



Komprimieren

Das JPEG-Format ist der Standard für digitale Bilder: Nicht nur jede Bildsoftware, sondern auch zahlreiche andere Programmtypen bis zur Software von Handys können JPEGs öffnen. Uwe Artmann erläutert die Funktionsweise.

Wenn es um Dateiformate geht, hört und liest man oft Sätze wie „Das Bild wurde als JPEG gespeichert.“ Diese Aussage ist nur bedingt richtig, da mit dem Begriff JPEG streng genommen kein Dateiformat beschrieben wird, sondern nur Arten der Komprimierung. Mit dieser Methode komprimierte Bilddaten können mittels verschiedener Dateiformate auf die Speicherkarte oder Festplatte geschrieben

werden – auch als .jpg. JPEG ist eine Abkürzung für „Joint Picture Expert Group“, ein vornehmlich japanischer Zusammenschluss von Experten der Bildindustrie, die verschiedene Methoden der Bildkompression in einem Standard niedergeschrieben haben. Abgesehen von Spezialfällen z. B. in der medizinischen Bildgebung handelt es sich eigentlich immer um eine verlustbehaftete Kompression von 8-Bit-Farbbil-

dern. Verlustbehaftet bedeutet in diesem Zusammenhang, dass die Reduzierung der Datenmenge mit einem Verlust an Bildinformation einhergeht. Dieser Verlust lässt sich steuern, wodurch der Benutzer selber entscheiden kann, ob die Datei möglichst klein werden soll oder ob die maximale Bildinformation erhalten bleiben soll.

Standardverfahren

Je stärker ein Bild komprimiert wird, desto mehr Bildinformationen gehen verloren. Das Ziel muss es daher sein, dass möglichst die Informationen eliminiert werden, die der Betrachter sowieso nicht, oder zumindest nur sehr schlecht wahrnehmen

kann. Bei der verlustbehafteten JPEG-Kompression bedient man sich zweier Eigenschaften des menschlichen, visuellen Systems. Zum einen ist für die Erfassung einer Szene und die Verarbeitung durch unser Sehzentrum die Helligkeitsverteilung viel wichtiger als die Farbinformation. Zum anderen können hohe Ortsfrequenzen, also feine Strukturen, schlechter erfasst werden als niedrige.

Das aus drei Farbkanälen (rot, grün, blau) bestehende Bild wird in seinen Helligkeitsanteil und zwei Farbkomponenten zerlegt. Da die Farbkomponenten für das Erfassen des Bildes durch den Beobachter weniger wichtig sind, werden Teile der Farbinformation

verworfen und die Farbkomponenten u und v unterabgetastet. Das in Digitalkameras übliche Verhältnis von Helligkeit zu Farbinformation ist 4:2:2, also eine Reduktion auf die Hälfte bei den Farbkomponenten.

Die eigentliche Kodierung wird auf Basis von Pixelblöcken vorgenommen. Helligkeits- und Farbkomponenten werden in 8x8 Pixel große Blöcke unterteilt und jeder Block einzeln kodiert und gespeichert.

Die JPEG-Kompression bedient sich der Diskreten Cosinus-Transformation (DCT). Diese mathematische Operation transformiert die einzelnen Pixelinformationen eines Blocks in einen gleich großen Satz von Koeffizienten. Jeder Koeffizient repräsentiert eine bestimmte Frequenz mit einer bestimmten Orientierung. Zwischen der Pixeldarstellung und den Transformationskoeffizienten kann ohne Verlust umgerechnet werden, hier gehen also noch keine Informationen verloren.

Dies erfolgt durch die Abbildung auf feste Werte mittels einer Quantisierungsmatrix (QM). Jeder DCT-Koeffizient wird durch sein Gegenüber in der QM geteilt, das gegen Null gerundete Ergebnis wird behalten. Je größer also ein Wert in der QM, desto ungenauer wird die Information gespeichert. Und der entsprechende DCT-Koeffizient muss größer sein, um nicht zu Null gerundet zu werden. Die QM kann durch einen Qualitätsfaktor in ihren Werten verändert werden. Dies ist der wichtigste Punkt, an dem Bildinformation gegen Dateigröße eingetauscht wird. Da für die Rekonstruktion die QM bekannt sein muss, wird sie mit in der Datei gespeichert.

Um möglichst effizient zu speichern, werden die Koeffizienten so umsortiert, dass die Positionen, die am wahrscheinlichsten hohe Werte enthalten, zuerst gelesen werden. Somit ergibt sich ein Zick-Zack-Muster über die

64 Werte, wobei der erste Wert der Mittelwert des Blocks ist, dann die niedrigen Frequenzen folgen und zum Schluss die hohen Frequenzen stehen. Somit stehen für Blöcke mit wenig Bildinformation am Ende der Kette viele Nullen. Gefördert wird dies dadurch, dass immer nur die Differenz zum vorhergehenden Wert gespeichert wird.

