Neue Methoden zur adaptiven Farbquantisierung für komprimierte Farbbildrepräsentation

Guido Vollbeding
Kapelan GmbH
Breite Straße 3
D-06108 Halle
e-mail: guido@jpegclub.org
URL: http://jpegclub.org
http://jpegclub.org/foveon/

Kurzfassung / Einleitung

Digitale Farbaufnahmesysteme gewinnen in Medien und Industrie eine wachsende Bedeutung. Während im qualitätsorientierten Videobereich Vollfarbaufnahmesysteme mit 3 CCDs und farbseparierendem Prisma verbreitet sind, dominieren selbst im anspruchsvollen Standbildbereich noch qualitätsmindernde 1-Chip-Geräte mit Farbmosaikfilter, da bis vor kurzem eine adäquate Technologie zur Vollfarbaufnahme mit den erforderlichen höheren Pixelauflösungen nicht verfügbar war.

Mit der Vorstellung des ersten 1-Chip-Vollfarbensensors X3 von Foveon im Jahre 2002 steht nun erstmalig eine digitale Aufnahmetechnik zur Verfügung, die Vollfarbaufzeichnung auch bei hochaufgelösten Standbildern ermöglicht, ohne die qualitätsreduzierende Farbmosaikfiltermethode zu verwenden, und ohne die mit 3-Chip-Systemen verbundenen enormen Kostenaufwände und anderen Probleme.

Die minderwertige Qualität bisheriger Farbsensoren mit Farbmosaikfilterprinzip führte zur Vernachlässigung der sorgfältigen Auswahl einer angemessenen Konfiguration bei der komprimierten Farbbildspeicherung mittels verbreiteter DCT-basierter Bildkompressionsverfahren, insbesondere JPEG.

In diesem Beitrag wird gezeigt, dass die in der Farbbildverarbeitung übliche und verbreitete Methode der Farbunterabtastung (Color Subsampling) im Kontext eines DCT-basierten Verarbeitungsprozesses verzichtbar ist, zugunsten einer neuen Methode der separierten Farbquantisierung. Es wird dargestellt und untersucht, wie mit der neuen Methode ein vergleichbarer Komprimierungsgrad mit wählbarer Luminanz- und Chrominanzquantisierung erreicht werden kann, mit verbessertem Adaptionsverhalten und ohne zusätzliche Subsampling-Artefakte im Vergleich zum bisherigen Verfahren. Die praktische Umsetzung der neuen Methode erfolgt mittels der verbreiteten JPEG-Software der Independent JPEG Group und ist relativ leicht anwendbar. Verschiedene Einsatzmöglichkeiten werden dargestellt und diskutiert.

Als Schlussfolgerung kann festgestellt werden, dass mittels der neuen Methode zur adaptiven Farbquantisierung die DCT-basierten Kompressionsverfahren durch Verzicht auf jegliche Farbunterabtastung strukturell vereinfacht und leichter anwendbar gemacht werden können. Die Option der Farbunterabtastung wird als vorübergehendes Feature beim Übergang von spatialer zu DCT-transformierter Verarbeitung erkannt, und eine spezifische Äquivalenz zwischen Reduktion im spatialen und DCT-Bereich wird aufgedeckt. Daneben wird durch Verzicht auf vorgefertigte Farbunterabtastung in vielen Bereichen ein höherer Qualitätsstandard ermöglicht, wie er für die mit dem Foveon X3 Sensor eingeleitete Zukunft der Vollfarbsensorsysteme angemessen ist.

1 Digitale Bilddarstellung / Bildrepräsentation (räumlicher Bereich)

Abbildung 1 zeigt eine (vereinfachte) schematische Darstellung der digitalen Bildformate im räumlichen Bereich (spatial domain). Ausgangspunkte sind das 1-kanalige monochromatische Graustufenbild (Abb. 1.1) und das 3-kanalige tri-chromatische Farbbild (Abb. 1.2), denen direkt eine technisch-physikalische Umsetzung aufnahmeseitig und ausgabeseitig entspricht.

