Bilddatenkompression mit geplanten Verlusten und hoher Rate

Rekonstruktion von Bildern aus Helligkeitskanten

Volker Aurich Uwe Daub
Abteilung für Informatik, Mathematisches Institut
Heinrich-Heine-Universität Düsseldorf
{aurich,daub}@cs.uni-duesseldorf.de

Zusammenfassung

Es wird ein Verfahren zur Kompression von Grauwertbildern beschrieben, das sehr hohe Kompressionsraten erreicht, ohne Artefakte zu erzeugen, die eine maschinelle Weiterverarbeitung erschweren. Gleichzeitig wird demonstriert, daß in geeignet extrahierten Helligkeitskanten genügend viel Information enthalten ist, um damit aus den Grauwerten einer groben Unterabtastung ein dem Ausgangsbild sehr ähnliches Bild rekonstruieren zu können.

1 Einführung

Verlustfreie Kompressionsverfahren erreichen bei Bilddaten meist nur geringe Kompressionraten. Daher wurden für Bilddaten spezielle Kompressionsverfahren entwickelt, die auf Kosten von Informationsverlusten höhere Kompressionsraten gestatten. Die Art der Informationsverluste ist bei den gängigen Verfahren eine eher zufällige Folge der gewählten Methode (DCT, Wavelets, Fraktale) und nicht gezielt geplant. Bei hohen Kompressionsraten treten verfahrensbedingte Artefakte auf wie Blockstrukturen, wellenartige Muster und Verschmieren von Helligkeitskanten, die unerwünscht sind und die maschinelle Weiterverarbeitung der Bilder erschweren.

Die vorliegende Arbeit beschreibt hingegen ein Kompressionsverfahren, das den unvermeidlichen Informationsverlust seiner Natur nach a priori so festlegt, daß eine maschinelle Weiterverarbeitung sogar erleichtert wird. Der Informationsverlust besteht im wesentlichen darin, daß feine Bilddetails unterhalb einer vorgegebenen Größenordnung und Kontrasthöhe eliminiert werden, während die gröberen Bildstrukturen sehr exakt erhalten bleiben.

Dementsprechend gliedert sich der Kompressionsvorgang ganz grob in zwei Schritte: einen, der durch gezielte Einführung von Verlusten eine sogenannte Abstraktion des Ausgangsbildes berechnet, und einen zweiten, der dieses abstrahierte Bild mit Standardverfahren verlustfrei komprimiert. Im ersten Schritt werden zunächst die feineren Bildstrukturen mit einem kantenerhaltenden Glättungsverfahren eliminiert. Anschließend wird mit einem nichtlinearen Differenzenfilter ein Ternärbild mit den Werten `+`, `-`, `0` gewonnen, in dem die Helligkeitskanten des geglätteten Bildes durch Vorzeichenwechsel beim Übergang von einem Pixel zu seinem Nachbarn kodiert sind. Dieses Ternärbild zusammen mit 1024 Grauwerten, die durch Abtastung des geglätteten Bildes in einem groben Gitter (32×32 Abtastpunkte) gewonnen werden, stellt die verlustbehaftete Abstraktion des ursprünglichen Bildes dar. Die in der Abstraktion enthaltene Information kann durch geschickte Anwendung von Standardkompressionsverfahren sehr stark, aber verlustfrei komprimiert werden.

Auch der Dekompressionsvorgang ist mehrstufig. Der wesentliche Schritt ist eine neuartige, kantenerhaltende Interpolation, die die großflächigeren Teile eines Bildes aus seiner Abstraktion rekonstruiert. Anschließend werden die bisher rekonstruierten Grauwerte unter Beachtung der Kanten soweit wie möglich ausgedehnt. Um schließlich auch in den

Abbildung 1: Testbild Sailboat (Nr. 14): Original und Rekonstruktion (rechts, 0.37 bpp).

restlichen Bereichen Grauwerte zu definieren, wird ein kantenüberspringender Prozeß verwendet, der die in dem Ternärbild enthaltene Information über die Richtung der Helligkeitsgradienten in den Kanten benutzt.

