

**Vortrag**

## **Digitale Bildrepräsentation**

**vom  
6. November 2014  
sowie  
6. November 2015**

**durch  
Guido Vollbeding  
Independent JPEG Group (IJG)  
Institut für Angewandte Informatik e.V. (InfAI) Leipzig**

**im Rahmen der Lehrveranstaltung**

## **Mikroskopische Bildverarbeitung**

**durch  
Ulf-Dietrich Braumann, Prof. Dr.-Ing.  
Hochschule für Technik, Wirtschaft und Kultur (HTWK) Leipzig  
Interdisziplinäres Zentrum für Bioinformatik, Universität Leipzig**

**Dreiteiliger Aufbau:**

- 1. Die 4 Grundgesetze (Fundamenteigenschaften) der digitalen Bildrepräsentation mittels DCT**
- 2. Prozessgliederung des JPEG-Verfahrens**
- 3. Nutzungsmöglichkeiten fortgeschrittener JPEG-Funktionen und Ausblick**

## Grundlagen

Bei der Verarbeitung und Übertragung von digitalen Bild- und insbesondere Filmdaten fällt eine erhebliche Datenmenge an. Diese Tatsache hat zur Forderung nach effizienten Verfahren zur Komprimierung der anfallenden Datenmenge auf ein vernünftiges technisch handhabbares Maß geführt.

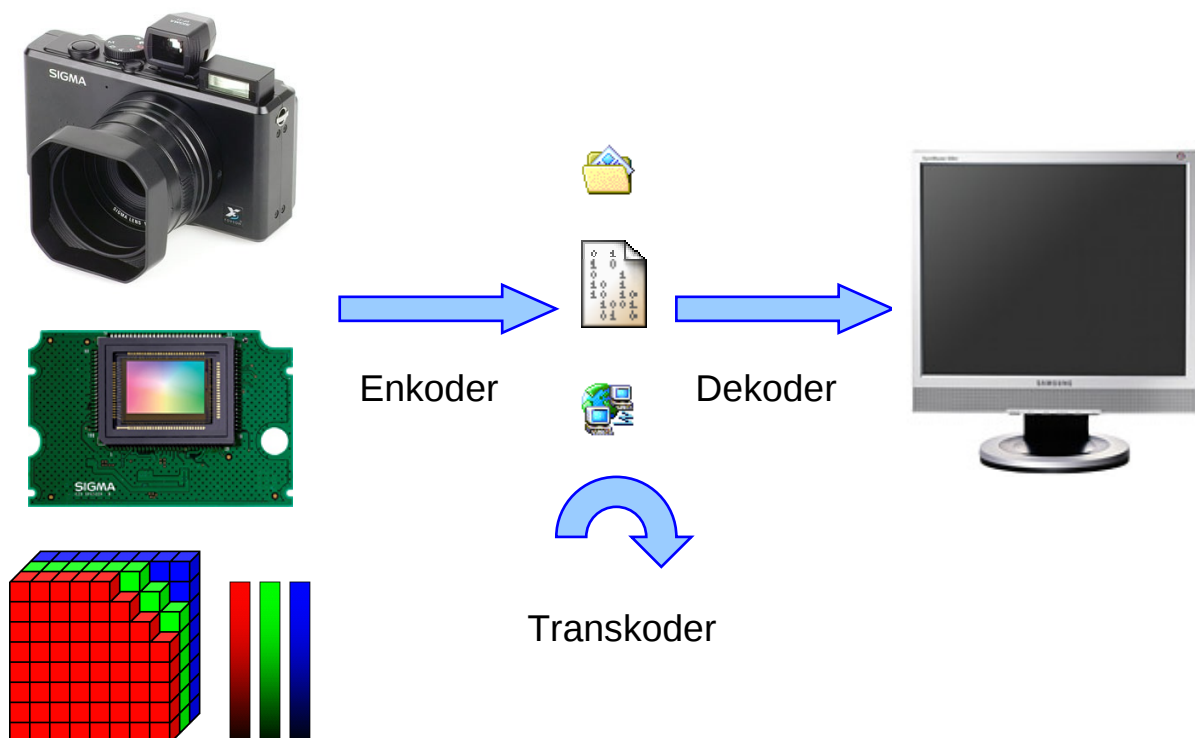
Der Kernbaustein aller Bild- und Filmkompressionsverfahren ist der sogenannte „Codec“. Diese Bezeichnung ist die Zusammensetzung aus den beiden Worten „Coder“ (Enkoder) und „Decoder“ (Dekoder).

Der Enkoder bezeichnet diejenige Funktion, die eine in Rohform (Pixel-Werte) vorliegende (Bild-)Datenmenge in einen komprimierten Datenstrom umwandelt (transformiert). Der Enkoder wird also im bzw. nachfolgend am Aufnahmegerät eingesetzt.

Der Dekoder ist die Funktion, welche auf Empfänger-(Wiedergabe-)Seite den komprimierten Datenstrom wieder in Pixel-Werte umwandelt, die auf dem Anzeigegerät als Bild dargestellt werden können.

Neben den beiden Grundfunktionen Enkoder und Dekoder gibt es noch Funktionen die mit dem Begriff „Transkoder“ (Umkodierer) bezeichnet werden. Dadurch wird die komprimierte Datenmenge von einer Form in eine andere mit neuen Eigenschaften umgewandelt, wobei es beim komprimierten Format bleibt.

Diese Zusammenhänge werden in folgender Abbildung veranschaulicht:



Der Codec selbst wiederum setzt sich aus verschiedenen Komponenten zusammen. Die Kernkomponente des Codecs ist eine Funktion, die als „Transformation“ bezeichnet wird. Dabei werden die Pixelwerte in eine für die komprimierte Darstellung und Übertragung geeignete Form umgewandelt.

Im JPEG-Bereich sowie einigen verbreiteten Filmformaten (MPEG-1, MPEG-2, DVD, DVB) wird als Transformation die sogenannte Diskrete Cosinus-Transformation benutzt, abgekürzt DCT.

Der Grund, warum gerade die DCT zur Transformation der Pixelwerte in eine geeignete komprimierbare Form verwendet wird, ist bisher nicht ausreichend verstanden, und aus diesem mangelnden Verständnis ergeben sich alle Fehlentwicklungen, die wir heute in der Bild- und Filmkodierung beobachten können. Diese Fehlentwicklungen zeichnen aus durch schlechtere Qualität und Performance, mangelhafte Flexibilität (Skalierbarkeit) sowie Inkompatibilität (Unverträglichkeit) der verschiedenen Verfahren.