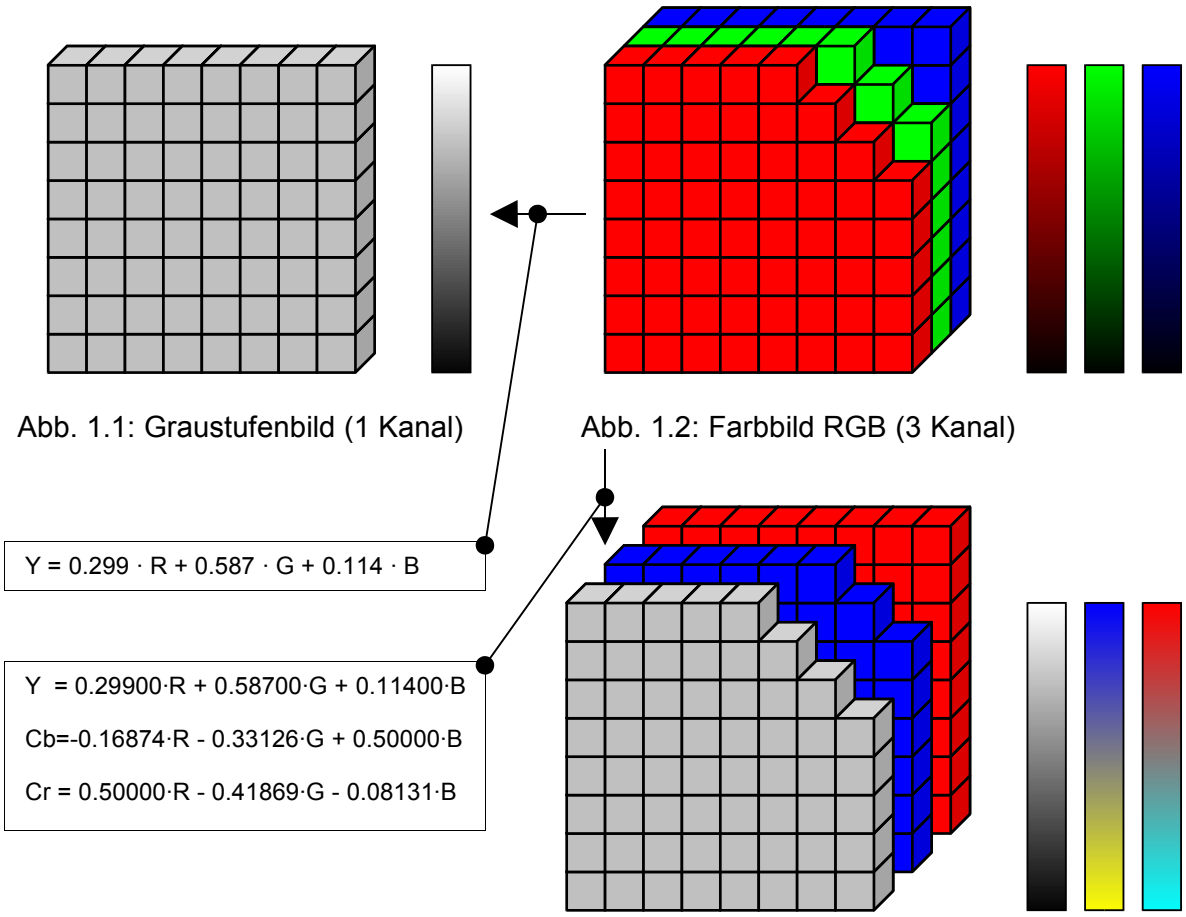


Abb. 1.1: Graustufenbild (1 Kanal)

Abb. 1.2: Farbbild RGB (3 Kanal)

Abb. 1.3: Farbbild YCbCr (3 Kanal)

