

# **Automatische Erkennung des optimalen Kompressionsverfahrens für Bilddateien in PDFs**

**Bachelor-Thesis**  
zur Erlangung des akademischen Grades B.Sc.

**Robin Werner**  
1996219



Hochschule für Angewandte Wissenschaften Hamburg  
Fakultät Design, Medien und Information  
Department Medientechnik

Erstprüfer: Prof. Dr. Andreas Plaß

Zweitprüfer: Philip Chinery

Hamburg, 19. 11. 2015

# Inhaltsverzeichnis

<b>1</b>	<b>Anforderungsanalyse</b>	<b>6</b>
1.1	Zielsetzung . . . . .	6
1.2	Voraussetzungen und Rahmenbedingungen . . . . .	7
<b>2</b>	<b>Vorüberlegungen</b>	<b>9</b>
2.1	Welche Kompressionsverfahren stehen zur Verfügung? . . . . .	9
2.1.1	Kann eines oder mehrere Kompressionsverfahren komplett abgeschlossen werden? . . . . .	11
2.1.2	Welches Eingangsformat haben die Bilder? . . . . .	12
2.2	Wie bestimmt man das optimale Verhältnis zwischen Qualität und Dateigröße? . . . . .	12
2.2.1	Wie misst man die Qualität von Bildern? . . . . .	13
2.2.2	Welche Bilder sollen für die Messungen verwendet werden? . . . . .	14
<b>3</b>	<b>Forschungsfragen</b>	<b>17</b>
3.1	Ist JPEG 2000 grundsätzlich dem herkömmlichem JPEG überlegen? . . . . .	17
3.1.1	Allgemein . . . . .	17
3.1.2	Kompressionsartefakte . . . . .	17
3.2	Welche Bildeigenschaften lassen sich schnell bestimmen und zur Bestimmung des optimalen Verfahrens verwenden? . . . . .	18
<b>4</b>	<b>Verfahren zur automatischen Bestimmung des Kompressionsverfahrens</b>	<b>20</b>
4.1	Vorhandene Verfahren . . . . .	20
4.1.1	Automatische Auswahl des Kompressionsverfahrens in Adobe Distiller . . . . .	20
4.2	Durchgeführte Versuche . . . . .	21
4.2.1	Versuch 1: Qualitätsunterschiede zwischen JPEG/JPEG 2000 . . . . .	21
4.2.2	Versuch 2: Komprimierung mit verschiedenen Verfahren und anschließender Vergleich . . . . .	22
4.2.3	Versuch 3: Aus Metadaten . . . . .	22
4.2.4	Versuch 4: Mittels Entropie/Farbinformationen . . . . .	23
4.2.5	Versuch 5: Kombination aus Metadaten, Entropie und Farbinformationen . . . . .	23
4.2.6	Versuch 6: Wiederholung von Versuch 3 unter Berücksichtigung der in Versuch 4 gewonnenen Erkenntnisse . . . . .	24

<b>5</b>	<b>Auswertung</b>	<b>25</b>
5.1	Messergebnisse . . . . .	25
5.1.1	Versuch 1: Qualitätsunterschiede zwischen JPEG/JPEG 2000	25
5.1.2	Versuch 2: Komprimierung mit verschiedenen Verfahren und anschließender Vergleich . . . . .	25
5.1.3	Versuch 3: Aus Metadaten . . . . .	26
5.1.4	Versuch 4: Mittels Entropie/Farbinformationen . . . . .	26
5.1.5	Versuch 5: Kombination aus Metadaten, Entropie und Farbinformationen . . . . .	27
5.1.6	Versuch 6: Wiederholung von Versuch 3 unter Berücksichtigung der in Versuch 4 gewonnenen Erkenntnisse . . . . .	27
5.2	Bewertung der Messergebnisse . . . . .	27
5.3	Vergleich mit Adobe Distiller . . . . .	29
<b>6</b>	<b>Fazit</b>	<b>30</b>
6.1	Ergebnis . . . . .	30
6.2	Probleme . . . . .	31
6.3	Ausblick . . . . .	32
<b>A</b>	<b>Anhang 1: Messwerte für Farbbilder aus Versuch 4</b>	<b>33</b>
<b>B</b>	<b>Anhang 2: Messwerte für Graustufenbilder aus Versuch 4</b>	<b>38</b>
	<b>Abbildungsverzeichnis</b>	<b>44</b>
	<b>Tabellenverzeichnis</b>	<b>45</b>
	<b>Literaturverzeichnis</b>	<b>46</b>

## Abstract

The processing speed and data capacities of modern computers are constantly rising. Nevertheless there is still a demand for storing documents at minimum file size, without decreasing their quality significantly. Especially mobile devices benefit from small file sizes, as they usually only have a limited amount of data volume for transmitting data (Sangtae 2012). There are also often file size restrictions when sending files as email attachment. The file format PDF is very suitable for this use case, as it allows individual elements of the document to be compressed with different algorithms in order to create a compact file without decreasing its quality too much, which can be viewed with free software on all major operating systems. This raises the question, how much quality loss is acceptable. When creating a PDF, the user is often confronted with selecting a compression algorithm for the images contained in the document. This has 3 major disadvantages:

The document contains multiple images which shouldn't all be compressed by the same algorithm, but the user can only select a single algorithm to compress all of the documents images.

Even advanced users can't always recognize which algorithm will work best.

If there are a lot of images to compress, the manual selection of the compression algorithm can be very time consuming.

Therefore it is desirable to relieve the user of this decision and to automatically detect the compression algorithm leading to the best result. A procedure is developed which can detect the optimal compression algorithm for 89% of tested images, using only metadata.

## Zusammenfassung

Die Rechengeschwindigkeiten und Speicherkapazitäten moderner Computer steigen stetig. Dennoch besteht nach wie vor ein Bedarf, Dokumente möglichst platzsparend speichern zu können, ohne dabei deren Qualität signifikant zu verschlechtern. Vor allem bei Mobilgeräte ist oftmals eine kleine Dateigröße vorteilhaft, da hier in der Regel nur ein begrenztes Datenvolumen für das Versenden und Empfangen von Daten zur Verfügung steht (Sangtae 2012). Auch für den Versand per E-Mail Anhang gibt es oftmals eine maximal zugelassene Dateigröße. Das Dateiformat PDF ist dafür besonders geeignet, da hier verschiedene Elemente des Dokuments mit unterschiedlichen Verfahren komprimiert werden können, um eine Datei geringer Dateigröße ohne zu großen Qualitätsverlust zu erzeugen, welche auf allen gängigen Betriebssystemen mit kostenloser Software betrachtet werden kann. Dies wirft die Frage auf, ab wann der Qualitätsverlust „zu groß“ ist. Bei der Erstellung der PDFs muss der Nutzer die Entscheidung, welches Kompressionsverfahren für die im Dokument enthaltenen Bilddaten verwendet werden soll, häufig selbst treffen. Dies hat 3 wesentliche Nachteile:

In dem Dokument sind mehrere Bilder enthalten, welche nicht alle mit dem gleichen

Verfahren komprimiert werden sollten. Der Nutzer kann aber nur ein Verfahren für alle Bilder einstellen.

Selbst erfahrene Nutzer können nicht immer erkennen, welches Verfahren besser geeignet ist

Wenn viele Bilder komprimiert werden sollen, stellt die manuelle Auswahl des besser geeigneten Verfahrens einen erheblichen Zeitaufwand dar.

Daher ist es wünschenswert, dem Nutzer die Entscheidung abzunehmen und automatisch zu erkennen, welches der Kompressionsverfahren das beste Ergebnis liefert. Es wird hierzu ein Verfahren entwickelt, welches alleine anhand von Metadaten in der Lage ist, das optimale Kompressionsverfahren für 89% der untersuchten Bilder korrekt zu erkennen.

# 1 Anforderungsanalyse

## 1.1 Zielsetzung

Das Ziel der automatischen Erkennung des optimalen Kompressionsverfahrens soll sein, dem Ersteller einer PDF eine Entscheidung abzunehmen: Für die möglicherweise enthaltenen Bilddaten das Kompressionsverfahren und dessen Parameter so auszuwählen, dass eine PDF mit möglichst geringer Dateigröße und guter Qualität entsteht. Dies ist wünschenswert, da Anwender, welche eine solche PDF erstellen möchten, vor folgenden Problemen stehen können:

- Der Anwender möchte sich nicht mit den technischen Details befassen und so wenig Einstellungen wie möglich vornehmen.
- Der Anwender kennt eines oder mehrere der möglichen Kompressionsverfahren nicht.
- Das Kompressionsverfahren lässt sich nur für das gesamte Dokument einstellen, dieses enthält jedoch verschiedene Bilder, welche nicht alle mit demselben Verfahren komprimiert werden sollten.
- Der fortgeschrittene Anwender kann zwar einschätzen, welches Kompressionsverfahren am besten geeignet ist, möchte aber viele Bilder verarbeiten, so dass die manuelle Auswahl des Kompressionsverfahrens einen nicht unerheblichen Zeitaufwand bedeuten würde.
- Für manche Bilder ist die manuelle Auswahl des am besten geeigneten Kompressionsverfahrens selbst für fortgeschrittene Benutzer schwierig.

Die Anwender können auch konkrete Anforderungen an die erstellte PDF haben. Diese sind abhängig vom Verwendungszweck. So kann es z.B. für Archivierungszwecke erforderlich sein, dass die PDF eine möglichst hohe visuelle Qualität besitzen soll. Für den Versand per E-Mail oder diverse Online Anwendungen kann eine möglichst geringe Dateigröße im Vordergrund stehen.

**Der relevante Anwendungsfall** Da das Dateiformat PDF mit der Absicht entwickelt wurde, Dokumente unabhängig vom verwendeten Betriebssystem ansehen und drucken zu können (Adobe 2006)[23], handelt es sich oftmals auch um einen einfachen Austausch von Informationen, bei dem es keine fest definierten Kriterien für

Qualität oder Dateigröße gibt. Es ist jedoch anzunehmen, dass bei dem Informationsaustausch möglichst alle Informationen erhalten bleiben sollen und eine möglichst kleine Dateigröße wünschenswert ist. Die hier entwickelte automatische Erkennung des optimalen Kompressionsverfahrens konzentriert sich auf diesen Anwendungsfall. Wenn Vorgaben an eine minimale Qualität oder maximale Dateigröße existieren, ist die Auswahl des Kompressionsverfahrens direkt von diesen abhängig und kann nicht automatisiert erkannt werden.

Es gibt bereits Software, die eine automatisch Erkennung des Kompressionsverfahrens zumindest theoretisch bietet (Ghostscript, Adobe Acrobat), es besteht jedoch nur sehr wenig Dokumentation darüber, wie diese arbeitet. Ein eigenes Verfahren ist daher wünschenswert.

Es muss zunächst geklärt werden, was das „optimale Verhältnis“ zwischen Dateigröße und Qualität ist. Außerdem muss untersucht werden, wie sich die Qualität objektiv/algorithmisch messen lässt. Es soll ermittelt werden, welche Bildeigenschaften sich am besten dazu eignen, das Kompressionsverfahren automatisch zu bestimmen. Dabei soll möglichst wenig Rechenleistung beansprucht werden. Die Bilder sollen so stark wie möglich komprimiert werden, ohne die Qualität dieser wesentlich zu reduzieren.

## 1.2 Voraussetzungen und Rahmenbedingungen

Die automatische Erkennung soll in eine Software zur Erstellung von PDF-Dateien integriert werden. Der Nutzer der Software hat bei Bedarf die Möglichkeit, das Kompressionsverfahren manuell einzustellen. Die automatische Erkennung des Kompressionsverfahrens wird immer dann verwendet, wenn der Nutzer keine solche Einstellung vorgenommen hat. Die Software läuft auf Windows Betriebssystemen und dient dazu, bestehende Dokumente in das PDF-Format zu konvertieren. Die Konvertierung erfolgt hierbei über einen virtuellen Drucker. Der Druckvorgang ist hierbei größtenteils identisch mit dem herkömmlichen Drucken auf Papier, mit dem Unterschied, dass der Anwender anstelle des Papiers am Ende eine PDF erhält. Dadurch ergeben sich folgende Einschränkungen und Rahmenbedingungen:

- Die zu komprimierenden Bilder können einen beliebigen Inhalt haben.
- Es stehen nur die im PDF-Standard vorgesehen Kompressionsverfahren (siehe 1.2.1) zur Verfügung
- Um etwaige Lizenzgebühren zu vermeiden sollen für die Erkennung des Kompressionsverfahrens nur Bibliotheken und Hilfsprogramme verwendet werden, deren Lizenz eine freie Verwendung zulässt
- Da ein Dokument eine beliebige Anzahl an Bildern haben kann und die dem Nutzer zur Verfügung stehende Rechenleistung unbekannt ist, soll die Erkennung möglichst wenig Rechenleistung beanspruchen.

## *1 Anforderungsanalyse*

- Die Bilddaten wurden bereits durch den Druckvorgang komprimiert
- Da die automatische Erkennung direkt in die Software integriert werden soll, ist die zu verwendende Programmiersprache vorgegeben

## 2 Vorüberlegungen

### 2.1 Welche Kompressionsverfahren stehen zur Verfügung?

Eine PDF Datei besteht aus verschiedenen Objekten. Bilder können innerhalb einer PDF nicht direkt eingebunden werden, sondern bestehen aus einem Stream-Objekt<sup>1</sup>, welches die Bilddaten enthält. Das Stream-Objekt hat einen Header, in welchem unter anderem angegeben werden kann, mit welchem Verfahren die Bilddaten dekomprimiert werden sollen (Adobe 2006: 60). Die Bilddaten werden komprimiert in das Stream-Objekt geschrieben und beim Auslesen mit dem im Header angegebenen Filter dekomprimiert. Kompressionsverfahren können in 2 Kategorien eingeteilt werden:

Verlustbehaftete Kompressionsverfahren werfen Teile der ursprünglichen Information zu Gunsten einer erhöhten Kompressionsrate. Dies ist besonders dann effektiv, wenn das zu komprimierende Material einen analogen Ursprung hat und zuvor digitalisiert wurde. Die meisten verlustbehafteten Verfahren erlauben dem Nutzer verschiedene Qualitätsstufen einzustellen, wobei eine höhere Komprimierung einer niedrigeren Qualität entspricht und umgekehrt eine höhere Qualität einer niedrigeren Komprimierung (Nelson 1996: 2).

Verlustfreie Kompressionsverfahren komprimieren Daten auf eine solche Art und Weise, dass bei der Dekompression alle Informationen exakt rekonstruiert werden können. In der Regel erreichen verlustfreie Kompressionsverfahren niedrigere Kompressionsraten als verlustbehaftete Verfahren.

Folgende Kompressionsverfahren sind laut Adobe (Adobe 2006: 39) für Bilddaten im PDF Standard vorgesehen:

- JPEG( Joint Photographic Experts Group, verlustbehaftet) für Graustufen und Farbbilder

JPEG ist die gebräuchliche Bezeichnung des ISO Standards (ISO/IEC 10918), welcher 2 Methoden zur Kompression definiert: Eine auf diskreter Kosinus-Transformation<sup>2</sup> basierende Methode für eine verlustbehaftete Kompression und eine verlustfreie Methode (Wallace 1992: 1), welche hier jedoch nicht verwendet werden kann. JPEG steht im Nachfolgendem immer für die verlustbehaftete, auf DCT basierende Kompression.

---

<sup>1</sup>Ein Stream-Objekt ist eine Sequenz von Bytes (Adobe 2006: 60).

