



Fotis Jannidis / Hubertus Kohle /
Malte Rehbein (Hg.)

Digital Humanities

Eine Einführung



J.B. METZLER



J.B. METZLER

Fotis Jannidis / Hubertus Kohle / Malte Rehbein (Hg.)

Digital Humanities

Eine Einführung

Mit Abbildungen und Grafiken

J.B. Metzler Verlag

Die Herausgeber

Fotis Jannidis ist Professor für Computerphilologie und Neuere Deutsche Literaturgeschichte an der Julius-Maximilians-Universität Würzburg.

Hubertus Kohle ist Professor für Mittlere und Neuere Kunstgeschichte an der Ludwig-Maximilians-Universität München.

Malte Rehbein ist Professor für Digital Humanities an der Universität Passau.



Gedruckt auf chlorfrei gebleichtem, säurefreiem und alterungsbeständigem Papier

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

ISBN 978-3-476-02622-4

ISBN 978-3-476-05446-3 (eBook)

Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

J.B. Metzler, Stuttgart

© Springer-Verlag GmbH Deutschland, 2017

Einbandgestaltung: Finken & Bumiller, Stuttgart (Foto: shutterstock)

Satz: primustype Hurler GmbH, Notzingen

Druck und Bindung: Ten Brink, Meppel, Niederlande

J.B. Metzler ist Teil von Springer Nature

Die eingetragene Gesellschaft ist Springer-Verlag GmbH Deutschland

www.metzlerverlag.de

info@metzlerverlag.de

III Digitale Objekte

12 Digitalisierung

Erste Grundlage der computer-gestützten Untersuchung kultureller Artefakte ist ein digitales **Abbild** des (analogen) Originals. Handelt es sich hierbei um ein optisches, d. h. auf Licht basierendes, für die sehende Betrachtung erstelltes Abbild, so sprechen wir bei diesem Transformationsprozess vom Analogen ins Digitale von einer **Bilddigitalisierung**. In der Regel erfolgt diese Transformation durch bildgebende Verfahren, die durch dafür entwickelte Geräte, in unseren Anwendungsfällen zumeist Digitalkameras und Scanner, vorgenommen werden.

Ein digitales Abbild eines Objekts nennen wir auch ein **Digitalisat** desselben. Digitalisate ersetzen heute häufig Funktionen in Forschung und Lehre, die früher häufig den Originalen vorbehalten waren. Dennoch ist die Nutzung von Abbildern an Stelle der Originale keine Neuerung des digitalen Zeitalters. Gerade im Bereich der Lehre, aber auch der Forschung verzichtet man bereits seit langem aus pragmatischen und Kostengründen oft auf das Studium des Originals und greift stattdessen auf Abbilder, etwa in Form von Dia-Positiven, Mikrofilmen oder Reproduktionen zurück. Diese Abbilder erfüllen dann als **Surrogat** (von lat. *surrogare*: an die Stelle eines anderen wählen) den Zweck des Originals.

Mit der Einführung der **Bilddigitalisierung** ändern sich aber die Rahmenbedingungen maßgeblich. Frühere Formen von Abbildungen waren selbst gegenständlich (materiell), und der Zugang zu ihnen war an ihre physische Existenz und Zugänglichkeit gebunden. Um sie zu verbreiten, mussten wiederum Kopien dieser Abbilder (oder weitere Abbildungen des Originals) angefertigt werden. Hierdurch entstehen direkte Kosten (Material, Arbeitszeit); außerdem besteht bei analoger Reproduktion die Gefahr des Informationsverlustes durch Verschlechterung der Qualität bei wiedergeholter Kopie, wie man sich bei der Xerox-Kopie leicht vorstellen kann.

Digitalisate sind an sich immateriell; als Rastergrafiken (s. Kap. 12.1) sind sie lineare Abfolgen von Bits. Zwar erfolgt ihre Speicherung in der Regel auf einem physikalischen Medium, doch ist diese Speicherung grundsätzlich medienunabhängig möglich. Es ist denkbar (freilich nicht sehr sinnvoll), die Bitfolge (Bitstream) eines Digitalisats auf Papier zu notieren und auf diese Art weiterzugeben und zu bewahren. Grundsätzlich ist Übertragung oder Kopie eines Digitalisats verlustfrei und identisch möglich und macht einen wesentlichen Vorteil der Digitalisierung gegenüber analogen Verfahren der Reproduktion aus.

Neben der verlustfreien Kopie und neuen Formen des Zugangs sind die Möglichkeiten der computergestützten Vernetzung und Verarbeitung von Digitalisaten der größte Fortschritt der Digitalisierung im Vergleich zu traditionellen Methoden der Reproduktion und der Arbeit am und mit dem Original. Häufig ist die hier angesprochene Bilddigitalisierung aber nur der erste Schritt oder besser: eine grundlegende Voraussetzung für eine weitere Verarbeitung. In Kapitel 12.4 werden Verfahren der Transkription erläutert, mittels derer der in digitalen Bildern transportierte Text erfasst und extrahiert und seinerseits zu digitalen Daten wird. Für Forschung, die auf Textmaterial beruht, und dies gilt zumindest anteilig für alle Disziplinen der Kultwissenschaften, ist dieser Schritt elementar. Grundlage für diese Extraktion von Text ist aber die Schaffung eines digitalen Abbildes des Trägers dieser Texte, sei er ein Buch, eine historische Zeitung, eine Handschrift auf Papier, Pergament oder Papyrus oder eine Inschrift auf einem Grabstein usw.

Analog-Digital-Wandlung: Durch Digitalisierung erzeugen wir ein digitales Ab-

bild von einem in der Regel materiellen, analogen Objekt. Diese Transformation erfolgt Technologie-gestützt durch sogenannte **Analog-Digital-Wandler**, die mittels physikalischer Sensoren das Objekt messen und die Messdaten in einer geeigneten Form speichern. Gängige Beispiele für Analog-Digital-Wandler sind Digitalkameras, die ein analoges Signal (Licht) in eine digitale Rastergrafik umsetzen, digitale Diktiergeräte, die akustische Signale in digitale Audiodaten transformieren, oder TV-Tuner für den Empfang analoger Fernsehsignale am Computer.