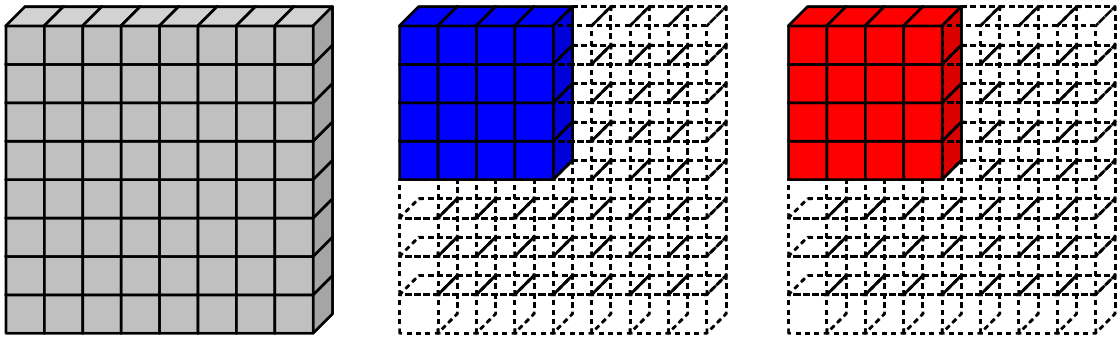


Abb. 1.4: Farbbild YCbCr (3 Kanal) mit Farbunterabtastung (Color Subsampling)

Abb. 1: Schematische Darstellung der Bildformate (räumlicher Bereich – spatial domain)

Tabelle 1 zeigt die wesentlichen Bild-Parameter bei der digitalen Darstellung. Für die verbreitetsten praktischen Anwendungen haben sich sinnvolle Maximalwerte herausgestellt, deren Überschreitung keine nennenswerten Verbesserungen der wahrnehmbaren Bildeigenschaften mehr bringen würden (z.B. 24 Bit Farbtiefe für RGB-„True-Color"-Format; 1024 x 768 Pixel für RGB-Monitor-Betrachtung). Dennoch ist der Platzbedarf zur Speicherung bzw. Bandbreitenbedarf zur Übertragung von digitalen Bildern in diesem Modell noch so hoch, dass Interesse an einer möglichst reduzierten Darstellungsweise besteht. Während die Farbtiefenauswahl meist relativ feststehend bzw. begrenzt ist, gibt es insbesondere bei der räumlichen Auflösung eine größere Vielfalt bei der Parameterwahl, so dass hierbei oft auch Methoden der Anpassung eines Bildes an unterschiedliche Auflösungen gefragt sind, was aufgrund der Vielzahl und unterschiedlichen Charakteristik der entsprechenden Verfahren gewisse Probleme mit sich bringen kann.

Bild-Parameter:	**Möglichkeiten/Methoden der Änderung/Reduktion**	**mittels**
- räumliche Auflösung	Reskalierung/Resampling	Rekonstruktion/Interpolation
- Bittiefe/Farbtiefe	Quantisierung/Dithering	

Tab. 1: Bild-Parameter und Möglichkeiten der Änderung

Bei der Verarbeitung von RGB-Farbbildern gemäß Abb. 1.2 fällt etwa der 3-fache Aufwand im Vergleich zum Graustufenbild (Abb. 1.1) an, da jeder der 3 RGB-Farbkanäle gleichartig und mit vergleichbaren Parametern entsprechend der einen Komponente beim Graubild behandelt wird.

Für verschiedene Bildverarbeitungs- und Bildübertragungsprozesse hat es sich deshalb als vorteilhaft erwiesen, das technische RGB-Farbbild in ein YCbCr-Farbbild mit getrennter (separierter) Luminanz (Y) und Chrominanz (Cb und Cr) zu überführen (Abb. 1.3).
Dies ermöglicht es, den Chrominanz-Anteil des Bildes (Komponenten Cb und Cr) mit anderen Parametern und Verfahren darzustellen und zu verarbeiten, die gegenüber der Luminanz reduziert sein können, ohne dadurch den empfindlicheren Luminanz-Anteil zu beeinflussen.

Dieser Sachverhalt wird in Abb. 1.3 durch die Trennung der 3 Farbkanäle im Vergleich zur engen Haftung in Abb. 1.2 verdeutlicht.
Man beachte, dass die in Abb. 1.2 und 1.3 rot und blau dargestellten Farbkomponenten jeweils unterschiedliche Bedeutung haben (Farbwertbalken rechts). Im YCbCr-Farbmodell handelt es sich dabei um Farbpolaritäten.

Eine verbreitete Methode zur Reduzierung des Umfangs der Farbkomponenten wird in Abb. 1.4 veranschaulicht. Dabei wird die räumliche Auflösung der Farbkomponenten jeweils je Dimension auf die Hälfte (flächenmäßig auf ein Viertel) im Vergleich zur Luminanz-Komponente reduziert.