<sup>2</sup>Diese wird im Nachfolgenden mit DCT bezeichnet.

## 2 Vorüberlegungen

Die Bilddaten werden für die Komprimierung zunächst in Blöcke von  $8 \cdot 8$  Pixeln zerlegt, anschließend wird auf die Blöcke eine DCT angewandt. Anschließend werden die DCT Koeffizienten quantisiert<sup>3</sup>. Sowohl bei der Durchführung der DCT als auch der anschließenden Quantisierung gehen auf Grund der endlichen Genauigkeit Informationen verloren (Wallace 1992: 4).

- JPEG 2000 (verlustfreie und verlustbehaftete Kompression möglich) für Graustufen und Farbbilder

JPEG 2000 ist eine Zusammenfassung mehrerer ISO Standards. Im Folgenden steht JPEG 2000 immer für die in ISO/IEC 15444-1 definierte Wavelet-basierte Kompression, siehe (ISO 2000).

- CCITT Group 3/Group 4 (verlustfrei) für binäre Bilder

Die CCITT Group 3/Group 4<sup>4</sup> Standards definieren Kompressionsverfahren für binäre Bilder. Diese Bilder zeichnen sich dadurch aus, dass jeder Pixel nur 2 mögliche Werte annehmen kann, 1 oder 0. Beide Verfahren verwenden eine Kombination aus Lauflängen- und Entropiekodierung. Bei der Lauflängenkodierung werden identische, aufeinander folgende Werte so zusammengefasst, dass an Stelle der sich wiederholenden Werte ein einzelner Wert und die Anzahl seiner Wiederholungen angegeben wird (aus 00000 wird  $5 \cdot 0$ ). Bei der anschließenden Entropiekodierung werden den statistisch am häufigsten auftretenden Werten die kürzesten Codewörter zugeordnet, während seltener auftretende Werte längere Codewörter zugeordnet bekommen.

- RLE (Run Length Encoding, verlustfrei) für binäre Bilder

Lauflängenkodierung, ein allgemeines Verfahren zur Kompression von Binärdaten.

- JBIG2 (Joint Bi-level Image Experts Group 2, verlustfreie und verlustbehaftete Kompression möglich) für binäre Bilder

Bei der JBIG2 Kompression werden sich wiederholende Symbole in eine Tabelle eingetragen, jedem Symbol wird ein Codewort zugeordnet. Dadurch muss jedes der sich wiederholenden Symbole nur einmal gespeichert vollständig gespeichert werden und kann anschließend durch sein Codewort ersetzt werden. Bei der Dekompression werden die Codewörter wieder durch die in der Tabelle hinterlegten Symbole ersetzt. Je nachdem, ob nur vollständig identischen oder auch stark ähnlichen Symbolen das gleiche Codewort zugeordnet wird, handelt es sich um eine verlustfreie oder verlustbehaftete Kompression.

- LZW (Lempel-Zip-Welch, verlustfrei) für Text, Grafiken und Bilder

---

<sup>3</sup>Wertebereiche werden zu einzelnen Werten zusammengefasst.

<sup>4</sup> offiziell CCITT T.4 / T.6 <https://www.itu.int/rec/T-REC-T/en>

LZW komprimiert Daten, indem sich wiederholende Zeichenfolgen durch Codewörter ersetzt und in einer dynamisch erzeugten Tabelle hinterlegt werden.

- Flate (verlustfrei) für Text, Grafiken und Bilder

Offiziell als Deflate bezeichnet ([Deutsch 1996](#)). Deflate verwendet eine Kombination aus LZ77 Algorithmus und anschließender Huffman Kodierung. Der LZ77 Algorithmus verwirft redundante Daten und ersetzt diese durch einen kompakteren Verweis<sup>5</sup>.

### 2.1.1 Kann eines oder mehrere Kompressionsverfahren komplett ausgeschlossen werden?

Die JBIG2 verlustbehaftete Kompression ist unsicher, da abhängig von der Implementation schwerwiegende Substitutionsfehler auftreten können ([KOST 2014: 2](#)). Diese lassen sich nicht offensichtlich als Fehler erkennen, sondern verfälschen das Dokument z.B. durch Zahlendreher. Auch eine verlustfreie JBIG2 Komprimierung kann unter Umständen problematisch sein, da der Anwender nicht erkennen kann, dass verlustfrei komprimiert wurde und im Zweifelsfall davon ausgehen muss, dass Zahlen oder Buchstaben vertauscht worden sind. Dadurch ist für Dokumente, die mit JBIG2 komprimierte Bilder enthalten, laut BSI <sup>6</sup> keine Rechtssicherheit mehr gegeben ([BSI 2013](#)). Daher sollte dieses Verfahren nur dann verwendet werden, wenn sich der Nutzer über die Konsequenzen eindeutig bewusst ist und kommt als Auswahlmöglichkeit für ein automatisiertes Verfahren nicht in Frage.

RLE wird nicht näher untersucht, da dies lediglich für binäre Bilder vorgesehen ist und die ebenfalls für binäre Bilder vorgesehene CCITT Group 4 (T.6) Kompression eine Kombination aus RLE und anderen Verfahren verwendet, so dass diese in fast allen Fällen bessere Ergebnisse liefern sollte.

Der Group 3 (T.4) Standard wurde dafür entwickelt, um „eine A4 Seite innerhalb von ungefähr einer Minute über eine Telefonleitung [frei aus dem Englischen übersetzt]“ ([CCITT:1](#)) übermitteln zu können. Auf Grund der Störanfälligkeit der damaligen analogen Telefonleitungen sind hier zusätzliche Informationen zur Fehlerkorrektur enthalten. Die später entwickelte Group 4 Kompression erzeugt kleinere Dateien, da hier keine Informationen zur Fehlerkorrektur mehr enthalten sind. Die Group 4 Kompression ist daher in einer digitalen Umgebung immer vorzuziehen ([Murray & VanRyper 1996](#)).

Deflate und LZW werden in der Regel nur dort eingesetzt, wo eine verlustfreie Kompression erforderlich ist. Da Deflate in der Regel „*wesentlich kompakter* [frei aus dem Englischen übersetzt]“ ([Adobe 2006: 71](#)) komprimiert als LZW und keine

---

<sup>5</sup>Dies ist eine stark vereinfachte Darstellung, eine detaillierte Darstellung wäre an dieser Stelle zu umfangreich und ist bei Bedarf unter folgender Quelle nachzulesen: ([Ziv & Lempel](#)).

<sup>6</sup>Bundesamt für Sicherheit in der Informationstechnik

nennenswerten Nachteile hat, ist Deflate immer zu bevorzugen. Da die hierzu vorhandene Quelle diesbezüglich keine weiteren Angaben macht, wird diese Behauptung in Versuch 2 zusätzlich bestätigt.

### 2.1.2 Welches Eingangsformat haben die Bilder?

Es wird vorausgesetzt, dass die zu komprimierenden Bilder zuvor in das PNG<sup>7</sup> Format konvertiert worden sind. Dies ist dadurch bedingt, dass die zu komprimierenden Bilddateien in der Endanwendung von einem virtuellen XPS-Drucker kommen werden. XPS steht für XML<sup>8</sup> Paper Specification. Das in erster Linie von Microsoft entwickelte Dateiformat wird von Microsoft folgendermaßen beschrieben: „Ein XPS-Dokument ... ist ein Dokumentformat, mit dem Sie den Dokumentinhalt anzeigen, speichern, freigeben, digital signieren und schützen können. Ein XPS-Dokument ist mit einem elektronischen Blatt Papier vergleichbar“ (Microsoft) XPS sieht für Bilder 4 mögliche Dateiformate vor: PNG, JPEG, TIFF<sup>9</sup> und WPD<sup>10</sup> Images. JPEG scheidet aus, da XPS nur die verlustbehaftete, auf DCT basierende Variante zulässt, das heißt bei der Konvertierung nach JPEG gehen in jedem Fall Bildinformationen verloren (Wallace 1992). Verlustbehaftete Verfahren sollten, sofern es möglich ist, stets nur in der letzten Stufe der Produktion verwendet werden, um bis dahin die volle Qualität beibehalten zu können. Die anderen 3 Formate sind grundsätzlich geeignet, da sie alle verlustfreie Kompression verwenden oder zumindest verwenden können. PNG bietet den Vorteil, dass die Bilddaten dort immer mit Deflate komprimiert werden, so dass aus diesen schnell abgelesen werden kann, wie gut die Deflate Kompression den jeweiligen Bildinhalt komprimiert, weiteres dazu in 3.2. Laut Microsoft können nur PNG oder JPEG direkt als Zielformat für Bilder in XPS Dateien eingestellt werden (Microsoft), so dass die anderen beiden Formate TIFF und WPD nicht verwendet werden können, da es unklar ist, wie eine derartige Implementierung zu erfolgen hat oder ob diese überhaupt möglich ist.

## 2.2 Wie bestimmt man das optimale Verhältnis zwischen Qualität und Dateigröße?

Das optimale Verhältnis zwischen Qualität und Dateigröße ist zunächst abhängig vom Verwendungszweck des erzeugten Bildmaterials. Während teilweise eine verlustfreie Komprimierung benötigt wird, z.B. wenn die Bilder archiviert oder zu medizinischen Zwecken verwendet werden sollen, gibt es auch Anwendungsgebiete, bei denen die Dateigröße wichtiger ist als eine perfekte Reproduktion des Bildes, z.B. für Websites und mobile Endgeräte mit geringem Speicherplatz. Da der Anwender abgesehen

---

<sup>7</sup>Portable Network Graphics, ein offenes Dateiformat für Bilddateien

<sup>8</sup>Extensible Markup Language

<sup>9</sup>Tagged Image File Format

<sup>10</sup>Windows Media Photo

von der automatischen Erkennung ebenfalls die Möglichkeit hat, das zu verwendende Kompressionsverfahren manuell einzustellen, soll die automatische Erkennung nur dann eine verlustfreie Kompression auswählen, wenn diese kleinere Dateigrößen erzielt als eine verlustbehaftete. Wenn eine verlustfreie Kompression grundsätzlich erforderlich ist, muss dies vom Anwender eingestellt werden. Ansonsten soll die Kompression so hoch wie möglich sein, ohne dass die wahrgenommene Qualität des Bildes zu stark nachlässt. Die Hauptschwierigkeit ist, dass zwar viele Untersuchungen darüber existieren, welcher Algorithmus die menschliche Wahrnehmung von Bildfehlern am besten widerspiegelt, es jedoch keine allgemein gültige Definition darüber gibt, wann die Qualität eines Bildes „gut genug“ oder visuell verlustfrei ist. Dies ist auch einleuchtend, da die wahrgenommene Qualität eines Bildes stets auch von wechselnden äußeren Einflüssen beeinflusst wird, wie z.B. dem Betrachtungsabstand, dem wiedergebenden Medium oder den Lichtverhältnissen. Selbst wenn diese Parameter definiert werden können, hängt es immer noch stark von dem Verwendungszweck ab, ab wann die Qualität als ausreichend anzusehen ist. Da PDF Dateien jedoch einen beliebigen Verwendungszweck haben können, ist es unmöglich, hier eine wissenschaftlich fundierte Definition für ausreichende Qualität festzulegen.

Daher wird für das zu entwickelnde Verfahren festgelegt, dass die Qualität der Bilder mindestens genauso hoch sein soll, wie die Qualität der in der Standardeinstellung von Ghostscript<sup>11</sup> erzeugten Bilder. Diese werden in der Standardeinstellung mit Qualität 75% erzeugt. Die Ghostscript Dokumentation bezieht sich hier auf /glqq the widely used IJG quality scale/grqq ([Ghostscript](#)).

### 2.2.1 Wie misst man die Qualität von Bildern?

Da bis auf JPEG und JPEG 2000 alle anderen für diesen Verwendungszweck in Betracht kommenden Verfahren verlustfrei arbeiten, ist bei diesen die Qualität des komprimierten Bildes identisch mit der des Originals. Um eine einfache und objektive Möglichkeit zu haben, die Qualität von Bildern messen zu können, welche mit JPEG oder JPEG 2000 verlustbehaftet komprimiert, ist ein algorithmisches Verfahren von Vorteil. Subjektive Befragungen sind wesentlich Zeitaufwändiger. Diese müssen sorgfältig geplant werden, um eine Verfälschung der Ergebnisse durch mögliche äußere Einflüsse wie Ermüdung der Subjekte, Lichtverhältnisse und Betrachtungsabstand zu vermeiden. Außerdem können die Befragungen nicht je nach Bedarf wiederholt werden.

Für eine algorithmische Bestimmung der Qualität gibt es im Wesentlichen 2 verschiedene Ansätze: Der „Full Reference“ Ansatz vergleicht die Qualität des komprimierten (oder in sonstiger Art und Weise veränderten) Bildes mit der des Originals. Dies geschieht anhand von vorher definierten Metriken und misst den Qualitätsunterschied.

Der „No Reference“ Ansatz ermittelt die Qualität des Bildes anhand von vorher

---

<sup>11</sup>Ghostscript ist eine Anwendung, die unter anderem Druckdaten direkt in PDF umwandeln kann.

definierten Metriken und misst die absolute Qualität des Bildes. Für diese Untersuchung wird ein Full Reference Verfahren verwendet, da hier die der Qualitätsverlust und nicht die absolute Qualität untersucht wird.

Das nach bester Kenntnis zum Zeitpunkt dieser Arbeit vorhandene Verfahren ist FSIM<sup>12</sup>, da dieses Verfahren am stärksten mit den Ergebnissen subjektiver Befragungen übereinstimmt (Zang 2011). FSIM basiert auf der Annahme, dass die menschliche Wahrnehmung und Interpretation von Bildern am stärksten von sogenannten „low-level features“ wie beispielsweise im Bild vorhandenen Kanten abhängt. Zusätzlich berücksichtigt FSIM, dass das menschliche Auge aus kontrastreichen Bildregionen mehr Informationen herausziehen kann als aus Regionen mit geringerem Kontrast (Zang 2011).

Die vorhandene MATLAB Implementation<sup>13</sup> stellt sich jedoch als zu langsam und instabil heraus, um diese hier sinnvoll verwenden zu können.

Daher wird das SSIM Verfahren verwendet, für welches eine C# Implementation<sup>14</sup> zur Verfügung steht. Diese läuft wesentlich schneller und stabiler als die MATLAB Implementation von FSIM. SSIM misst den Qualitätsunterschied anhand der Änderungen von Kontrast, Helligkeit und Struktur gegenüber einem Referenzbild (Wang 2004).

### 2.2.2 Welche Bilder sollen für die Messungen verwendet werden?

Um das zu entwickelnde Verfahren testen zu können, wird eine eigene Bildsammlung zusammengestellt, welche aus verschiedenen bereits vorhandenen Bildsammlungen und eigens zu diesem Zweck ausgewählten oder erstellten Bildern besteht. Bei der Auswahl werden öffentlich zugängliche Bilder und Bildsammlungen bevorzugt, um eine unabhängige Überprüfbarkeit der Versuche zu gewährleisten. Außerdem werden nur Bilder verwendet, welche zuvor nicht bereits mit einem verlustbehaftetem Verfahren komprimiert worden sind.

Das von der Endanwendung zu verarbeitende Bildmaterial ist beliebig und enthält neben herkömmlichen Fotos auch Bildschirmfotos, Logos oder allgemein Abbildungen mit vielen gleichfarbigen Flächen. Daher werden die Bilder so ausgewählt, dass sowohl Fotos als auch Bilder mit einem größeren Anteil an gleichfarbigen Flächen enthalten sind.

