



MultiMediaPraktikum

Universität Tübingen Wilhelm-Schickard-Institut für Informatik



Wintersemester 2002/ 2003
Aufgabe 2 - MPEG-Video



Marc-Oliver Pahl
Ulrike Schaal
David Eißler

Inhaltsverzeichnis

MPEG-1	4
Farbformate	4
RGB	4
YUV	4
YIQ	5
Subsampling	6
4:4:4	6
4:2:2	6
4:2:0	6
Dekodierung des MPEG-1-Datenstroms	7
Unterschiede zwischen MPEG-1 und MPEG-2?	9
Welche Vor- und Nachteile besitzt MPEG-2 im Vergleich zu anderen Formaten wie MJPEG („Motion JPEG“) und H.261?	9
MPEG-2 unterstützt mehrere Ebenen („Levels“) und Profile. Welche sind das und wo werden sie eingesetzt?	9
MPEG2	8
CodeComments	10
main (avi2mpg1.c):	10
avi2m1v.c:	10
putseq.c:	10
Praxis	11
Praxis	12
Praxis	13

MPEG-1 Farbformate

RGB

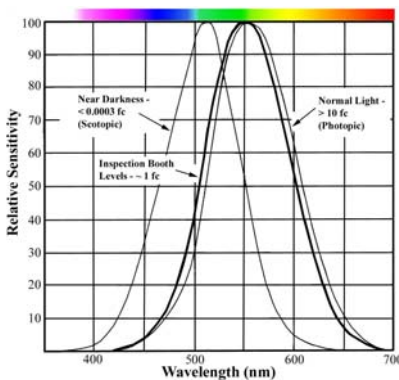
Beim RGB-Farbformat werden die Farben als drei-Komponenten-Vektor gespeichert; und zwar entsprechend ihren Anteilen an den Grundfarben [R]ot, [G]rün, [B]lau (additive Farbmischung).

Das Format wird vor allem bei Bildern auf dem Computer verwendet, da dort das Bild eben gerade aus den drei Grundfarben zusammengesetzt wird.

YUV

Beim YUV-Format wird die Farbe als Luminanz (=Helligkeit) [Y] und Chrominanz (=Farbdifferenzsignal) [UV] gespeichert. Bei der Gewichtung der Farbanteile für das Luminanzsignal hat man die Empfindlichkeitsfunktion des Auges zugrunde gelegt:

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$



Grün wird vom Auge am intensivsten wahrgenommen, weshalb es auch am Meisten zur Luminanz (auch für Schwarzweißdarstellung) beiträgt. Im U- und V-Teil des Vektors wird der Unterschied von Blau bzw. Rot zur Luminanz gespeichert.

Die Norm wurde im Zusammenhang mit der Fernsehnorm PAL (auch bei SECAM) entwickelt und hat den Vorteil, dass auftretende Störsignale dem Betrachter nur wenig auffallen, weil man sie eben gar nicht so stark wahrnehmen kann. Aus demselben Grund sind sich U- und V-Kanal auch sehr gut für das Subsampling geeignet.

Ein großer Vorteil dieses Formats für das Fernsehen ist auch, dass es abwärtskompatibel zu Schwarzweißfernsehern ist. Diese stellen nämlich einfach das Y-Signal (die Luminanz) dar.

Aus RGB errechnet es sich folgendermaßen:

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$

$$U = B - Y$$

$$V = R - Y$$

YIQ

Dieses Format wurde mit der NTSC-Fernsehnorm (USA/ Japan) entwickelt und hat eigentlich dieselben Zielsetzungen, wie das YUV-Format.

Entsprechend ist das Y-Signal auch gleich. Nur I und Q entsprechen nicht U und V sondern sind zu den U- und V-Vektoren im RGB-System um 33° gedreht:

$$I = 0.877(R-Y) \cos(33^\circ) - 0.492(B-Y) \sin(33^\circ)$$

$$Q = 0.877(R-Y) \sin(33^\circ) - 0.492(B-Y) \cos(33^\circ)$$

Dadurch sind die Farben dann in den I- und Q-Kanälen allerdings „vermischt“, was dazu führt, dass Übertragungsfehler (durch Subsampling verstärkt) zu Farbverschiebungen (also nicht nur Farbe stärker oder schwächer sondern andere Farbe!) führen.