Das analoge Signal (von gr. *analogos*: verhältnismäßig) als Input der Digitalisierung besitzt unendlich viele Werte in beliebig feiner Abstufung; es ist zeit- und wertekontinuierlich, d. h. zwischen zwei Messpunkten, seien sie auch noch so dicht beieinander, befinden sich unendlich viele weitere Werte. Das digitale Signal (von lat. *digitus*: Finger) als Ziel der Digitalisierung ist hingegen zeit- und wertediskret, es stehen nur endlich viele, definierte Werte zur Verfügung. Zur Diskretisierung des analogen Signals wird dessen Messbereich daher in eine bestimmte Anzahl aneinander-grenzender Intervalle (Wertebereiche) unterteilt und jedem dieser Intervalle ein diskreter Zahlwert (Code) zugewiesen (Quantisierung). Die Größe dieses Intervalle wird **Auflösung** der Digitalisierung genannt (im Schaubild ist dies die y-Achse); je kleiner die Intervalle, desto höher oder besser die Auflösung, denn es werden mehr Details des analogen Signals in das digitale Abbild übernommen. Die Häufigkeit, mit der das analoge Signal pro Zeiteinheit gemessen (abgetastet) wird, ist das zweite Maß für die Qualität der Digitalisierung, sie wird **Abtastrate oder Samplingrate** genannt (im Schaubild ist dies die x-Achse).

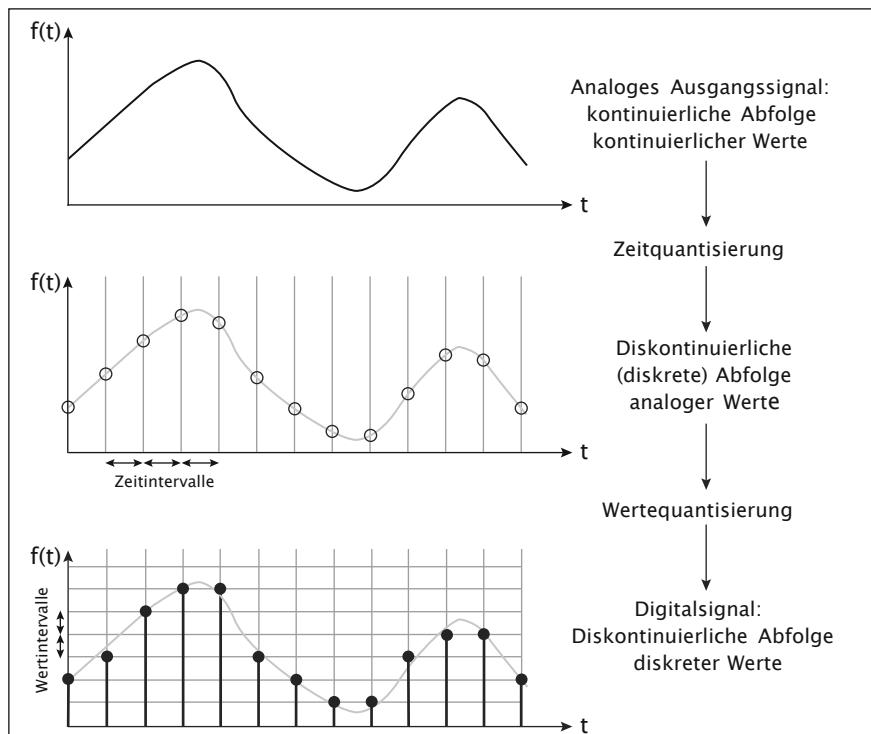


Abb. 44 Analog-/Digitalwandlung mit zeitlicher- und Wertequantisierung des analogen Signals

Grundsätzlich kann man jedes analoge Signal digitalisieren, das durch physikalische Sensoren gemessen werden kann. Für die Digital Humanities sind dabei vor allem akustische und optische Sensoren von Interesse, die Ton- bzw. Bilddigitalisierung ermöglichen. Die Bilddigitalisierung wird im Folgenden behandelt. Für bildbasierte Quellen z. B. in der Kunstgeschichte oder Geschichte hat dies naheliegend eine hohe Bedeutung. Aber auch als Grundlage zur Erzeugung von elektronischen Texten auf Basis materieller Textträger ist die Bilddigitalisierung ein wichtiger Schritt, denn Text an sich ist eine immaterielle Größe, die von seinem Schreiber oder Drucker auf einem Trägermaterial (Papier, Pergament, Papyrus, Stein), das als Bild betrachtet werden kann, manifestiert wurde. Der Text muss daher zunächst von seinem Träger (Dokument, Inschrift) abstrahiert werden, was der Mensch durch das Lesen implizit tut. Dem Computer aber muss dieser Schritt explizit beigebracht werden, bevor Texte als Zeichenketten digital vorliegen, die verarbeitet werden können. Hierauf wird in Kapitel 12.4 eingegangen.

12.1 | Grundlagen digitaler Bilder

Wir kennen zwei prinzipielle Möglichkeiten, um Bilder als digitale Daten zu repräsentieren. Die erste ist ihre Repräsentation als zweidimensionale Matrix der Bildpunkte (Pixel), aus denen sich das Bild zusammensetzt (Rastergrafik, Bitmap), die zweite ist eine Erklärung, wie das Bild auf Basis elementarer Zeichenroutinen erzeugt werden kann (Vektorgrafik). Beide Repräsentationsarten digitaler Bilder haben in den Digital Humanities ihre Relevanz; durch Digitalisierung werden jedoch vornehmlich Rastergrafiken erzeugt.

Rastergrafiken

In der Rastergrafik wird jeder Bildpunkt durch einen Wert beschrieben, der den Farbwert des Bildpunktes repräsentiert. Die Adressierung einzelner Bildpunkte erfolgt über zweidimensionale Koordinaten (x- und y-Wert); als Referenzpunkt wird in der Regel der Ursprung des Koordinatensystems links oben angesetzt. Abbildung 45 zeigt eine schwarz-weiße Rastergrafik, die nur die Werte 0 (für weiß) und 1 (für schwarz) kennt.

