

Proseminar Datenkomprimierung
Dr. U. Tamm

JPEG - Kompression

WS 2002/03

Torsten Zichner

Inhaltsangabe:

1. Einleitung
2. JPEG – Kompression
 - 2.1. Konvertierung des Bildes in ein geeignetes Farbmodell
 - 2.2. DCT – Diskrete Cosinus Transformation
 - 2.3. Quantisierung
 - 2.4. Beispiel
 - 2.5. Kodierung der Koeffizienten
3. JPEG – Dekodierung

1. Einleitung

JPEG – Joint Photographic Expert Group (seit 1988 Zusammenarbeit zwischen CCITT - International Telegraph and Telephone Consultative Committee - und ISO/IEC - International Organization for Standardization/International Electrotechnical Commission) erarbeiteten einen Standard zur Kompression von digitalen Standbildern. 1993 wurde dieser unter der Bezeichnung ISO 10918 als Standard definiert.

JPEG verfolgte das Ziel, eine Bildkompression mit

- Akzeptabler Komplexität
- Unabhängigkeit von der Bildbeschaffenheit
- Vom Benutzer beeinflussbarer Bildqualität/Kompression

zu schaffen.

2. JPEG - Kompression

- ist eine verlustbehaftete Kompression. Dies bedeutet nicht zwangsläufig eine Qualitätsminderung. Es können auch nur für das Auge unsichtbare Informationen entfernt werden (hohe Frequenzen, Farbinformationen)
- Der Grundvorgang („Baseline“-Codec) bei dem JPEG-Kompressionsverfahren setzt sich aus folgenden Schritten zusammen:



- Konvertierung des Bildes in ein geeignetes Farbmodell (nicht im JPEG-Standard vorgeschrieben, wird aber im Regelfall durchgeführt)
- DCT – Diskrete Cosinus Transformation
- Quantisierung
- Kodierung der Koeffizienten

2.1. Konvertierung des Bildes in ein geeignetes Farbmodell

Der JPEG-Standard schreibt kein Farbmodell vor. Einige eignen sich jedoch mehr als andere. In der Regel YUV-Farbraum (technisch entstanden, Farbfernsehen/Übertragung). Helligkeit und Kontrast werden in einem einzigen Kanal (Luminanz) gespeichert. Die Farbwerte in zwei anderen Kanälen (Chrominanz).

$$Y := 0.299 * R + 0.587 * G + 0.114 * B$$
$$U := B - Y$$
$$V := R - Y$$

Das Auge kann Helligkeitsunterschiede weit besser wahrnehmen als Farbunterschiede. Die Farbkanäle können daher in ihrem Informationsgehalt verringert werden. 2 oder 4 benachbarte Pixel in den Farbkanälen können gemittelt, und zu einem einzigen zusammengefasst werden (4:2:2 oder 4:1:1 **Subsampling**). So wird die Datenmenge verringert, ohne für das Auge erkennbare Unterschiede zu erzeugen.

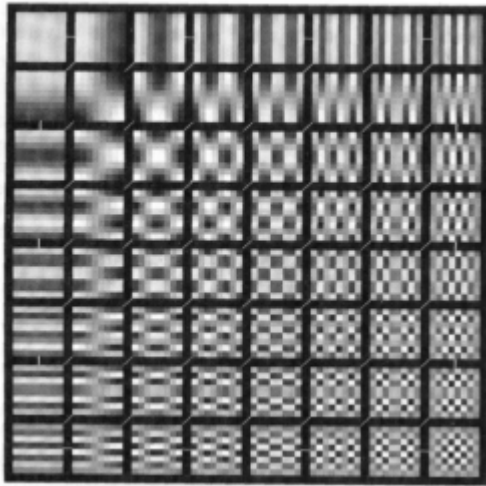
2.2. DCT – Diskrete Cosinus Transformation

- Spezielle Form der Fourier-Transformation
- DCT ist kein Kompressionsverfahren
- Sie wird in vielen Vorstufen verwendet (jpg, mp3)
- Bild wird in 8x8 Matrizen aufgeteilt (alle folgenden Schritte, arbeiten mit diesen Matrizen – es wird keine Gesamtbildbetrachtung vorgenommen)
- Frequenzanalyse der Helligkeitsschwankungen (trennt grobe und feine Strukturen eines Bildes)
- Hohe Frequenz – schneller Helligkeitswechsel; niedrige Frequenz – langsamer Helligkeitswechsel (jeweils von links nach rechts und oben nach unten)

$$F(u, v) = \frac{1}{4} C(u)C(v) \left[\sum_{x=0}^7 \sum_{y=0}^7 f(x, y) * \cos \frac{(2x+1) u \pi}{16} \cos \frac{(2y+1) v \pi}{16} \right] \begin{matrix} \text{mit} \\ x, y=0 \dots 7 \\ \text{und} \\ u, v=0 \dots 7 \end{matrix} \quad \begin{matrix} C(u), C(v) = 1/\sqrt{2} \text{ f\u00fcr } u, v = 0 \\ C(u), C(v) = 1, \text{ sonst} \end{matrix}$$

x, y bezeichnen die Koordinaten innerhalb des zu kodierenden Blockes; u, v sind die Koordinaten der Basisfunktionen der DCT

