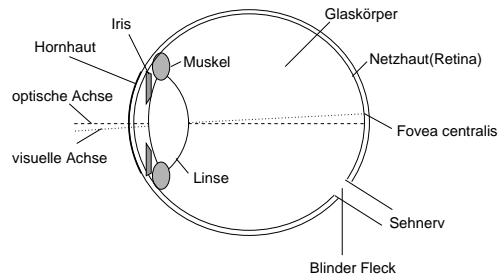


Netzhaut und Sehnerven

Die visuelle Wahrnehmung beim Menschen wird in erster Linie durch den Aufbau des Auges bestimmt. Das Licht fällt durch Pupille und Linse auf die Netzhaut (Retina), wo die einfallende Lichtenergie in Nervenreize umgesetzt wird. Die Pupille fungiert dabei wie eine Blende, die ihren Durchmesser entsprechend der einfallenden Lichtmenge verändert (Adaption). Die Linse selbst ist durch eine spezielle Muskulatur verformbar und so in der Lage, eine scharfe Abbildung auf die Netzhaut zu ermöglichen (Akkommodation).



In der Retina befinden sich zwei verschiedene Typen von Rezeptoren: Die **Zapfen**, sie sind sowohl helligkeits- als auch farbempfindlich und befähigen uns zur Farb- und Detailwahrnehmung. Ihre Zahl beträgt etwa 6 Millionen. Sie sind im zentralen Bereich der Retina (Makula) angesiedelt und arbeiten bei normaler Tageshelligkeit. Die **Stäbchen** dagegen können zwar keine Farben unterscheiden, dafür sind sie aber wesentlich helligkeitsempfindlicher und werden bei geringer Beleuchtung aktiv. Ihre Zahl beträgt pro Auge etwa 120 Millionen, wobei im Zentrum der Retina (Fovea) selbst keine Stäbchen vorkommen.

Das vom menschlichen Auge wahrnehmbare Lichtspektrum reicht von einer Wellenlänge von 780 nm (rot) bis 390 nm (violett). Entsprechend ihrer unterschiedlichen Empfindlichkeitsmaxima unterscheidet man drei verschiedene Typen von Zapfen, die über unterschiedlich lichtempfindliche Sehpigmente verfügen. Typ A besitzt ein Maximum bei Grün (langwelliges Pigment, 558 nm), Typ B bei Gelb-Rot (mittelwelliges Pigment, 531 nm) und Typ C bei blau-violett (kurzwelliges Pigment, 419 nm). Das Gehirn kombiniert die Reize der drei unterschiedlichen Zapfentypen zu neuen Farben.

Die Retina mit ihren unterschiedlichen Sehzellen weist zudem zwei markante Regionen auf: die **Fovea centralis**, die Stelle mit der größten Sehschärfe, etwa in der Mitte der Netzhaut, aber nicht exakt auf der optischen Achse. An dieser Stelle befinden sich ausschließlich Zapfen, ihre Dichte ist an dieser Stelle am höchsten. Dazu kommt der sogenannte **blinde Fleck**, die Stelle, an der der Sehnerv mit ca. 1,5 Millionen Nervenfasern das Auge verläßt, so daß das dort eintreffende Licht nicht wahrgenommen werden kann. Das Verhältnis zwischen den optischen Rezeptoren und den Nervenfasern beträgt etwa 80:1, d.h. die Vorverarbeitung der visuellen Information erfolgt hier bereits mit einem hohen Komprimierungsfaktor.

Das menschliche Auge ist in der Lage bis zu 1.000 Helligkeitswerte zu unterscheiden. Bei der Grafikdarstellung im Computer wird dieser Bereich infolge nicht-linearer Grauwertskalierung reduziert, so daß in der Regel 8 Bit pro Bildpunkt für die Darstellung von Schwarz-Weiß-Bildern ausreicht. Eine Ausnahme macht hier die medizinische Bild-datenverarbeitung, die auf eine höhere Auflösung angewiesen ist.

Abb. 3.14. Optische Wahrnehmung beim Menschen

automatischen Bildauswertungsverfahren. Da der Informationsgehalt des ursprünglichen Bildes nicht weiter reduziert wird, ist die Komprimierungsrate hier nicht so hoch.