Ein Nachteil dieser Methode besteht darin, dass zur Rück-Umwandlung in ein RGB-Farbbild zwecks Ausgabe in der Nenn-Auflösung (Luminanz Y) eine partielle Auflösungsanpassung (Vergrößerung des Farbanteils) erfolgen muss, was gewisse Probleme mit sich bringen kann (s.o.).

Die hier besprochene Vielfalt der Bilddarstellung im räumlichen Bereich kann man konkret sehr gut am Beispiel der möglichen Bildformate im DCAM/IIDC-Standard [1] zur paketorientierten Übertragung digitaler Bildinformationen im IEEE-1394-Protokoll nachvollziehen.

2 DCT-basierte Bilddarstellung/Bildrepräsentation

Abbildung 2 zeigt die Verfahrensweise bei digitalen Bildformaten mittels DCT (Diskrete Cosinus-Transformation), wobei jede Komponente üblicherweise in 8x8-Pixel-Blöcke zerlegt und die Transformation jeweils blockweise unabhängig durchgeführt wird.

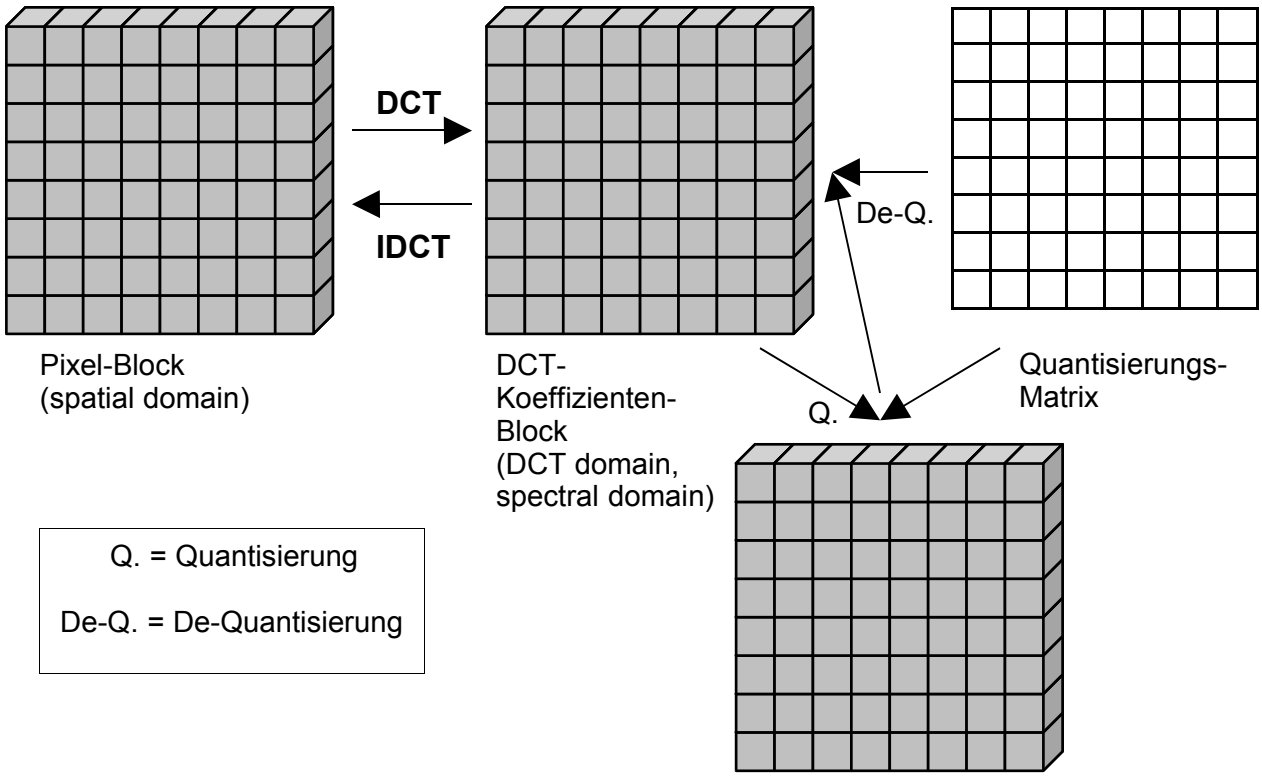


Abb. 2: Digitale Bildformate mit DCT (Diskrete Cosinus-Transformation)

Die (eindimensionale) NxN Diskrete Cosinus-Transformation ist gegeben durch [2]:

DCT:

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

mit

$$C(u) = \frac{1}{\sqrt{2}} \quad für \quad u = 0$$
$$C(u) = 1 \qquad für \quad u > 0$$
$$s(x) = 1 - D\ Pixelwert$$
$$S(u) = 1 - D\ DCT - Koeffizient$$

Die (eindimensionale) NxN Inverse Diskrete Cosinus-Transformation ist gegeben durch:

IDCT:

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

DCT-basierte Bildformate finden breite Anwendung in den JPEG [2, 3], MPEG [4, 5] und DV [6] Standards.