Aus RGB errechnet es sich folgendermaßen:

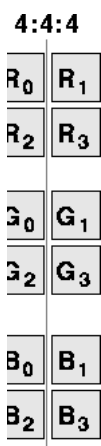
$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.30 & 0.59 & 0.11 \\ 0.60 & -0.28 & -0.32 \\ 0.21 & -0.52 & 0.31 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Subsampling

Wie zuvor erwähnt, nimmt das Auge Helligkeitsänderungen stärker wahr, als Farbschwankungen. Das macht man sich beim Subsampling zunutze:

Farbinformation wird ungenauer gespeichert als Helligkeitsinformation.

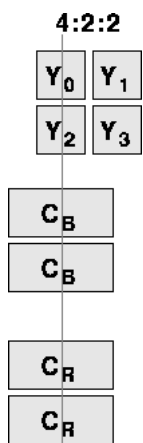
Subsampling bedeutet „Unterabtastung“. D.h. von unseren drei Komponenten des Farbvektors (s.o.) werden nicht pro Pixel jeweils alle drei sondern zum Beispiel pro 2x2-nur ein U- und V-Wert übertragen. Dieser ist dann für alle vier Pixel gleich.



4:4:4

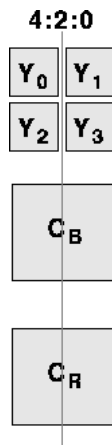
„kein Subsampling“

Bei diesem Subsamplingformat werden für vier Pixel vier Y-, vier U- und vier V-Werte gespeichert.



4:2:2

- horizontales Subsampling um Faktor 2
- je zwei aufeinanderfolgende Farbinformationen werden zusammengefasst
- jede Zeile enthält doppelt so viele Y-Samples wie Cb oder Cr Samples
- Einsparung von 33% Bandbreite

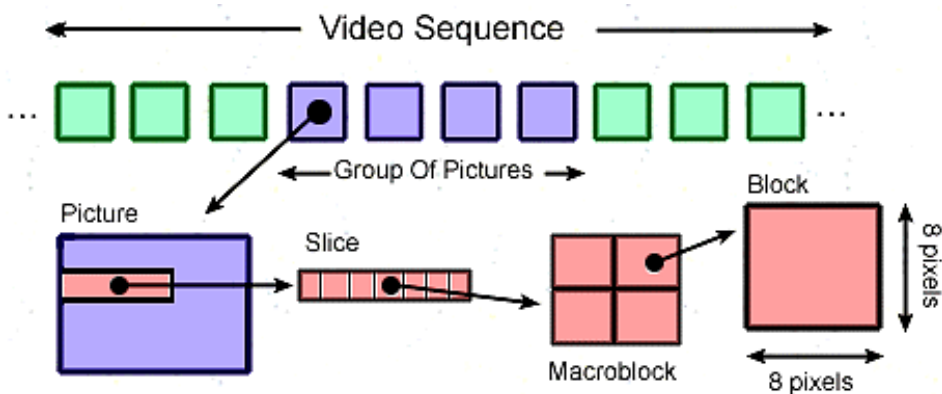


4:2:0

- horizontale und vertikale Zusammenfassung von Farbinformationen, jeweils um Faktor 2
- je vier Farbinformationen werden zusammengefasst
- auf 4 Pixel Luminanzinformation kommt je ein Pixel Farbinformation
- Einsparung von 50% Bandbreite

Der MPEG-1 Datenstrom setzt sich aus Layern zusammen. Diese werden beim Entpacken auseinander genommen und dann dekodiert.

Dekodierung des MPEG-1-Datenstroms



Sequence-Layer...

dient zur Steuerung des Zwischenspeichers bei der Dekodierung.

Group of Pictures-Layer...

beinhalten die einzelnen kodierten Bilder und beginnt mit einem I-Frame (damit die Dekodierung starten kann, muss sie bei einem Vollbild beginnen). Die einzelnen Bilder innerhalb der GOP (ca. 0,5 sec) lassen sich nur aus der gesamten GOP rekonstruieren (B-/ P-Frames).