Exkurs 3: JPEG – Komprimierungsprozeß

Der JPEG-Komprimierungsprozeß verläuft in vier Stufen:

- **Farbkonversion und Sampling**

Im ersten Teil der JPEG-Komprimierung wird jede Farbkomponente (rot, grün, blau) anteilig zerlegt in die Bestandteile **Luminanz** (Helligkeit) und **Chrominanz** (Farbinformation). JPEG erlaubt einen höheren Verlust an Information im Bereich der Chrominanz, da das menschliche Auge hier weniger sensibel ist. Zwar tragen alle Farbkomponenten im RGB-Modell Helligkeitsinformation, doch grün beeinflusst die Helligkeitswahrnehmung stärker als blau oder rot. Mit einem in der Norm **CCIR/ITU 601** festgelegten Verfahren wird die RGB-Information in die Komponenten Y (entspricht der Helligkeit), Cb (entspricht der Blaukomponente) und Cr (entspricht der Rotkomponente) transformiert. Bei diesem YCbCr-Farbmodell handelt es sich um eine spezielle Variante des YUV-Farbmodells (vgl. Abb. 3.8).

$$\begin{pmatrix} +0.299 & +0.587 & +0.114 \\ +0.1687 & -0.3313 & +0.5 \\ +0.5 & -0.4187 & -0.081 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix}$$

Im Bildzerlegungs-Prozeß (**Sampling**) werden die Anteile Cb und Cr mit einer geringeren Auflösung bestimmt als die Y-Komponente. Eine typische Samplingrate spricht der Y-Komponente z.B. eine vierfach höhere Auflösung zu als den beiden anderen Komponenten (4:1:1). So können etwa vier (r,g,b)-Pixel, die ursprünglich einen Speicherplatz von 4·3 Byte = 12 Byte belegen auf 4 (Y) + 1 (Cb) + 1 (Cr) = 6 Byte reduziert werden. Je niedriger die Samplingauflösung, desto höher die dadurch bereits erzielte Komprimierung und desto kürzer die Gesamtkomprimierungszeit.

- **Diskrete Cosinus Transformation (DCT)**

Mit Hilfe der DCT werden Intensitätsdaten in Frequenzdaten transformiert, mit denen eine Aussage darüber möglich ist, wie schnell die Intensität von Farbe und Helligkeitsinformation im Bild variiert. In der JPEG-Kodierung wird das Bild in Bildblöcke von 8×8 Pixeln unterteilt. Im weiteren Verlauf wird jede Bildkomponente (Y,Cb,Cb) separat abgearbeitet, wobei die einzelnen Komponenten eine unterschiedliche Anzahl von Bildblöcken besitzen können (z.B. aufgrund der unterschiedlichen Samplingraten). Die Datenpunkte in einem 8×8 Pixel großen Block starten links oben (0,0) und enden rechts unten (7,7). Der Bildpunkt (x,y) besitzt die Bildinformation f(x,y). Die DCT erzeugt einen neuen 8×8-Block (u,v) über die Transformation

$$F(u, v) = \frac{1}{4} C(u) C(v) \left[\sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \right]$$

mit

$$C(z) = \begin{cases} \frac{1}{\sqrt{2}} & \text{wenn } z = 0 \\ 1 & \text{sonst.} \end{cases}$$

Das Ergebnis dieser Transformation ist ähnlich wie bei einer Fouriertransformation eine Frequenzraummatrix, die die Frequenz- und Amplitudenverteilung für den betrachteten Raum angibt, d.h. die 64 Funktionswerte der ursprünglichen, von den beiden räumlichen

Dimensionen x und y abhängigen Matrix werden in ihr Spektrum transformiert, indem die DCT einen Basiswechsel auf eine Basis von 64 orthogonalen, diskreten Signalen durchführt. Große, regelmäßig gefärbte Flächen im Bild schlagen sich daher als niedrige Frequenzanteile nieder, feine Details in den hohen Frequenzanteilen. Dabei enthält $F(0,0)$ den Durchschnittswert der 8×8 -Matrix, $F(1,0)$ gibt den Grad an, bei dem sich die Bildwerte nur langsam verändern (niedrige Frequenzen) und $F(7,7)$ den Grad, bei dem sich die Bildwerte in beide Richtungen sehr rasch verändern (hohe Frequenzen). In einem natürlichen Bild sind scharfe Linien und abrupte Farbwechsel eher selten. Somit konzentriert sich der größte Teil des Ausgangssignals auf die niedrigen Frequenzen, viele der einzelnen Koeffizienten sind deshalb sehr klein, d.h. nahezu gleich Null.