Das beschriebene Verfahren erreicht bei Bildern sehr unterschiedlicher natürlicher Szenen eine Kompressionrate zwischen 95 % und 99 %. Wie an Bildausschnitten deutlich zu sehen ist (siehe Abbildungen 10 und 6), enthalten die rekonstruierten Bilder keine Artefakte im Sinne von zusätzlichen Strukturen, die in den Ausgangsbildern nicht vorhanden waren, sondern sind eher idealisierte Abstraktionen. Alle Verarbeitungsschritte zur Berechnung von Abstraktion und Rekonstruktion sind lokale Bildoperationen im Ortsbereich, die sich massiv parallel durchführen lassen.

2 Abstraktion

Der erste Schritt des Kompressionsverfahrens dient dazu, im Bildsignal f kontrolliert Informationsverluste einzuführen. Dies geschieht mit dem kantenerhaltenden Glättungsverfahren aus [AW95]. Es entfernt alle Details unterhalb einer vorgegebenen Größenordnung, erhält aber die Kanten gröberer Strukturen sehr exakt. Gegenüber anderen kantenerhaltenden Glättungsverfahren hat es den Vorteil, daß es aus einer Kette lokaler Filteroperationen besteht, die eine massiv parallele Implementierung gestatten.

$$\text{Glättungsoperator} = G_{8\sigma,\tau} \circ G_{4\sigma,2\tau} \circ G_{2\sigma,4\tau} \circ G_{\sigma,8\tau} \circ T_\sigma$$

T_σ ist ein linearer Gaußscher Tiefpaß mit Varianz σ^2. Die anderen Filterstufen sind nichtlineare Modifikationen Gaußsscher Tiefpaßfilter und haben die Gestalt

$$G_{\sigma,\tau} f(p) = \frac{1}{N(p)} \sum_{q \in U(p)} e^{-\frac{||p-q||^2}{2\sigma^2}} e^{-\frac{|f(p)-f(q)|^2}{2\tau^2}} f(q), \quad N(p) = \sum_{q \in U(p)} e^{-\frac{||p-q||^2}{2\sigma^2}} e^{-\frac{|f(p)-f(q)|^2}{2\tau^2}}$$

wobei $U(p)$ eine Umgebung von p ist und $N(p)$ zur Normierung des Gesamtgewichtes dient. Die Werte von σ und τ bestimmen die maximale Größenordnung bzw. den maximalen Kontrast der Details, die eliminiert werden.

Die Grauwerte des so geglätteten Signals werden auf einem groben Gitter (32×32) abgetastet. Zur Extraktion der Kanteninformation wird das geglättete Signal durch ein

(a) Original (b) geglättet (c) Kanten (d) Abtastwerte

Abbildung 2: Der Weg zur „Abstraktion“.

DOG-ähnliches Differenzenfilter (siehe [Weu94]) geschickt, das durch Einfügung einer Nichtlinearität unempfindlich gegen Rauschen gemacht wurde. Das Ergebnissignal hat längs der Helligkeitskanten des geglätteten Bildes auf der helleren Seite einen Saum positiver Werte und auf der dunkleren Seite einen Saum negativer Werte; ansonsten sind seine Werte infolge der Nichtlinearität exakt 0. Durch Vorzeichenbildung ensteht daraus ein Ternärbild, in dem jeder Wechsel von + nach − oder umgekehrt nicht nur die Existenz einer Kante zwischen zwei Pixeln, sondern auch die Richtung der Helligkeitsänderung angibt. Dieses Ternärbild zusammen mit den abgetasteten Grauwerten wird als Abstraktion des Ausgangsbildes bezeichnet (siehe Abbildung 2).

3 Rekonstruktion

Der wesentliche Schritt der Dekompression ist ein Algorithmus, der zu jedem abstrahierten Bild eine vollständige Grauwertverteilung rekonstruiert. Die grundlegende Idee hierbei ist, den Grauwert eines Pixels p durch geeignete Interpolation aus den Grauwerten der von dort aus sichtbaren Abtastpunkte zu mitteln; dabei heiße ein Pixel q von p aus sichtbar, wenn die Verbindungsstrecke der Pixelmittelpunkte keine der in dem Ternärbild kodierten Kanten kreuzt. Diese Kanten bilden also die Grenzen des Bereichs, über den gemittelt wird. Daher sind sie dann auch im rekonstruierten Bild deutlich ausgeprägt und scharf.