Die Bilder der Datenbank „DatasetDBD“<sup>15</sup> werden komplett übernommen, da diese eine Mischung aus Text auf einfarbigem Hintergrund und beliebigen Motiven (Fotos) enthalten. Für solche Bilder lässt sich das optimale Kompressionsverfahren schwer manuell einschätzen, da der einfarbige Hintergrund voraussichtlich besser mit

---

<sup>12</sup> „A Feature Similarity Index for Image Quality Assessment“ (Zang 2011)

<sup>13</sup> <http://www4.comp.polyu.edu.hk/~cslzhang/IQA/FSIM/FSIM.htm>, letzter Zugriff 10.10.2015

<sup>14</sup> [www.lomont.org/Software/Misc/SSIM/SSIM.html](http://www.lomont.org/Software/Misc/SSIM/SSIM.html), letzter Zugriff 15.10.2015

<sup>15</sup> <http://www.lrde.epita.fr/cgi-bin/twiki/view/Olena/DatasetDBD>, letzter Zugriff 8.10.2015

## 2 Vorüberlegungen

Deflate komprimiert wird, während der restliche Bildinhalt voraussichtlich besser mit JPEG komprimiert werden kann.

Dies gilt auch für Bildschirmfotos, die sowohl viele gleichfarbige Flächen als auch komplexe Bilder enthalten. Daher werden zusätzlich 5 solcher Bildschirmfotos hinzugefügt.

Die „Kodak Lossless True Color Image Suite“<sup>16</sup>, welche aus 25 unkomprimierten Farbbildern besteht, wird ebenfalls komplett übernommen.

Außerdem werden die ersten 125 Bilder der UCID - Uncompressed Colour Image Database<sup>17</sup> verwendet. Hierbei handelt es sich um Fotos von verschiedenen Objekten und Szenarien.

Des Weiteren werden 10 Clip Arts hinzugefügt, welche aus vergleichsweise wenigen Farben bestehen und insgesamt wenig komplex sind.

Die Bilder werden so gewählt, dass möglichst viele verschiedene Arten von Bildmaterial enthalten sind. Insgesamt umfasst die Bildsammlung 291 Bilder und hat eine Dateigröße von 265 MB.

**Binäre Bilder** Binäre Bilddaten zeichnen sich dadurch aus, dass jeder Pixel nur 2 mögliche Werte besitzt, 0 oder 1. Da die Metadaten die Farbtiefe angeben, mit welcher das Bild codiert wurde, diese jedoch nicht zwangsläufig der minimal notwendigen Farbtiefe entspricht, müssen für die korrekte Erkennung eines Binärbildes im schlimmsten Fall die Farbinformationen aller Pixel untersucht werden. Sobald mehr als 2 verschiedene Farben gefunden werden, kann die Erkennung jedoch abgebrochen werden.

Da die JBIG2 Kompression und CCITT Group 3 ausgeschlossen werden können, sollten binäre Bilder mit CCITT Group 4 komprimiert werden, da dieses das einzige überbleibende Verfahren ist, welches speziell für die Kompression von binären Bildern entwickelt wurde und in anderen Versuchen die besten Ergebnisse erzielt hat, siehe (Khan 2011). Binäre Bilder werden daher nicht weiter untersucht und sind in der Bildsammlung nicht enthalten.

**Graustufenbilder** Graustufenbilder sind Farbbildern nicht gänzlich unähnlich, da Farbbilder im RGB Farbraum sich stets aus je 3 Graustufenbildern zusammensetzen, welche die Helligkeitsinformationen für Rot, Grün und Blau repräsentieren (Wallace 1992: 3). Graustufenbilder benötigen nur ein Drittel des Speicherplatzes eines entsprechenden RGB Bildes. Da bei der JPEG Kompression die Reduzierung der Chrominanz „*einen effektiven Weg der Datenreduktion* [frei aus dem Englischen übersetzt]“ (Skrodas 2001: 41) darstellt, ist hier die Überlegung angebracht, dass JPEG schlechter funktionieren könnte. Da Graustufenbilder keine Chrominanz enthalten, kann diese

---

<sup>16</sup><http://r0k.us/graphics/kodak/>, letzter Zugriff 8.10.2015

<sup>17</sup><http://homepages.lboro.ac.uk/~cogs/datasets/ucid/data/ucid.v2.tar.gz>, letzter Zugriff 8.10.2015

## *2 Vorüberlegungen*

nicht reduziert werden. Daraus folgt, dass Verfahren, welche sich diese Reduzierung zu Nutze machen, für Graustufenbilder weniger effektiv sein könnten als für Farbbilder.

# 3 Forschungsfragen

## 3.1 Ist JPEG 2000 grundsätzlich dem herkömmlichem JPEG überlegen?

### 3.1.1 Allgemein

JPEG 2000 ist ein wesentlich neuerer Standard als JPEG und unter anderem mit dem Ziel entwickelt worden, zu dem Zeitpunkt der Entwicklung bestehende Kompressionsverfahren sowohl Qualitativ als auch in Bezug auf die Kompressionsrate zu übertreffen ([Christopoulos 2000](#)).

Es hat sich jedoch gezeigt, dass die subjektiv wahrgenommene Qualität der komprimierten Bilder bis zu einer Kompressionsrate von 20:1 für JPEG komprimierte Bilder höher ist als für jene, welche mit JPEG 2000 komprimiert wurden ([Ebrahimi 2004](#)). Daher wird in 3.2.1 gemessen, welche unterschiedlichen Auswirkungen die beiden Kompressionsverfahren auf die Qualität der komprimierten Bilder haben. Es wird untersucht, ob die Parameter der JPEG 2000 Kompression so gewählt werden können, dass die Qualität der komprimierten Bilder höher und die Dateigröße niedriger ist als bei einer JPEG Kompression mit 75% Qualität.<sup>1</sup> Die Einstellung wird so gewählt, damit die anfangs definierte minimale Qualität erreicht wird.

### 3.1.2 Kompressionsartefakte

Bei einer verlustbehafteten Kompression treten sowohl bei JPEG als auch bei JPEG 2000 Kompressionsartefakte auf. Kompressionsartefakte sind Bildfehler, welche direkt aus der verlustbehafteten Kompression resultieren. JPEG 2000 verwendet eine Wavelet-basierte Kompression ([ISO 2000](#)), welche typischerweise Unschärfe und Gibbs-Ringing-Artefakte<sup>2</sup> verursacht. JPEG verwendet eine DCT basierte Kompression, die typischerweise Blockartefakte verursacht ([Marziliano 2004](#)).

Abbildung 3.1 soll die Unterschiede der Kompressionsartefakte von JPEG und JPEG 2000 verdeutlichen. Die Kompressionsartefakte sind jeweils rot umrandet. Bei der JPEG 2000 Kompression ist in der linken Bildhälfte gut die Unschärfe zu erkennen, im Schriftteil werden Gibbs-Ringing-Artefakte sichtbar. Die Unschärfe könnte

---

<sup>1</sup>In diesem und allen nachfolgenden Fällen bezieht sich „JPEG 75% Qualität“ auf die Einstellung des verwendeten GDI+(Graphics Device Interface) JPEG Encoders, in der Annahme, dass sich dieser wie Ghostscript ebenfalls auf die „IJG quality scale“ bezieht.

<sup>2</sup>Gibbs-Ringing-Artefakte sind Linien, welche parallel zu scharfen Kanten auftreten([Hornak](#)).

jedoch unter Umständen die wahrgenommene Qualität verbessern, indem sie die bei der Digitalisierung entstandenen Artefakte korrigiert, wie z.B. beim Scannen eines Dokumentes entstandene Artefakte (KOST 2008). Bei der JPEG Kompression werden Blockartefakte sichtbar, wobei diese den gesamten Schriftbereich umranden, an den markierten Stellen wird die  $8 \cdot 8$  Pixel Blockstruktur am besten sichtbar.

## 3.2 Welche Bildeigenschaften lassen sich schnell bestimmen und zur Bestimmung des optimalen Verfahrens verwenden?

Eine gute Methode, um schnell Informationen über das Bildmaterial zu erhalten, sind häufig die Metadaten der Bilddateien, da diese ausgelesen werden können, ohne die gesamte Datei zu laden. So enthalten PNGs in ihrem Header unter anderem Metadaten zu den Abmessungen des Bildes und der Farbtiefe (libpng 2003), welche sich eignen, um erste Schlüsse über das Bild zu ziehen.

Leider geben die Informationen zur Farbtiefe jedoch nicht die kleinstmögliche Farbtiefe an, sondern die Farbtiefe, mit welcher das Bild kodiert wurde. Die vom XPS-Drucksystem erzeugten PNGs, welche die zu entwickelnde Methode als Input erhalten wird, werden immer mit 24 Bits pro Pixel „true color“ erzeugt, auch wenn das Bild mit einer geringeren Farbtiefe originalgetreu dargestellt werden könnte. Die einzig verwertbaren Metadaten aus dem PNG Header, die für die zu entwickelnde Methode von Bedeutung sein könnten, sind die Angaben zu den Abmessungen des Bildes. Um zuverlässige Informationen zu anderen Bildeigenschaften zu erhalten, ist es erforderlich, das gesamte Bild zu untersuchen.

Da die Entropie zumindest für verlustfreie Kompressionsverfahren einen Anhaltspunkt bietet, wie stark das Bild maximal komprimiert werden kann (Shannon 1948, Verdu 2009), wird in Versuch 4 untersucht, ob diese für eine Bestimmung des optimalen Verfahrens verwendet werden kann.

Es wird ebenfalls untersucht, ob die Anzahl der im Bild vorhandenen Farben zur Bestimmung des optimalen Kompressionsverfahrens verwendet werden kann. Dies geschieht aus der Überlegung heraus, dass die Wahrscheinlichkeit, sich innerhalb des Bildes wiederholende Pixelfolgen zu finden, bei einer geringen Anzahl voneinander verschiedener Farben höher sein könnte als bei einer großen Anzahl verschiedener Farben.

Außerdem wird untersucht, wie viel Zeit für die Erfassung der jeweiligen Bildeigenschaft erforderlich ist und welche der Eigenschaften sich am besten zur Bestimmung des optimalen Kompressionsverfahrens verwenden lassen.

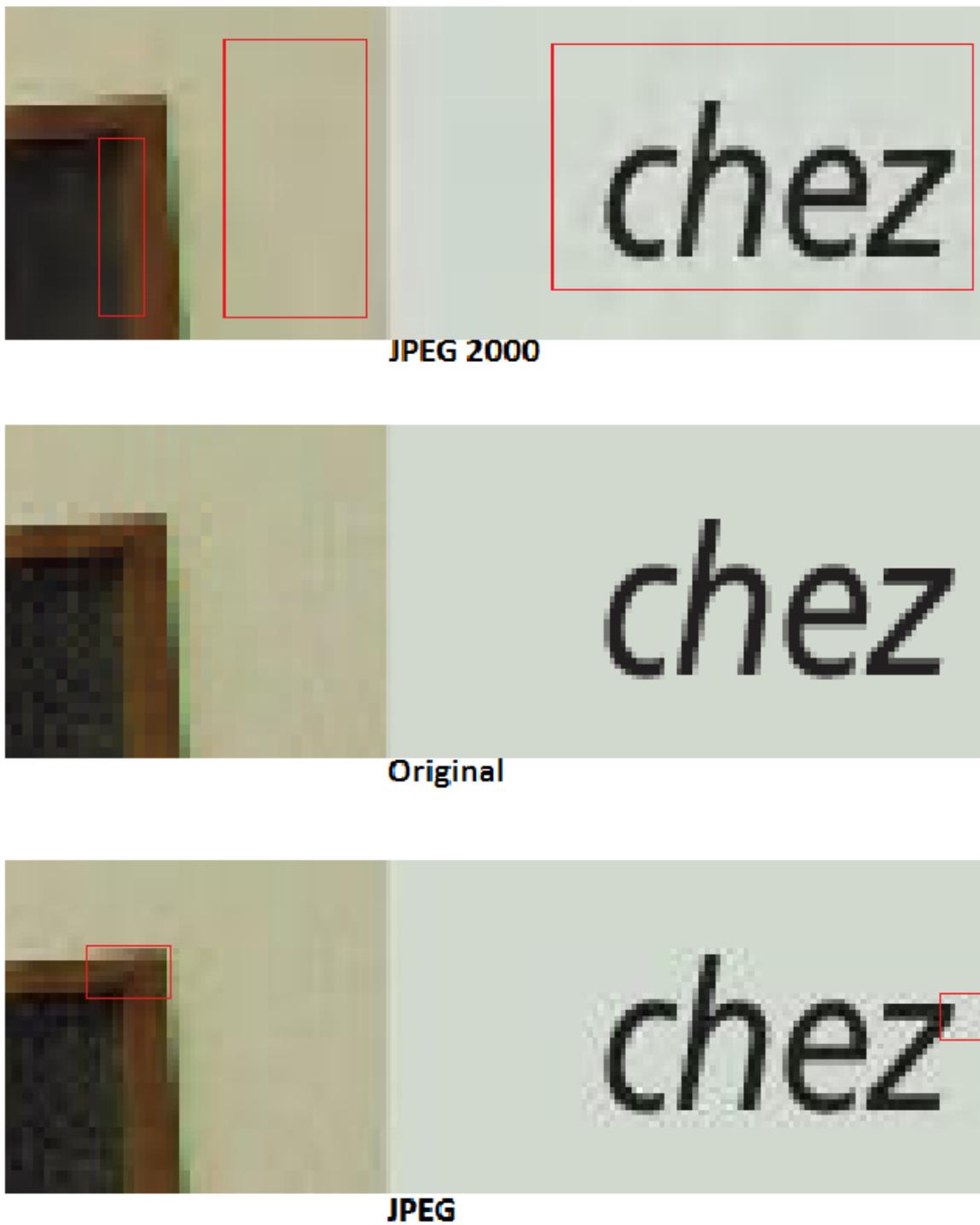


Abbildung 3.1: Typische Kompressionsartefakte

# 4 Verfahren zur automatischen Bestimmung des Kompressionsverfahrens

## 4.1 Vorhandene Verfahren

Es gibt bereits diverse Verfahren, um automatisch das Kompressionsverfahren für Bilder zu bestimmen oder diese nach festgelegten Kriterien automatisch zu sortieren. Die vorhandenen Verfahren sind jedoch für den Verwendungszweck dieser Untersuchung ungeeignet, da sie entweder zu rechenaufwändig sind oder andere Eingangs- oder Zielformate voraussetzen. Dennoch lassen sich vereinzelt Ansätze daraus verwenden, so haben (Edmunson 2013) ein Verfahren zur Kategorisierung von JPEG komprimierten Bildern entwickelt, welches ausschließlich die Metadaten auswertet und so wesentlich schneller arbeitet, als vergleichbare Verfahren. Daher wird in Versuch 3 untersucht, ob die Metadaten auch für das zu entwickelnde Verfahren verwendet werden können. Auch die Kombination der Auswertung von Metadaten und anderer Bildinformationen wurde von (Edmunson 2013) inspiriert, da sich in deren Untersuchung herausstellte, dass die Metadaten alleine keine ausreichend zuverlässige Bestimmung ermöglichen, jedoch in Kombination mit anderen Bildeigenschaften zu guten Ergebnissen führen können. Vorhandene Software<sup>1</sup> zur Erstellung von PDFs verfügt teilweise auch über eine automatische Erkennung des zu verwendenden Kompressionsverfahrens, diese ist jedoch nicht ausreichend dokumentiert, so dass nicht vollständig beurteilen lässt, ob die für die Erkennung verwendeten Parameter sinnvoll sind.