- **Quantisierung**

Als nächstes werden die Koeffizienten der 8×8 -Matrix quantisiert, d.h. die Werte $F(u,v)$ werden durch einen Wert $Q(u,v)$ aus einer durch das JPEG-Komitee vorgegebenen Quantisierungstabelle dividiert und auf die nächste ganze Zahl aufgerundet:

$$S(u,v) = \left\lfloor \frac{F(u,v)}{Q(u,v)} + 0.5 \right\rfloor.$$

Die Quantisierungstabelle benutzt eine feinere Quantisierung für Koeffizienten niedriger Frequenzen und eine gröbere für höhere Frequenzen. Hohe Frequenzen, die bereits einen Koeffizienten $F(u,v)$ nahe Null besitzen, werden durch das Verfahren zu Null quantisiert. Die vorgegebene Quantisierungstabelle kann durch eine eigene, die dann in der JPEG-Datei mit abgespeichert werden muß, ersetzt werden. Eine entsprechende Dekodierung erfolgt durch Multiplikation von $S(u,v)$ mit den entsprechenden Koeffizienten aus der Quantisierungstabelle $Q(u,v)$, wodurch die ursprünglichen Werte $F(u,v)$ allerdings nur annähernd und nicht exakt wiederhergestellt werden können. In der Quantisierung liegt der eigentliche, verlustbehaftete Charakter der JPEG-Kodierung begründet. Das Verfahren trägt den neurophysiologischen Eigenschaften des menschlichen Sehens Rechnung. Unschärfe Kanten werden weniger kritisch wahrgenommen als Fehler in der Helligkeitswahrnehmung.

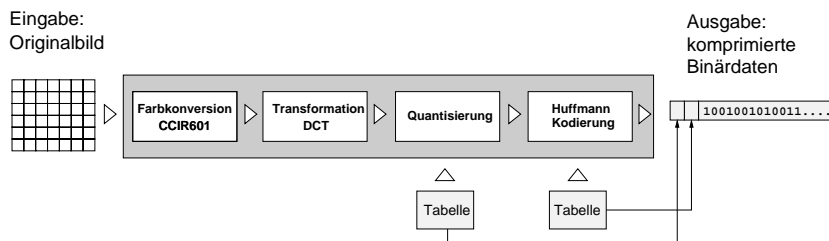


Abb. 3.15. JPEG Verarbeitungsschritte

- **Eigentliche Komprimierung**

Für die eigentliche Komprimierung werden die quantisierten Daten folgendermaßen vorbereitet: Die $F(0,0)$ -Koeffizienten der 8×8 -Matrizen werden jeweils als Differenzwert von einem Block zum nächsten Block gespeichert, da sie für aufeinanderfolgende Blöcke meist

sehr ähnlich sind. Die restlichen Blöcke von $F(1,0)$ bis $F(7,7)$ werden in einer Zick-Zack-Anordnung nacheinander geschrieben, d.h. die Koeffizienten niedriger Frequenzen kommen zuerst und die typischerweise zu Null quantisierten Anteile hoher Frequenz folgen am Ende, was eine effiziente Komprimierung ermöglicht. Diese Komprimierung erfolgt letztendlich durch eine modifizierte **Huffmann-Kodierung** bzw. über eine arithmetische Kodierung. Für die arithmetische Kodierung ist allerdings der Erwerb einer Lizenz notwendig, da dieses Verfahren von IBM, AT&T und Mitsubishi patentiert wurde. Für die hier verwendete Huffmann-Kodierung werden die Differenzen aufeinanderfolgender Koeffizienten als Ausgangswerte kodiert.

Weiterführende Literatur:

ISO/IEC IS 10918-1: Digital Compression and Coding of Continuous-tone Still Images, Part 1: Requirements and Guidelines.

ISO/IEC IS 10918-2: Digital Compression and Coding of Continuous-tone Still Images, Part 2: Compliance Testing.

W. B. Pennebaker, J. L. Mitchell: JPEG Still Image Data Compression Standard, van Nostrand Reinhold, New York NJ, USA, 1993.

Exkurs 4: JPEG – Dateiaufbau

Das Dateiformat für die Speicherung von JPEG-kodierten Bilddaten ist das JFIF-Format (**JPEG File Interchange Format**). JFIF-Dateien enthalten die eigentlichen Grafikdaten in der eben beschriebenen Form zusammen mit Informationen, die zum Entpacken dieser Daten notwendig sind. Die einzelnen Segmente der JFIF-Datei werden durch Markierungsblöcke (2 Byte, Inhalt 0xFF) voneinander getrennt (vgl. Abb. 3.16). Sie umfassen insbesondere:

- **Start of Image (Sol)**
Markierung (2 Byte, 0xFFD8) zur Kennzeichnung des Beginns der zum Bild gehörenden Information. Steht in der Regel am Beginn der Datei.
- **End of Image (Eol)**
Analog zur Sol Markierung gibt Eol das Ende des Bildes an (2 Byte, 0xFFD9). Eol steht daher im Normalfall am Ende der Datei und trägt keine weitere Information.
- **Application (APP0)**
Direkt nach dem Sol folgt ein Application Marker (2 Byte, 0xFFE0), gefolgt von den 4 Zeichen "JFIF". Desweiteren folgen eine Versionsnummer, Angaben über die Bildgröße und Pixeldichte. Optional kann ein eventuell vorhandenes Thumbnail-Bild mit zugehöriger Größeninformation folgen.
- **Quantization Table (QT)**
Die Quantisierungstabelle startet mit einem führenden (2 Byte, 0xFFDB). Es folgen die 64 Einträge der Quantisierungstabelle und Informationen bzgl. deren Präzision (8 oder 16 Bit Präzision).
- **Start of Frame i (SoFi)**
Markierung (2 Byte, 0xFFC*i*), die den eigentlichen Beginn von Bild *i* anzeigt. Danach folgen Informationen über Datengenauigkeit, Bilddimensionen, sowie Komponenten. Die Komponentendaten setzen sich aus einer Komponentenummer, horizontalem und vertikalem Sampling-Faktor und der Nummer der benutzten Quantisierungstabelle zusammen. Es können beliebig viele Komponenten hintereinander folgen.
- **Define Huffmann Table (DHT)**
DHT beschreibt eine Huffmantabelle und startet mit dem DHT-Marker (2 Byte, 0xFFC4). Zuerst kommt ein Index für die Huffmantabelle, danach Zählwerte und zuletzt die Inhalte. Die Zählwerte geben an, wieviele Codes mit einer bestimmten Anzahl von

Bits in der Tabelle vorkommen. Der i-te Zählwert gibt die Anzahl der Codeworte mit i Bit an.

- **Start of Scan (SoS)**
Nach der SOS Markierung (2 Byte, 0xFFDA) beginnt der Bitstrom, der die eigentlichen Bilddaten wiedergibt. Vorher werden noch die Anzahl der Komponenten, sowie die Nummern der zugehörigen Huffmanstabellen angegeben.

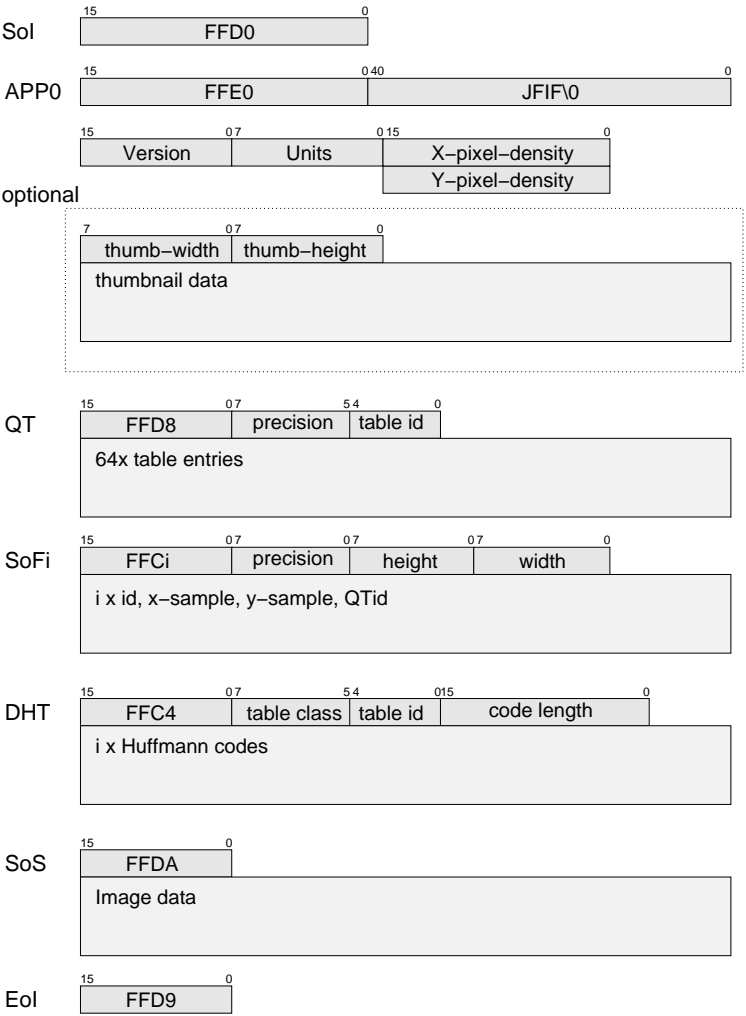


Abb. 3.16. JPEG Dateiformat – JPEG File Interchange Format

WWW

Kommunikation, Internetworking, Web-Technologien

Meinel, C.; Sack, H.

2004, XLII, 1178 S. In 2 Bänden, nicht einzeln erhältlich.,

Hardcover

ISBN: 978-3-540-44276-9