Die Skalierbarkeit ist ein wichtiges Merkmal in der Anwendung von Bild- und Filmkodierungsverfahren. Es gibt einen unmittelbaren Zusammenhang zwischen den allgemeinen DCT-Funktionen, der Komprimierbarkeit von Bilddaten sowie deren Skalierbarkeit. Erst durch das Verständnis und die Ausnutzung dieser grundlegenden Zusammenhänge können die Anforderungen an einen Bild- und Film-Codec optimal erfüllt werden.

Diese grundlegenden Zusammenhänge können in Form von 4 Grundgesetzen oder Fundamenteigenschaften ausgedrückt werden, welche nachfolgend näher erläutert werden. Zentral ist dabei in jedem Fall die Eigenschaft der Skalierbarkeit, welche in folgender Abbildung anschaulich ausgedrückt wird:



# Die 4 Grundgesetze (Fundamenteigenschaften) der digitalen Bildrepräsentation mittels DCT

## 1. Das diskrete Skalierungsgesetz

In bisher verbreiteten Anwendungen zur digitalen Bild- und Filmkodierung (JPEG, MPEG, DVD, DVB) wird im Kern die Diskrete Cosinus-Transformation (DCT) mit der Größe 8 verwendet. Das bedeutet, es werden 8 gegebene Pixelwerte in sogenannte Koeffizienten umgerechnet, welche als Schwingungsmuster mit unterschiedlichen Frequenzen gedeutet werden können:

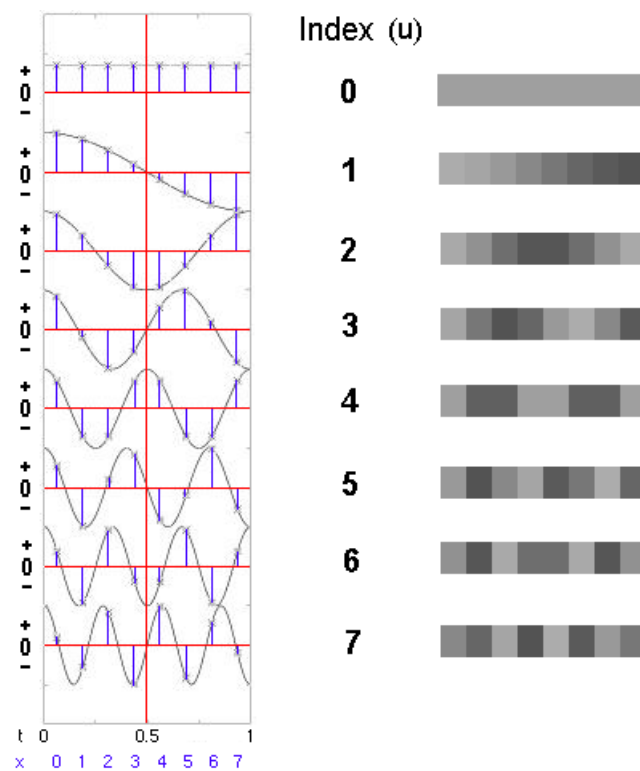


Abbildung: Schwingungsmuster (Basisfunktionen) der 8-Punkt DCT

In der Abbildung sieht man auf der linken Seite die kontinuierlichen (stetigen) Cosinuswellen als Grundgerüst, wovon jeweils 8 diskrete Stützpunkte genommen werden. Auf der rechten Seite sieht man die entsprechende Darstellung als Hell-Dunkel-Intensitäten.

Die Transformation bedeutet nun, dass wir eine beliebige 8-Punkt-Folge als Kombination solcher Schwingungsmuster repräsentieren können. Es kann jeder 8-Punkt-Folge eindeutig eine solche Schwingungskombination zugeordnet werden (berechnet im Encoder mittels DCT Algorithmus), und umgekehrt kann aus der Schwingungskombination wieder eindeutig die ursprüngliche Pixelfolge bestimmt werden (errechnet im Dekoder mittels IDCT, Inverse DCT).

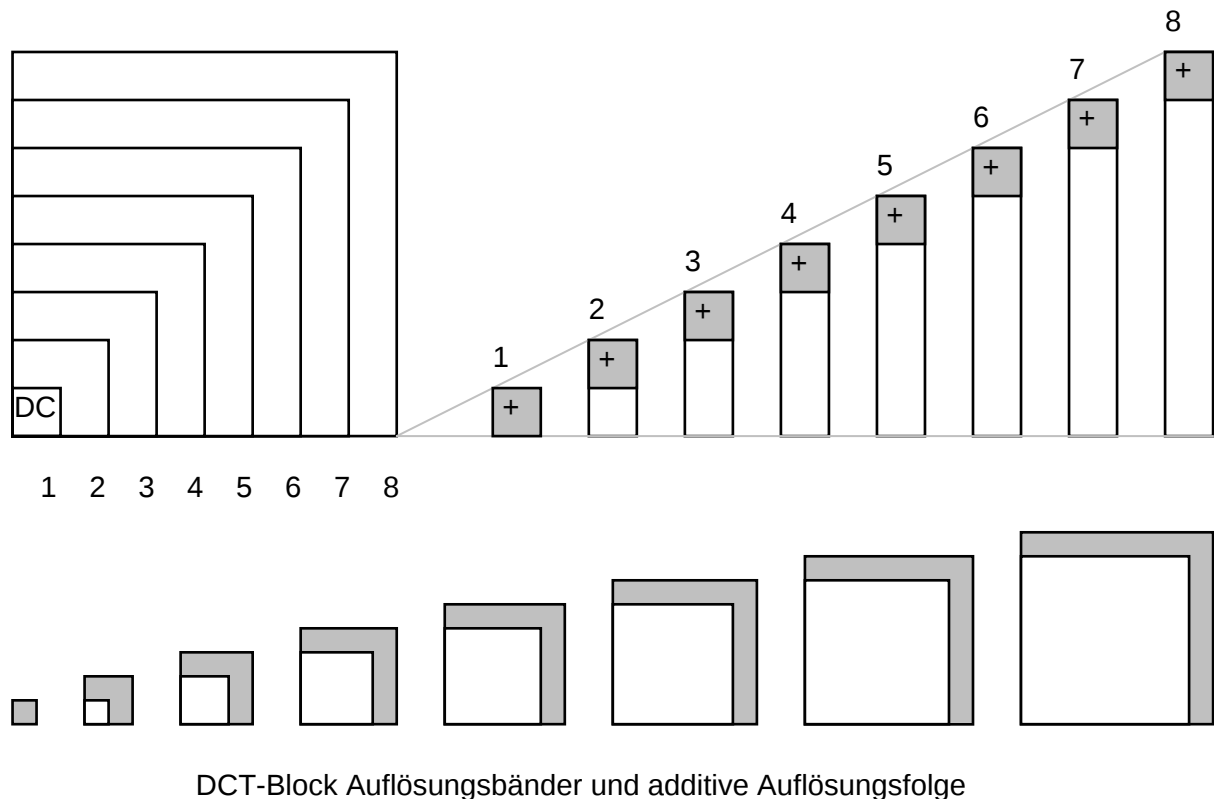
Durch mehrmalige Anwendung der Transformation auf die Zeilen und Spalten in einem zweidimensionalen 8x8 Pixelblock und durch die Unterteilung eines beliebigen Pixelbildes in solche 8x8 Pixelblöcke wird das komplette Bild verarbeitet.