### 4.1.1 Automatische Auswahl des Kompressionsverfahrens in Adobe Distiller

In der Standardeinstellung unterscheidet die automatische Auswahl des Kompressionsverfahrens in Adobe Distiller für Farb- und Graustufenbilder zwischen JPEG oder Deflate. Es ist nur ansatzweise beschrieben, nach welchen Kriterien dort entschieden wird. Laut Handbuch (Adobe 2015) wird JPEG für alle „8-Bit Bilder mit flüssigen Farbänderungen“ verwendet, „Für alle anderen Bilder wird Flate verwendet“ Optional kann die automatische Erkennung darauf eingestellt werden, für „Niederfrequente

---

<sup>1</sup>z.B. Adobe Distiller, Ghostscript

Bilder“ eine verlustbehaftete JPEG 2000 Kompression durchzuführen und für „hochfrequente Bilder“ eine verlustfreie JPEG 2000 Kompression. Es wird hierbei keine Unterscheidung zwischen Graustufen und Farbbildern getroffen. Für Binärbilder wird keine automatische Erkennung vorgenommen (Adobe 2015: 30).

## 4.2 Durchgeführte Versuche

### 4.2.1 Versuch 1: Qualitätsunterschiede zwischen JPEG/JPEG 2000

**Ansatz** Es soll untersucht werden, wie die Parameter der JPEG 2000 Kompression gewählt werden müssen, um die gesamte Bildsammlung stärker zu komprimieren als JPEG in der Einstellung 75% Qualität. Gleichzeitig soll die Qualität der komprimierten Bilder mindestens genauso hoch sein wie die der JPEG komprimierten Bilder. Zur Messung der Qualität wird der SSIM Algorithmus verwendet, dieser misst den Qualitätsunterschied der JPEG und JPEG 2000 komprimierten Bilder gegenüber den verlustfrei komprimierten Originalen.

Um untersuchen zu können, ob die JPEG 2000 Kompression bei gleicher oder besserer Qualität stärker komprimiert als JPEG, wird zunächst die gesamte Bildsammlung mit JPEG in der Einstellung 75% Qualität komprimiert und die resultierende Dateigröße aller komprimierten JPEG Dateien addiert. Anschließend wird mit dem SSIM Algorithmus die Qualität für jedes JPEG gemessen und der Mittelwert über die komprimierte Bildsammlung gebildet. Danach werden die Bilder mit JPEG 2000 in verschiedenen Qualitätsstufen komprimiert. Für jede der Qualitätsstufen wird der Speicherplatzbedarf für die komprimierte Bildsammlung sowie der Mittelwert der mit dem SSIM Algorithmus gemessenen Qualität bestimmt. Anschließend wird ermittelt, welches die niedrigste JPEG 2000 Qualitätsstufe ist, für welche die Qualität der Bilder im Mittel gleich oder besser ist als unter Verwendung des JPEG Kompressionsverfahrens in der Einstellung 75% Qualität und der Speicherplatzbedarf verglichen.

#### Versuchsaufbau

Die JPEG Kompression wird direkt aus Visual Studio 2013 unter Verwendung des GDI+ JPEG Encoders durchgeführt. Der für die Qualität zuständige Parameter wird auf 75 gesetzt. Für die JPEG 2000 Kompression wird die Magick.NET Bibliothek verwendet, welche einen Wrapper für ImageMagick für .NET darstellt.<sup>2</sup> Für die Messung der Qualität wird die SSIM Implementation für C# von Chris Lomont verwendet.

Die erzeugten JPEG 2000 Bilder werden hierfür ohne weiteren Qualitätsverlust in PNGs umgewandelt, da der SSIM Algorithmus die JPEG 2000 komprimierten Bilder nicht direkt verarbeiten kann. Hierfür wird ebenfalls Magick.NET verwendet.

---

<sup>2</sup><http://magick.codeplex.com/>, letzter Zugriff 20.10.2015

## 4.2.2 Versuch 2: Komprimierung mit verschiedenen Verfahren und anschließender Vergleich

**Ansatz:** Dieser Versuch erzeugt unter den festgelegten Rahmenbedingungen eine optimal komprimierte Bildsammlung, welche als Referenz für die anschließenden Versuche verwendet wird. Die Erkennung muss in den anderen Versuchen schneller erfolgen, um sinnvoll einsetzbar zu sein. Wenn die Erkennung langsamer ist als in diesem Versuch, ist sie auf jeden Fall schlechter, da die hier verwendete Erkennung in Bezug auf die Dateigröße der komprimierten Bildsammlung bereits das optimale Ergebnis liefert und daher nur noch in der Geschwindigkeit überboten werden kann

Von jedem Bild der Sammlung wird je eine Kopie mit Deflate, LZW und JPEG mit 75% Qualität sowie JPEG 2000 mit Qualität<sup>3</sup> 38 komprimiert und anschließend das komprimierte Bild mit der kleinsten Dateigröße behalten, die restlichen Bilder werden verworfen. Unter der Voraussetzung, dass keine anderen Verfahren oder Parameter verwendet werden dürfen, entsteht so eine komprimierte Bildsammlung mit optimaler Größe. Da Graustufenbilder sich in Bezug auf die Komprimierung anders verhalten könnten als Farbbilder, wird der Versuch für beide Kategorien jeweils einmal durchgeführt. Für Farbbilder wird zusätzlich die für die Erkennung des zu verwendenden Kompressionsverfahrens benötigte Zeit gemessen.

**Versuchsaufbau** Für jedes Bild werden die 4 unterschiedlich komprimierten Kopien in den Arbeitsspeicher geschrieben, die Kopie mit dem geringsten Speicherplatzbedarf wird auf die Festplatte kopiert. Es wird die Zeit gemessen, welche für das Schreiben in den Arbeitsspeicher und den anschließenden Vergleich des Speicherplatzbedarfs benötigt wird.

Anschließend werden alle Referenzbilder mit Magick.NET in Graustufenbilder mit 8 Bit pro Pixel umgewandelt und der Versuch erneut durchgeführt.

## 4.2.3 Versuch 3: Aus Metadaten

**Ansatz** Wenn die Bilder bereits als PNG vorliegen, also schon mit Deflate komprimiert wurden, lässt sich aus deren Abmessungen und der Dateigröße direkt das Kompressionsverhältnis zwischen Rohdaten und komprimiertem Bild errechnen.

Vermutung: Ab einem zu bestimmenden Kompressionsverhältnis könnte Deflate grundsätzlich stärker komprimieren als JPEG 2000.

Aus den Metadaten werden die Abmessungen des Bildes in Pixeln ausgelesen. Da das Bild bereits als PNG vorliegt und somit schon mit Deflate komprimiert wurde ([libpng 2003](#)), lässt sich aus dem Verhältnis zwischen Abmessungen und Dateigröße der Kompressionsfaktor annähernd bestimmen. Das unkomprimierte RGB Bild hat

---

<sup>3</sup>Einstellung des verwendeten Magick.NET Encoders für JPEG 2000. Laut Handbuch handelt es sich bei dieser Angabe um das Spitzen-Signal-Rausch-Verhältnis in dB. <http://www.imagemagick.org/script/jp2.php>, letzter Zugriff 5.11.2015

abzüglich Header oder Metadaten eine Dateigröße von

$$3 \cdot \text{Höhe} \cdot \text{Breite}$$

Byte, wobei Höhe und Breite jeweils in Pixeln gemessen werden. Der Wert ist deswegen ungenau, da die Header und Metadaten ignoriert werden. Diese haben für die verwendeten Bilder eine weitgehend einheitliche Größe und sind sehr klein<sup>4</sup>, so dass sie vernachlässigt werden können. Es soll untersucht werden, wie gut sich das optimale Kompressionsverfahren anhand des Kompressionsverhältnisses zwischen Rohdaten und Deflate komprimierten Daten bestimmen lässt.

**Versuchsaufbau** Jedes Bild der Sammlung wird in das PNG Format mit 24 Bit pro Pixel verlustfrei umgewandelt, anschließend wird die Abmessungen in Pixeln sowie die Dateigröße  $D$  in Byte aus den Metadaten ausgelesen und der Kompressionsfaktor mit

$$\frac{3 \cdot \text{Höhe} \cdot \text{Breite}}{D}$$

ermittelt. Alle Bilder mit Kompressionsfaktor  $< x$  werden mit JPEG 2000 Qualität 38 komprimiert und gespeichert, Bilder mit Kompressionsfaktor  $> x$  mit Deflate.  $x$  wird solange verändert, bis der Wert gefunden ist, für welchen die Dateigröße der komprimierten Bildsammlung am kleinsten ist.

### 4.2.4 Versuch 4: Mittels Entropie/Farbinformationen

**Ansatz** Vermutung: Bilder mit einer niedrigen Entropie könnten sich besser mit Deflate besser komprimieren lassen als mit JPEG 2000. Es soll untersucht werden, ob die Entropie und Farbinformationen ausreichend sind, um automatisch das optimale Kompressionsverfahren auswählen zu können.

**Versuchsaufbau** Mit Hilfe von Magick.NET werden die Entropie sowie die Anzahl der im Bild vorhandenen Farben für jedes Bild berechnet und zusammen mit dem Kompressionsfaktor und dem in Versuch 2 ermittelten optimalen Kompressionsverfahren in eine Tabelle geschrieben. Anhand der Tabelle werden die jeweiligen Grenzwerte für Entropie und Farbinformationen so gewählt, dass so wenig Fehlerkennungen wie möglich auftreten. Es wird zusätzlich die Zeit gemessen, welche für die Messung der Eigenschaften und die Auswahl des Kompressionsverfahrens benötigt wird.

### 4.2.5 Versuch 5: Kombination aus Metadaten, Entropie und Farbinformationen

**Ansatz** Beobachtung: Die Metadaten lassen sich verwenden, um den Kompressionsfaktor zu berechnen und eine Prognose darüber zu erstellen, ob das Bild mit Deflate

---

<sup>4</sup>Die Bildsammlung benötigt insgesamt 5 von 265 MB für die Metadaten.

oder JPEG 2000 stärker komprimiert werden kann. Durch die zusätzliche Messung der Entropie und Farbinformationen kann die Genauigkeit verbessert werden, da so eine zusätzliche Unterscheidung zwischen JPEG und JPEG 2000 möglich ist.

Anhand der Tabelle aus Versuch 4 werden die Grenzwerte für Kompressionsfaktor, Entropie und Farbinformationen so gewählt, dass die Anzahl der Fehlerkennungen minimal ist.

**Versuchsaufbau** Für alle Bilder werden unter Verwendung vom Magick.NET der Kompressionsfaktor, die Entropie und die Anzahl der vorhandenen Farben gemessen, anschließend wird das Kompressionsverfahren anhand der im Ansatz festgelegten Kriterien ausgewählt und eine komprimierte Kopie des Bildes gespeichert.

#### 4.2.6 Versuch 6: Wiederholung von Versuch 3 unter Berücksichtigung der in Versuch 4 gewonnenen Erkenntnisse

**Ansatz** Aus der in Versuch 4 erzeugten Tabelle lässt sich ablesen, dass eine genauere Festlegung des Grenzwertes für den Kompressionsfaktor sowie eine zusätzliche Unterscheidung zwischen JPEG und JPEG 2000 die Genauigkeit der Erkennung weiter verbessern sollte.

**Versuchsaufbau** Jedes Bild der Bildsammlung wird in das PNG Format mit 24 Bit pro Pixel verlustfrei umgewandelt, anschließend die Abmessungen in Pixeln sowie die Dateigröße  $D$  in Byte aus den Metadaten ausgelesen und der Kompressionsfaktor mit

$$\frac{3 \cdot \text{Höhe} \cdot \text{Breite}}{D}$$

ermittelt. Anschließend wird das Kompressionsverfahren unter Berücksichtigung der Ergebnisse aus Versuch 4 ausgewählt.

# 5 Auswertung

## 5.1 Messergebnisse

### 5.1.1 Versuch 1: Qualitätsunterschiede zwischen JPEG/JPEG 2000

Mit einer Qualitätseinstellung von Qualität 40 oder weniger lassen die Bilder mit JPEG 2000 stärker komprimieren als mit JPEG 75% Qualität. Die höchste Kompression, bei welcher die Qualität laut SSIM noch mindestens so hoch ist wie bei JPEG, kann mit JPEG 2000 Qualität 38 erreicht werden.

**Tabelle 5.1:** Qualitätsunterschiede zwischen JPEG und JPEG 2000

	SSIM	Dateigröße (Summe aller Dateien in MB)
PNG (Deflate)	1	265
JPEG 2000 Qualität 41	nicht gemessen	148
JPEG 2000 Qualität 40	0,9970	136
JPEG 2000 Qualität 39	0,9965	124
JPEG 2000 Qualität 38	0,9956	113
JPEG Qualität 75%	0,9949	139
JPEG 2000 Qualität 37	0,9944	103

Unter Verwendung jener Qualitätseinstellung lässt sich der Speicherplatzbedarf für die vorliegende Bildsammlung gegenüber der herkömmlichen JPEG Kompression um 18,7% reduzieren und zusätzlich eine etwas bessere Qualität erreichen.

### 5.1.2 Versuch 2: Komprimierung mit verschiedenen Verfahren und anschließender Vergleich

Für die unter oben genannten Voraussetzungen optimal komprimierte Bildsammlung ergibt sich daraus für Farbbilder folgende Verteilung: 25 der Bilder wurden mit Deflate komprimiert. 126 Bilder wurden mit JPEG 2000 komprimiert 140 Bilder wurden mit JPEG komprimiert. Keines der Bilder wurde mit LZW komprimiert, LZW scheint Deflate also grundsätzlich unterlegen zu sein. Die Dateigröße der komprimierten Bildsammlung beträgt 104 MB. Die Dauer für die Erkennung der Kompressionsverfahren beträgt hier 11 Minuten und 24 Sekunden.

Nachdem die Farbbilder in 8-Bit Graustufen umgewandelt wurden, beträgt die Dateigröße der Bildsammlung vor Optimierung 124 MB. Nach der Optimierung beträgt die Dateigröße 76 MB, wobei nun 68 der Bilder mit Deflate, 162 der Bilder mit JPEG 2000 und 61 der Bilder mit JPEG komprimiert wurden.

### 5.1.3 Versuch 3: Aus Metadaten

Der aus den Abmessungen des Bildes und der Dateigröße errechnete Kompressionsfaktor lässt sich gut verwenden um abzuschätzen, ob das Bild stärker mit Deflate oder JPEG 2000 komprimiert wird. Ab einem Kompressionsfaktor von 25 lassen sich die RGB Bilder bis auf wenige Ausnahmen mit Deflate am stärksten komprimieren.

Wenn der Kompressionsalgorithmus alleine anhand des Kompressionsfaktors gewählt wird und man festlegt, dass alle Bilder mit einem Kompressionsfaktor  $\geq 25$  mit Deflate und alle Bilder mit einem Kompressionsfaktor  $< 25$  mit JPEG 2000 komprimiert werden sollen, beträgt der Speicherplatzbedarf für die komprimierte Bildsammlung 112 MB. Die Dauer für die Erkennung der optimalen Kompressionsverfahren beträgt 0.08 Sekunden (Die Dauer variiert bei mehrmaliger Wiederholung des Versuches leicht, was an der leicht unterschiedlichen Auslastung des Rechners durch im Hintergrund laufende Prozesse liegt).