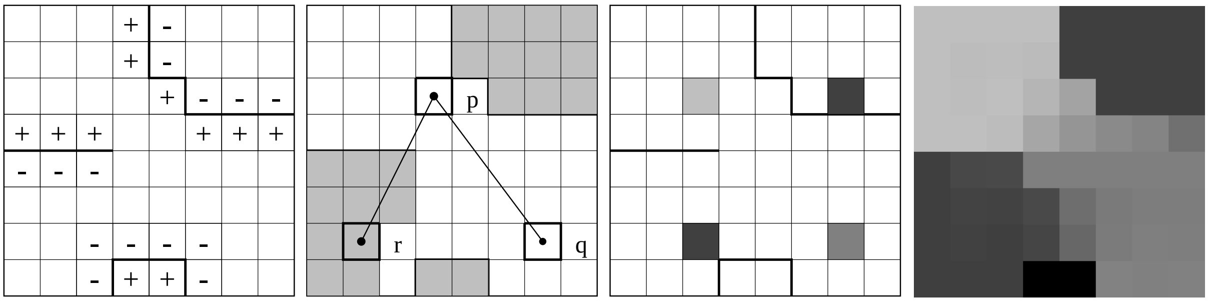
Die kantenerhaltende Grauwertinterpolation wurde auf der Grundlage der Shepard-Interpolation (siehe [Sh68]), die allerdings noch keine Interpolationsbarriere kennt, entwickelt.

Zur Interpolation wird der folgende Operator I verwendet:

$$I(p) = \begin{cases} f_i & \text{falls } p = p_i \\ \Big(\sum\limits_{p_i \in \mathcal{S}(p)} \frac{f_i}{||p - p_i||^2} \Big) \Big/ \Big(\sum\limits_{p_i \in \mathcal{S}(p)} \frac{1}{||p - p_i||^2} \Big) & \text{sonst} \end{cases}$$

Dabei bezeichnen die f_i die Abtastwerte an den Stützstellen p_i. Aus der Information über den Kantenverlauf wird der Sichtbarkeitsbereich $\mathcal{S}(p)$ für jedes einzelne Pixel p bestimmt. Dieser Sichtbarkeitsbereich bestimmt so die Grenzen der Interpolation. (Der Sichtbarkeitsbereich $\mathcal{S}(p)$ des Pixels p ist in Abbildung 3(b) weiß markiert.)

Weil im Sichtbarkeitsbereich eines Pixels kein Abtastpunkt des abstrahierten Bildes liegen muß, wird das Interpolationsverfahren unter ständiger Vergrößerung des Bereichs bereits definierter Grauwerte iteriert. Um schließlich auch noch in Gebieten, die keinen Abtastwert enthalten, Grauwerte festzulegen, wird ein kantenüberspringender Schritt angeschlossen. Hierbei wird die im abstrahierten Bild enthaltene Information über die Richtung des Helligkeitsgradienten benutzt und die Graustufe um eine Konstante erhöht bzw.



(a) Kanten (b) Bereich $\mathcal{S}(p)$ (c) Abtastwerte (d) Interpolation

Abbildung 3: Interpolation mit Sichtbarkeitsbereich.

Abbildung 4: Kantenbild, Quadtreeausgleich und erzielte Verkürzung.

erniedrigt, je nachdem ob in dem Ternärbild ein Übergang von `-` nach `+` oder umgekehrt stattfindet. Dadurch entsteht zwar fast immer eine Abweichung des Grauwerts eines Pixels von dem im Originalbild; es entstehen jedoch keine nennenswerten zusätzlichen Strukturen, weil der Grauwert in der richtigen Ordnungsrelation zu denen der Pixel in seiner Umgebung steht.

4 Verlustfreie Kompression

Die in der Abstraktion eines Bildes enthaltenen Abtastwerte werden durch zeilenweisen Durchlauf serialisiert und als Textstring verlustfrei komprimiert. Dafür wird eine arithmetische Kodierung 2. Ordnung gewählt, weil sie im Vergleich mit mehreren Standardverfahren die höchste Kompressionsrate erreicht.