In der bisherigen Anwendung wird die DCT immer mit gleicher Größe im Encoder und Dekoder benutzt, und zwar mit dem Wert 8. Es werden also immer 8 Pixelwerte enkodiert, und aus den komprimierten Daten werden 8 Pixelwerte dekodiert.

Das diskrete Skalierungsgesetz (erste Fundamenteigenschaft der digitalen Bildrepräsentation mittels DCT) besagt nun, dass die durch die DCT berechneten kodierten Koeffizienten nicht nur wie bisher als Schwingungswerte gedeutet werden können, sondern direkt als Repräsentanten für die Bilddetails in der entsprechenden Auflösung: Der erste Koeffizient (Index 0) repräsentiert das Bild in der 1/8-Auflösung, der erste und zweite Koeffizient zusammen repräsentieren das Bild in der 2/8-(=1/4-)Auflösung, der erste, zweite und dritte Koeffizient zusammen repräsentieren das Bild in der 3/8-Auflösung, usw.... Jeder Koeffizient fügt genau die Details in der jeweiligen Auflösungsstufe hinzu. Die Pixelwerte der jeweiligen Auflösung können direkt durch eine DCT der entsprechenden Größe berechnet werden. Dazu wurden entsprechend optimierte Algorithmen entwickelt, die diesen Rechenprozess effizient mit hoher Performance durchführen.

Aus den mit bisherigen Verfahren gegebenen 8 DCT-Koeffizienten eines kodierten Bildes können also mittels SmartScale-Dekoder direkt 8 kleinere Auflösungsstufen berechnet werden, was in den folgenden Abbildungen veranschaulicht wird:

### Interpretation der DCT als „Bildpyramide“



**DCT :** | Auflösungs-Separation mit **arithmetischer** Progression  
**Subband Coding / Wavelet :** | Auflösungs-Separation mit **geometrischer** Progression

Eigenschaften von Bildtransformationen



Abbildung: Ähnlichkeit unterschiedlicher Auflösungsgrößen (Zoom-Stufen) des gleichen Bildes als grundlegendes Merkmal für Bildkodierungsverfahren.



Abbildung: Auflösungsfolge in geometrischer Progression – Kennzeichen vieler Bildkodierungsverfahren



Abbildung: Auflösungsfolge in arithmetischer Progression – Kennzeichen DCT-basierter Bildkodierungsverfahren.  
Die geometrische Folge erscheint als Untermenge der arithmetischen Folge (schwarz umrandete Bilder).  
Dazwischen gibt es die zusätzlichen Stufen der arithmetischen Folge.



Dabei ergibt sich eine optimale Qualität der erhaltenen Resultate sowie ein deutlicher Performance-Gewinn im Vergleich zu herkömmlichen Vorgehensweisen.

Die Komprimierbarkeit mit dem DCT-Verfahren kann anhand dieses Zusammenhangs einfach dadurch erklärt werden, dass damit die Bilddetails (=Koeffizienten) höherer Auflösungsstufen stärker reduziert (normale Komprimierung mittels Quantisierung der Koeffizienten) oder ganz weggelassen (im Skalierungsfall) werden.

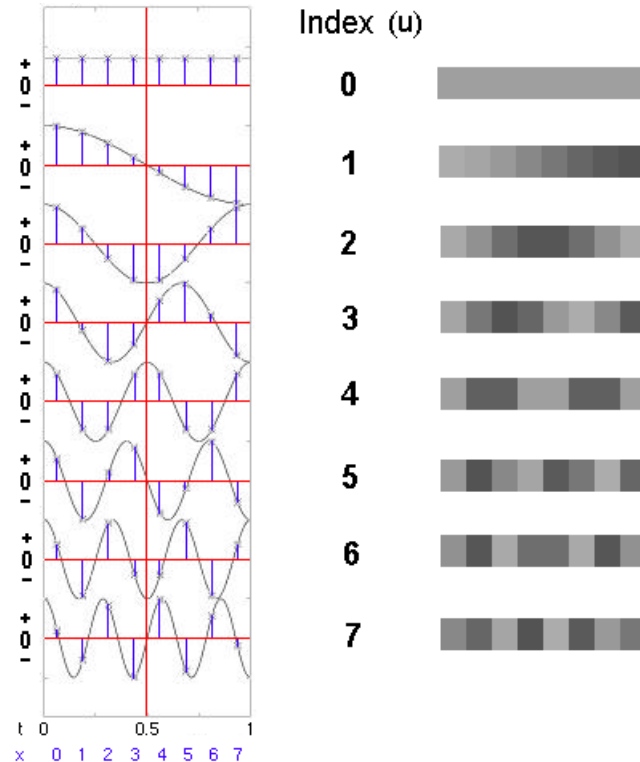
Neben den kleineren Auflösungsstufen können mittels erweiterter DCT auch direkt größere Auflösungsstufen berechnet werden. So können aus den 8 normalen Koeffizienten auch 9, 10, 11, 12, 13, 14, 15 oder 16 Pixelwerte berechnet werden, was in einer Bildskalierung von 112.5%, 125%, 137.5%, 150%, 162.5%, 175%, 187.5% und 200% resultiert.

Diese Option kann in Anwendungen zur noch stärkeren Datenreduktion als mit gewöhnlichen Mitteln benutzt werden.

Die Skalierungsoptionen können umgekehrt auch im Enkoder durch erweiterte DCT-Funktionen genutzt werden, und es werden Anwendungen zur Transkodierung ermöglicht.

## 2. Das stetige Skalierungsgesetz (Harmoniegesetz)

Die gegebene Abbildung der Schwingungsmuster legt eine weitere Interpretation nahe:



Die berechneten diskreten (digitalen) Schwingungswerte stehen in unmittelbarem Zusammenhang mit den zugrundeliegenden harmonischen Cosinuswellen.

Man kann sagen, dass durch die DCT eine stetige, kontinuierliche, analoge Repräsentation des digitalen Bildbereiches gegeben wird.