Eine Lauflängenkodierung reduziert die Daten weiter. Es bedeutet, dass nur die Anzahl gleicher Werte gespeichert wird und nicht alle Werte einzeln. Stehen am Ende einer Wertefolge also zwanzig mal die „0“, wird nur „20x0“ gespeichert. Die verbleibenden Werte werden mittels der Huffman-Kodierung komprimiert – einer verlustfreien, statistischen Methode zur Datenreduzierung.

Dateiformat

Die nach JPEG-Standard komprimierten Bilddaten können nun in eine Datei gespeichert werden. Heutige Digitalkameras aller Hersteller benutzen ein Dateiformat, welches im Standard CP-3461 des japanischen Industrieverbandes JEITA (Japan Electronics and Information Technology Association) spezifiziert ist. Dieser baut auf dem Dateiformat Exif 2.1 (Exchangeable Image File Format) der gleichen Organisation auf.

MetaDaten

Neben den Bilddaten werden verschiedene Metainformationen gespeichert. Der Standard legt fest, welche Daten sich in der Datei befinden müssen, und gibt zusätzlich die Möglichkeit, proprietäre, also nicht standardisierte, Daten in die Datei zu schreiben. Es wird hier in einen öffentlichen und einen privaten Bereich unterschieden. Im öffentlichen Bereich befinden sich verschiedene Metadaten wie zum Beispiel die wichtigsten Aufnahmeparameter und Angaben zum Kameramodell. Der private Bereich ist nicht definiert und bei

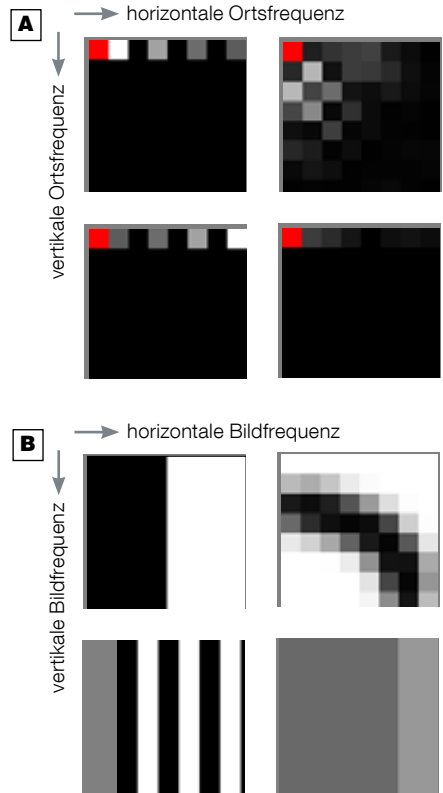
Zwischen-speicherung

Das Ergebnis von wiederholter Speicherung eines Bilds ohne und mit verschiedenen starker JPEG-Kompression. Wird das Bild in jedem Schritt komprimiert, werden schnell Artefakte sichtbar. Daher sollten Zwischenergebnisse in der Bildverarbeitung verlustfrei gespeichert werden.



Diskrete Cosinus-Transformation (DCT)

Die Koeffizienten der DCT beschreiben den Bildinhalt mittels der Stärke und Verteilung von horizontalen und vertikalen Ortsfrequenzen. Jeder einzelne Koeffizient (A) entspricht einer so genannten Basisfunktion im Bild (B), wobei der Wert oben links dem Mittelwert des Blocks entspricht.



ZigZag

Zick-Zack Sortierung – Die Koeffizienten werden nach der Quantisierung einem Zick-Zack Muster folgend ausgelesen. Somit wird die Wahrscheinlichkeit erhöht, dass viele Nullen aneinander gereiht werden. Oben links sind die niedrigen Ortsfrequenzen, unten rechts die hohen.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

den meisten Herstellern auch nicht dokumentiert, wodurch es vorkommen kann, dass Programm A mehr Information aus einem Bild lesen kann als Programm B. Ein Programm, welches so ziemlich alle Informationen aus Ihren Dateien extrahieren kann, ist „ExifTool“ von Phil Harvey. (<http://www.sno.phy.queensu.ca/~phil/exiftool/>)

JPEG-Artefakte

Im Gegensatz zu Dateien mit unkomprimierten Daten sind Bilddateien mit JPEG-Daten in ihrer Größe abhängig vom Bildinhalt. Enthält ein Block von 8x8 Pixeln eine ebene Fläche, so muss nur ein Koeffizient gespeichert werden, da alle anderen zu Null werden. Je mehr Flächen sich also im Bild befinden, desto kleiner die Datei. Hierbei muss man bedenken, dass das Rauschen der Kamera verhindert, dass Flächen wirklich in 64 gleiche Pixelwerte innerhalb eines Blocks resultieren. Daher beobachtet man oft, dass die von Kameras ausgegebenen Dateien größer werden, je höher die eingestellte Empfindlichkeit (ISO). Oft sind auch Dateien einer Spiegelreflexkamera kleiner als die einer mit ebenso vielen Pixeln ausgestatteten Kompaktkamera.