Rastergrafiken werden üblicherweise anhand von zwei Werten gemessen: ihrer Bildgröße und ihrer Farbtiefe.

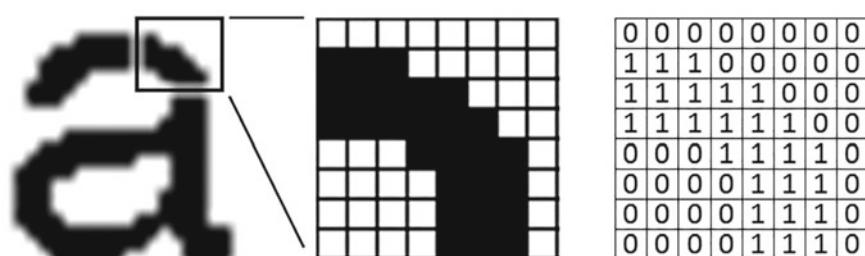


Abb. 45 Schematisches Beispiel einer monochromen Rastergrafik

Die **Bildgröße** wird entweder als Wertepaar beschrieben, das die Anzahl der Bildpunkte in der Horizontalen und in der Vertikalen (bzw. der Spalten- und Zeilenzahl der Tabellierung) repräsentiert, oder sie wird als Zahlenwert angegeben, der durch Multiplikation der Spalten- und Zeilenzahl entsteht. Bei letzterem wird zwar die Gesamtzahl der Pixel ermittelt, aber das Seitenverhältnis geht verloren. So hat der Ausschnitt rechts in Abbildung 45 etwa eine Größe von 8×8 bzw. 64 Pixeln; die gleiche Pixelzahl würde aber zum Beispiel auch mit einem Bild der Größe 4×16 erreicht.

Die **Farbtiefe** bestimmt, wie viele Helligkeits- und Farbwerte zur Beschreibung jedes einzelnen Pixels zur Verfügung stehen. Das Beispiel oben kennt nur zwei Farbwerte, schwarz und weiß, und hat damit eine Farbtiefe von 1, da mit einem Bit zwei Zustände repräsentiert werden können (mathematisch gilt: $\log_2 2 = 1$). In der Regel haben digitale Bilder, die uns in den Geistes- und Kulturwissenschaften begegnen, jedoch eine deutlich größere Farbtiefe. Schwarz/weiß-Bilder besitzen genau einen **Farbkanal**; man spricht hierbei in der Ausdifferenzierung der Farb- bzw. Helligkeitswerte von **Graustufen**. Bei einer Farbtiefe von 8 Bit sind es $2^8 = 256$, bei 16 Bit sind es $2^{16} = 65.536$ verschiedene Werte, die jeder Pixel annehmen kann. Digitale schwarz/weiß-Bilder begegnen uns in den Geistes- und Kulturwissenschaften nicht nur bei der Digitalisierung originär schwarz/weißer Objekte (etwa Fotografien), sondern auch aus früheren Digitalisierungskampagnen, als etwa Farbscanner, vor allem großformatige, deutlich teurer waren als gegenwärtig und der Speicherbedarf von Farbbildern die damals zur Verfügung stehenden Speicherkapazitäten gesprengt hätte (so noch in den 1990er Jahren).

Heute ist die Bilddigitalisierung in Farbe das gängige Verfahren. Zur Repräsentation der Farben solcher Bilder existieren verschiedene Modelle, von denen hier nur das **RGB-Farbmodell** vorgestellt werden soll, bei dem die Farben aus den drei Grundfarben (Primärfarben) Rot, Grün und Blau (RGB) additiv zu Mischfarben zusammengesetzt werden. Mischt man Rot und Grün in gleichem Verhältnis, so erhält man die Mischfarbe Gelb; Rot und Blau ergeben Magenta; Blau und Grün ergeben Cyan. Die Addition aller drei Grundfarben führt zu Weiß; schaltet man alle aus, so bleibt Schwarz. Das RGB-Farbmodell ist damit an das menschliche Auge angelehnt, das drei Rezeptoren für den kürzeren, mittleren und längeren Wellenbereich besitzt.

Bei der Repräsentation eines Bildes mittels des RGB-Farbmodells werden daher drei Farbkanäle verwendet, für die entsprechend den schwarz/weiß-Bildern gilt, dass die Anzahl der Bits die unterschiedlichen Abstufungen des Tonwerts pro Farbkanal angeben. Da die kleinste Einheit für die Farbtiefe 1 Bit pro Farbkanal ist, benötigt ein Farbbild mindestens 3 Bit und erlaubt damit $2^3 = 8$ verschiedene Farbwerte, diese sind: Schwarz, Rot, Grün, Blau und die Mischfarben Cyan, Magenta, Gelb und Weiß. Für gegenwärtige Digitalisierungsprojekte sind Farbtiefen von 24 Bit (8 Bit je Kanal) und 48 Bit (16 Bit je Kanal) üblich. Sie erlauben eine Differenzierung von 16.277.216 bzw. 281.474.976.710.656 Farbwerten. Neben dem RGB-Modell existieren andere Farbmodelle, etwa das subtraktive CMYK-Modell, das von Druckern verwendet wird.

Da das Produkt aus Bildgröße und Farbtiefe den Speicherbedarf einer Rastergrafik determiniert, können die Datenmengen, die bei Digitalisierungsprojekten entstehen, gewaltig sein. Trotz des im Moore'schen Gesetz beschriebenen ständigen Anstiegs von Rechenleistung und (physischer) Speicherkapazität, verursachen diese Datenmengen Kosten und Arbeit, die sich nicht nur aus den reinen Kosten für die Anschaffung von Speichermedien, sondern insbesondere aus deren Betrieb inkl. Sicherung

und ggf. notwendiger Migration der Daten in neue Formate sowie dem Zugänglichmachen (Übertragungszeit) dieser Daten ergeben.