Picture-Layer...

enthält ein Einzelbild (P-/ B-/ I-Frame).

Slice-Layer...

enthält mehrere Makroblöcke des kodierten Bildes (die Größe ist variabel). Im Slice-Layer wird die DCT-Qualität skaliert.

Macroblock-Layer...

beinhaltet einen Makroblock.

Block-Layer...

entspricht einem Block

Frametypen:

Im MPEG-Datenstrom gibt es drei verschiedene Arten, wie Frames kodiert werden: I-, P- und B-Frames.

Die I-(Intra coded)-Frames enthalten am meisten Bildinformationen. In ihnen ist das gesamte Bild kodiert.

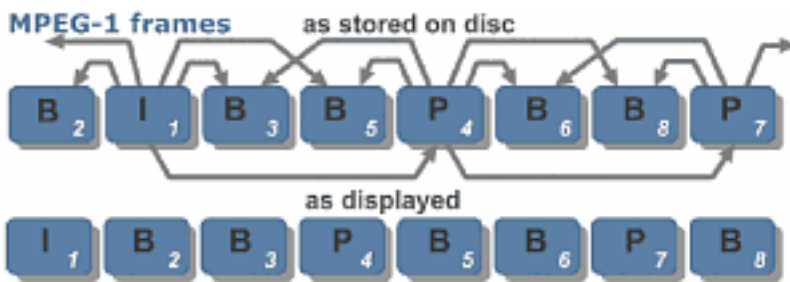
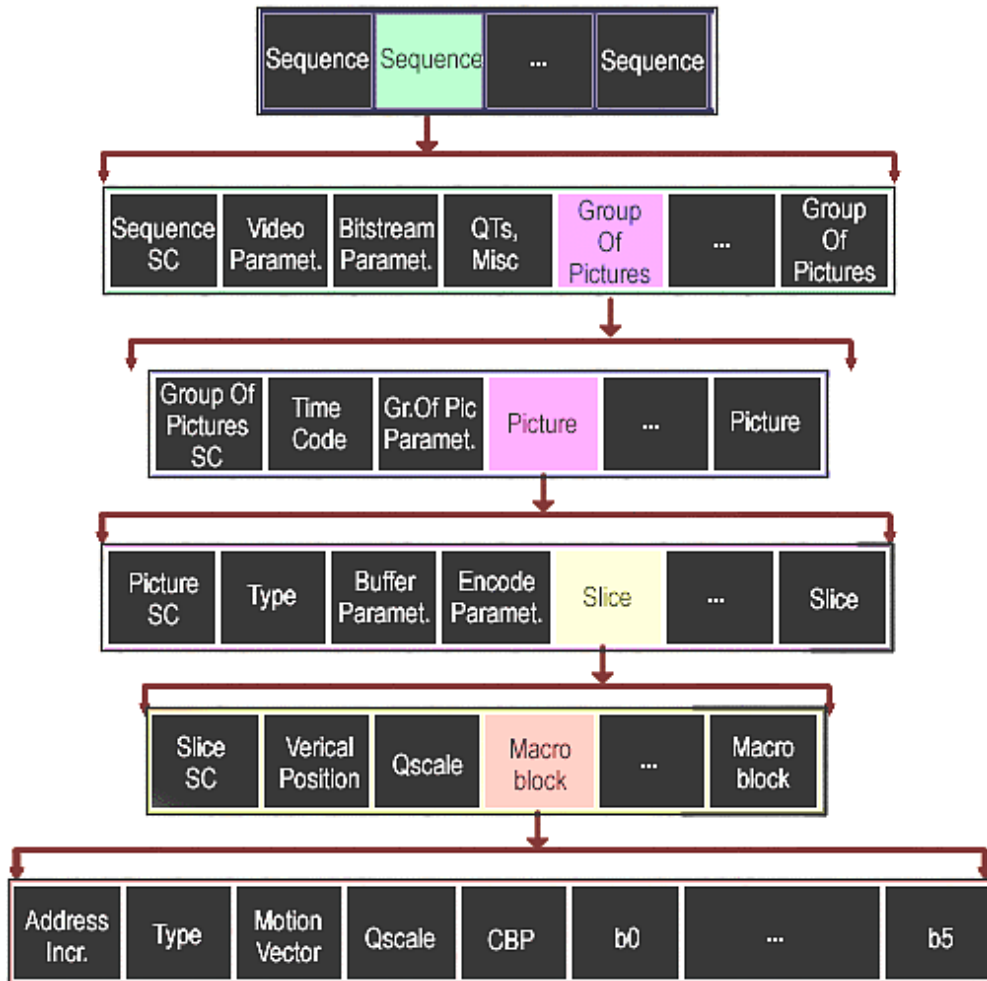
I-Frames