Ein Vorteil des Einsatzes der DCT besteht darin, dass eine flexible Datenreduktion im DCT-Bereich mittels einfacher Quantisierung der DCT-Koeffizienten möglich ist.
Mittels unterschiedlicher Quantisierungsmatrizen ist somit auch eine separate Reduzierung der Chrominanz- und Luminanz-Komponenten im YCbCr-Farbmodell möglich, ohne dass zu diesem Zweck auf Methoden der Farbunterabtastung (Color Subsampling) im spatialen Bereich zurückgegriffen werden müsste.

Der JPEG-Standard [3] enthält separate Quantisierungstabellen für Luminanz und Chrominanz, die als Ausgangsbasis für weitere Skalierungen empfohlen werden:

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. 2: JPEG Luminanz- und Chrominanz-Quantisierungstabellen

3 Quantisierung statt Subsampling

Während der JPEG-Standard praktisch die Verwendung beliebiger Quantisierungstabellen separat für jede einzelne Farbkomponente bietet, werden in der praktischen Anwendung meist Einschränkungen bei der Konfiguration der Quantisierungsparameter gemacht.

Es ist sogar so, dass die Auswahl der Quantisierungsparameter meist auf einen einzigen skalaren Wert („Qualitäts"- oder „Komprimierungs"-Wert) beschränkt wird, der sowohl die Luminanz- als auch die Chrominanz-Quantisierungstabelle gleichermaßen skaliert.
Eine stärkere Chrominanz-Reduzierung ist dann nur noch mittels Color-Subsampling-Parametrisierung möglich. Verbreitet ist hierbei ein horizontales und/oder vertikales Color Subsampling mit Faktor 2.

Viele JPEG-Anwendungen benutzen die Software-Bibliothek der Independent JPEG Group (IJG) [7]. Das Tool *cjpeg* der IJG-Software-Distribution bietet zwei Optionen zur Einstellung des Reduktionsgrades: *-quality* als Skalierungswert für die Quantisierung (Wertebereich 0 bis 100, wobei 100 minimalem Verlust mit Quantisierungstabellen mit allen Einträgen gleich 1 entspricht), und *–sample* zur Subsampling-Parametrisierung. Es ist zu beachten, dass als Voreinstellung für das Subsampling der Wert 2x2 gewählt wurde, was einem horizontalen und vertikalen Color Subsampling mit Faktor 2 und damit entsprechenden Farbverlusten entspricht.

Halten wir fest: Um den Reduktionsgrad bei JPEG einzustellen, können bzw. müssen entsprechend der IJG-Software zwei Parameter festgelegt werden:

a) der *–quality* Wert für Quantisierung, und
b) der *–sample* Parameter für Subsampling.

Man beachte auch, dass eine variable Subsampling-Option einen nicht unerheblichen Anteil der Komplexität eines JPEG-Codecs ausmacht sowie weitere Entscheidungen betreffs der Wahl geeigneter Resampling-Verfahren (siehe Kap. 1, Tab. 1) erfordert.

Es stellt sich daher die Frage, ob man nicht auch völlig ohne Subsampling-Option auskommen kann, wenn dies nicht durch systembedingte Vorgaben beim Ein-/Ausgabe-Bildformat erforderlich ist.