Bildkompression: Daher wendet man Kompressionsalgorithmen an, um die Größe der Dateien im Speichermedium zu reduzieren. Dabei ist zwischen verlustfreier und verlustbehafteter Kompression zu unterscheiden. Die verlustbehaftete Kompression bemüht sich, als überflüssig betrachtete Information aus den Daten zu eliminieren und somit die Datenmenge zu verkleinern. Was überflüssig sein kann, lässt sich am besten am Beispiel der Kompression von Musikdateien erläutern. Hier wird zu Grunde gelegt, dass das menschliche Ohr in der Lage ist, Töne nur innerhalb eines bestimmten Frequenzbereichs wahrzunehmen und die Töne auch nur bis zu einem bestimmten Grad auszudifferenzieren. Damit können Nuancen in der Frequenzdifferenzierung zusammengefasst und Frequenzen außerhalb des Bereichs subjektiver menschlicher Wahrnehmung, die aber dennoch von den Instrumenten produziert und von Aufnahmegeräten erfasst wurden, eliminiert werden.

Für die Kompression von Bilddaten gilt analog: Es wird zusammengefasst bzw. weggelassen, was vom menschlichen Auge nicht mehr ausdifferenziert werden kann bzw. was über die Aufnahmefähigkeit der menschlichen Bildwahrnehmung hinausgeht (man sagt, die Kompression ist dann »visuell verlustfrei«). Im Gegensatz zur Kompression von Audiodaten, ist dies allerdings nicht explizit formuliert, sondern ist geleitet von der Frage, was für einen bestimmten Anwendungsfall tolerabel ist bzw. ob in einem bestimmten Kontext der Informationsverlust unmerklich und ästhetisch erträglich bleibt. So wird bei der Anwendung von Kompressionsverfahren für Bilder in der Regel nicht davon ausgegangen, dass der menschliche Betrachter sich diese Bilder mit einer Lupe anschaut oder beliebig auf dem Bildschirm vergrößert.

Bei der **verlustfreien Kompression** geht keine Information verloren, da die Daten lediglich durch Zusammenfassung von Redundanzen anders organisiert werden. Weist zum Beispiel ein Bild eine große weiße Fläche oder ein anderes, wiederkehrendes Muster auf, so wird der Kompressionsalgorithmus dies erkennen und lediglich die Information über Gestalt und Position dieser Fläche abspeichern anstatt jeden einzelnen Pixel. Dies ist zwar mit Rechenaufwand verbunden, erfordert in der Regel aber deutlich weniger Speicherplatz, denn eine Anweisung etwa der Form »wiederhole den folgenden Pixel 10.000 Mal« ist kürzer als 10.000 Mal den gleichen Pixel zu speichern. Da diese Operation umkehrbar ist, d. h. die Beschreibung von Gestalt und Position der fraglichen Fläche wieder in exakt die ursprüngliche Rasterung überführt werden kann, ist die Kompression verlustfrei.

Zwar können verlustfrei komprimierte Bilder wieder exakt (eben verlustfrei) in das Original zurückgeführt werden, dies geht aber nur dann, wenn der Kompressionsalgorithmus offengelegt, dokumentiert und auch in der Zukunft programmiert werden kann bzw. die zur Dekompression notwendige Software zur Verfügung steht. Dies ist gerade bei kommerziellen Kompressionsalgorithmen nicht unbedingt der Fall; selbst bei offengelegten Standards stellt dies einen Unsicherheitsfaktor für die Zukunft dar, weswegen die Praxisregeln »Digitalisierung« der Deutschen Forschungsgemeinschaft DFG die auf Dauer ausgerichtete Erstellung und Speicherung eines unkomprimierten **Bildmasters** empfehlen, von dem dann auch in der Zukunft andere Formate erzeugt werden können.

	Formate (Auswahl)	Übertragung / Speicher	Anwendungen
Verlustbehaftete Kompression	jpeg, jpeg2000, gif, png	Netz (statische Bilder)	Betrachtung mit menschlichem Auge
Verlustfreie Kompression	tiff	Netz (Kacheltechnik)	Bildanalyseverfahren
keine Kompression	tiff	Speicherung als Digitalisierungsmaster	Langzeitarchivierung

Tab. 8 Bildkompression (Überblick)

Das Bildformat TIFF

Unter der Menge der zur Verfügung stehenden Bildformate, hat sich das **Tagged Image File** Format (kurz: TIFF, auch TIF) als Quasi-Standard durchgesetzt und wird von der DFG für den digitalen Bildmaster empfohlen. Die Möglichkeiten, die TIFF bietet, sind relativ komplex (so lässt sich etwa TIFF als Container für andere Bildformate verwenden oder es können verlustbehaftete Kompressionsalgorithmen unterstützt werden), und nicht alle werden von allen Programmen unterstützt. Als kleinstes gemeinsames Vielfaches, das alle Programme, die TIFF erzeugen oder verarbeiten, unterstützen sollen, ist das *Baseline TIFF* spezifiziert, das verlustfreie Kompression, Kachelung und das RGB-Farbmodell bis zu einer Farbtiefe von 24 Bit unterstützt.

TIFF dokumentiert einige technische Parameter des Bildes, wie etwa die Bildgröße(n) und ob und wenn ja welcher Kompressionsalgorithmus verwendet wurde, im sogenannten Image File Directory, eine Art Metadaten des digitalen Bildes. Diese Angaben sind auch bei der Beurteilung der Qualität von Digitalisaten von Interesse. Über die Baseline hinaus bietet TIFF sogenannte *Extensions*, deren Unterstützung aber nicht notwendigerweise durch jede Software gewährleistet ist, ebenso wie die Verwendung von *private Tags* zur Beschreibung von Metadaten innerhalb der Bilddatei.