Nach der Umwandlung in Graustufen wird die Bildsammlung am stärksten komprimiert, wenn alle Bilder mit einem Kompressionsfaktor  $> 12$  mit Deflate und der Rest mit JPEG 2000 komprimiert werden. Die so komprimierte Bildsammlung hat eine Dateigröße von 77,9 MB.

### 5.1.4 Versuch 4: Mittels Entropie/Farbinformationen

Für die vorliegende Bildsammlung werden alle RGB Bilder, welche eine Entropie<sup>1</sup> von 0,184 oder weniger besitzen, mit Deflate stärker komprimiert als mit JPEG 2000. Allerdings haben nur 9 der 25 Bilder, welche mit Deflate stärker komprimiert werden, eine Entropie  $< 0,184$ . Bilder mit einer Entropie  $> 0,86$  werden größtenteils stärker mit JPEG komprimiert als mit JPEG 2000, ab einer Entropie von 0,96 werden alle vorliegenden Bilder mit JPEG am stärksten komprimiert. Aus der Anzahl der vorhandenen Farben lassen sich für die Bildsammlung nur begrenzt sinnvolle Schlüsse ziehen. Alle Bilder der Bildsammlung mit weniger als 9177 Farben werden mit Deflate am stärksten komprimiert, alle Bilder der Sammlung mit mehr als 450130 Farben lassen sich mit JPEG am stärksten komprimieren. Die unter diesen Kriterien komprimierte Bildsammlung hat eine Dateigröße von 105 MB. Die Dauer für die Erkennung der Kompressionsverfahren beträgt 3 Minuten und 18 Sekunden.

Nachdem die Bilder in Graustufen umgewandelt wurden, ist der Kompressionsfaktor für die Wahl des Kompressionsverfahrens ein besserer Indikator als die Entropie

---

<sup>1</sup>Es ist nicht näher dokumentiert, wie ImageMagick die Entropie misst. Es handelt sich hierbei vermutlich um die relative Entropie, siehe (Shannon 1948: 24-25).

oder die Anzahl der vorhandenen Farben. Durch Auswertung der erzeugten Tabelle lassen sich keine sinnvollen Kriterien bestimmen, um zwischen JPEG und JPEG 2000 zu unterscheiden. Da JPEG 2000 die Gesamtheit der Graustufenbilder stärker komprimieren kann als JPEG, wurde lediglich zwischen Deflate und JPEG 2000 Kompression gewählt. Alle Bilder der Bildsammlung, welche weniger als 197 Farben beinhalten oder eine Entropie  $< 0,34$  besitzen, werden mit Deflate komprimiert, der Rest mit JPEG 2000. Die komprimierte Bildsammlung benötigt in diesem Fall 78,2 MB Speicherplatz.

### 5.1.5 Versuch 5: Kombination aus Metadaten, Entropie und Farbinformationen

Aus dem vorherigen Versuch kann ermittelt werden, dass sich diejenigen Bilder sich besten mit Deflate komprimieren lassen, welche entweder eine Entropie  $\leq 0,184$ , einen Kompressionsfaktor  $> 32$  oder weniger als 9177 Farben beinhalten. Alle Bilder der Bildsammlung, welche entweder eine Entropie  $> 0,967$ , einen Kompressionsfaktor  $< 1.82$  oder mindestens 450130 Farben enthalten, lassen sich mit JPEG am stärksten komprimieren. Für den Rest soll JPEG 2000 verwendet werden.

Die Graustufenbilder werden mit Deflate komprimiert, wenn Sie entweder weniger als 196 Farben, einen Kompressionsfaktor  $> 11,76$  oder eine Entropie  $< 0,34$  besitzen. Bilder, welche einen Kompressionsfaktor  $< 1,178$  besitzen, werden mit JPEG komprimiert. Der Rest wird mit JPEG 2000 komprimiert. Durch die Kombination der 3 Bildeigenschaften Entropie, Farbinformationen und Kompressionsfaktor kann das Ergebnis für Farbbilder verbessert werden. Die Dateigröße der komprimierten Bilddaten beträgt nun für die Farbbilder 104 MB, die Differenz zu der unter den anfangs festgelegten Rahmenbedingungen optimal komprimierten Bildsammlung beträgt  $< 0,5\%$  (511 KB).

Für die Graustufenbilder konnte das Ergebnis aus Versuch 3 nicht verbessert werden, die Dateigröße der komprimierten Bildsammlung beträgt hier 77,9 MB.

### 5.1.6 Versuch 6: Wiederholung von Versuch 3 unter Berücksichtigung der in Versuch 4 gewonnenen Erkenntnisse

Die Ergebnisse aus Versuch 3 konnten leicht verbessert werden, die Dateigröße der komprimierten Bildsammlung beträgt nun 107 MB (112.293.909 Bytes)

## 5.2 Bewertung der Messergebnisse

Die Ergebnisse aus Versuch 1 sind nur unter der Annahme relevant, dass die Qualitätseinstellung „JPEG 75%“ in Ghostscript und dem verwendeten GDI+ Encoder

identische Ergebnisse liefert. Da der JPEG Standard jedoch keine exakte Implementation sondern vielmehr Mindestanforderungen an den Encoder definiert, ist dies nicht zwangsläufig gegeben. Auch der JPEG 2000 Standard definiert keine konkrete Implementation des Encoders. Die gewonnene Erkenntnis ist also strenggenommen nur für die tatsächlich verwendeten Encoder gültig: Der JPEG 2000 Encoder aus Magick.NET komprimiert Bilder mit der Einstellung Qualität=38 besser, als der JPEG Encoder aus GDI+ mit der Einstellung 75% Qualität. Zudem korreliert der SSIM Algorithmus nicht zu 100% mit den Ergebnissen subjektiver Befragungen und definiert Qualität anhand der Ähnlichkeit von in 2 Bildern vorhandenen Strukturen (Wang 2004: 611). Dies ist zu tolerieren, da es kein algorithmisches Verfahren gibt, welches zu 100% mit den Ergebnissen subjektiver Befragungen übereinstimmt und der SSIM Algorithmus bessere Ergebnisse liefert als die meisten Algorithmen zur Messung von Bildqualität (Ponomarenko 2009).

Da in Versuch 2 die Bilder jeweils mit jedem der vorgesehenen Kompressionsverfahren komprimiert werden und anschließend nur das komprimierte Bild mit der kleinsten Dateigröße behalten wird, sind die Ergebnisse hier unter Vernachlässigung der Metadaten in jedem Fall korrekt und dienen als Referenz für die weiteren Versuche.

Versuch 3 liefert ein etwas schlechteres Ergebnis als die beiden nachfolgenden, was vor allem an der Art und Weise liegt, mit welcher dieser durchgeführt wurde. Der Grenzwert für die Unterscheidung zwischen Deflate und JPEG 2000 wurde hier nur in ganzzahligen Schritten verändert. Dies hat den Hintergrund, dass für jeden Schritt sämtliche Bilder einmal komprimiert werden mussten und somit eine kleinere Schrittgröße den Aufwand um ein Vielfaches vergrößert hätte. Außerdem wurde keine Unterscheidung zwischen JPEG und JPEG 2000 durchgeführt. Durch Auswertung der in Versuch 4 erzeugten Tabelle wird deutlich, dass durch eine genauere Festlegung des Grenzwertes und eine zusätzliche Unterscheidung zwischen JPEG und JPEG 2000 ein besseres Ergebnis erzielt werden kann, daher wird dies in Versuch 6 nachgeholt.

Versuch 4 zeigt, dass die Messung der Entropie und der Anzahl der enthaltenen Farben ebenfalls eine gute Möglichkeit bietet, ein geeignetes Kompressionsverfahren auszuwählen. Diese Methode liefert auf Grund der Vernachlässigung des Kompressionsfaktors in Bezug auf die Dateigröße geringfügig schlechtere Ergebnisse, bietet jedoch den Vorteil, dass sie für ein beliebiges Eingangsformat funktioniert, während die Methode aus Versuch 5 voraussetzt, dass die Bilddaten bereits mit Deflate komprimiert wurden. Die Methode ist jedoch im Vergleich zu der alleinigen Auswertung der Metadaten wesentlich langsamer, da hier die Bilder vollständig geladen und alle Pixel ausgewertet werden.

Die in Versuch 5 verwendete Kombination aus Kompressionsfaktor, Entropie und Farbwerten liefert in Bezug auf die Dateigröße beinahe das in Versuch 2 ermittelte optimale Ergebnis. Mit der Methode aus Versuch 2 lässt sich die Bildsammlung zwar 0,5% stärker komprimieren, der Zeitaufwand ist jedoch um ein vielfaches größer. Dies liegt vor allem daran, dass die JPEG 2000 Kompression zeitaufwändig ist und jedes Bild mehrfach komprimiert werden muss, um das optimale Verfahren zu bestimmen. Die Zeit wurde hier nicht gemessen, da diese mit der in Versuch 4 gemessenen Zeit

übereinstimmen sollte, wenn man die 0.08 Sekunden für die zusätzliche Auswertung der Metadaten vernachlässigt.

Die Methode aus Versuch 6 bietet den besten Kompromiss aus Genauigkeit und Geschwindigkeit der Erkennung. Die resultierende Dateigröße ist hier zwar 3% weiter von der optimalen Größe entfernt als in Versuch 5, die Erkennung ist dafür jedoch etwa 2500-mal schneller.

### 5.3 Vergleich mit Adobe Distiller

Die automatische Erkennung des Kompressionsverfahrens für Bilder in Adobe Distiller funktioniert nach anderen Kriterien, daher ist ein direkter Vergleich nicht möglich. Die dortige automatische Erkennung ist darauf ausgelegt, dass die verlustbehaftete JPEG oder JPEG 2000-Kompression besonders für Bilder mit „*scharfen Kanten*“ (Adobe 2015: 10) ungeeignet ist und orientiert sich daher an den im Bild vorhandenen Frequenzen. Das dortige Verfahren muss zunächst konfiguriert werden und unterscheidet dann entweder zwischen Deflate und JPEG, oder zwischen JPEG 2000 (verlustbehaftet) und JPEG 2000(verlustfrei)(Adobe 2015: 10). In der Standardeinstellung für Farbbilder und Graustufenbilder wird zwischen JPEG und Deflate unterscheiden. Es erfolgt keine unterschiedliche Erkennung für Farb- und Graustufenbilder.

# 6 Fazit

## 6.1 Ergebnis

Unter den Anfangs festgelegten Rahmenbedingung lässt sich eine Methode zur automatischen Erkennung des Kompressionsverfahrens entwickeln, welche dieses nahezu optimal auswählt. Für diejenigen Bilder, bei denen das optimale Kompressionsverfahren falsch erkannt wird, liefert das falsch erkannte Kompressionsverfahren nur vernachlässigbar schlechtere Ergebnisse, als das optimale Verfahren. Die Abweichung zu der optimalen Dateigröße beträgt hierbei weniger als 0,5%, wenn zur Auswahl des Kompressionsverfahrens die Entropie, die Anzahl der im Bild enthaltenen Farben und der Kompressionsfaktor zwischen Rohdaten und Defalte komprimierten Bilddaten verwendet wird. Diese Methode ist jedoch wesentlich langsamer als eine alleinige Auswertung der Metadaten.

Anhand der reinen Metadaten lässt sich das optimale Kompressionsverfahren für Farbbilder etwa 3%<sup>1</sup> schlechter bestimmen als unter Berücksichtigung einer Kombination verschiedener Eigenschaften. Da die Metadaten jedoch etwa 2500-mal schneller ausgelesen werden können, als die Entropie oder Farbinformationen, ist diese Methode zu bevorzugen, wenn die Geschwindigkeit der Erkennung von Bedeutung ist.

Für Graustufenbilder lässt sich das optimale Kompressionsverfahren am besten bestimmen, wenn lediglich der Kompressionsfaktor zwischen Rohdaten und Deflate komprimierten Daten für die Bestimmung verwendet wird. Die Entropie und Anzahl verschiedener Grautöne geben in diesem Fall keinen Aufschluss über das optimale Kompressionsverfahren und verlangsamen die Erkennung unnötig.

Die JPEG 2000-Kompression liefert für viele Bilder in Bezug auf Dateigröße und Qualität bessere Ergebnisse als die JPEG-Kompression. Vor allem Bilder, welche eine Mischung aus einfarbigen Flächen und Fotos enthalten, werden mit JPEG 2000 stärker komprimiert. Lediglich 4 von 125 Bildern des DatasetDBD, welches ausschließlich aus Bildern mit einer Mischung aus Text auf einfarbigem Hintergrund und Fotos besteht, können mit JPEG stärker komprimiert werden als mit JPEG 2000.

Es gibt allerdings auch Bilder, welche sich mit JPEG stärker komprimieren lassen. Besonders Bilder mit einer sehr hohen Entropie lassen sich mit JPEG stärker komprimieren als mit JPEG 2000.

Die Erkennung des optimalen Kompressionsverfahrens wird dadurch etwas vereinfacht, dass manche der erlaubten Kompressionsverfahren von Anfang an ausgeschlossen werden können. Die LZW Kompression z.B. ist nur aus historischen Gründen

---

<sup>1</sup>In Bezug auf die resultierende Dateigröße aller komprimierten Bilder.

verfügbar, da Deflate in den ersten Versionen des PDF-Standards noch nicht vorgesehen war.

## 6.2 Probleme

Das optimale Kompressionsverfahren für ein Bild lässt sich nicht ohne Einschränkungen definieren, da es nicht nur vom Bildinhalt, sondern letztlich auch vom Verwendungszweck des komprimierten Bildes abhängig ist. Es ist zunächst erforderlich, entweder eine Definition für ausreichende Qualität zu finden, oder eine Höchstgrenze für den Speicherplatzbedarf festzulegen. Wenn man sich dazu entscheidet, eine ausreichende Qualität zu definieren, muss man entweder zeitaufwändige Personenbefragungen durchführen, oder sich auf einen der vorhandenen Algorithmen zur Bestimmung der Qualität verlassen. Die Definition der ausreichenden Qualität hat jedoch gravierende Auswirkungen auf die Auswahl des optimalen Kompressionsverfahrens: Wenn man die Ergebnisse aus Versuch 1 und Versuch 2 betrachtet stellt man fest, dass bei einer etwas geringeren Anforderung an die Qualität die gesamte Bildsammlung unter alleiniger Verwendung von JPEG 2000 auf eine geringere Dateigröße reduziert werden kann als die in Versuch 2 ermittelte optimale Größe.

Die gewonnenen Erkenntnisse bauen größtenteils darauf auf, dass der SSIM Algorithmus die Qualitätsunterschiede zwischen JPEG und JPEG 2000 korrekt identifizieren kann. Diese SSIM Werte stehen jedoch teilweise im Konflikt mit den Ergebnissen der Untersuchung von (Ebrahimi 2004), in welcher festgestellt wurde, dass JPEG 2000 erst ab einer Kompressionsrate von 20:1 qualitativ bessere Ergebnisse liefert als JPEG. Da die dort verwendete Bildsammlung jedoch nur 29 Referenzbilder enthält, ist eine allgemeine Gültigkeit nicht gegeben.

Zudem ist auch problematisch, dass nicht alle JPEG Encoder qualitativ gleichwertige Bilder produzieren und der verwendete JPEG Encoder aus GDI+ nicht ausreichend dokumentiert ist.

