

# Bildkompression

Johannes Dambacher

Otto-von-Guericke-Universität Magdeburg  
johannes.dambacher@st.ovgu.de

**Zusammenfassung.** Ohne Bildkompression wäre unser alltäglicher Umgang mit Bildern nicht möglich. Sie erlaubt es mittels Handys und Digitalkameras Bilder zu machen und diese über das Internet zu verbreiten. Die Bilder selbst finden vielfältige Anwendung in Wissenschaft, Industrie, Medizin, Kunst und nicht zuletzt im Privaten. Dieser Text wird kurz auf verschiedene Möglichkeiten der verlustlosen und verlustbehafteten Kompression eingehen, verwendete Standards und Formate beschreiben sowie deren Vor und Nachteile darstellen. [1]

## 1 Introduction

Ohne Bildkompression wäre der digitale Bildaustausch über das Internet ebenso wie Bau günstiger Digitalkameras und Handys mit Kamerafunktion teurer und ein vielfaches an Speicher und Datenübertragung notwendig. Dies liegt an der Menge Speicher die nötig ist um Bildinformation zu speichern. Ziel der Bildkompression ist es den Speicherbedarf eines Bildes möglichst zu reduzieren, ohne dabei jedoch Information zu verlieren. Dies ist möglich da Bilder redundante Information enthalten und die menschliche Wahrnehmungsfähigkeit begrenzt ist. Im Falle der Kompression ist zwischen verlustloser und verlustbehafteter Kompression zu unterscheiden. Bei ersterer besteht zwischen decodiertem Bild und Original kein Unterschied, bei letzterem enthält das decodierte Bild weniger Information als das Original. Dies stellt bei Bildern und Fotos die für die menschliche Betrachtung gedacht sind meist kein Problem dar, da sich für Menschen kein Unterschied feststellen lässt.

Als Beispiel für die Notwendigkeit und den Nutzen der Bildkompression sei ein einfarbiges Bild ohne Rauschen gegeben. Wird dieses Pixelweise in einer Matrix mit Seiten  $M$  und  $N$  abgespeichert, enthält es redundante Information, da sich keine Pixel unterscheiden. Dennoch bedarf es für ein Binärbild  $M \cdot N$  Bits zur Speicherung und für ein Farbbild schon  $M \cdot N \cdot \text{drei Byte}$ , obwohl es ausreicht nur die Seitenverhältnisse und die Farbe zu speichern um das Bild ohne Informationsverlust wiederherstellen zu können.

Enthält das Bild Linien, reicht es diese über Start-, Endpunkt, Farbe und Breite zu definieren, anstatt Pixelweise die gesamte Matrix. Auf ähnliche Weise lassen sich Kurven und geometrische Figuren ebenso wie Farbverläufe codieren. Diese Art der Speicherung findet bei sogenannten Vektorgraphiken Verwendung. Im Gegensatz zur Speicherung als Matrix bedarf es bei der Speicherung als Vektorgrafik weniger Speicherplatz,

jedoch ist die Decodierung aufwändiger. Als Vorteil kommt auch die Skalierbarkeit der Vektorgrafik hinzu. Verwendet werden Vektorgrafiken aus diesem Grund fast ausschließlich für digital erzeugte Bilder wie beispielsweise Diagramme, Logos oder Zeichensätze.

Folgend wird auf die Kompression und Codierung von Fotos und Bildern eingegangen. Bei diesen ist eine Speicherung als Vektorgrafik ungünstig, da weder gleichmäßige Farbflächen und Verläufe auftreten und selbst einfache geometrische Objekte wie Linien durch Rauschen und Artefakte gestört sind. Hinzu kommt das ein Foto erst auf Linien und Farbverläufe untersucht werden muss, bevor es anhand dieser codiert werden kann. Die Suche nach Codierbaren Formen ist jedoch rechen aufwändig und nicht in akzeptabler Zeit auf Geräten wie Handys und Kameras zu realisieren. Dennoch ist auch in Fotos Information redundant vorhanden. Betrachtet man ein Bild mit blauem Himmel tragen die einzelnen Pixel innerhalb des Himmels wenig Information.