Eine unkomprimierte Bilddatei kann in andere Formate übertragen werden, die anderen Anwendungszwecken dienen als die der Langzeitarchivierung, wofür sich die unkomprimierte Rastergrafik anbietet. Cohen und Rosenzweig (2005, 97–98) führen etwa das Beispiel von Dorothea Langes Fotografie *Migrant Mother* an, deren Digitalisat die Library of Congress in Washington im unkomprimierten TIFF-Format archiviert, das auf Grund seiner Größe von 6.849×8.539 Pixeln und einer Farbtiefe von 8 Bit (256 Graustufen) 55 Megabyte beansprucht. Online wird jedoch eine Fassung der *Migrant Mother* dem Benutzer zur Betrachtung angeboten, die im JPEG-Format lediglich 163 Kilobyte beansprucht, etwa 0,12 % der ursprünglichen Größe, die sich, neben dem geringeren Speicherbedarf auf dem Webserver (im Gegensatz zum internen Archivserver), vor allem in der deutlich schnelleren Ladezeit und dem Traffic der Website bemerkbar macht.



Abb. 46 Qualitätsverlust im jpeg durch Kompression

Die Joint Photographic Experts Group (**JPEG**, sprich: »JAY-Peg«) beschreiben Methoden zur in der Regel verlustbehafteten Kompressionen für digitale Bilder, die unter anderem im Grafikformat JPEG File Interchange Format (JFIF) Verwendung finden, die gemein gebräuchlich nicht ganz korrekt ›JPEG-Bilder‹ genannt werden. JPEG erlaubt verschiedene Kompressionsgrade, bei denen, vereinfacht gesagt, eine höhere Komprimierung eine kleinere Dateigröße bewirkt, aber auch einen größeren Informations- und damit je nach Anwendungsfall höheren Qualitätsverlust nach sich zieht (s. Abbildung 46). JPEG wurde vor allem zur Speicherung von Fotografien oder fotoähnlicher Bildern entwickelt.

Das verbesserte Nachfolgeformat JPEG 2000, ebenfalls von der Joint Photographic Experts Group, findet seit einiger Zeit ebenfalls Verwendung im Bereich der kulturwissenschaftlichen Digitalisierung, konnte sich bislang noch nicht durchsetzen. Andere verbreitete Bildformate sind etwa das Graphics Interchange Format (GIF) und Portable Network Graphics (PNG), die auf Grund ihrer hohen, verlustfreien, Kompressionsrate neben JPEG die am häufigsten verwendeten Bildformate im Internet sind. Eine Alternative zur Archivierung von Bildmastern stellen sie wie auch JPEG und JPEG 2000 jedoch nicht dar.

Vektorgrafiken

Beispiel Vektorgrafik / SVG

Um etwa ein Rechteck zu zeichnen, können folgende Informationen beschrieben werden: eine Anweisung, dass es sich um das grafische Primitiv vom Typ Rechteck handelt; die Position der linken oberen Ecke des Rechtecks; seine Breite und Höhe; Linienstärke und -farbe, mit der es gezeichnet werden soll; die Füllung des Rechtecks.

Das folgende Code-Beispiel im XML-basierten SVG-Format (scalable vector graphics) zeichnet mit einer schwarzen Linie der Stärke 20 Pixel ein grün ausgefülltes Rechteck, das 500 Pixel breit und 200 Pixel hoch ist und um jeweils 100 Pixel vom Koordinatenursprung nach rechts unten versetzt wird:

```
<svg>
    <rect x="100" y="100" width="500" height="200" fill=
        "green" stroke="black" stroke-width="20px"/>
</svg>
```

Aus solchen Anweisungen lassen sich komplexe Grafiken bis zu fotorealistischen Szenen erzeugen.

Typische Anwendungsbereiche in den Digital Humanities für Vektorgrafiken sind etwa zwei- und dreidimensionale Modelle oder die Darstellung von Geoinformationen. Im Gegensatz zu Rastergrafiken beschreiben Vektorgrafiken, wie ein Bild auf Basis einfacher Zeichenroutinen, sogenannter **grafischer Primitive**, zusammengesetzt ist. Dazu zählen u. a. Punkte, Linien, Polygone, Kreise und Ellipsen, außerdem Informationen über Farbe, Strichform und Füllung sowie Operationen auf den grafischen Primitiven wie etwa Drehungen und Spiegelungen. Grundlage zur Beschreibung der Zeichenroutinen ist ein zweidimensionales Koordinatensystem und die so-

genannte z-Ordnung, die festlegt, welches Objekt der Grafik bei Überlappung über dem anderen erscheint. Neben zweidimensionalen existieren auch dreidimensionale Vektorgrafiken.

Raster- und Vektorgrafiken im Vergleich

Für Computerbildschirme ist die Vektorgrafik das ältere Verfahren. Ähnlich einem Oszilloskop setzten frühe Vektorbildschirme jedes Bild per Kathodenstrahlen aus einzelnen Linien (daher der Begriff »Vektor«) zusammen, die Befehle hierzu (Ursprung des Vektors, Länge und Richtung) erhielten sie vom Prozessor. Eine Rastergrafik ermöglichte hingegen den zeilenweisen Aufbau eines Bildes.

Ein wesentlicher Unterschied zwischen den beiden Repräsentationsarten von digitalen Bildern tritt bei der Skalierung, d. h. bei der Größenänderung des Bildes zu Tage. Bei Rastergrafiken wird durch die Skalierung die Bildauflösung geändert. Das heißt, dass aus einer vorgegebenen Rastergrafik ein neues Bild mit einer höheren (oder niedrigeren) Anzahl von Bildpunkten erzeugt wird. Möchte man beispielsweise eine Rastergrafik der Bildgröße 100×100 auf 200×200 vergrößern, erzeugt man dadurch viermal so viele Bildpunkte. Da bei Rastergrafiken der Pixel die kleinste Informationseinheit bildet, müssen die Pixel des Originalbildes für die Vergrößerung wiederholt oder interpoliert werden. Es entstehen im Grunde größere Pixel. Möchte man das ursprüngliche Bild etwa auf 50×50 verkleinern, erzeugt man nur noch ein Viertel der Bildpunkte, wodurch Bildpunkte und damit Information unweigerlich verloren geht. Daher ist die Skalierung von Rastergrafiken in der Regel mit einem Qualitätsverlust verbunden. Bei der Skalierung einer Vektorgrafik hingegen werden die grafischen Primitiven, aus denen sich die Vektorgrafik zusammensetzt, durch geometrische Transformation gestreckt. Die Skalierung von Vektorgrafiken ist damit verlustfrei und umkehrbar.