Tatsächlich kann man die IDCT (Inverse DCT im Dekoder) unmittelbar als eine stetige Interpolationsformel interpretieren. Das heißt, die stetigen Zwischenwerte können unmittelbar mittels einer stetig interpretierten IDCT-Formel aus den gegebenen diskreten Koeffizienten berechnet werden.

Dieser Zusammenhang birgt das Potenzial zu einer beliebigen (freien) Skalierbarkeit direkt aus den kodierten Bilddaten.

Er ist aber zunächst auch wichtig zum Verständnis der erfolgreichen Verwendung des DCT-Ansatzes in den verbreiteten Anwendungen im Vergleich zu anderen Versuchen:

Die DCT bewirkt eine stetige Repräsentation des Bildes innerhalb eines Pixelblocks. Man kann sagen, dass mittels DCT die Unstetigkeiten (Diskontinuitäten) zwischen den diskreten (digitalen) Pixelwerten innerhalb eines Blockes zu den Blockgrenzen hin verschoben werden. Daher ist die Blockbildung im Zusammenhang mit dem DCT-Ansatz kein Nachteil, solange man überhaupt mit digitaler Bildrepräsentation in Form von diskreten (digitalen) Pixelwerten arbeitet. Jeder Versuch die durch die Blockbildung eventuell hervorgerufenen Artefakte durch andere Verfahren vollkommen vermeiden zu wollen muss daher scheitern. Die DCT behandelt diesen Effekt schon in optimaler Weise, und mit andersartigen Verfahren können nur andersartige Artefakte hervorgerufen werden. Die grundlegende Tatsache der durch die Digitalisierung entstehenden Diskretisierung in Form von Pixelwerten kann nicht umgangen werden.

## Mathematische Definitionen der FDCT und IDCT

Ausgangspunkt für die mathematischen Definitionen der FDCT und IDCT sind kontinuierliche Cosinus-Schwingungen mit unterschiedlicher Frequenz (siehe obige Abbildung der Schwingungsmuster):

$$\cos tu\pi \quad ; \quad u = 0, 1, \dots, 7 \quad ; \quad 0 < t < 1$$

$u$  ist der diskrete Index mit wachsender Frequenz,  $t$  ist die kontinuierliche Variable im Intervall 0 bis 1.

So ergibt sich für  $u = 0$  eine konstante Funktion, für  $u = 1$  ergibt sich eine halbe Cosinus-Vollschwingung, für  $u = 2$  eine ganze Cosinus-Vollschwingung usw.

Die kontinuierliche Variable  $t$  wird nun statt im ganzen Intervall 0 bis 1 an den diskreten Stellen

$$t = \frac{2x+1}{16} \quad ; \quad x = 0, 1, \dots, 7$$

abgetastet. So entsteht innerhalb des Intervalls (Block bei JPEG) eine äquidistante Abtastung mit Abstand  $1/8$  zwischen den Punkten, an den Intervalgrenzen (Blockgrenzen bei JPEG) vorn und hinten ein Abstand von  $1/16$  zum Randpunkt. Auf diese Weise entsteht durch die Aneinanderreihung von Blöcken bei JPEG ein gleichmäßiger Abstand ( $1/8$ ) zwischen den Pixeln über das gesamte Bild.

Mit diesen Vorbetrachtungen lassen sich die nun folgenden eigentlichen Definitionen leichter verstehen:

Eindimensionale 8-Punkt FDCT:

$$S(u) = \frac{C(u)}{2} \sum_{x=0}^7 s(x) \cos \frac{(2x+1)u\pi}{16} \quad ; \quad u = 0, 1, \dots, 7$$

Eindimensionale 8-Punkt IDCT:

$$s(x) = \sum_{u=0}^7 \frac{C(u)}{2} S(u) \cos \frac{(2x+1)u\pi}{16} \quad ; \quad x = 0, 1, \dots, 7$$

wobei

$$C(u) = \frac{1}{\sqrt{2}} \quad \text{für } u = 0$$

$$C(u) = 1 \quad \text{für } u > 0$$

$$s(x) = 1 - D \text{ Pixelwert}$$

$$S(u) = 1 - D \text{ DCT - Koeffizient}$$

Zweidimensionale 8x8 FDCT:

$$S(v, u) = \frac{C(v)}{2} \frac{C(u)}{2} \sum_{y=0}^7 \sum_{x=0}^7 s(y, x) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16}$$

Zweidimensionale 8x8 IDCT:

$$s(y, x) = \sum_{v=0}^7 \frac{C(v)}{2} \sum_{u=0}^7 \frac{C(u)}{2} S(v, u) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16}$$

wobei

$$C(u) = \frac{1}{\sqrt{2}} \quad \text{für } u = 0$$

$$C(u) = 1 \quad \text{für } u > 0$$

$$C(v) = \frac{1}{\sqrt{2}} \quad \text{für } v = 0$$

$$C(v) = 1 \quad \text{für } v > 0$$

$$s(y, x) = 2 - D \text{ Pixelwert}$$

$$S(v, u) = 2 - D \text{ DCT - Koeffizient}$$

Erweiterungen:

1-D FDCT, N=1...16:

$$S(u) = \frac{8}{N} \frac{C(u)}{2} \sum_{x=0}^{N-1} s(x) \cos \frac{(2x+1)u\pi}{2N} \quad ; \quad u = 0, 1, \dots, N-1$$

1-D IDCT, N=1...16:

$$s(x) = \sum_{u=0}^{N-1} \frac{C(u)}{2} S(u) \cos \frac{(2x+1)u\pi}{2N} \quad ; \quad x = 0, 1, \dots, N-1$$

wobei

$$C(u) = \frac{1}{\sqrt{2}} \quad \text{for } u = 0$$

$$C(u) = 1 \quad \text{for } u > 0$$

$$s(x) = 1 - D \text{ sample value}$$

$$S(u) = 1 - D \text{ DCT - coefficient}$$

2-D FDCT,  $N=1\dots 16$ :

$$S(v, u) = \frac{8}{N} \frac{8}{N} \frac{C(v)}{2} \frac{C(u)}{2} \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} s(y, x) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N}$$

2-D IDCT,  $N=1\dots 16$ :

$$s(y, x) = \sum_{v=0}^{N-1} \frac{C(v)}{2} \sum_{u=0}^{N-1} \frac{C(u)}{2} S(v, u) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N}$$

wobei

$$C(u) = \frac{1}{\sqrt{2}} \quad \text{for } u = 0$$

$$C(u) = 1 \quad \text{for } u > 0$$

$$C(v) = \frac{1}{\sqrt{2}} \quad \text{for } v = 0$$

$$C(v) = 1 \quad \text{for } v > 0$$

$s(y, x)$  = 2-D sample value

$S(v, u)$  = 2-D DCT-coefficient

#### NOTE:

- (1) Coefficient values with index 8 and greater need not be calculated in the FDCT case, and are set to zero in the IDCT case, because no more than 8x8 coefficients per block are stored in the encoded data.
- (2) There are different scaling factors compared to the normal (unrelated) mathematical N-point or NxN-point DCT definitions respectively. In our case the calculation always refers to the standard 8-point or 8x8-point DCT, so that the scaling factors are the same as those in the IDCT case, and are multiplied by 8/N per dimension in the FDCT case.