Die Notwendigkeit und der Wunsch Bilder zu Komprimieren entstand mit der Verbreitung leistungsstarker Computer und Farbmonitoren. Digital entstandene Bilder lassen sich bereits mit Methoden der Text- und Datenkomprimierung codieren. Hinzu kommt die Möglichkeit diese als Vektorgrafik zu speichern, wodurch es erst mit der Einführung digitaler Bildsensoren notwendig wurde speziell für digitale Fotos Kompressionsverfahren zu entwickeln. Zu dieser Zeit bereist verfügbare Codierverfahren eignen sich nicht für die Bildkompression. Bei der Entropiecodierung ist es die Symbolmenge in Farbbildern, bei der Lauffängencodierung werden die Runs aufgrund des Bildrauschens im Extremfall nur noch ein Pixel lang, wodurch sogar codiert mehr Speicher verwendet wird als uncodiert. Im Zuge dessen entstanden verschiedene Formate zur Speicherung von digitalen Fotos. Verwendet wurden neben frequenzbasierten Verfahren wie der Diskreten Cosinus oder Fourier Transformation, auch Fraktale Kompressionsarten und andere Methoden. 1992 wurde von der Joint Photographic Experts Group (JPEG) - Komitee nach sechs Jahren arbeit erstmals ein Internationaler Standard zur Komprimierung digitaler Bilder vorgestellt. Obwohl darin keine Spezifikation zur Speicherung der codierten Daten gemacht wird, ist der Standard inzwischen meist verbreiteter Standard unter den verlustbehafteten Kompressionsverfahren.

Ein neuerer Standard des JPEG-Komitees zur verlustbehafteten Kompression stellt JPEG2000 dar.

WebP?

Ein weiteres weit verbreitetes Format stellt das Portable Network Graphics (PNG) Format dar. Es ist das meist verbreite verlustfrei komprimierende im Internet verwendete Format.

Im medizinischem Bereich hat sich der DICOM-Format etabliert. Es stellt einen offenen Standard dar, welcher zur Speicherung und zum Austausch medizinischer Bilder verwendet wird. Zusätzlich zur reinen Bildinformation lassen sich Metainformationen wie Patienten- und Klinikdaten speichern. Da der DICOM-Standard keine Kompression unterstützt, wird auf diesen nicht weiter eingegangen.

In Teil 2 werden grundlegende Kompressionsarten welche auch in der Bildkompression Verwendung finden beschrieben, während sich Teil 3 mit den speziell in der Bildkompression verwendeten Verfahren beschäftigt. In

Teil 4 wird abschließend anhand gängiger Bildformate das Zusammenspiel der vorgestellten Kompressionsarten erläutert und Teil 5 gibt einen kurzen Ausblick über das Thema.

## 2 Entropie-,Laufängen- und Prädiktionscodierung

Als grundlegende Verfahren zur Bilddatenkompression eignen sich auch Verfahren welche auch in anderen Gebieten zur Datenkompression eingesetzt werden. Folgend werden deshalb Entropie-, Laufängen- und Prädiktionscodierung vorgestellt. Diese finden - meist kombiniert - bei allen verbreiteten Formaten Verwendung.