Die in der Abstraktion enthaltene Kanteninformation in Gestalt des Ternärbildes ist meist wesentlich umfangreicher als die Grauwertinformation und erfordert eine andere Art der Kodierung. Um die Korrelation zwischen den Pixeln des Ternärbildes auszunutzen, wird es als ein Quadtree dargestellt, der durch einen Preorder-Durchlauf serialisiert und als Textstring mit den vier Symbolen `+`, `-`, `0` und `>` dargestellt wird. Weil für die Kanteninformation die Position der Übergänge zwischen `0` und `+` bzw. zwischen `0` und `-` unbedeutend ist, kann man diese Übergänge in geschickter Weise so verrücken, daß der durch die Serialisierung entstehende Textstring um 30% bis 50% kürzer wird (siehe Abbildung 4). Außerdem wird durch diesen Ausgleich der Quadtreegrenzen bewirkt, daß die Häufigkeiten der vier Symbole des Textstrings sehr unterschiedlich werden. Dadurch wird die nachfolgende verlustfreie Kompression des Textstrings begünstigt. Ausführliche Untersuchungen zeigten, daß eine arithmetische Kodierung 4. Ordnung eine deutlich höhere Kompressionsrate erreicht als andere übliche Verfahren wie LZ77 oder Huffmankodierungen.

Es wurden etliche völlig unterschiedliche, natürliche Grauwertbilder auf diese Weise komprimiert (siehe Abbildung 13). Dabei fiel die Datenrate von 8 bit pro Pixel im Ori-

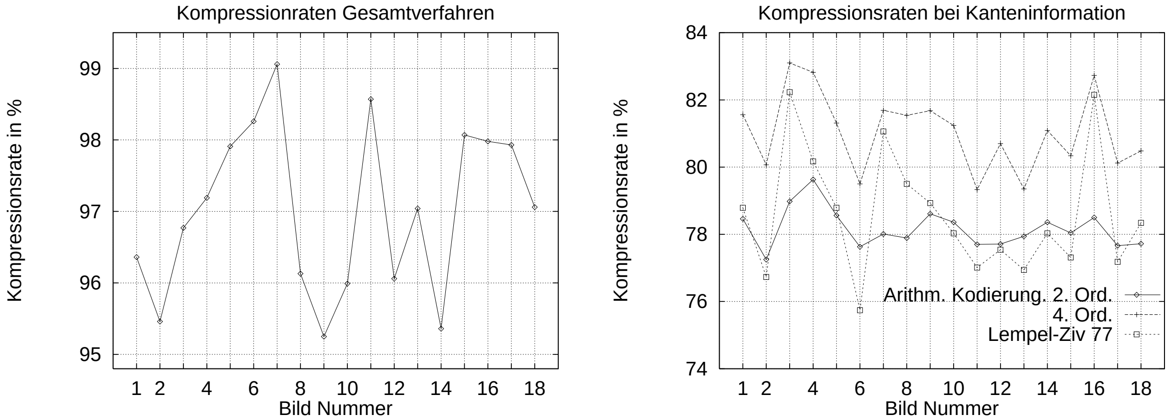
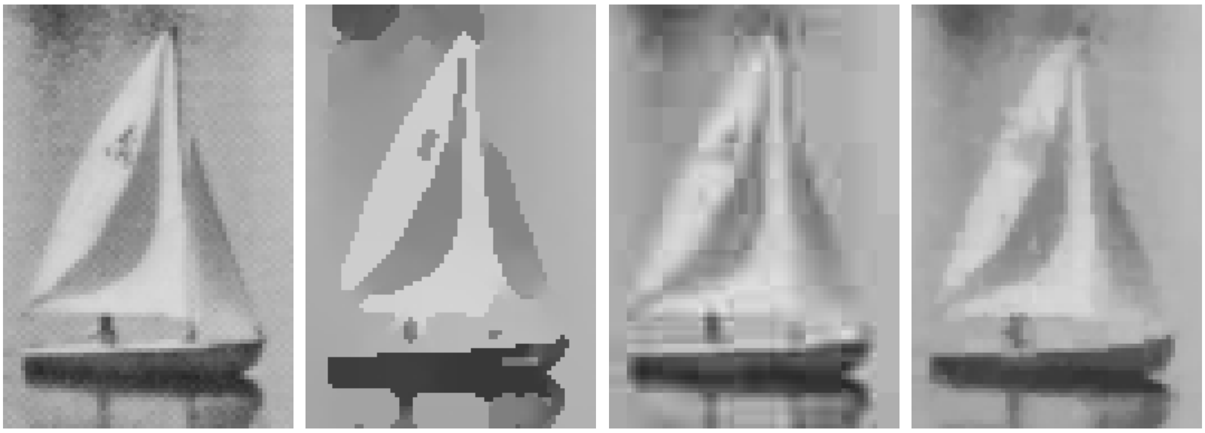


Abbildung 5: Kompressionsrate für das Gesamtverfahren und für die Kanteninformation.