Als Ansatzpunkt betrachten wir die getrennten Luminanz- und Chrominanz-Quantisierungstabellen. Wenn wir die stärkere Chrominanz-Reduktion nicht durch Color-Subsampling erreichen wollen, so bleibt uns noch eine entsprechend angepasste Skalierung der Chrominanz-Quantisierungstabelle. Diese könnten wir, ähnlich dem *cjpeg –sample* Parameter, in der Form

cjpeg –quality q1,q2

angeben, wobei q1 der Skalierungswert für die Luminanz und q2 der Skalierungswert für die Chrominanz bedeutet. Dies wäre eine kompatible Erweiterung der bisherigen *–quality* Option: Die Angabe eines einzelnen Wertes q1 bedeutet, dass intern wie bisher mit q2=q1 gerechnet wird. Die *–sample* Option brauchen wir nun effektiv nicht mehr benutzen, vorausgesetzt wir definieren als Voreinstellungsparameter 1x1 statt 2x2.

Wir haben also effektiv den Subsampling Parameter durch einen **separaten Chroma-Qualitäts-(Quantisierungs-)Selektor** ersetzt. Der separate Chroma-Selektor bietet eine feinere Steuerung der Reduktion der Farbkomponenten als die wenigen (groben) möglichen Subsampling-Werte.

4 Interpretation der DCT als „Bildpyramide“

Als grundlegende Ursache der Eignung der DCT für Zwecke der Bildkodierung wird häufig ihre „Dekorrelations“-Eigenschaft angegeben.
Dabei wird jedoch ein wichtiger Aspekt übersehen, der bisher offenbar noch nicht erkannt wurde, obwohl er im Nachhinein sehr anschaulich und einleuchtend wirkt:
Jeder nxn (n=1...8) Teil-Block der 8x8 DCT kann unmittelbar als eine auf n/8 der linearen Auflösung reduzierte Bildversion interpretiert werden. Die Bildpunkte der auflösungsreduzierten Bildversion ergeben sich dabei direkt durch Anwendung einer nxn IDCT gemäß Formel (2) mit N=n auf die gemäß Formel (1) mit N=8 ermittelten Original-8x8-DCT-Koeffizienten! [8, 9]

Diese praktisch sehr nützliche Eigenschaft ist in Abbildung 3 veranschaulicht. Der DCT-Block ist dazu mit dem DC-Koeffizienten in der unteren linken Ecke orientiert. Jedes „Spektralband“ mit einer bestimmten horizontalen oder vertikalen Frequenz fügt dabei genau die Details der zugehörigen Auflösungsstufe hinzu. Mit anderen Worten, die DCT stellt eine direkte Separation des Bildes in eine Folge mit schrittweise zunehmender Auflösung („Bildpyramide“) dar.
Eine derartige Eigenschaft wird gewöhnlich Verfahren mit „Subband Coding“ bzw. „Wavelet“ als Vorteil zugeschrieben. Während jedoch mittels Subband Coding oder Wavelet nur eine Auflösungs-Separation in *geometrischer* Folge erreicht wird, liegt bei der DCT sogar eine feinere *arithmetische* Auflösungsfolge vor, d.h. Details werden in additiver Schrittfolge (1/8 Bildgröße) statt in multiplikativer Schrittfolge (üblicherweise mit Faktor 2) hinzugefügt.

Dieses Resultat ist in Tabelle 3 zusammengefasst, und es erklärt uns nun auch den Zusammenhang zwischen Quantisierung im DCT-Bereich und Auflösung im spatialen Bereich: Indem wir die höherfrequenten Koeffizienten stärker (gröber) quantisieren, entfernen wir äquivalent Details der entsprechenden Auflösungsstufe. Oder andersherum: Setzen wir beispielsweise alle Koeffizienten *ausserhalb* des 4x4 Teil-DCT-Blocks auf Null (gröbste Quantisierung), so können wir immernoch ‚exakt' ein Bild der halben Auflösung rekonstruieren, da nur die Details der höheren Auflösungsstufen entfernt sind. Das Null-Setzen (spezielle Art der Quantisierung) höherwertiger „Frequenzbänder" (gemäß Abb. 3) entspricht also unmittelbar einem Subsampling im spatialen Bereich!