Unter Entropiecodierung versteht man die Verwendung unterschiedlich langer Codewörter für die zu codierenden Symbole in Abhängigkeit der Symbolhäufigkeit. Dem am häufigst vorkommendem Grauwert wird das möglichst kürzeste Codewort - ein Bit - vergeben, dem seltenst auftretendem Grauwert das längste Benötigte. Formel zur Symbolhäufigkeit. Damit ist diese Art der Codierung dem Morsecode ähnlich, da auch dort oft vorkommende Buchstaben kurze Codewörter haben. Jedoch muss die Codierung im Gegensatz zum Morsecode eindeutig sein, da sonst zusätzliche Zeichen zur Trennung der Codewörter benötigt werden. Durch die Notwendigkeit der Rundung auf Ganze Bits, oder Zahlen ist die Mathematisch erreichbare Codelänge in der Realität nicht erreichbar, aber annähernd gut. Eine beweisbar optimale Lösung im Binärraum lässt sich mittels Huffmancodierung erzeugen. Mittels Arithmetischer Codierung kann eine bessere Kompressionsrate erreicht werden, da diese im Zahlen- statt Bitraum arbeitet. Nachteil der Arithmetischen Codierung ist ihre vergleichsweise aufwendige Implementierung und Langsamkeit gegenüber der Huffmancodierung. Da die Huffmancodierung praktisch kaum schlechtere Ergebnisse erzielt als die Arithmetische Codierung, wird zumeist auf die Huffmancodierung zurückgegriffen. Entropiecodierung stellt eine Möglichkeit dar, Daten aller Art unter möglichst geringem Speicheraufwand zu codieren und kann zusätzlich zu anderen Komprimierungen verwendet werden. Einziger Nachteil ist Notwendigkeit einer Codierungstabelle welche zu jedem Bild generiert und gespeichert werden muss um bei der Decodierung die Symbole zu dereferenzieren zu können. Im Gegensatz zur Entropiecodierung wird die Laufängencodierung verwendet um häufig auftretende Symbolfolgen statt einzelner Symbole zu codieren. Gerade in Bildern treten häufig Flächen auf in welchen sich die Grauwerte über eine längere Folge nicht oder nur sehr gering voneinander unterscheiden. Eine solche Folge lässt sich einfach speichern indem lediglich der Grauwert und Länge der Folge vermerkt werden. Ein solches Tupel wird gängiger Weise als Run bezeichnet. Auf diese Weise kann ein einfarbiges Bild bereits mit Angabe der Seitengröße und der Farbe des einzigen Runs gespeichert werden, statt die gesamte Bildmatrix zu speichern. Dieser Ansatz funktioniert jedoch nur bei digital erstellten Bildern gut, welche weder Rauschen noch Farbverläufe enthalten. Auf Bildern die Rauschen enthalten - wie es bei Fotos der Fall ist - werden die Runs im schlechtesten Fall nur je einen Pixel lang, so dass eine

Laufängencodierung den Speicherbedarf sogar erhöht statt zu verringern. Dies kann verhindert werden indem die Codierung nicht auf Grauwert-, sondern auf Biteebene durchgeführt wird, da bei einer Grauwertänderung nicht immer alle Bits kippen. Bei Verwendung des Grey- statt Binärcodes wird dieser Effekt weiter verbessert, da durch den Greycode sicher gestellt ist, dass eine Grauwertänderung zu nur genau einem Bit kipp führt. In diesem Fall bleiben immer sieben von acht Runs erhalten. Eine derartige Verwendung der Laufängencodierung führt bei Bildern mit einfarbigen Flächen bereits zu guten Ergebnissen, Farbverläufe stellen jedoch weiterhin ein Problem dar, da die Runs trotz Verwendung des Greycodes und Codierung auf Biteebene immer wieder abbrechen. Dies kann umgangen werden indem die Laufängencodierung nicht auf den Originalgrauwerten selbst, sondern der Differenz zum Vorhergegangenen Grauwert geschieht. Dazu wird am Anfang jeder Bildzeile der Grauwert gespeichert, danach nur noch die Differenz zum Vorgänger. Bei gleichmäßigen linearen Farbverläufen hat dies zur Folge das die Werte gleich sind und dadurch längere Runs möglich werden. Diese Art der Kodierung stellt bereits fast eine Art der Prädiktionscodierung dar. Bei der Prädiktionscodierung wird anhand des bereits dekodierten Bildes, noch während der Dekodierung, eine Vorhersage über den Grauwert des nächsten Pixels getroffen. Bei der Codierung muss nur der Fehler zwischen Erwartetem und Tatsächlichem Grauwert gespeichert werden. Die meisten Prädiktionscodierer verwenden zur Vorhersage nicht nur die aktuelle, sondern auch darüber liegenden Bildzeilen. Dies ermöglicht es auch an Kanten eine Vorhersage der zu erwartenden Grauwerte. Bei guter Abschätzung lassen sich die Fehler mittels Laufängen- und Entropiecodierung besser komprimieren als die original Grauwerte.