(a) Original (b) KdA (c) JPEG (d) YF

Abbildung 6: Sailboat Ausschnitte

ginalbild auf 0.08 bis 0.38 bit pro Pixel; das entspricht einer Kompressionrate zwischen 99 % und 95 % (siehe Abbildung 5).

5 Vergleich mit anderen Verfahren

Das oben beschriebene Kompressionsverfahren, das wir KdA (Kompression durch Abstraktion) nennen, wurde mit einigen bekannten Verfahren verglichen, und zwar mit (Baseline-) JPEG, „Efficient Pyramid Image Coder (EPIC)“ (siehe [AS90]) und YF (benannt nach dem Autor Yuval Fisher; siehe [Fi94]), die auf der diskreten Kosinustransformation, Wavelets bzw. fraktaler Kompression beruhen. Diese Verfahren liefern für einen menschlichen Betrachter exzellente Ergebnisse, solange die Kompressionsrate nicht extrem hoch ist (bis etwa 90 %).

Da es kein adäquates Maß für den Informationsgehalt eines Bildes gibt, wurde der Vergleich durch subjektive Betrachtung verschiedener Bilder durchgeführt. Als Testbilder wurden 18 Bilder völlig unterschiedlicher, natürlicher Szenen verwendet. Der Parameter τ wurde immer zu $\tau = 32$ gewählt und der Parameter σ variierte zwischen $\sigma = 0.5$ und $\sigma = 1.5$. Die Parameter der anderen Kompressionsverfahren wurden für jedes Bild so eingestellt, daß sie eine ähnlich hohe Kompressionsrate erzielten wie das KdA-Verfahren (95 % bis 99 %). In allen Fällen erzeugten die Vergleichsverfahren Artefakte, die zwar für einen menschlichen Betrachter nicht allzu störend sind, solange das Bild nur recht klein reproduziert wird, die aber eine maschinelle Weiterverarbeitung erschweren. Dies wird deutlich, wenn man sich vergrößerte Ausschnitte der dekomprimierten Bilder ansieht. In

Abbildung 7: Testbild Lena (Nr. 10): Original und Rekonstruktion (rechts, 0.32 bpp).

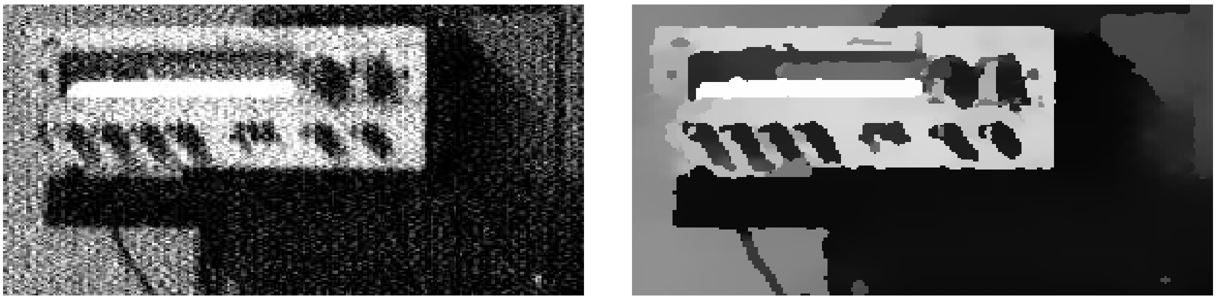


Abbildung 8: Testbild Radio (Nr. 13): Original und Rekonstruktion (rechts, 0.24 bpp).

den mit KdA komprimierten Bildern dagegen sind die aus dem Originalbild extrahierten Kanten explizit und unverfälscht enthalten, so daß jede auf Kanten basierende Weiterverarbeitung verlustfrei möglich ist.

Literatur

[AS90] Adelson, Edward H.; Simoncelli, Eero P. *Subband Image Coding with Three-tap pyramids.* Picture Coding Symposium 1990, Cambridge MA. (EPIC-Code `ftp://whitechapel.media.mit.edu/pub/epic.tar.Z`)

[AW95] Aurich, Volker; Weule, Jörg. *Non-Linear Gaussian Filters Performing Edge Preserving Diffusion.* In: Mustererkennung 1995. Springer Verlag 1995 (Reihe: Informatik Aktuell). (Proc. 17. DAGM-Symposium.)