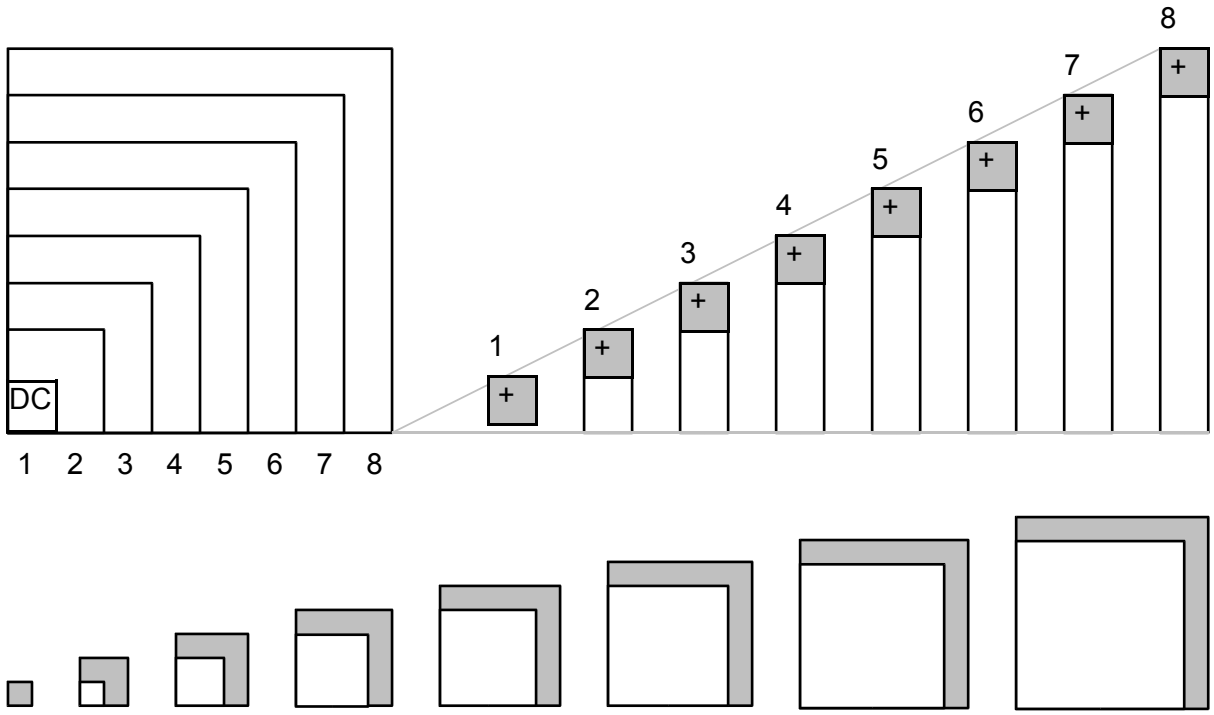


Abb. 3: DCT-Block Auflösungsbänder und additive Auflösungsfolge

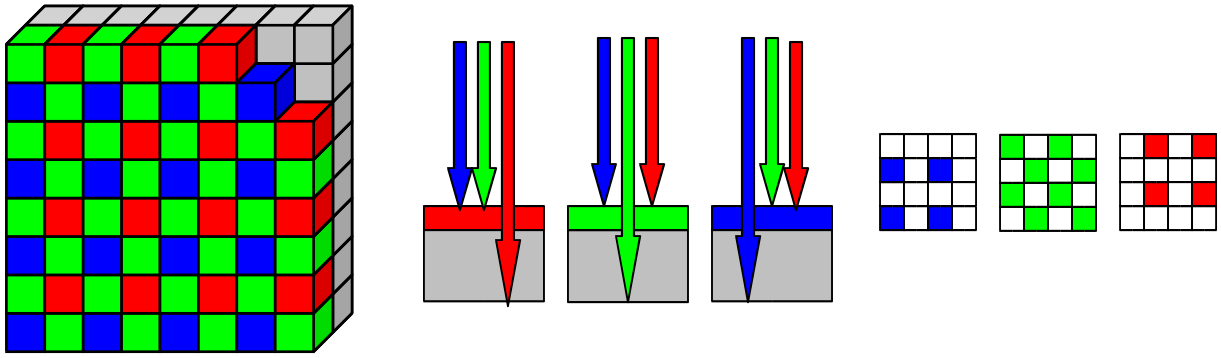
DCT :	Auflösungs-Separation mit ***arithmetischer*** Progression
Subband Coding / Wavelet :	Auflösungs-Separation mit ***geometrischer*** Progression