Ein zweiter Unterschied zeigt sich im Speicherbedarf. Bei unkomprimierten Rastergrafiken berechnet sich der Speicherbedarf aus dem Produkt aus Bildgröße und Farbtiefe. Er ist grundsätzlich unabhängig vom Bildinhalt und von daher gut kalkulierbar. Kompressionsalgorithmen für Rastergrafiken beruhen in ihrer Effektivität hingegen sehr wohl auf dem dargestellten Bild. Bessere Kompressionsraten und damit niedrigeren Speicherbedarf erreichen Bilder mit vielen wiederkehrenden Mustern. Der Speicherbedarf von Vektorgrafiken ist unabhängig von Bildgröße und Farbtiefe, sondern steigt mit der Komplexität des dargestellten Bildes.

Die Wiedergabe von Vektorgrafiken auf dem Bildschirm führt zu einem Rechenaufwand, denn das Bild muss aus den grafischen Primitiven, die in der Vektorgrafik beschrieben sind, für jede Wiedergabe errechnet werden (**Rendering**). Dieser Rechenaufwand ist abhängig von der Komplexität der Grafik und daher zunächst unbekannt. Rastergrafiken haben in der Regel einen geringeren Wiedergabeaufwand, der zudem konstant und inhaltsunabhängig ist.

Eine Umwandlung einer Rastergrafik in eine Vektorgrafik (**Vektorisierung**) ist in der Regel sehr aufwendig, obwohl es hierfür viele Anwendungsgebiete gibt. Eine Vektorgrafik lässt sich hingegen jederzeit durch Rendering in eine Rastergrafik beliebiger Größe umwandeln, was auch Voraussetzung für die Anzeige auf modernen Bildschirmen ist.

Rastergrafiken und Vektorgrafiken lassen sich auf unterschiedliche Art und Weise manipulieren. Bei Rastergrafiken kommen hierfür in der Regel Bildbearbeitungspro-

gramme zum Einsatz, die eine Fülle von Möglichkeiten bieten, die Grafiken zu verändern, die allerdings – naturgemäß – stets auf Rechenoperation auf den digitalen Rasterdaten beruhen. Möchte man ein Detail aus einem digitalen Foto entfernen, so ist hierfür in der Regel ein recht hoher kreativer Aufwand sowie Rechenaufwand erforderlich, insbesondere müsste in diesem Fall der Hintergrund des zu entfernenden Details aus nicht vorhandener Information hochgerechnet werden. Bei Vektorgrafiken würde man hingegen schlicht die Anweisungen, die dieses Detail beschreiben, streichen.

12.2 | Bilddigitalisierung

Digitalisierung ist ein bildgebendes Verfahren und bedeutet die Erzeugung eines digitalen Abbildes eines oder mehrerer realer Objekte. Im Allgemeinen messen bildgebende Verfahren physikalische Größen der Objekte, die sie abbilden (siehe oben). Bei den optischen Verfahren wie die der Bilddigitalisierung ist dies das Licht, das von den Objekten reflektiert wird. Bildgebende Verfahren an sich sind lange bekannt; schon Aristoteles beschrieb das Prinzip einer Lochkamera, bei der Licht durch eine Öffnung eingefangen wurde und auf einer Projektionsfläche ein (auf dem Kopf stehendes) Abbild erzeugte. Man denke zudem etwa an die im 13. Jahrhundert von Astronomen zur Beobachtung eingesetzte *Camera obscura* (lat. dunkle Kammer), an Fernrohre oder Mikroskope. Diese Instrumente erzeugen ein optisches Abbild der Realität, halten allerdings das Bild nicht dauerhaft fest, sondern projizieren es für die unmittelbare Betrachtung durch das menschliche Auge. Diese Bilder dienten, im Falle der Lochkamera, unter anderem als Vorlage für die Anfertigung von realitätsgetreuen Zeichnungen (Abbildern) auf Papier.

Mit der Erfindung der Fotografie in der ersten Hälfte des 19. Jahrhunderts, durch die das durch die Optik der Kamera eingefangene, reflektierende (transiente) Licht eine chemische Reaktion auf lichtempfindlichem Fotopapier erzeugt, welche dauerhaft (persistent) fixiert werden kann, wurde der manuelle Prozess des Übertragens des durch Licht hervorgebrachten Abbildes der Realität in ein anderes Medium (Papier) erstmalig automatisiert.

Digitalisierung funktioniert nun ähnlich dieser analogen Fotografie, nur dass kein Fotopapier verwendet wird, sondern Bildsensoren das Signal, also das durch die Optik des Instruments (Kamera, Scanner) reflektierte Licht der realen Objekte, erfassen und in digitale Daten umwandeln, die wiederum als Rastergrafiken dauerhaft auf digitalen Medien gespeichert werden können.

Verfahren

Für die Bilddigitalisierung in den Kulturwissenschaften kommen vor allem zwei Gruppen bildgebender Instrumente zum Einsatz: Digitalkameras und Scanner. Dennoch darf man nicht vergessen, dass auch vor dem digitalen Zeitalter Verfahren angewandt wurden, um Abbilder realer Objekte ortsunabhängig bzw. ohne direkten Zugang zum Objekt bzw. Besitz derselben in Forschung und Lehre zu verwenden. Zeichnungen, (analoge) Fotografien oder Fotokopien können gedruckt und somit in Buchform verbreitet werden. Insbesondere Fotografien halten in Vervielfältigung als Abzug oder Dia-Positiv Einzug in Bildarchive, oder es werden hochwertige optische

Reproduktionen etwa von mittelalterlichen Handschriften in den Tafelwerken der Paläographie verwendet.