Hierdurch wird eine 8x8 Ortsmatrix, in eine 8x8 Frequenzmatrix umgewandelt. Der Informationsgehalt bleibt der gleiche (die Informationen werden nur in anderer Form dargestellt).



Betrachtet man die Basisfunktionen der 2-dimensionalen DCT, die eigentlich ein Produkt zweier eindimensionaler DCT ist, stellt man fest, dass sich in der linken oberen Ecke die einzige konstante Basisfunktion befindet. Der Koeffizient dieser Funktion, deren Frequenz in beiden Dimensionen Null ist, hei\u00dft DC-Koeffizient (direct-current terms), w\u00e4hrend die \u00fcbigen 63 Koeffizienten AC-Koeffizienten (alternating-current terms) hei\u00dfen. Bis zu diesem Zeitpunkt wurde noch keine Kompression der Daten erreicht, sie wurden lediglich f\u00fcr die anschließende Kompression konditioniert

2.3. Quantisierung

in den meisten Bildern sind viele Koeffizienten nahezu 0. Setzt man diese Koeffizienten zu 0, dann ist dies nach der Dekompression kaum bemerkbar! Weiterhin sind einige Koeffizienten wichtiger als andere!

- So ist z.B. die Grundhelligkeit und der Helligkeitsverlauf am wichtigsten (niedrige Frequenzen)
- Weniger wichtig ist die Textur des Bildes (mittlere Frequenzen)
- Sehr feine hochfrequente Details sind nahezu nicht beobachtbar

Die Koeffizienten werden also durch unterschiedlich starke Quantisierung „gleich wichtig“ gemacht. Die Quantisierung erfolgt mit einer Quantisierungsmatrix. Jeder Koeffizient der DCT-Matrix wird durch einen entsprechenden Wert (aus der Quantisierungsmatrix) geteilt. Es gibt keine vorgeschriebenen Standardtabellen f\u00fcr Quantisierungsmatrizen. Somit l\u00e4sst sich die Bildqualit\u00e4t \u00fcber die Quantisierung steuern. Dies hat zur Folge, dass die genutzte Quantisierungsmatrix im Bild mit transportiert werden muss, um es an anderer Stelle wieder dekodieren zu k\u00f6nnen. Dies erh\u00f6ht zwar den Speicherbedarf, aber erh\u00f6ht auch die Flexibilit\u00e4t des Standards und wird somit ohne weiteres geduldet.

2.4. Beispiel

Ein Beispiel f\u00fcr Ablauf von DCT und Quantisierung

95	88	87	95	88	95	95	95
143	144	151	151	153	170	183	181
153	151	162	166	162	151	126	117
143	144	133	130	143	153	159	175
123	112	116	130	143	147	162	189
133	151	162	166	170	188	166	128
160	168	166	159	135	101	93	98
154	155	153	144	126	106	118	133

Bildmatrix

98	95	91	89	90	95	101	106
140	143	149	156	163	167	168	167
146	149	154	159	159	151	137	126
149	142	136	137	145	156	163	166
119	117	118	125	140	157	170	176
137	147	160	170	172	166	157	150
166	167	164	152	132	112	99	93
151	153	150	139	125	118	119	123

rekonstruierte Bildmatrix

↓

↑

93	2	-8	-7	3	1	1	-2
-38	-58	11	17	-3	5	5	-3
-84	63	-1	-17	2	7	-4	-0
-51	-37	-10	13	-10	5	-1	-4
-85	-42	50	-8	18	-5	-1	1
-63	66	-13	-1	2	-6	-2	-2
-16	14	-37	18	-12	4	3	-3
-53	31	-7	-10	23	-0	2	2

DCT-Koeffizienten

→

31	0	-1	0	0	0	0	0
-7	-8	1	1	0	0	0	0
-12	7	0	-1	0	0	0	0
-5	-3	0	0	0	0	0	0
-7	-3	3	0	0	0	0	0
-4	4	0	0	0	0	0	0
-1	0	-1	0	0	0	0	0
-3	1	0	0	0	0	0	0

quantisierte DCT-Koeffizienten

3	5	7	9	11	13	15	17
5	7	9	11	13	15	17	19
7	9	11	13	15	17	19	21
9	11	13	15	17	19	21	23
11	13	15	17	19	21	23	25
13	15	17	19	21	23	25	27
15	17	19	21	23	25	27	29
17	19	21	23	25	27	29	31

Quantisierungsmatrix

2.5. Kodierung der Koeffizienten

Die DC-Koeffizienten (Matrixinhalt an Position 1,1) benachbarter Blöcke unterscheiden sich nur wenig und werden daher als Differenz zum Vorgängerblock übertragen.