### 3. Das Gesetz der Acht

Das „Gesetz der Acht“ gibt eine Erklärung, warum die Größe 8 die angemessene Standard- und Bezugs-Größe für die DCT ist. Stichwort: Größenverhältnis zwischen Vollbilddarstellung und Vorschaubilddarstellung.

### 4. Das Netzgesetz (Infrastrukturgesetz, Strom-Analogiegesetz)

In den bisherigen Erklärungen war immer von den 8 „DCT-Koeffizienten“ die Rede, die als Schwingungsmuster gedeutet werden können.

In der Fachsprache haben diese Koeffizienten besondere Namen, die auf einen weiteren wichtigen Zusammenhang hinweisen.

Der erste Koeffizient (Index 0 in der Abbildung) heißt DC-Koeffizient (DC = Direct Current = Gleichstrom), also Gleichstrom-Koeffizient. Dieser beschreibt den Durchschnittswert der Pixel innerhalb eines Blocks, und daher ist auch unmittelbar klar und auch bekannt, dass der

DC-Koeffizient einer 1/8-Skalierung der Pixelfolge bzw. des Bildes entspricht (bei DCT mit der Standardgröße 8).

Die anderen 7 Koeffizienten heißen AC-Koeffizienten (AC = Alternating Current = Wechselstrom).

Praktisch kann man sagen, dass die Bildkodierung mittels DCT die Bildinformationen in eine überwiegende Wechselstromdarstellung überführt. Es gibt im eindimensionalen Falle 1 DC-Koeffizienten und 7 AC-Koeffizienten, im zweidimensionalen Fall (8x8 Block) sogar 1 DC-Koeffizienten und 63 AC-Koeffizienten.

Und es ist diese Wechselstromdarstellung, welche all die vorteilhaften Eigenschaften wie Komprimierbarkeit, variable Qualität und Skalierung bei guter Performance gewährleistet. Diese Tatsache steht in voller Analogie zum Energieverbundnetz, welches ebenfalls auf Wechselstrombasis funktioniert. Das einheitliche „Format“ des Wechselstromes gewährleistet vor allem die einfache Transformierbarkeit für die verschiedenen Anwendungen und zur verlustarmen Übertragung. Auch Gleichstromkomponenten lassen sich innerhalb des Wechselstromnetzes betreiben, aber nur als untergeordneter Spezialfall. Das universelle Format für Austausch und Übertragung von Energie ist Wechselstrom.

Diese Tatsache erscheint heute selbstverständlich, das war aber nicht immer so.

Das Wechselstromnetz musste sich erst über Jahrzehnte hinweg gegenüber dem Gleichstromansatz durchsetzen. Kein geringerer als der bedeutende amerikanische Erfinder Thomas Alva Edison hat lange Zeit das Gleichstromnetz propagiert, bevor er am Ende seines Lebens zu der Einsicht kam, dass dies der größte Fehler gewesen sei, den er gemacht habe. Letztendlich erfolgreicher Gegenspieler von Edison in dem sogenannten „Stromkrieg“ am Ende des 19. Jahrhunderts war der Industrielle und Ingenieur George Westinghouse, der mit Hilfe der Entwicklungen des genialen Erfinders Nikola Tesla den Wechselstromansatz verfolgte und durchsetzte.

Beim Stromnetz haben wir also den Wechselstrom als universelles Austausch-/Übertragungs-Format. Dieses gewährleistet eine flexible Transformierbarkeit in unterschiedliche Spannungen für verschiedene Zwecke und Anwendungen. Kraftwerke und Generatoren müssen am Ende Wechselstrom in dem festgelegten Format in das Netz einspeisen. Verbrauchergeräte müssen zur Energieversorgung Wechselstrom aus dem Verbundnetz in dem gegebenen Format beziehen und mittels Transformatoren auf die spezielle Anwendung anpassen.

Ganz ähnlich muss nun der multimediale Informationsaustausch auf ein geeignetes Universalformat gestützt werden, welches die verschiedenen Anwendungsanforderungen erfüllen kann.

Nach allen bisherigen Erkenntnissen ist das erweiterte DCT-Format mit SmartScale-Funktion die gesuchte Lösung, welche die gegebenen Anforderungen erfüllt:

Flexible Transkodierbarkeit bezüglich Auflösung und Datenmenge (entspricht Transformierbarkeit des Wechselstroms mit unterschiedlichen Spannungen und Energiemengen), Verfügbarkeit flexibler und effizienter Dekoder (entspricht Verbrauchergeräten) und Enkoder (entspricht Kraftwerken/Generatoren).

## **Vergleich mit alternativen Ansätzen**

Nach der Darstellung und dem richtigen Verständnis der zuvor erläuterten 4 Grundsätze zur Bildkodierung kann man sich nun leicht die Irrtümer der anderweitig verfolgten Ansätze klarmachen.

Da wäre zunächst der mit „JPEG 2000“ verfolgte Ansatz der Wavelet-Verfahren, die seit dem Jahre 2000 versuchen nennenswerte Anwendungsbereiche zu erreichen.

Die sogenannten „Wavelets“ sind Basisfunktionen, die, in Unkenntnis der oben dargestellten Skalierungseigenschaften der DCT-Basisfunktionen, speziell für den Zweck der Skalierung eingeführt wurden. Im Vergleich zu den Skalierungsmerkmalen der DCT sind aber die Skalierungseigenschaften der Wavelets minderwertig, da sie nur Skalierungen mit Zweierpotenzen gewähren, d.h. auf die Hälfte, ein Viertel, ein Achtel usw.... Zwischenwerte wie im erweiterten DCT-Verfahren sind damit nicht unmittelbar möglich. Weiterhin gibt es keine direkte Möglichkeit zur Vergrößerung.

Des Weiteren gibt es bei Wavelets keine dem Zweiten Grundgesetz entsprechende Eigenschaft, das heißt keinen harmonischen Zusammenhang, und die Performance ist schlecht, das heißt die Verfahren sind sehr rechenaufwändig und speicherintensiv.