Mikroformen (Mikrofiche, Mikrofilm) sind stark verkleinerte analoge Abbildungen von gedruckten Vorlagen auf Polyester-basiertem Filmmaterial. Da man davon ausgeht, dass die Haltbarkeit dieses Materials bis zu 500 Jahre beträgt, werden Mikroformen seit ihrer Erfindung im 19. Jahrhundert vor allem für **revisionssichere Langzeitarchivierung** verwendet (Backup-Funktion). Stellvertretend als für das Original zu benutzendes Surrogat leisten sie zudem bei der Sekundärnutzung (Überprüfen von Stellen, Vorlage zur editorischen Transkription, etc.) einen Beitrag, um ein wertvolles Objekt wie etwa eine licht- und fettempfindliche mittelalterliche Handschrift zu schützen. Kritiker der Digitalisierung von Kulturgut stellen die nachgewiesene, lange Haltbarkeit von Mikroformen immer wieder der noch nicht abschließend beantworteten Frage nach der Langzeitverfügbarkeit digitaler Daten gegenüber. Dabei betonen sie auch die technologische Unabhängigkeit der Mikroformen und stellen dies als einen wesentlichen Vorteil im Vergleich zur schnelllebigen Entwicklung von Computertechnologie dar.

Neben ihrer archivalischen Funktion bieten Mikroformen außerdem die Möglichkeit der (allerdings qualitätsmindernden) Vervielfältigung. So kann zum Beispiel eine ganze Jahresausgabe einer Tageszeitung in Kleinstformat verschickt und gelagert werden. Mikroformen sind im Vergleich zur abgebildeten Vorlage sehr klein (daher der Name). Daher ist zum Betrachten ein spezielles Lesegerät erforderlich. Da Mikroformen weiterhin Verwendung finden und längst nicht alle Mikroformen digitalisiert wurden, finden sie sich nebst ihrer Lesegeräte in Bibliotheken und Archiven.

Mikroformen können durch spezielle Mikrofilm- oder Mikrofiche-Scanner digitalisiert und damit digitale Daten ohne Rückgriff auf das (unter Umständen nicht mehr existierende Original) erzeugt werden. Dieser Prozess ist weitgehend automatisierbar. Dabei kann allerdings die Qualität dieser Digitalisierung bestenfalls genauso gut sein wie die ursprüngliche Verfilmung auf Mikroform, die jedoch unter bestimmten Umständen wiederum von besserer Qualität sein könnte als das Original heute, das seit dem Zeitpunkt der Mikroverfilmung beschädigt oder verloren gegangen sein könnte.

Digitalkameras arbeiten im Wesentlichen wie analoge Kameras, nur dass an Stelle des lichtempfindlichen Films ein elektronischer Bildsensor (meist CCD-Sensoren: Charge-coupled Device) tritt. Der Standardaufbau einer Fotokamera besteht aus einer Öffnung (Blende), vor der eine oder mehrere optische Linsen (Objektiv) angebracht sind, die das einfallende Licht bündeln und das Scharfstellen (Fokussieren) des gesamten Bildes oder eines Ausschnittes ermöglichen.

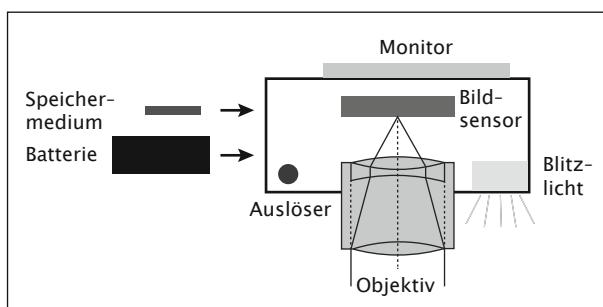


Abb. 47 Schematischer Aufbau einer Digitalkamera

Die wesentlichen Parameter, die entweder von der Kamera automatisch berechnet oder vom Benutzer manuell eingestellt werden, sind die Fokussierung (Anpassung der Optik an die Entfernung zwischen Objekt und Kamera) sowie die Größe der Blendenöffnung (beeinflusst die Schärfentiefe) und die Verschlusszeit (regelt die Dauer der Belichtung des lichtempfindlichen Films oder Bildsensors). Darüber hinaus bestimmt die Brennweite des Objektivs den Aufnahmewinkel.

Durch Betätigen des Auslösers der Kamera öffnet sich die Blende in der eingestellten Weite und lässt das Licht durch das Objektiv für die Dauer der eingestellten Verschlusszeit auf den elektronischen Bildsensor. Grundsätzlich werden zwei Sensorarten unterschieden: Flächensensoren für handelsübliche Digitalkameras und Zeilensensoren für spezielle Scannerkameras.

One-Shot-Sensoren registrieren gleichzeitig die drei Grundfarben der additiven Farbmischung (Rot, Grün und Blau). Der Chip an sich arbeitet monochrom, d. h. er kann nur hell oder dunkel, nur schwarz oder weiß sehen. Daher wird eine Farbfilterschicht benötigt, die das Licht in seine Spektralbestandteile zerlegt. Three-Shot-Sensoren nehmen die drei Grundfarben nacheinander auf. Sie benötigen daher jeweils einen Filter für rot, grün und blau und setzen das Farbbild anschließend zusammen. Zeilensensoren werden in speziellen Scannerkameras eingesetzt, die, ähnlich wie ein Flachbettscanner, das Bild zeilenweise abtasten.

Scanner: Scanner sind allgemein Datenerfassungsgeräte, die Objekte systematisch, nach vorgegebenen Mustern oder Rastern abtasten (von engl. *to scan*: untersuchen) und das in diesem Verfahren durch Bildsensoren aufgenommene analoge Signal in digitale Form umwandeln. Scanner existieren in vielfältigen Formen und Anwendungsbereichen (so etwa in der Medizin), im Kontext der Geisteswissenschaften interessieren uns vornehmlich (allerdings nicht ausschließlich) zunächst die verschiedenen Ausprägungen der Bildscanner.

Scanner arbeiten in der Regel nach folgendem Prinzip: Die Bildvorlage wird beleuchtet, und das reflektierte Licht wird über eine Stablinse, welche das reflektierte Licht bündelt und das Streulicht eliminieren soll, an einen optoelektronischen Zeilensensor geleitet. Die analogen Lichtsignale werden vom Analog-Digital-Wandler pixelweise in Digitalsignale umgewandelt, während gleichzeitig entweder die Vorlage oder die Sensoroptik schrittweise senkrecht zum Sensor zur nächsten Zeile bewegt wird. Ein frühes Beispiel für einen solchen Scanner ist das Fax-Gerät, bei dem allerdings die Datenübertragung auf analogem Signal erfolgte. Bei der Abtastung mit einem Flächensensor kann die gesamte Vorlage oder flächenhafte Teile der Vorlage gleichzeitig gescannt werden.