Tab. 3: Eigenschaften von Bildtransformationen

5 Vollfarbensensor (Foveon X3) statt Farbmosaiksensor (Bayer)

Die traditionelle Methode der Farbunterabtastung zwecks Reduktion des Farbdatenumfangs erweist sich als besonders nachteilig bei der Speicherung von True-Color-Vollfarbenbildern, wie sie von digitalen Kameras aufgenommen werden, die mit dem neuartigen Foveon-X3-Sensor [10] ausgerüstet sind. In herkömmlichen Digitalkameras mit Farbmosaikprinzip, meist im Bayer-Schema angeordnet (Abb. 4), wird pro Bildpixel nur eine Farbkomponente gemessen, während die anderen zwei Farbkomponenten herausgefiltert werden, wodurch bereits ein hoher Anteil (2/3) der Farbinformation verlorengeht. Der Foveon-X3-Sensor beruht auf der Tatsache, dass das Halbleitermaterial (Silizium) unterschiedliche Farbtöne in unterschiedlichen Tiefen absorbiert und somit durch eine 3-schichtige Anordnung der Photorezeptoren volle 3 Farbwerte pro Pixel gemessen werden können. Abbildung 4 stellt die Unterschiede beider Verfahren schematisch gegenüber.

Bayer Mosaik Aufnahme



Foveon X3 Aufnahme

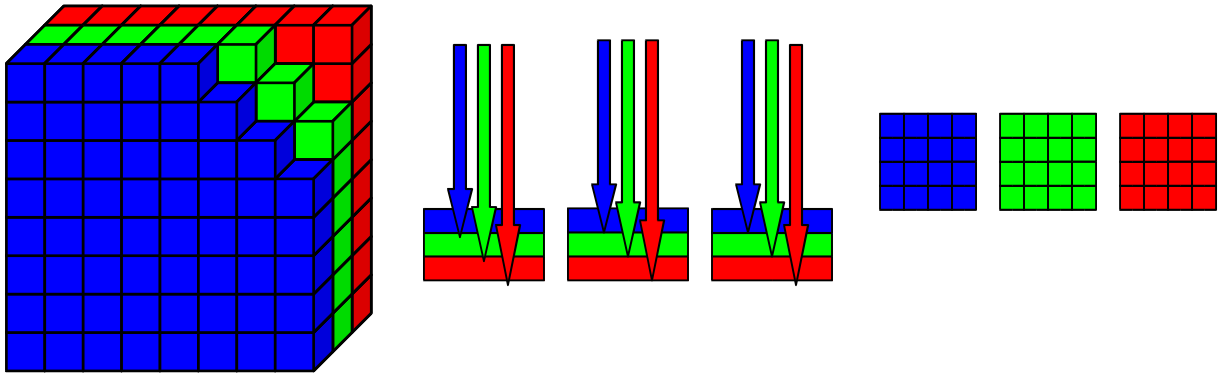


Abb. 4: Schematische Gegenüberstellung Farbmosaik- und Vollfarben-Aufnahme

6 Literatur / Veröffentlichungen

[1] IIDC 1394-based Digital Camera Specification
Version 1.30 July 25, 2000
Instrumentation and Industrial Control Working Group (II-WG) of the 1394 Trade Association, Digital Camera Sub Working Group (DC-SWG)

[2] William B. Pennebaker, Joan L. Mitchell:
"JPEG Still Image Data Compression Standard",
Van Nostrand Reinhold, 1993, ISBN 0-442-01272-1.

[3] JPEG Spezifikation ISO/IEC IS 10918-1 und CCITT Recommendation ITU-T T.81

[4] MPEG-1 Spezifikation ISO/IEC 11172

[5] MPEG-2 Spezifikation ISO/IEC 13818

[6] DVCAM Format Overview
http://www.sony.ca/dvcam/pdfs/dvcam%20format%20overview.pdf

[7] The Independent JPEG Group's software
http://www.ijg.org

[8] Andrew B. Watson, Ames Research Center, Moffett Field, California
"Ideal Shrinking and Expansion of Discrete Sequences" (1986)
http://vision.arc.nasa.gov/publications/IdealShrinking.pdf

[9] Rakesh Dugad, Narendra Ahuja
"A Fast Scheme for Image Size Change in the Compressed Domain" (2001)
http://vision.ai.uiuc.edu/~dugad/research/dct/index.html

[10] Foveon, Inc. – X3 Technology
http://www.foveon.net