P-(Predicted)-Frames enthalten nur Differenzinformationen zu einem vorhergesagten Wert und sind daher ohne vorigen I-Frame (bzw. in der Folge dann auch P-Frame) nutzlos.

P-Frames

B-(Bidirectionally predicted)-Frames enthalten sogar nur Differenzinformationen zu dem aus dem vorherigen und dem nachfolgenden I-Frame (bzw. P-Frame) vorhergesagten Werten und brauchen daher beide.

B-Frames



Beim Dekodieren werden die Bilder zuerst in der Reihenfolge, in der sie im Stream ankommen (oben) dekodiert und anschließend in die richtige Reihenfolge zum Anzeigen (unten) gebracht.

MPEG2

Unterschiede zwischen MPEG-1 und MPEG-2?

Fest definierte Sequenzen von Bildarten, die eine minimale Ende-zu-Ende-Verzögerung bei gegebener Datenrate ermöglichen.

MPEG-2 unterstützt niedrigere Audio-Abtastraten.

MPEG-2 unterstützt interlaced video.

MPEG-2 unterstützt höhere Bandbreite.

MPEG-2 unterstützt höhere Bildauflösung.

MPEG-2 unterstützt Skalierung.

Welche Vor- und Nachteile besitzt MPEG-2 im Vergleich zu anderen Formaten wie MJPEG („Motion JPEG“) und H.261? **MPEG2**

MPEG-2 unterstützt mehrere Ebenen („Levels“) und Profile. Welche sind das und wo werden sie eingesetzt?

Profile:

- simple profile: nicht skalierbar, keine B-Frames, 4:2:0
- main profile: nicht skalierbar, B-Frames, 4:2:0
- SNR-skalierbar: SNR-skalierbar, B-Frames, 4:2:0
- räumlich skalierbar: SNR- und räumlich skalierbar, B-Frames, 4:2:0
- high profile: SNR- und räumlich skalierbar, B-Frames, 4:2:0, 4:2:2

Skalierung:

Die Bilder werden in verschiedenen Qualitätsstufen kodiert. Dabei werden unterschieden:

- a) räumliche Skalierung (unterschiedliche Bildgrößen)
- b) Skalierung der (Bildwiederhol-)Rate
- c) Amplitudenskalierung (Auflösung der Quantisierung der DCT-Koeffizienten)

Levels:

- high level
- high-1440-level
- main-level
- low level

Vor allem das main-Level findet seine Anwendung in der Übertragung digitaler Videos über Kabel, Satellit und andere Broadcast-Kanäle 2-80MB/ s, digitale Speicherung und andere Kommunikationsanwendungen.

main und high profile werden auch für hochqualitatives Fernsehen verwendet.

Profile	Namen der Profile	Simple Profile	Main Profile	SNR-skalierbares Profil	Räumlich skalierbares Profil	High Profile	
	Eigenschaften der Profile	keine B-Frames	B-Frames				
		4:2:0					4:2:0 oder 4:2:2
Nicht skalierbar		SNR-skalierbar	SNR- oder räumlich skalierbar				
Level-Ebenen	High Level 1920 Pixel/Zeile 1152 Zeilen		≤ 80 MBit/s			≤ 100 MBit/s	
	High-1440 Level 1440 Pixel/Zeile 1152 Zeilen		≤ 60 MBit/s		≤ 60 MBit/s	≤ 80 MBit/s	
	Main Level 720 Pixel/Zeile 572 Zeilen	≤ 15 MBit/s	≤ 15 MBit/s	≤ 15 MBit/s		≤ 20 MBit/s	
	Low Level 352 Pixel/Zeile 288 Zeilen		≤ 4 MBit/s	≤ 4 MBit/s			