150	80	20	4	1	0	0	0
92	75	18	3	1	0	0	0
26	19	13	2	1	0	0	0
3	2	2	1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

ZIG-ZAG-SCAN und Run-Length- Encoding:

Die Koeffizienten werden im Zick-Zack angeordnet, was gleichzeitig einer Anordnung nach Ihrer Wichtigkeit entsprechend der visuellen Wahrnehmung entspricht.

Jetzt macht man sich die längeren Nullketten, die dann von einem von Null verschiedenen AC-Koeffizienten gefolgt werden, zu Nutzen. Anstatt nämlich jeden Koeffizienten einzeln zu speichern, wird jeder AC-Koeffizient, der nicht Null ist, in Kombination mit der Anzahl der Nullen, die ihm vorausgegangen sind, angegeben (Run-Length-Encoding). Alle Nullen, denen kein AC-Koeffizient mit Wert ungleich Null folgt, werden weggelassen. Das „End of ...“ Symbol wird dann bereits vor dem 64. Wert gesetzt.

Die Entropie-Kodierung ist der letzte Komprimierungsschritt. Die durch RLE zusammengefassten Ketten werden meistens mittels Huffman kodiert.

Die Quantisierungstabellen und die Tabellen für die Huffman-Codierung sind nicht fest vom JPEG-Standard vorgegeben, was einerseits eine Flexibilität bei der Kompression bedeutet, andererseits aber auch, dass diese Tabellen immer mit dem komprimierten Bild mitgeschickt werden müssen.

Dies ist die generelle Funktionsweise des JPEG-Algorithmus, der auf der DCT basiert. Da der Standard jedoch nicht allzu genau festgelegt ist, bleibt viel Raum für verschiedene Implementierungen.

DCT-basierte Methoden

Sequentielles Kodieren:

Bei dieser auf DCT-basierenden Methode werden die 8x8 Blöcke von links nach rechts und Reihe für Reihe von oben nach unten gelesen. So baut sich dann auch das Bild beim Dekodieren auf.



Progressives Kodieren:

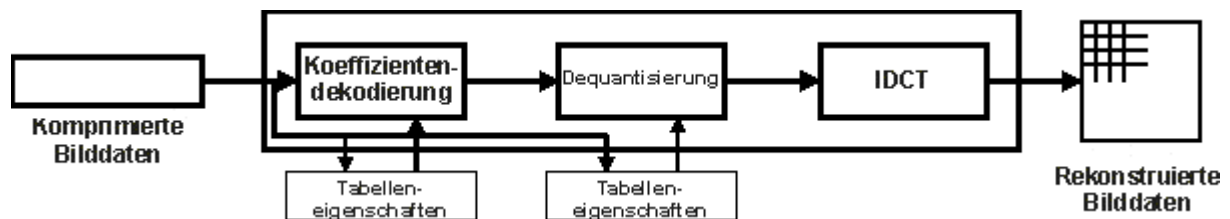
Diese Methode arbeitet die Blöcke in der gleichen Reihenfolge wie oben ab, wenn auch in mehreren Durchläufen. Um dies zu realisieren, muss zwischen Quantisierung und Kodierung noch ein Speicher (memory buffer) eingebaut werden.



Die Koeffizienten jedes transformierten und quantisierten Blocks werden dort gespeichert und in den verschiedenen Durchläufen, je nach Wichtigkeit kodiert. So kann das Bild beim Dekodieren in schlechter Qualität aufgebaut werden. Diese Qualität verbessert sich dann in mehreren Schritten bis zum gewünschten Ergebnis.

JPEG - Dekodierung

Bei der Dekodierung werden alle Schritte des Kodierens in invertierter Reihenfolge rückgängig gemacht.



Da diese quantifizierten DCT-Koeffizienten nun aber nur gerundete Werte sind, erhält man auch nach Multiplikation mit den entsprechenden Einträgen der mitgelieferten Quantisierungstabelle nur gerundete DCT-Koeffizienten (Verlust an Information).

Nach Anwendung der inversen Diskreten-Cosinus-Transformation kann das Bild jetzt rekonstruiert werden.

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