Zur Bildung von Artefakten kommt es vor allem dann, wenn die Koeffizienten der DCT zurück transformiert werden und bei der Kompression zu viele Informationen verloren gegangen sind. Die jeweils 64 Koeffizienten werden in 8x8-Pixel-Blöcke umgerechnet. Wenn man die mögliche Unterabtastung des Farbsignals in Betracht zieht, können diese Störungen im späteren Bild sogar 16x16 Pixel groß sein. Sichtbar werden Störungen im Bild immer dann, wenn innerhalb eines Blocks ein Sprung von dunkel zu hell abgebildet werden soll, da hierfür die Koeffizienten aller Frequenzen benötigt werden. Wird zu stark komprimiert,

ist die Information über die hohen Frequenzen sehr ungenau oder fehlt sogar. Dann kommt es dazu, dass der Sprung als Welle dargestellt wird.

Problematisch sind auch feine Linien, die schmaler als ein Block sind und sich über mehrere Blöcke ziehen. Da durch solche Linien mehrere Blöcke in ihrem Mittelwert unterschiedlich stark betroffen sind, werden diese sichtbar. Problemfälle treten im besonderen Maße bei Texten in Pixelbildern und bei Grafiken auf, was man manchmal bei schlecht bearbeiteten Bildern im Internet beobachten kann. Für Texte und Grafiken bieten sich andere Methoden zu Kompression an, mittels JPEG-Algorithmen sollten nur Bilder komprimiert werden.

Umgang mit JPEG

Jeder Speichervorgang führt zu einem erneuten Verlust an Informationen im Bild. Je nach Einstellung können diese verschwindend gering sein, doch wird ein Bild häufiger gespeichert, werden auch hier Artefakte sichtbar. Daher sollte man für seine Zwischenspeicherungen in der digitalen Bildbearbeitung möglichst ein Dateiformat wählen, welches die Bilddaten unkomprimiert aufnehmen kann. Um den Effekt der kumulierten Artefakte zu visualisieren, haben wir ein Bild immer wieder gespeichert, ausgelesen und erneut abgespeichert. Vorm erneuten Speichern, wurde das gesamte Bild leicht verschoben.

Als Referenz benutzen wir hier TIFF, das Bild wurde nicht für jeden Speichervorgang komprimiert. Im Vergleich dazu, werden auch in der besten JPEG-Qualität bei 50 Durchgängen Störungen sichtbar. Nimmt man eine reduzierte Qualitätsstufe, die bei einmaligem Speichern noch kaum Artefakte zeigt, hat man sich bereits nach zehn Durchgängen deutliche Spuren der Kompression eingefangen.

LEXIKON

■ Huffman-Kodierung

Die Huffman-Kodierung geht auf den Amerikaner David H. Huffman zurück. Der Pionier der Computertechnik stellte dieses Verfahren bereits 1952 vor. Die Methode vermeidet Zahlenwerte direkt mittels Bits zu beschreiben. Ein Wert zwischen 0 und 255 würde normalerweise mit 8 Bits kodiert, 64 Werte bräuchten somit 512 Bit. Ähnlich wie im Morse-Alphabet werden nach Huffman statistisch wahrscheinlichere Werte kürzer kodiert als unwahrscheinlichere Werte. Die zuerst erstellten Informationen in einem so genannten Binärbaum sorgen dafür, dass die häufigsten Werte mit weniger Bits beschrieben werden müssen als weniger häufige Werte. Wie stark hierdurch komprimiert werden kann, ist wieder abhängig von der statistischen Verteilung der Werte, je mehr gleiche Werte im Datensatz, desto weniger Bits werden zur Beschreibung benötigt.

■ Diskrete Cosinus-Transformation (DCT)

Die DCT ist eine spezielle Form der mathematischen Transformationsmethoden um zwischen verschiedenen Beschreibungsförmungen zu wechseln. Man unterscheidet hier

zwischen dem Ortsraum und dem Frequenzraum. Im Ortsraum wird jeder Position ein Wert zugeschrieben, es handelt sich also zum Beispiel um ein Pixelbild. Im entsprechenden Frequenzraum wird der gleiche Inhalt mittels der Anteile von Frequenzen beschrieben. Durch die Überlagerung von verschiedenen Cosinus-Funktionen lassen sich alle Informationen darstellen. Um eine Fläche zu beschreiben, reicht der Mittelwert. Die Frequenzinformationen werden zu Null. Um eine Kante darzustellen, werden Cosinus-Funktionen aller Frequenzen überlagert.

■ Quantisierung

Der Begriff Quantisierung stammt vom lateinischen Begriff „quantitas“ – Menge. Hierbei werden Werte aus einem Wertebereich in einen abgestuften Wertebereich abgebildet. Die Größe des Bereichs, der auf den gleichen Wert abgebildet wird, ist bei der JPEG-Kompression in der Quantisierungsmatrix festgehalten. Bei einem Quantisierungswert von 1 wird jeder Wert erhalten. Bei einem Wert von 10 würden alle Werte zwischen 0 und 9 zu „0“, alle Werte zwischen 10 und 19 zu „10“ usw.