Die hier vorgestellten einfachen Codierungsmethoden werden meist kombiniert um eine besser Kompression zu erreichen.

### 3 Transformationscodierung

Die Transformationscodierung versucht mittels Transformation der Bildmatrix, unwichtige Information besser von wichtiger Information abzugrenzen, um anschließend die für das Original unwichtigen Informationsteile zu löschen. Eine anschauliche Art der Transformationscodierung stellt die Transformation aus dem RGB-Farbraum (Rot, Grün, Blau; in Kameras und Bildschirmen verwendet) in den YCbCr-Farbraum (Helligkeit, Blauanteil, Rotanteil; in Fernsehtechnik und JPEG-Standard verwendet) dar. Während im RGB-Raum Farb- und Helligkeitsinformation in allen Kanälen gemeinsam vorkommen sind diese im YCbCr-Raum voneinander getrennt. Die Transformation lässt sich mittels einer Matrixmultiplikation durchführen und ist umkehrbar, wodurch sie sich ohne Einschränkung einsetzen lässt. Betrachtet man nach der Transformation lediglich den Y-Kanal erhält man ein schwarzweiß Bild wie es in früheren Fernsehapparaten Verwendung fand. Der Y-Kanal enthält die gesamte Helligkeitsinformation des RGB-Bildes. Dadurch lässt sich das YCbCr-Signals sowohl für Schwarzweißfernseher als auch Farbfernseher verwenden. Bei Farbbildschirmen muss lediglich der Cr- und Cb-Kanal

hinzu geschaltet werden. Dies stellt eine Möglichkeit der Datenreduktion dar. Denn gibt man sich mit einem Schwarzweißbild zufrieden, entfallen zwei der drei Kanäle und man reduziert die Datenmenge bereits auf ein Drittel. Solch drastische Reduktion ist jedoch selten akzeptabel, weshalb oft auf die Schwäche der menschlichen Wahrnehmung zurückgegriffen wird. Anstatt für jedes Pixel die Farbinformation zu speichern wird in den Cb- und Cr-Kanälen nur jedes zweite, vierte, oder achte Pixel gespeichert. Dies wird als Unterabtastung der Farbinformation bezeichnet. Bei der Decodierung werden die fehlenden Farbwerte aus den vorhandenen interpoliert. Das rekonstruierte Bild lässt sich von Menschen trotz erheblicher Reduzierung der Farbinformation nicht vom Original unterscheiden. Grund ist die Eigenschaft menschlicher Wahrnehmung Helligkeit besser aufzulösen als Farbe. Andere in der Bildkompression verwendete Transformationen transformieren nicht innerhalb des Farbraumes sondern verwenden andere Basisvektoren für die Bildmatrix. Dies führt zu einer Transformation aus dem Orts- in den Frequenzraum. Da in der Bildkompression allem voran die Diskrete Kosinus Transformation *DCT* sowie die Diskrete Wavelet Transformation *DWT* Anwendung finden, wird im weiteren nur auf diese eingegangen. Es sei jedoch angemerkt, dass die *DCT* aus der Fouriertransformation hervorging. Bei der *DCT* wird das Bild aus dem Ortsraum in den Frequenzraum transformiert. Dabei werden aus den Ortsbasen der Bildmatrix Frequenzbasen in Cosinusform. Ein Wert in der Matrix bildet nicht länger einen Grauwert eines Pixels ab, sondern den *DCT*-Koeffizienten und damit genau eine Welle, welche über das gesamte Bild wirkt. Die *DCT*-Koeffizienten, beschreiben die Amplitude der zugehörigen Welle, während die Richtung und Frequenz aus der Position in der Matrix hervorgehen. Auch der Wertebereich umfasst nun nicht länger nur Ganzzahlen zwischen 0 und 255, sondern alle Reellwertigen Zahlen.