Außerdem stellt die große Vielfalt der möglichen Bildinhalte eine Herausforderung dar: Selbst die 291 Bilder, für welche im Zusammenhang dieser Ausarbeitung Versuche durchgeführt wurden, sind nicht genug, um ein komplett allgemeingültiges Auswahlverfahren zu entwickeln, da die Ergebnisse der verschiedenen Kompressionsverfahren extrem abhängig vom eigentlichen Bildinhalt sind und 291 Bilder nur einen Bruchteil aller möglichen Bildinhalte darstellen.

JPEG 2000 funktioniert dort besonders gut, wo die ursprüngliche JPEG Kompression Schwachstellen hat, ist aber langsamer und wird nicht nativ von dem .NET Framework<sup>2</sup> unterstützt, erfordert also zusätzliche Software von Drittanbietern. Außerdem ist JPEG 2000 erst ab PDF Version 1.5 verfügbar, also nur begrenzt abwärtskompatibel (Adobe 2006). Daher macht es Sinn, die Verwendung der JPEG 2000 Kompression als mögliche Option für das automatische Auswahlverfahren gesondert ein- und ausschalten zu können.

---

<sup>2</sup>Das .NET Framework ist die Entwicklungsplattform der Endanwendung.

## 6.3 **Ausblick**

Es ist denkbar, dass in Zukunft zusätzliche Kompressionsverfahren in den PDF-Standard aufgenommen werden; sowohl Deflate als auch JPEG 2000 waren in der ursprünglichen Fassung nicht enthalten und sind erst später hinzugefügt worden. Wünschenswert ist hier ein Kompressionsverfahren, welches für jegliche Art von Bildmaterial ein besseres Ergebnis liefert als die derzeit vorhandenen Verfahren. Dies hätte den Vorteil, dass weder der Nutzer noch die Software das Kompressionsverfahren auswählen müsste.

# **A Anhang 1: Messwerte für Farbbilder aus Versuch 4**

# A Anhang 1: Messwerte für Farbbilder aus Versuch 4

Bild	Entropie	Farben	Kompressionsfaktor	optimales Kompressionsverfahren
3	0,764639104	5	154,6500932	deflate
136_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,17435063	123	52,98592166	deflate
106_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,267939196	164	26,26924398	deflate
1194985138747548807no_cycles.svg.thumb	0,255543973	295	5,558643691	deflate
40_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,227248682	321	26,91134868	deflate
122_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,175477146	334	31,35453254	deflate
1194984884404348229protection_obligatoire_10.svg.med	0,156098099	665	11,24297314	deflate
screenshot	0,371125785	881	45,03554297	deflate
heart-th	0,447130752	923	5,312566674	deflate
119543096047570093johnny_automatic_bending_bee.svg.med	0,1268251	1110	9,755962261	deflate
i25	0,802217668	1159	9,243151758	deflate
1194984554962989016xbill_01.svg.med	0,324972962	1178	8,553645525	deflate
1195429565207837224jeanmichel_VTT.svg.med	0,230734421	1263	4,574890215	deflate
green-bike-th	0,285602262	1832	2,004454343	deflate
1195431012993269747johnny_automatic_champagne_on_ice.svg.med	0,386116022	3060	3,998954811	deflate
green-falcon-md	0,326482453	5935	5,722337378	deflate
11949837181420514439mascot_enrique_meza_c_01.svg.med	0,421674543	6757	5,657803221	deflate
22_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,162833274	9176	30,85527526	deflate
120_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,317583958	9754	29,92973067	JPEG2000
3_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,134037202	9799	36,19947556	deflate
i02	0,757374344	11101	1,775375122	JPEG
i01	0,913938883	11304	1,459292156	JPEG
18_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,183956379	11318	27,82434549	deflate
128_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,302172565	11458	29,01310327	JPEG2000
i16	0,912125028	12132	1,992265003	JPEG2000
i17	0,906073842	14768	1,821630753	JPEG2000
wholescreen	0,381354917	15341	25,72460022	deflate
36_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,280712591	17231	17,55209115	JPEG2000
i10	0,916863304	18104	1,849482461	JPEG2000
54_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,152152999	18437	32,40909085	deflate
52_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,159155608	18642	27,97867476	deflate
11_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,290985813	20021	21,70288287	deflate
i20	0,758177686	20194	2,175107866	JPEG2000
i04	0,881197996	20475	1,900634808	JPEG2000
i12	0,908238542	21527	1,869430446	JPEG
i19	0,94877253	21789	1,63684954	JPEG
i06	0,938766453	22456	1,585689053	JPEG
126_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,294444325	22904	16,98183483	JPEG2000
i24	0,884857225	23015	1,708391866	JPEG
i09	0,926276386	23423	1,834246584	JPEG2000
i21	0,885508887	24201	1,619981708	JPEG
i03	0,889002574	27766	2,108788765	JPEG2000
i15	0,869671659	28201	1,881824447	JPEG
i11	0,90421672	29942	1,65417903	JPEG
i13	0,928355956	30195	1,286799821	JPEG
i07	0,91447876	30303	1,969066417	JPEG2000
78_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,281049717	32262	23,17829802	deflate
188_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,308134644	32750	23,08589162	JPEG2000
i08	0,94946662	34387	1,45044633	JPEG
grenze	0,454816191	38840	15,17523681	JPEG
ucid00054	0,924438321	39476	2,075186383	JPEG2000
i18	0,879366849	40022	1,475482176	JPEG
ucid00015	0,715009247	41709	2,393766234	JPEG
42_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,218570121	42721	31,6732064	JPEG2000
i22	0,905155197	42903	1,594907752	JPEG
test	0,48828463	44050	11,06207122	JPEG
i14	0,938666855	47719	1,49942039	JPEG
90_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,223327882	49032	26,35752752	JPEG2000
ucid00058	0,947243967	49625	2,026454845	JPEG2000
i05	0,938351068	51073	1,397445466	JPEG
i23	0,93837628	52698	1,941136397	JPEG2000
ucid00040	0,879193867	53071	1,843119364	JPEG
112_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,297845046	54835	19,64857165	JPEG2000
207_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,851051266	54876	13,76626006	JPEG2000
48_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,321895826	56715	20,26787643	JPEG2000
ucid00073	0,862506638	56717	1,79546859	JPEG
68_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,316504913	59714	21,13070084	JPEG2000
139_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,361255487	60288	24,22258402	JPEG2000
ucid00077	0,8598981	60604	1,935029231	JPEG
168_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,338196015	60666	22,678614	JPEG2000
ucid00047	0,92180745	60728	1,979534234	JPEG2000
ucid00065	0,862498865	62573	1,959665362	JPEG2000
ucid00062	0,906385414	62620	1,692537699	JPEG
190_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,34609973	63132	17,36562083	JPEG2000
ucid00045	0,923461046	63458	1,67408118	JPEG
ucid00035	0,918205747	64980	1,787594634	JPEG
76_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,317861804	65072	19,24318967	JPEG2000
ucid00048	0,938359521	65991	1,755919811	JPEG

# A Anhang 1: Messwerte für Farbbilder aus Versuch 4

Bild	Entropie	Farben	Kompressionsfaktor	optimales Kompressionsverfahren
ucid00081	0,887874212	66867	1,940778715	JPEG
ucid00072	0,951356444	68042	1,899606439	JPEG2000
ucid00009	0,912453732	69936	1,86786161	JPEG
140_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,254758922	70005	19,56935628	JPEG2000
30_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,237885955	70594	22,00124362	JPEG2000
142_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,355085909	70690	19,70639241	JPEG2000
ucid00050	0,947412325	70859	1,751082135	JPEG
ucid00001	0,883189478	71357	1,668073168	JPEG
92_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,403199972	71692	18,07449041	JPEG2000
ucid00018	0,883097422	72303	1,776016621	JPEG
ucid00067	0,903643254	73657	1,774595257	JPEG
66_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,369319779	74418	17,32486021	JPEG2000
187_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,535355537	74710	9,156156284	JPEG2000
107_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,432622832	74779	12,99584452	JPEG2000
ucid00059	0,926548373	74814	1,673126674	JPEG
131_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,292091486	74867	19,83498537	JPEG2000
ucid00028	0,94143696	75926	1,731369746	JPEG
ucid00039	0,965902422	76414	1,818205358	JPEG
ucid00057	0,904173668	76698	1,826114417	JPEG
ucid00071	0,924445927	77008	1,787405035	JPEG
ucid00080	0,901849476	77836	1,393625216	JPEG
98_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,282006255	78300	21,1853692	JPEG2000
34_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,293774255	78881	21,38073323	JPEG2000
74_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,293005968	79033	19,24341458	JPEG2000
ucid00029	0,941809341	79061	1,562700198	JPEG
ucid00064	0,901491675	80620	1,812946456	JPEG
124_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,228083716	80717	22,28799157	JPEG2000
ucid00063	0,916817616	81113	1,852264519	JPEG
176_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,376803552	81273	16,14794589	JPEG2000
ucid00042	0,954639038	84002	1,628952078	JPEG
64_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,282262577	84368	28,00256248	JPEG2000
ucid00056	0,901260653	84547	1,799714401	JPEG
ucid00004	0,879707695	84796	1,401007126	JPEG
16_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,257702805	85115	19,63107931	JPEG2000
116_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,368125531	85608	21,35191546	JPEG2000
ucid00041	0,944489758	86168	1,79818908	JPEG
ucid00002	0,885002119	86351	1,635546879	JPEG
ucid00033	0,912261287	87039	1,54660877	JPEG
96_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,333994345	87146	17,10963457	JPEG2000
ucid00049	0,959006041	88058	1,904138379	JPEG2000
ucid00043	0,932413839	88092	1,561500441	JPEG
ucid00034	0,907604321	89181	1,526001511	JPEG
ucid00051	0,96656724	89508	1,926786403	JPEG2000
7_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,290058192	89673	21,54298993	JPEG2000
149_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,312070107	89800	14,95842999	JPEG2000
ucid00012	0,923880507	90583	1,497471311	JPEG
ucid00082	0,950876061	90787	1,511958288	JPEG
178_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,425945244	91881	14,30730542	JPEG2000
ucid00061	0,938375288	92196	1,709387045	JPEG
ucid00030	0,936867444	92452	1,449273795	JPEG
ucid00044	0,896880505	92873	2,07773762	JPEG
ucid00046	0,932925815	92885	1,62583142	JPEG
84_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,284990499	93102	17,82739437	JPEG2000
114_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,465833812	93503	14,46912261	JPEG2000
ucid00103	0,915709215	94066	1,710393305	JPEG
ucid00074	0,960680728	94571	1,929862906	JPEG
ucid00031	0,950160092	95471	1,606226393	JPEG
ucid00075	0,905755925	95540	1,517001921	JPEG
ucid00070	0,961993132	96964	1,650891748	JPEG
ucid00008	0,923969235	97347	1,506928831	JPEG
ucid00032	0,915637117	98312	1,602007725	JPEG
ucid00027	0,919154447	98347	1,535592108	JPEG
ucid00105	0,838286255	99421	1,719047189	JPEG
15_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,570866318	99978	11,24848264	JPEG2000
ucid00060	0,942444309	100023	1,780278289	JPEG
ucid00021	0,935350059	100388	1,601051037	JPEG
104_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,353865851	101409	18,84523395	JPEG2000
101_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,381587602	102625	16,96203074	JPEG2000
165_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,505744132	103012	13,0141497	JPEG2000
ucid00076	0,952784043	103111	1,765712797	JPEG
ucid00053	0,973584715	103464	1,801627447	JPEG
ucid00011	0,942973276	103621	1,608315564	JPEG
157_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,812520261	103847	7,070128866	JPEG2000
70_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,379752551	104070	17,01930647	JPEG2000
ucid00066	0,916207003	104764	1,86978602	JPEG
ucid00016	0,894374294	105074	1,733858545	JPEG
ucid00013	0,936695956	105362	1,544167197	JPEG
ucid00109	0,890700893	105907	1,464171702	JPEG

## A Anhang 1: Messwerte für Farbbilder aus Versuch 4

Bild	Entropie	Farben	Kompressionsfaktor	optimales Kompressionsverfahren
ucid00010	0,925247462	105912	1,428844132	JPEG
ucid00121	0,953453206	106650	1,635224938	JPEG
ucid00123	0,910510793	106962	1,642314182	JPEG
ucid00124	0,928562332	107162	1,56654768	JPEG
ucid00104	0,891439912	107290	1,614067926	JPEG
ucid00106	0,913177554	107508	1,582078022	JPEG
ucid00014	0,934967028	107805	1,44127026	JPEG
100_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,389595144	107835	16,86650923	JPEG2000
ucid00083	0,940833175	108680	1,536856477	JPEG
102_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,314625015	109819	19,48441101	JPEG2000
118_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,305331334	110137	18,64097579	JPEG2000
ucid00005	0,93966001	111007	1,291010572	JPEG
ucid00079	0,931272701	112010	1,424207156	JPEG
ucid00055	0,953058684	112581	1,656404193	JPEG
72_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,352334054	112589	15,22601678	JPEG2000
ucid00112	0,872617754	113278	1,680410256	JPEG
ucid00026	0,923620783	114075	1,419139071	JPEG
ucid00087	0,913085266	114136	1,348360225	JPEG
145_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,347302398	114203	16,52117337	JPEG2000
ucid00078	0,960357546	114657	1,190796234	JPEG
ucid00090	0,93763843	114855	1,561980123	JPEG
ucid00119	0,919847982	114897	1,489593725	JPEG
ucid00098	0,938899412	115137	1,533248936	JPEG
ucid00052	0,949758369	115336	1,769972902	JPEG
ucid00116	0,934873344	116975	1,581314545	JPEG
ucid00038	0,904349551	117114	1,609873928	JPEG
ucid00122	0,961129942	117143	1,564391257	JPEG
ucid00125	0,955641386	117620	1,55922597	JPEG
14_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,575955369	119457	15,01382774	JPEG2000
ucid00068	0,930504416	119688	1,405959711	JPEG
ucid00017	0,953317502	120801	1,336893402	JPEG
halfdesk	0,769280761	121630	3,398697953	JPEG
ucid00088	0,959050501	123001	1,29684092	JPEG
86_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,372096542	123352	17,76310401	JPEG2000
ucid00037	0,953266325	123646	1,243265917	JPEG
ucid00006	0,929893746	124940	1,37893435	JPEG
ucid00114	0,939264027	125053	1,518353099	JPEG
ucid00023	0,945451278	126519	1,327254643	JPEG
ucid00111	0,937382256	126848	1,482603122	JPEG
129_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,400833358	126856	17,28263343	JPEG2000
ucid00003	0,979885335	126872	1,3942017	JPEG
ucid00120	0,967719904	127385	1,321938854	JPEG
ucid00019	0,932174614	128098	1,423495652	JPEG
ucid00084	0,939648378	128188	1,231905428	JPEG
146_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,36753577	128527	17,83401128	JPEG2000
ucid00085	0,965184078	128820	1,218329977	JPEG
ucid00099	0,96189934	129345	1,305659843	JPEG
ucid00115	0,947860114	129384	1,445006014	JPEG
ucid00020	0,943702273	129884	1,401916207	JPEG
ucid00022	0,9454255	129921	1,243113939	JPEG
32_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,329783198	130404	18,65373187	JPEG2000
ucid00108	0,958811966	131972	1,448366296	JPEG
ucid00100	0,956644427	132382	1,336838861	JPEG
162_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,41433541	133425	15,80641932	JPEG2000
ucid00024	0,953802736	133430	1,197495473	JPEG
57_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,499928392	134173	11,45420116	JPEG2000
ucid00093	0,962252597	134757	1,376649263	JPEG
ucid00117	0,973694609	135455	1,442765065	JPEG
ucid00118	0,944259247	135655	1,391662164	JPEG
133_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,371879626	136289	14,64220791	JPEG2000
158_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,565678991	136622	9,887709473	JPEG2000
110_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,354275315	136737	10,49592841	JPEG2000
ucid00069	0,95536114	137345	1,595494494	JPEG
134_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,525612689	137690	15,181512	JPEG2000
95_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,556995008	138703	10,21572897	JPEG2000
ucid00101	0,949414661	138813	1,478848661	JPEG
ucid00107	0,97117791	139449	1,406351931	JPEG
ucid00025	0,960966273	141311	1,351198917	JPEG
ucid00110	0,95386496	141316	1,424929699	JPEG
160_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,427505501	141843	13,8945542	JPEG2000
ucid00086	0,929834043	143016	1,226291213	JPEG
ucid00113	0,944857984	143679	1,386874838	JPEG
28_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,38835204	143733	15,128986	JPEG2000
152_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,479711006	144279	11,71069893	JPEG2000
200_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,398547342	144332	14,23719105	JPEG2000
ucid00094	0,985534504	144737	1,307674571	JPEG
ucid00089	0,957436208	145282	1,390353255	JPEG
ucid00036	0,913576501	146217	1,308695884	JPEG