Daneben werden heute Verfahren propagiert, die alle im Kern auf dem nachfolgend beschriebenen Ansatz basieren. Dazu gehören etwa „HD Photo“ bzw. „JPEG XR“ sowie MPEG-4/AVC („Advanced Video Coding“, H.264) und HEVC („High Efficiency Video Coding“, H.265).

Diese Verfahren zeichnen sich dadurch aus, dass sie zunächst gegen das Zweite Grundgesetz verstoßen, indem sie die Grundlage der harmonischen Basisfunktionen verlassen, mit der primären Absicht, Rechenaufwand einzusparen, insbesondere um die mit „High Definition“ (HD) verbundenen höheren Bildauflösungen (Pixelzahlen) technisch bewältigen zu können. Dies geschieht wiederum in Unkenntnis der dabei verlorengehenden wichtigen Grundeigenschaften.

So wird etwa, um den Aufwand zu reduzieren, zunächst die Transformationsgröße von 8 auf 4 reduziert. Damit geht erstmal unwissentlich ein Großteil der direkten Skalierungsmöglichkeiten verloren. Dann wird aber noch von der „reinen“ DCT ganz abgegangen (durch „Vereinfachung“ der in der Berechnung auftretenden konstanten Koeffizienten), um weitere Einsparungen beim Rechenaufwand zu gewinnen. Dadurch gehen letztlich die Skalierungseigenschaften völlig verloren – die so kodierten Bildinformationen sind fest an eine Auflösung gebunden, zumal noch eine Technik namens „Pixel-Prediction“ eingeführt wird, um den durch die Abkehr von der DCT hervorgerufenen Verlust an Kompressionseffekt zu kompensieren (und wodurch auch wieder zusätzlich Rechenaufwand an anderer Stelle hinzugefügt wird, der zuerst eingespart werden sollte).

Die Abkehr von der reinen DCT bei diesen Verfahren macht also unmittelbar die harmonischen Grundeigenschaften entsprechend dem Zweiten Grundgesetz zunichte, und infolgedessen gehen auch alle nützlichen Skalierungseigenschaften entsprechend dem Ersten Grundgesetz verloren.

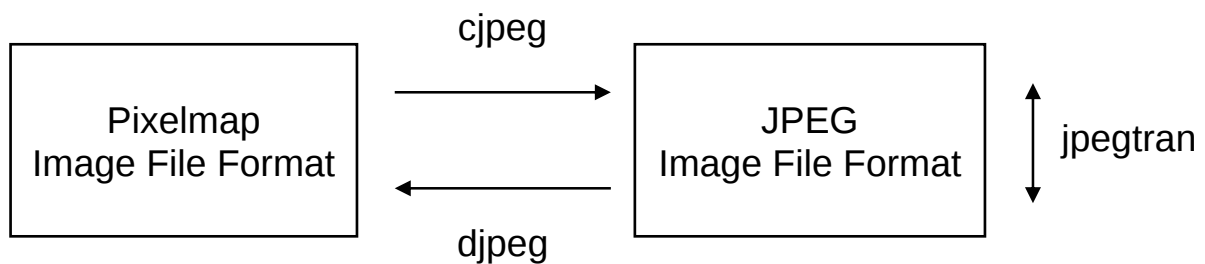
Zusätzlich muss der Rechenaufwand, der im Bereich der unmittelbaren Transformation eingespart wird, an anderer Stelle wieder hinzugefügt werden, um den Kompressionsverlust zu kompensieren. In der Praxis zeigt sich letztlich ein höherer Rechenaufwand, ähnlich wie bei den Wavelet-Verfahren.

## Prozessgliederung des JPEG-Verfahrens

Die grobe Struktur ist gegeben durch folgendes Schema:

### Independent JPEG Group

### JPEG Software Tools



Einen genaueren Überblick über den JPEG-Prozess gibt die folgende Abbildung:

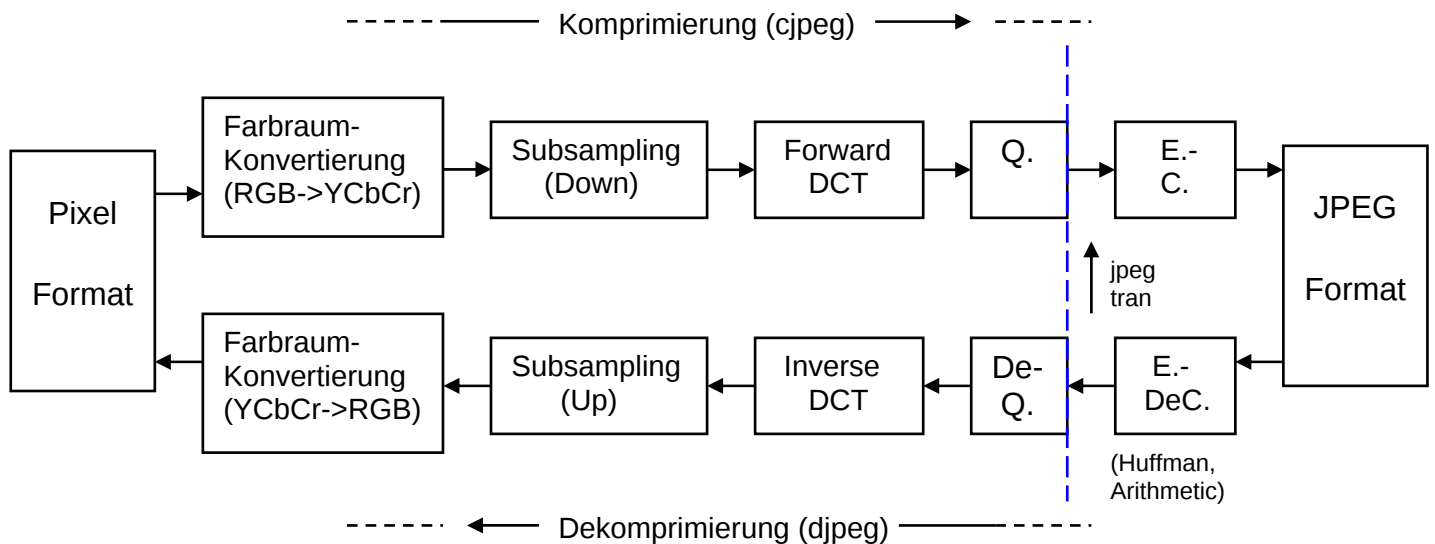


Abb.: JPEG Prozess Überblick

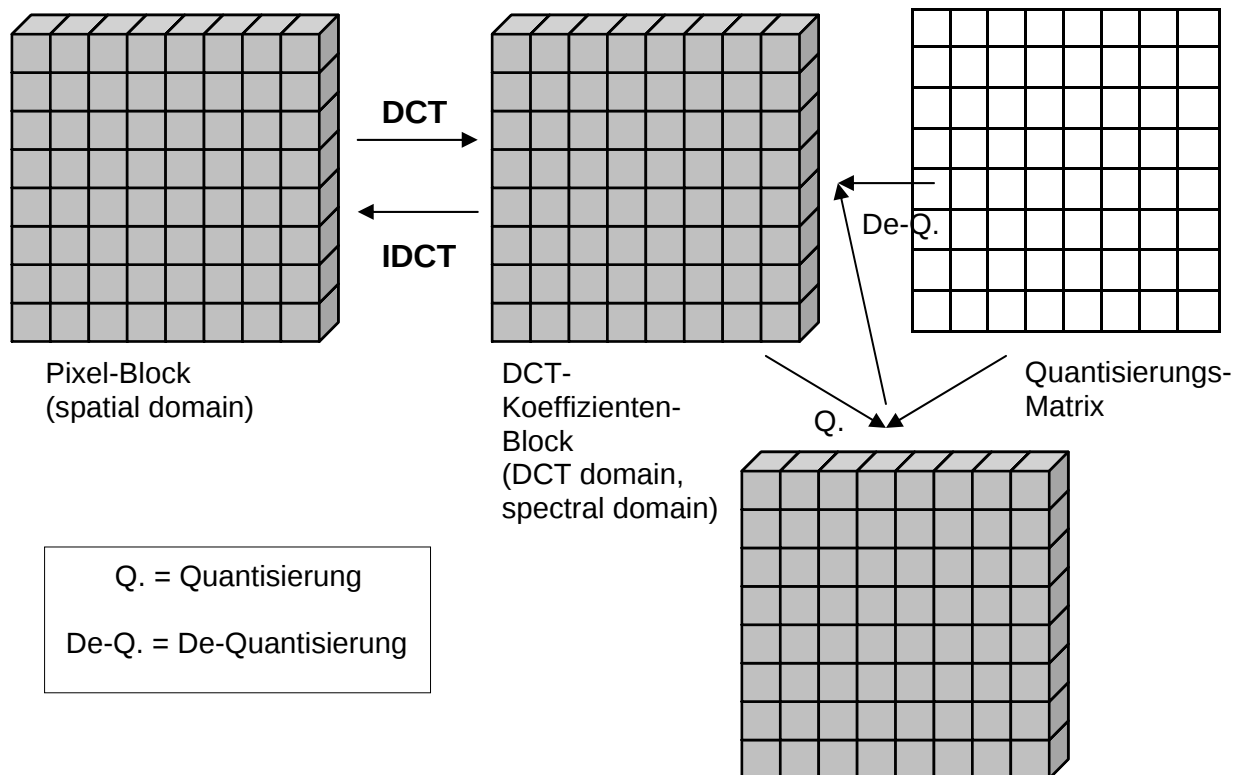
Die einzelnen Elemente der JPEG-Prozesskette weisen ein unterschiedliches „Verlustpotenzial“ auf.

Der Teilprozess der Entropie-Kodierung bzw. –Dekodierung am einen Ende der Prozesskette operiert vollkommen verlustfrei, während alle anderen Teile einem mehr oder weniger großen Verlustpotenzial unterliegen. Hier greift deswegen auch die Funktionalität des *jpegtran*-Tools und speziell die verlustfreie Drehfunktion an.

Der Hauptangriffspunkt zur „Verluststeuerung“ bei JPEG liegt im Teilprozess der Quantisierung, hierbei wird sozusagen „mutwillig“ durch Vorgabe der Verlustgrad bestimmt. Ein weiterer Angriffspunkt zur JPEG Verluststeuerung, der in praktischen Anwendungen leider oft unterschätzt oder übersehen wird, liegt im Teilprozess der Farbrunterabtastung (Color Subsampling).

Das eigentliche „Herz“ oder Zentrum des JPEG Prozesses stellt die Diskrete Cosinus-Transformation (DCT oder FDCT [Forward DCT]) bzw. ihr Gegenstück, die Inverse Diskrete Cosinus-Transformation (IDCT), dar. Hierbei werden Pixelwerte in „Schwingungswerte“ umgewandelt und umgekehrt. Diese Umwandlung geschieht „theoretisch“ verlustfrei, die Verlustfreiheit kann jedoch praktisch aufgrund begrenzter Rechengenauigkeit nicht hundertprozentig garantiert werden. Je nach Qualität und Anforderung der Implementierung sind hier mehr oder weniger geringfügige Abweichungen möglich. Ähnliches gilt für den Teilprozess der Farbraumkonvertierung.

Alle diese möglichen Verlustquellen, sowohl die vorgeblich gesteuerten (Quantisierung und Subsampling) als auch die inhärenten (DCT/IDCT, Farbkonvertierung), werden bei den verlustfreien *jpegtran*-Operationen und speziell also bei den verlustfreien Drehfunktionen umgangen!



|    |    |    |    |     |     |     |     |
|----|----|----|----|-----|-----|-----|-----|
| 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  |

Luminanz (table K.1)

|    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|
| 17 | 18 | 24 | 47 | 99 | 99 | 99 | 99 |
| 18 | 21 | 26 | 66 | 99 | 99 | 99 | 99 |
| 24 | 26 | 56 | 99 | 99 | 99 | 99 | 99 |
| 47 | 66 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |

Chrominanz (table K.2)

Tab.: JPEG Luminanz- und Chrominanz-Quantisierungstabellen

# Nutzungsmöglichkeiten fortgeschrittener JPEG-Funktionen

## Überblick

Neue Funktionen und Eigenschaften sind in folgenden Bereichen verfügbar:

- Verlustfreie Re-Komprimierung bereits vorhandener JPEG-Bilddateien mittels arithmetischer Kodierung
- Verlustfreie Komprimierung von Bilddaten
- Verlustfreie Umskalierung bereits vorhandener JPEG-Bilddateien und weitere verlustfreie Transformationsfunktionen (Ausschneiden, Drehen, etc.)
- Skalierte Enkodierung und Dekodierung
- Vergrößerter Farbumfang (Wide Gamut), um verschiedene und zukünftige Aufnahmegeräte und Ausgabemedien abzudecken
- Erweiterter Bereich der unterstützten Bittiefen

Diese Funktionen stehen mit den neuen Versionen 7, 8 und 9 der „Independent JPEG Group“ Software („libjpeg“ Paket) zur Integration in Anwendungsprogrammen zur Verfügung. Empfohlene Referenz-Applikationen, die den aktuellen Stand der JPEG-Entwicklung unterstützen, sind cPicture, PhotoLine und Directory Opus.