Innerhalb des Frequenzraumes lassen sich das Filter auf das Bild anwenden. Dabei haben beispielsweise Tiefpassfilter eine glättende, Rausch und Kanten unterdrückende Auswirkung, wodurch das Bild nach der Rücktransformation verwaschen und unscharf ist. Im Gegensatz dazu sind Hochpassfilter Kanten und Rausch erhaltend, es entfallen jedoch Flächen- und Farbverblaufinformationen. Die Komprimierung mittels *DCT* macht die Annahme, dass Hochfrequente Anteile weniger wichtig für das Bild als niederfrequente Anteile sind, da keine abrupten Grauwertänderungen vorhanden sind. Die hochfrequenten *DCT*-Koeffizienten werden nicht benötigt und deshalb mit dem Wert Null ersetzt. Da dies jedoch eine Kanten unterdrückende Wirkung zur Folge haben kann, wird statt einfachem Ersetzen mit Null gewöhnlich zuerst eine Quantisierung der *DCT*-Koeffizienten mittels Quantisierungstabelle durchgeführt. Durch die Division bleiben auch große hochfrequente Amplituden erhalten, welche bei der Rücktransformierung die Kanteninformation tragen.

Während die *DCT* im JPEG-Standard verwendet wird, findet die diskrete Wavelettransformation *DWT* im neueren JPEG2000 Verwendung. Grund sind Eigenschaften der *DWT* welche sie für die Kompression besser geeignet machen. Während eine Änderung an nur einem *DCT*-Koeffizienten bereits Auswirkungen auf das gesamte Bild hat, hat eine Änderung an den *DWT*-Koeffizienten nur an einem begrenztem

Ort im Bild Auswirkung. Daher der auch der Name Wavlet aus dem Französischen für 'kleine Welle'. Auch lassen sich mittels DWT viele Funktionen kompakter repräsentieren als mittels DCT. Wie auch bei der DCT werden zur weitem Kompression die Koeffizienten zuerst quantisiert und anschließend codiert. Ein weiterer Vorteil der DWT gegenüber der DCT ist die - zumindest analytisch - bessere Laufzeit in  $O(n)$  statt  $O(n \cdot \log n)$

Eine andere Methode der Bildkomprimierung bietet die Fraktale Kompression. Jedoch fand, noch findet diese in K Bildkompressionsstandards Verwendung. Idee der Fraktalenkompression ist die Selbstähnlichkeit eines Bildes. Vereinfacht gesagt wird dazu das Bild bei der Codierung in wenige Pixel große Blöcke zerlegt. Anschließend wird versucht die Blöcke mittels Transformation, Rotation, Skalierung eines anderen Blockes darzustellen. Grauwertunterschiede werden mittels Grundrechenarten beschrieben. Da das finden ähnlicher Blöcke komplex ist, findet die Fraktale Bildkompression derzeit keine breite Verwendung, obwohl es Formate dazu gibt und diese anfänglich von Browsern und Betriebssystemen unterstützt wurden. Eine interessante anwendungsmöglichkeit bietet sich bei der Archivierung von Bildern, da verglichen mit der DWT gute Kompressionsraten erreicht werden, das decodieren jedoch schneller geht.

#### **4 Verwendete Standarts (Stand der Technik?)**

Im diesem Abschnitt werden verschiedene bekannte Formate vorgestellt und deren Einsatzgebiet beschrieben.  
JPEG, JPEG200(WebP), PNG(Giff)

#### **5 Ausblick**

##### **Literatur**

1. Bäni, W.: Wavelets. Oldenburg Verlag, Oldenburg, 2. auflage edn. (2005)