# A Anhang 1: Messwerte für Farbbilder aus Versuch 4

Bild	Entropie	Farben	Kompressionsfaktor	optimales Kompressionsverfahren
ucid00102	0,954147262	148055	1,477548342	JPEG
ucid00007	0,960431122	149088	1,293133386	JPEG
164_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,420657209	149133	14,86762385	JPEG2000
8_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,357820296	149604	14,27519381	JPEG2000
ucid00092	0,933720724	150251	1,472693361	JPEG
20_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,46663481	150302	13,95634027	JPEG2000
141_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,45499685	151738	14,30249976	JPEG2000
ucid00096	0,960122957	152285	1,248730255	JPEG
ucid00097	0,9483294	152774	1,290733706	JPEG
46_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,312740876	153108	12,25475462	JPEG2000
ucid00091	0,969785552	153374	1,299260522	JPEG
88_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,355219626	156024	16,3451339	JPEG2000
184_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,603313393	157244	10,44617948	JPEG2000
182_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,549421663	158190	12,24090595	JPEG2000
12_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,401006444	158394	14,21522926	JPEG2000
154_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,339485744	158582	16,64321917	JPEG2000
62_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,353711863	159900	15,68606848	JPEG2000
ucid00095	0,978541657	160155	1,207878537	JPEG
180_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,509302925	161059	11,21958018	JPEG2000
94_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,420857821	162828	13,59182676	JPEG2000
151_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,474021616	166132	10,88423392	JPEG2000
44_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,313208941	166174	13,92634061	JPEG2000
202_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,391421097	167980	10,23746114	JPEG2000
174_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,414523311	168417	13,79144404	JPEG2000
60_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,429970897	172142	11,86542713	JPEG2000
172_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,417766764	174303	16,72041488	JPEG2000
198_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,437089753	181838	13,55284015	JPEG2000
166_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,462842059	185228	14,91512223	JPEG2000
80_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,343047775	187121	19,59040543	JPEG2000
127_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,547697928	191529	13,54919685	JPEG2000
138_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,442889086	193885	10,32020519	JPEG2000
150_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,463924123	194487	12,09883823	JPEG2000
196_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,61197817	199927	10,90243453	JPEG2000
123_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,54107343	200325	14,64005534	JPEG2000
121_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,514198905	203948	11,61749154	JPEG2000
132_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,439958036	206208	11,37365768	JPEG2000
82_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,423266428	207048	15,06727455	JPEG2000
186_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,499959203	214235	12,11926781	JPEG2000
130_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,419109995	222559	14,19416717	JPEG2000
170_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,506782902	227033	11,3205019	JPEG2000
177_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,575921363	227237	10,51046341	JPEG2000
65_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,746941645	231229	9,658043391	JPEG2000
50_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,495892555	237468	11,83310175	JPEG2000
194_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,449760957	245535	12,4280176	JPEG2000
148_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,658387655	257011	9,282841685	JPEG2000
206_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,558828229	258068	9,481956987	JPEG2000
91_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,700624729	270283	8,071941007	JPEG2000
204_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,585605358	273705	8,397888376	JPEG2000
77_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,612036597	292628	8,237687275	JPEG2000
171_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,640794997	318502	10,57743218	JPEG2000
192_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,598905977	346739	9,01566284	JPEG
81_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,627324387	411869	9,080497729	JPEG2000
43_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,84407183	423467	6,238585375	JPEG2000
125_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,916305508	436564	6,11882894	JPEG2000
137_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,829186153	450130	5,529310068	JPEG
210_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,600351893	451825	7,892160425	JPEG
103_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,826708308	602003	5,677223116	JPEG

## **B Anhang 2: Messwerte für Graustufenbilder aus Versuch 4**

B Anhang 2: Messwerte für Graustufenbilder aus Versuch 4

Bild	Entropie	Farben	Kompressionsfaktor	optimales Kompressionsverfahren
ucid00007	0,968771611	251	1,036705036	JPEG
ucid00024	0,956717306	256	1,042382856	JPEG
ucid00085	0,964950986	256	1,046566592	JPEG
ucid00037	0,964077271	256	1,05425492	JPEG
ucid00095	0,975217159	255	1,054283186	JPEG
ucid00094	0,985417508	255	1,05911062	JPEG
ucid00084	0,938901808	256	1,060310098	JPEG
ucid00078	0,966516812	256	1,060401599	JPEG
ucid00088	0,962030664	248	1,069370261	JPEG
i08	0,957321655	254	1,070272565	JPEG
ucid00086	0,927640195	252	1,071468277	JPEG
ucid00003	0,978637925	254	1,073023774	JPEG2000
ucid00096	0,956479618	255	1,075377269	JPEG
ucid00025	0,972453851	254	1,077634781	JPEG
ucid00023	0,947727115	256	1,080418082	JPEG
ucid00091	0,970779728	246	1,095150563	JPEG
i05	0,943957378	256	1,098657189	JPEG
ucid00097	0,944663888	256	1,098982672	JPEG
ucid00026	0,942676459	256	1,103349196	JPEG
ucid00102	0,953497398	251	1,104924187	JPEG2000
ucid00005	0,946344553	256	1,106080382	JPEG
ucid00022	0,958117196	256	1,107126768	JPEG
ucid00036	0,908028578	252	1,107631984	JPEG
ucid00019	0,923764251	253	1,111080971	JPEG
ucid00089	0,963997196	255	1,112797786	JPEG
ucid00017	0,954975591	256	1,113358627	JPEG
ucid00120	0,96804138	253	1,114223051	JPEG
i13	0,933467287	254	1,114677885	JPEG
ucid00020	0,939357649	256	1,123256053	JPEG
i14	0,946816178	252	1,127856815	JPEG
ucid00068	0,92219167	256	1,13069783	JPEG
ucid00107	0,969920333	254	1,133069768	JPEG
ucid00069	0,968949716	256	1,133317962	JPEG2000
ucid00100	0,960034362	253	1,137040807	JPEG
ucid00093	0,964542745	255	1,139287598	JPEG
ucid00113	0,945493154	255	1,139518709	JPEG
ucid00110	0,951475418	255	1,140047316	JPEG
ucid00099	0,962635712	254	1,140437479	JPEG
ucid00117	0,972269998	256	1,146241925	JPEG
ucid00014	0,939297627	256	1,147861422	JPEG
ucid00011	0,940865109	248	1,151734561	JPEG2000
ucid00108	0,955407395	256	1,155022912	JPEG2000
i18	0,880549503	255	1,156531253	JPEG
ucid00033	0,925873708	256	1,163312762	JPEG2000
ucid00118	0,940200288	254	1,164559961	JPEG
ucid00092	0,930113487	252	1,166729966	JPEG
i19	0,946542833	255	1,166875382	JPEG2000
ucid00006	0,937065836	256	1,166944641	JPEG
ucid00079	0,928445842	254	1,167554471	JPEG
ucid00013	0,944121631	256	1,170857204	JPEG
ucid00031	0,957754755	256	1,178344751	JPEG2000
ucid00122	0,960229512	254	1,183601228	JPEG2000
ucid00021	0,929308102	256	1,18377939	JPEG2000
ucid00115	0,945252305	256	1,184978031	JPEG
ucid00083	0,939170454	250	1,186787715	JPEG
ucid00111	0,935496765	254	1,187590606	JPEG2000
ucid00008	0,933245414	256	1,188186307	JPEG
ucid00052	0,957942291	256	1,18910615	JPEG2000
ucid00010	0,934105293	256	1,192459788	JPEG
ucid00055	0,963655607	242	1,194118326	JPEG2000
i01	0,912422549	221	1,194473809	JPEG
ucid00101	0,952490365	255	1,195054644	JPEG
ucid00087	0,913261937	252	1,195868764	JPEG
ucid00076	0,960698675	256	1,20533366	JPEG2000
ucid00114	0,935058212	256	1,208044289	JPEG2000
ucid00082	0,955855282	252	1,210290125	JPEG
ucid00075	0,90290402	251	1,210864076	JPEG
i22	0,918827407	251	1,211356467	JPEG
ucid00043	0,941969926	256	1,212686429	JPEG2000

## B Anhang 2: Messwerte für Graustufenbilder aus Versuch 4

Bild	Entropie	Farben	Kompressionsfaktor	optimales Kompressionsverfahren
ucid00004	0,883850965	256	1,213307579	JPEG
ucid00045	0,932019032	238	1,215760964	JPEG2000
ucid00042	0,963854361	245	1,21938785	JPEG2000
ucid00109	0,887631175	256	1,21982665	JPEG2000
ucid00030	0,938097268	256	1,22005374	JPEG
ucid00070	0,962669406	256	1,224239707	JPEG2000
ucid00080	0,902114538	256	1,225666889	JPEG
i11	0,905823881	255	1,230076204	JPEG2000
ucid00062	0,915104291	242	1,231154583	JPEG2000
ucid00125	0,951393295	255	1,232180796	JPEG2000
ucid00029	0,943903344	253	1,232621125	JPEG2000
ucid00124	0,922092977	253	1,236481642	JPEG2000
ucid00048	0,945009637	240	1,238381981	JPEG2000
ucid00038	0,904619668	241	1,243299986	JPEG2000
ucid00053	0,983474793	250	1,245544793	JPEG2000
ucid00059	0,942806559	252	1,245892082	JPEG2000
ucid00027	0,928481296	256	1,246800685	JPEG
i24	0,896392517	233	1,24913752	JPEG2000
ucid00028	0,94780435	256	1,25308638	JPEG2000
i23	0,94873121	240	1,254261509	JPEG2000
ucid00046	0,938718308	240	1,256265096	JPEG2000
i06	0,948840362	249	1,258734274	JPEG
ucid00090	0,945283205	255	1,260453129	JPEG2000
ucid00119	0,911327724	256	1,26229824	JPEG
ucid00032	0,92330231	256	1,263352696	JPEG2000
i04	0,887928548	250	1,269241198	JPEG2000
i17	0,90888599	255	1,269249392	JPEG2000
ucid00041	0,956370626	251	1,271071057	JPEG2000
ucid00116	0,928126772	253	1,271893336	JPEG2000
i10	0,931650463	245	1,272173153	JPEG2000
i09	0,933491416	248	1,274267456	JPEG2000
ucid00034	0,929387545	256	1,274275715	JPEG2000
ucid00074	0,963112576	251	1,274317011	JPEG2000
ucid00071	0,92497181	256	1,276600718	JPEG2000
ucid00121	0,950446567	253	1,284280936	JPEG2000
i21	0,90006609	252	1,285717088	JPEG2000
ucid00098	0,948187387	254	1,289587952	JPEG
ucid00049	0,970548617	240	1,290874949	JPEG2000
ucid00051	0,976074818	248	1,291629713	JPEG2000
ucid00123	0,91045494	256	1,293422628	JPEG2000
ucid00060	0,959713598	227	1,298016743	JPEG2000
ucid00067	0,903555288	256	1,303178938	JPEG2000
ucid00016	0,896765313	253	1,307347045	JPEG2000
ucid00061	0,949698092	209	1,307399206	JPEG2000
i12	0,913925413	238	1,314110405	JPEG2000
ucid00012	0,934161468	256	1,31638925	JPEG
ucid00066	0,924607991	225	1,320713398	JPEG2000
ucid00050	0,960471762	244	1,332149851	JPEG2000
ucid00106	0,910139656	254	1,335189574	JPEG2000
ucid00112	0,871567162	255	1,335951674	JPEG2000
i07	0,909745244	245	1,344070879	JPEG2000
ucid00057	0,914946081	250	1,346823858	JPEG2000
ucid00056	0,912512713	243	1,355104179	JPEG2000
i15	0,875079162	255	1,356216544	JPEG2000
ucid00002	0,918386498	256	1,358230917	JPEG2000
ucid00040	0,886836577	250	1,363354576	JPEG2000
ucid00001	0,903310608	253	1,376228475	JPEG2000
ucid00104	0,890288622	251	1,378931126	JPEG2000
ucid00081	0,913445489	216	1,380228297	JPEG2000
ucid00047	0,924460494	238	1,395630138	JPEG2000
ucid00039	0,967624245	241	1,401110296	JPEG2000
i02	0,746443298	249	1,407923004	JPEG2000
ucid00064	0,901462	220	1,414995754	JPEG2000
i16	0,914429862	238	1,419644598	JPEG2000
ucid00105	0,842568118	256	1,419654849	JPEG2000
ucid00073	0,861361001	256	1,426090741	JPEG2000
ucid00072	0,952322065	256	1,436846375	JPEG2000
ucid00054	0,942219611	249	1,439181325	JPEG2000
ucid00044	0,918097127	248	1,441619006	JPEG2000