[Fi94] Fisher, Yuval. *Fractal Image Compression: Theory and Application* Springer Verlag 1994. (Code `http://inls.ucsd.edu/y/Fractals/`)

[Sh68] Shepard, Donald. *A two dimensional interpolation function for irregulary-spaced data.* Proc. 23rd National Conference ACM, 1968, S. 517–524.

[Weu94] Weule, Jörg. *Iteration nichtlinearer Gauß-Filter in der Bildverarbeitung.* Dissertation Heinrich-Heine-Universität Düsseldorf 1994.

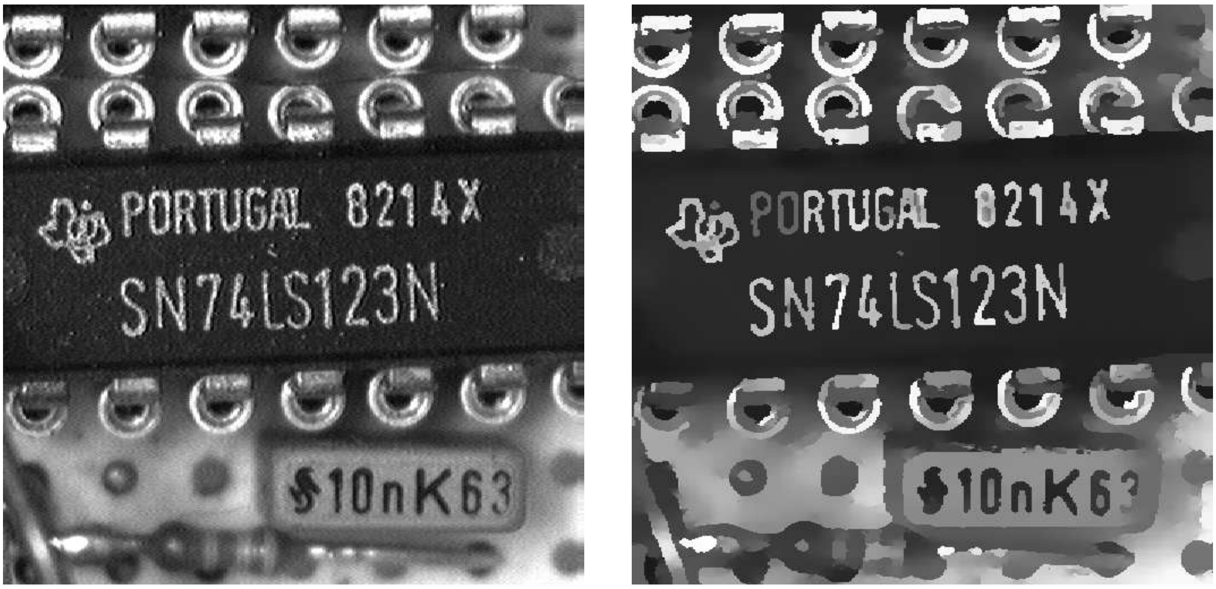


Abbildung 9: Testbild IC (Nr. 9): Original und Rekonstruktion (rechts, 0.38 bbp).

(a) Original

(b) KdA

(c) JPEG

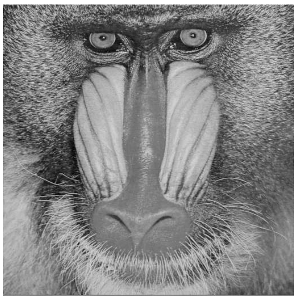
(d) YF

Abbildung 10: IC Ausschnitte

Abbildung 11: Zwei weitere Rekonstruktionen.

Abbildung 12: Ausschnittvergrößerungen

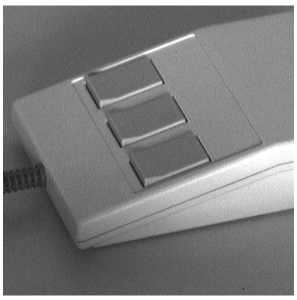
(a) Nr. 1 (0.29 bbp)



(b) Nr. 2 (0.36 bpp)

(c) Nr. 7 (0.08 bpp)

(d) Nr. 8 (0.31 bpp)



(e) Nr. 11 (0.11 bpp)

(f) Nr. 12 (0.32 bpp)

(g) Nr. 16 (0.16 bpp)

(h) Nr. 17 (0.17 bpp)

Abbildung 13: Einige der Testbilder.