Für Projektarbeiten stehen weiterhin folgende Möglichkeiten der Performance-Optimierung und andere Erweiterungen zur Verfügung:

- Gezielte starke Performance-Optimierung bei der verlustfreien Komprimierung
- Performance-Optimierung durch Nutzung von Prozessorbefehlssatzerweiterungen im Bereich der DCT-Funktionen bei Enkodierung und Dekodierung
- Performance-Optimierung für arithmetische Kodierung mit maschinennaher Programmierung
- Kodierung mit höheren Bittiefen („HDR“ – High Dynamic Range)
- Anwendung für Bewegtbild (Motion-JPEG) mit entsprechenden Anpassungen und Optimierungen

Langfristig geht die Entwicklung in Richtung einer Option für die direkte Kodierung im ursprünglichen „Radiance“-Bereich statt wie bisher üblich im Ausgabe-bezogenen Bereich. Erst damit werden echte „HDR“ (High Dynamic Range) sowie „AR“ (Augmented Reality) Anwendungen möglich.

Die direkte „Radiance“-Repräsentation im Kern des Codecs stellt die logische Fortsetzung der Entwicklung in der Bildkodierung dar.

Die innere Grundstruktur des ursprünglichen (Referenz-) JPEG-Verfahrens weist bereits die dafür nötigen Voraussetzungen auf, es besteht jedoch erheblicher Entwicklungsbedarf für die konkrete Umsetzung.

## Arithmetische Kodierung

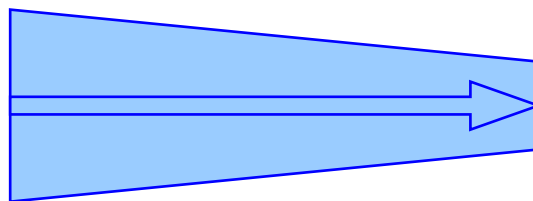
Durch die verlustfreie Umwandlung von vorhandenen (Huffman kodierten) JPEG-Bilddateien in arithmetisch kodierte JPEG-Bilddateien kann der Speicherplatzbedarf um 5 bis 20 Prozent reduziert werden.

Diese Funktion ist besonders für die Archivierung großer Bilddatenbestände interessant.

Für die Verwendung im Datenaustausch ist zu beachten, dass auf Empfängerseite ein entsprechender Dekoder basierend auf Version ab JPEG 7 erforderlich ist.

Neben den im Überblick genannten Referenz-Applikationen gehören dazu unter anderen das populäre Bildprogramm Irfanview sowie die Bildbearbeitungs- und Archivierungssoftware StudioLine Photo.

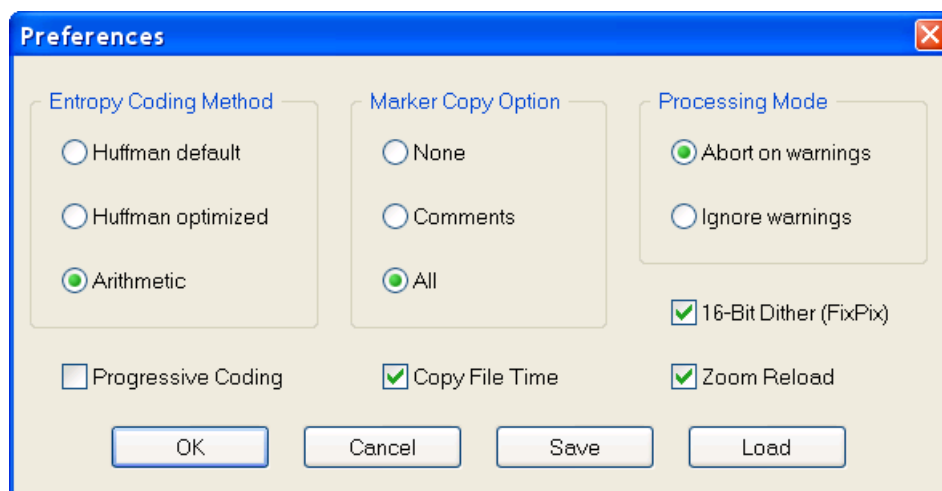
JPEG Huffman kodiert



JPEG arithmetisch kodiert



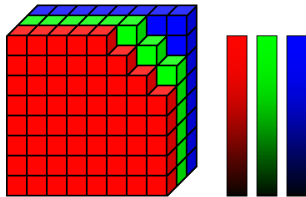
Das Programm Jpegcrop bietet unter anderem eine Funktion zur Durchführung der verlustfreien Konvertierung an Einzelbildern an:



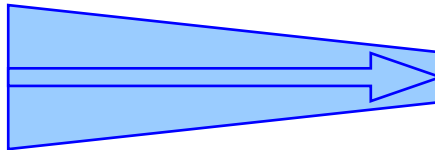
## Verlustfreie Komprimierung

Unkomprimiert vorliegende Bilddaten können seit der Version JPEG 8 verlustfrei komprimiert werden, und ab Version 9 mit verbesserter Rate, die anderen verfügbaren Verfahren überlegen ist (Komprimierung auf unter 50 Prozent für fotorealistische Bilder).

Bild unkomprimiert



JPEG komprimiert



## SmartScale

SmartScale bezeichnet ein neues Verfahren mit flexiblen Möglichkeiten zur Umskalierung bei Enkodierung, Dekodierung und Transkodierung. Dadurch kann die Bildgröße den jeweiligen Anforderungen flexibel angepasst werden.



## Erweiterter Farbumfang (Wide Gamut)

Durch die festgelegten drei Primärfarbenwerte im Standard-RGB-Farbmodell wird die Menge der darstellbaren Farbtöne auf einen Teilbereich des sichtbaren Spektrums begrenzt.

Eine erweiterte Kodierung entsprechend der im IEC 61966-2-1 Standard festgelegten Modi erlaubt nun auch die Darstellung zusätzlicher Farbtöne, um weitere Anwendungsbereiche abzudecken und die Möglichkeiten neuartiger Aufnahmegeräte und Ausgabemedien zu unterstützen.

Mit diesem Merkmal ist JPEG gegenwärtig der einzige Bildkodierungsstandard, der sämtliche in IEC 61966-2-1 definierten Farbkodierungen unterstützt und damit den größtmöglichen Farbtonumfang für alle Anwendungen bietet.