## B Anhang 2: Messwerte für Graustufenbilder aus Versuch 4

Bild	Entropie	Farben	Kompressionsfaktor	optimales Kompressionsverfahren
ucid00063	0,921126323	235	1,4441286	JPEG2000
ucid00103	0,921397425	253	1,46396819	JPEG2000
ucid00035	0,930840462	256	1,46949392	JPEG2000
ucid00009	0,932192758	255	1,473348171	JPEG2000
ucid00018	0,941158549	256	1,475692594	JPEG2000
i03	0,894788967	241	1,525737034	JPEG2000
ucid00058	0,950860481	224	1,527657558	JPEG2000
ucid00077	0,875320019	256	1,532193457	JPEG2000
ucid00065	0,862217038	229	1,544919496	JPEG2000
i20	0,794513041	255	1,731237617	JPEG2000
ucid00015	0,720317613	244	1,882172739	JPEG2000
halfdesk	0,779576477	256	2,961063218	JPEG2000
1194985138747548807no_cyclsvgthumb	0,27835168	221	3,924646782	deflate
103_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,827389773	256	4,015211499	JPEG2000
137_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,828614411	256	4,041392061	JPEG2000
heart-th	0,624079812	196	4,318418314	deflate
43_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,846442246	256	4,399993159	JPEG2000
green-bike-th	0,216777581	54	4,719101124	deflate
125_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,912883219	256	4,821492912	JPEG2000
157_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,826044698	256	5,057799104	JPEG2000
11949837181420514439mascot_enrique_meza_c_01svgmed	0,519301009	248	5,398162328	deflate
210_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,605057836	256	5,462538884	JPEG2000
77_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,615231642	256	5,479438342	JPEG2000
204_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,592943986	256	5,539836396	JPEG2000
91_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,710174157	256	5,613490348	JPEG2000
i25	0,804885316	251	5,661694408	deflate
65_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,74569562	256	6,134966472	JPEG2000
206_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,564086534	256	6,346633814	JPEG2000
148_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,658482974	256	6,405582391	JPEG2000
81_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,633918937	256	6,494854909	JPEG2000
192_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,602715143	255	6,74116557	JPEG2000
207_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,859607051	243	7,175732557	JPEG2000
187_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,553763584	256	7,193628074	JPEG2000
138_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,443506016	256	7,206184168	JPEG2000
158_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,569102924	256	7,2093332	JPEG2000
171_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,642942692	256	7,24163799	JPEG2000
110_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,356468681	256	7,263162817	JPEG2000
184_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,607133157	256	7,266939136	JPEG2000
202_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,394511581	256	7,369290208	JPEG2000
177_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,583525481	256	7,410638204	JPEG2000
95_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,558194031	256	7,508495492	JPEG2000
121_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,517393853	256	7,657531724	JPEG2000
1195431012993269747johnmy_automatique_champagne_on_icesvgmed	0,504763496	17	7,664741928	deflate
green-falcon-md	0,331470042	141	7,690140845	deflate
151_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,475727924	256	7,705474753	JPEG2000
1194984554962989016xbill_01svgmed	0,357075515	254	7,706275033	deflate
119542956520783724jeanmichel_VTTsvgmed	0,481758082	14	7,717802599	deflate
50_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,494954093	256	7,832107795	JPEG2000
57_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,502594402	256	7,891623274	JPEG2000
196_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,627083014	256	7,993381856	JPEG2000
170_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,50591089	256	8,157272464	JPEG2000
152_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,481049638	256	8,199561952	JPEG2000
150_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,471500842	256	8,200174515	JPEG2000
182_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,552672948	256	8,217150272	JPEG2000
186_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,503982822	256	8,293775703	JPEG2000
132_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,437906697	256	8,461097778	JPEG2000
15_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,610995437	256	8,505629366	JPEG2000
60_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,43229412	256	8,538491058	JPEG2000
46_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,314776298	256	8,760635864	deflate
107_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,436249351	256	8,760943528	JPEG2000
194_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,453617505	256	9,020253065	JPEG2000
test	0,506206352	255	9,135002974	JPEG2000
14_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,596277042	256	9,378701279	JPEG2000
20_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,472583095	256	9,470834024	JPEG2000
180_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,514098963	256	9,499499774	JPEG2000
200_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,402095105	256	9,500036928	JPEG2000
94_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,422866487	256	9,533205565	JPEG2000
165_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,514151023	256	9,588177882	JPEG2000
174_nouvel-obs_hbhn300_constructedPdf_Nouvelobs2402PDF	0,415376108	256	9,633495796	JPEG2000

## B Anhang 2: Messwerte für Graustufenbilder aus Versuch 4

Bild	Entropie	Farben	Kompressionsfaktor	optimales Kompressionsverfahren
44_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,314030519	256	9,647189521	deflate
127_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,52653198	256	9,694049931	JPEG2000
12_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,404219897	256	9,726578109	JPEG2000
160_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,430049434	256	9,749478025	deflate
133_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,375630097	256	9,79715333	JPEG2000
198_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,438880013	256	9,810170825	JPEG2000
28_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,396483679	256	9,851293944	JPEG2000
130_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,419149203	256	9,970161039	JPEG2000
178_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,426877449	256	10,06246234	JPEG2000
8_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,357500599	256	10,10358679	JPEG2000
82_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,430786433	256	10,46256223	JPEG2000
126_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,295540416	256	10,49203127	deflate
141_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,475025247	256	10,50761936	JPEG2000
114_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,485748805	256	10,51373867	JPEG2000
162_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,416586004	256	10,57122495	deflate
119543096047570093johnny_automatic_bending_beesvgmed	0,155306965	251	10,60857794	deflate
123_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,578253184	256	10,6316495	JPEG2000
166_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,466684467	256	10,76576184	JPEG2000
88_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,35979795	256	10,80300193	deflate
62_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,356061342	256	10,90990437	JPEG2000
154_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,340470922	256	11,12871009	JPEG2000
172_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,421648098	256	11,24933316	JPEG2000
176_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,376116648	256	11,4251174	JPEG2000
70_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,380955001	256	11,42558628	deflate
164_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,422098284	256	11,44377195	JPEG2000
149_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,316338559	256	11,44983609	deflate
100_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,389357924	256	11,48252728	deflate
72_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,352071422	256	11,54298358	JPEG2000
145_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,350073301	256	11,54596238	deflate
grenze	0,473925967	255	11,55633826	deflate
134_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,53917882	251	11,65778115	JPEG2000
129_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,399555246	256	11,71313161	JPEG2000
66_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,372354411	256	11,76144377	deflate
96_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,337086808	256	11,81921056	deflate
32_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,340698964	256	12,17034802	deflate
146_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,37025323	256	12,27426867	deflate
36_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,283254253	245	12,28503892	deflate
84_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,284967475	254	12,30387723	deflate
101_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,382759633	256	12,30843941	JPEG2000
118_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,305263014	256	12,39444321	deflate
104_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,357665823	256	12,4854055	deflate
86_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,378275874	256	12,59410679	deflate
76_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,321676768	256	12,66165062	JPEG2000
74_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,299073783	256	12,74703209	deflate
92_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,406415261	256	12,81647014	deflate
190_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,348841402	256	12,9679894	deflate
102_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,321670857	256	13,02055806	deflate
112_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,305153549	256	13,0815736	deflate
142_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,355986652	256	13,27068822	deflate
131_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,293382566	256	13,29789104	JPEG2000
16_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,275555209	256	13,41960724	deflate
80_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,34238388	256	13,42122622	JPEG2000
140_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,25880472	256	13,61263844	deflate
34_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,298917546	256	13,98093498	deflate
30_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,243051453	256	14,01843161	deflate
68_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,317690295	256	14,2052705	deflate
98_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,283957515	256	14,31058837	deflate
1194984884404348229protection_obligatoire_10svgmed	0,402679829	17	14,37240498	deflate
7_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,290770522	256	14,38050819	deflate
124_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,230522231	256	14,46538524	deflate
11_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,291931843	255	14,6605999	deflate
116_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,374066965	256	14,69155129	deflate
48_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,327218117	256	14,83908251	deflate
78_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,303745503	256	15,05911621	deflate
168_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,339359569	256	15,79993743	deflate
139_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,370477468	256	16,02638244	deflate
188_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,311046713	256	16,19005589	deflate
64_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,286176673	256	16,58127033	JPEG2000
106_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,269989892	116	17,27790592	deflate

B Anhang 2: Messwerte für Graustufenbilder aus Versuch 4

Bild	Entropie	Farben	Kompressionsfaktor	optimales Kompressionsverfahren
90_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,233563495	256	17,56396213	deflate
52_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,159515816	251	17,75851863	deflate
40_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,268820286	174	17,9178019	deflate
22_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,173339863	256	18,06723128	deflate
18_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,194124291	227	18,21941595	deflate
128_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,304084189	256	18,39138555	deflate
120_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,318170665	256	19,88236289	deflate
122_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,174298776	174	20,06207475	deflate
54_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,152389652	256	20,30012921	deflate
42_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,220530071	256	20,75135236	deflate
3_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,134286756	254	20,7818404	deflate
wholescreen	0,399238341	256	22,46343841	deflate
136_nouvel-obs_hbhnr300_constructedPdf_Nouvelobs2402PDF	0,173316903	99	30,80497379	deflate
screenshot	0,376331375	230	32,80154905	deflate
3	0,755201876	5	131,6571429	deflate

# Abbildungsverzeichnis

3.1	Typische Kompressionsartefakte . . . . .	19
-----	--	----

# Tabellenverzeichnis

5.1	Qualitätsunterschiede zwischen JPEG und JPEG 2000 . . . . .	25
-----	---	----

# Literaturverzeichnis

- Adobe Systems Incorporated „PDF Reference, sixth edition“ [http://www.images.adobe.com/content/dam/Adobe/en/devnet/pdf/pdfs/pdf\\_reference\\_1-7.pdf](http://www.images.adobe.com/content/dam/Adobe/en/devnet/pdf/pdfs/pdf_reference_1-7.pdf), 2006, letzter Zugriff: 10. 09. 2015
- Adobe System Incorporated „Adobe® PDF Creation Settings/Adobe® Acrobat® DC SDK“ [http://www.images.adobe.com/content/dam/Adobe/en/devnet/acrobat/pdfs/AcrobatDC\\_PDFCreationSettings.pdf](http://www.images.adobe.com/content/dam/Adobe/en/devnet/acrobat/pdfs/AcrobatDC_PDFCreationSettings.pdf), 2015, letzter Zugriff 29.10.2015
- Bundesamt für Sicherheit in der Informationstechnik „BSI TR-03138 Ersetzen des Scannen (RESISCAN)“[https://www.bsi.bund.de/DE/Publikationen/TechnischeRichtlinien/tr03138/index\\_htm.html](https://www.bsi.bund.de/DE/Publikationen/TechnischeRichtlinien/tr03138/index_htm.html), 2013, letzter Zugriff; 15.09.2015
- Christopoulos et al. „The JPEG2000 still image coding system: an overview.“ *Consumer Electronics, IEEE Transactions on 46.4 (2000): 1103-1127*.
- Deutsch, Peter: „DEFLATE Compressed Data Format Specification version 1.3“ <http://www.w3.org/Graphics/PNG/RFC-1951>, letzter Zugriff 10.11.2015
- Ebrahimi et al. „JPEG vs. JPEG 2000: an objective comparison of image encoding quality.“ *Optical Science and Technology, the SPIE 49th Annual Meeting*. International Society for Optics and Photonics, 2004.
- Edmundson, David, and Gerald Schaefer „Sorting JPEG images at a glance.“ *Proceedings of the 4th ACM Multimedia Systems Conference*. ACM, 2013.
- Artifex „Details of Ghostscript output devices“ <http://ghostscript.com/doc/current/Devices.htm>, letzter Zugriff 30.10.2015
- Hornak, Joseph P. „The Basics of MRI“ <https://www.cis.rit.edu/htbooks/mri/chap-11/chap-11.htm>, letzter Zugriff 11.11.2015
- International Standardization Organization „Information technology — JPEG 2000 image coding system“ *ISO/IEC 15444-1* first edition, 2000
- Kerr, Douglas A. *Chrominance subsampling in digital images*. The Pumpkin,(1), November (2005).

- Khan, Azam. „Algorithm Study and Matlab Model for CCITT Group4 TIFF Image Compression.“ 2011 <http://www.diva-portal.org/smash/get/diva2:398511/FULLTEXT02>, letzter Zugriff 5.10.2015
- Koordinisationsstelle für die dauerhafte Archivierung elektronischer Unterlagen „Bericht zum Thema Dateiformat und Datenkomprimierung im Projekt "Konservierung Planarchiv"des StAZH“ <http://kost-ceco.ch/cms/download.php?9eebfdd2b8ed2c8716a29d2bc4a47936>, letzter Zugriff 16.11.2015
- Koordinisationsstelle für die dauerhafte Archivierung elektronischer Unterlagen „Analyse JBIG2 Komprimierung Fehlerhafte Xerox Scans“ <http://kost-ceco.ch/cms/download.php?16af9af3981293488688ad6640de5901>, 2014, letzter Zugriff 15.09.2015
- libpng „Portable Network Graphics (PNG) Specification (Second Edition) Information technology-Computer graphics and image processing-Portable Network Graphics (PNG): Functional specification. ISO/IEC 15948:2003 (E)“ <http://www.w3.org/TR/2003/REC-PNG-20031110>, 2003, letzter Zugriff 22.09.2015
- Marziliano, Pina, et al. „Perceptual blur and ringing metrics: application to JPEG2000“ *Signal processing: Image communication* 19.2 (2004): 163-172.
- Microsoft „Print Schema Specification “ <https://msdn.microsoft.com/en-us/windows/hardware/gg463385.aspx>, letzter Zugriff 29.09.2015
- Microsoft „XPS F.A.Q “ <http://windows.microsoft.com/de-de/windows/xps-documents-faq#1TC=windows-7>, letzter Zugriff 12.11.2015
- Murray, James D. & VanRyper, William *Encyclopedia of graphics file formats* O'Reilly & Associates, 1996
- Nelson, Mark, and Jean-Loup Gailly. *The data compression book*. Vol. 2. New York: M& T Books, 1996.
- Ponomarenko, Nikolay, et al. „TID2008-a database for evaluation of full-reference visual quality assessment metrics.“ *Advances of Modern Radioelectronics* 10.4 (2009): 30-45.
- Ha, Sangtae, et al. „Tube: time-dependent pricing for mobile data.“ *ACM SIGCOMM Computer Communication Review* 42.4 (2012): 247-258.
- Sayood, Khalid *Lossless Compression Handbook* Academic Press, 18.12.2002
- SHANNON, CE. *A Mathematical Theory of Communication*.
- Skodras et al. „The JPEG 2000 still image compression standard.“ *Signal Processing Magazine, IEEE* 18.5 (2001): 36-58.

THE INTERNATIONAL TELEGRAPH AND TELEPHONE CONSULTATIVE COMMITTEE: „SERIES T: TERMINAL EQUIPMENT AND PROTOCOLS FOR TELEMATIC SERVICES“ *STANDARDIZATION OF GROUP 3 FACSIMILE APPARATUS FOR DOCUMENT TRANSMISSION*  
[https://www.itu.int/rec/dologin\\_pub.asp?lang=e&id=T-REC-T.4-198811-S!!PDF-E&type=items](https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-T.4-198811-S!!PDF-E&type=items), letzter Zugriff 9.11.2015

Verdu, Sergio. „Fifty years of Shannon theory.“ *IEEE Transactions on information theory* 44.6 (1998): 2057-2078.

Wallace, Gregory K. „The JPEG still picture compression standard.“ *Consumer Electronics, IEEE Transactions on* 38.1 (1992): xviii-xxiv.

Wang, Zhou, et al. „Image quality assessment: from error visibility to structural similarity.“ *Image Processing, IEEE Transactions on* 13.4 (2004): 600-612.

Wang, Zhou, Hamid R. Sheikh, and Alan C. Bovik. „No-reference perceptual quality assessment of JPEG compressed images.“ *Image Processing. 2002. Proceedings. 2002 International Conference on. Vol. 1. IEEE, 2002.*

Zhang, Lin, et al. „FSIM: a feature similarity index for image quality assessment.“ *Image Processing, IEEE Transactions on* 20.8 (2011): 2378-2386.

Ziv J., Lempel A.: „A Universal Algorithm for Sequential Data Compression“ *IEEE Transactions on Information Theory, Vol. 23, No. 3, pp. 337-343.*

Ich versichere, die vorliegende Arbeit selbstständig ohne fremde Hilfe verfasst und keine anderen Quellen und Hilfsmittel als die angegebenen benutzt zu haben. Die aus anderen Werken wörtlich entnommenen Stellen oder dem Sinn nach entlehnten Passagen sind durch Quellenangaben eindeutig kenntlich gemacht.

Ort, Datum

Robin Werner