten, quantisierten DCT mit 8×8 Pixel großen Einzelblöcken (vier 8×8 -Blöcke ergeben einen 16×16 -Block). **Jan Kleinert** 

Aufbau eines MPEG-Streams



Abstammung: I-Bilder enthalten ein komplettes Ausgangsbild. P-Bilder werden aus früheren, B-Bilder aus früheren und späteren I- oder P-Bildern errechnet.

aufwendig, weshalb 16×16 Pixel große Blöcke (Macroblock) zusammengefaßt werden. Für jeden Block berechnet ein Algorithmus die örtliche Verschiebung von einem Bild zum anderen. Weiter wird die inhaltliche Differenz zwischen Block 1 und Block 2 festgestellt. Dabei ergibt sich ein sogenanntes Fehlerbild, das gespeichert wird. Der Aufwand zur Speicherung von Bewegungsvektor und Fehlerbild ist um so kleiner, je geringer die Unterschiede zwischen zwei Frames sind.

Das geschilderte Verfahren bezeichnet man als blockbasierte Bewegungskompensation (Motion Compensation). Diese aufwendige Berechnung wird durch MPEG-Encoder ausgeführt.



Internet: Mehr Details zu MPEG in englischer Sprache unter <http://www.vol.it/MPEG> und <http://www.crs4.it/HTML/LUIGI/MPEG/mpegfaq.html>

Literatur: CHIP 2/96, Filme unter Druck, Seiten 108 und 109

Software: Xing-Player <http://www.xingtech.com>