Beim **Flachbettscanner** wird das Objekt mit der zu scannenden Seite nach unten auf eine Glasplatte gelegt. Zum Scannen fährt ein beweglicher Spiegel in einem »Bett« unter dieser Glasplatte durch (das Objekt selbst wird nicht bewegt) und wirft das Licht auf einen fest installierten Spiegel, der es wiederum durch die Linse auf den Bildsensor wirft. Beim Flachbettscanner werden die Objekte einzeln auf die Vorlage gelegt, weshalb Flachbettscanner (zweidimensionale) Objekte unterschiedlichster Größe und Beschaffenheit, je nach Bauart bis zu einem Format von DIN A0 scannen können.

Buchscanner: Da die Optik des Flachbettscanners keine Schärfentiefe besitzt, muss die Vorlage ganz flach auf der Glasplatte aufliegen, um ein durchgängig scharfes Bild zu erreichen. Dies bereitet beim Scannen von Büchern auf Grund der Buchbindung häufig Probleme. Einen anderen Ansatz verfolgt der **Buchscanner**, der speziell auf die Digitalisierung von Büchern spezialisiert ist und daher in den Digital Humanities eine große Rolle einnimmt. Buchscanner verwenden im Gegensatz zum Flach-

bettscanner eine oder zwei Digitalkameras (trotz des Namens »Scanner«), die zu Beginn des Scanprozesses auf die Lichtverhältnisse sowie auf die Größe der Buchseiten eingestellt werden und pro Scanvorgang eine Einzel- bzw. eine Doppelseite des Buches ablichten. Die einfachste Vorrichtung hierbei ist die sogenannte **Buchwippe**, bei der das Buch an der gewünschten Stelle aufgeschlagen und mechanisch fixiert bzw. unter eine Glasplatte gedrückt wird, um Wölbungen der Seite und damit Unschärfen in der Aufnahme zu vermeiden. Buchwippen erfordern häufig einen Öffnungswinkel des zu scannenden Buchs von 90 bis 180 Grad, was zur Belastung von Heftung, Bundmaterial, Einbandrücken und Gelenken des Buches führen kann. Als eine Institution, die konservatorisch tätig ist und viele historisch wertvolle Altbestände besitzt, hat die Herzog-August-Bibliothek in Wolfenbüttel daher die Buchwippe zu einem System weiterentwickelt, das einen Öffnungswinkel von 45 Grad erlaubt (*Wolfenbüttler Buchspiegel*). Dabei wird ein Spiegel in das Objekt eingeführt, das Spiegelbild anstatt der Buchseite digitalisiert und das Abbild anschließend umgerechnet.

Das Weiterblättern bei einfachen Buchscannern, bei Buchwippe und Buchspiegel erfolgt von Hand. Bei umfangreichen Digitalisierungsprojekten kommen zudem vollautomatische Buchscanner, sogenannte **Scan-Roboter**, zum Einsatz. Bei Scan-Robotern mancher Bauart wird die Seite nach dem Scanvorgang angesaugt und automatisch umgeblättert. Luftströme werden hierbei auch verwendet, um das Papier für die Aufnahme zu glätten und ersetzen Glasplatte oder mechanische Fixierungen. Scan-Roboter anderer Bauart verwenden andere Mechanismen zum Glätten und Umblättern der Seite. Im überwachten Betrieb der Digitalisierungspraxis schaffen moderne Scan-Roboter mehrere Bücher pro Tag.

Parameter

Der Digitalisierung liegen technische Parameter zu Grunde, die wir auch als Parameter einer Modellierung auffassen können, denn jedes Digitalisat ist als Abbild ein Modell seiner Vorlage. Daher gelten auch bei der Digitalisierung die allgemeinen Grundsätze der Modellbildung. Insbesondere müssen wir uns vor Augen führen, dass jede Digitalisierung ein zweckgebundener, reduzierender und subjektiver Vorgang ist.

Die Aussagekraft eines Digitalisats und letztendlich sogar dessen Verwendbarkeit sind im wissenschaftlichen Kontext an die Parameter seiner Entstehung gebunden. Wird etwa die Farb-Fotografie einer Straßenszene in Graustufen digitalisiert, kann man das Digitalisat nicht zur Analyse von Modefarben verwenden (das Original aber sehr wohl). Hier greift das Verkürzungsprinzip der Datenmodellierung. Wird dem Benutzer zusätzlich die Information vorenthalten, dass die Vorlage kein schwarz-weiß-Bild war, sondern lediglich mit Graustufen gescannt wurde, kann dies zu problematischen Fehlinterpretationen führen. Daher ist es elementar, bei der Nutzung digitaler Daten die Parameter ihrer Entstehung zu kennen und kritisch zu prüfen, ob und wie die Daten für den jeweiligen Forschungszweck geeignet sind. Dies ist nicht immer einfach, und viele der Eignungskriterien digitaler Daten für die eigene Forschungsfrage drücken sich nicht in simplen technischen Parametern der Digitalisierung aus. Cohen und Rosenzweig führen aus:

»[...] the quality of the digital image rests on the quality of the original, the digitizing method employed, the skill of the person doing the digitizing, and the degree to which the digital copy has adequately ›sampled‹ the analog original« (Cohen/Rosenzweig 2005, 95).

Liegt eine Unschärfe im Digitalisat einer Fotografie an der Fokussierung oder Verwacklung der Kamera, mit der das ursprüngliche Bild erzeugt wurde, also am Fotografen, oder an der Fokussierung oder Verwacklung der Digitalkamera, mit der die Fotografie digitalisiert wurde? Ist der Text einer Randglosse in einem mittelalterlichen Codex wirklich nicht lesbar, oder ist er lediglich auf Grund der starken Wölbung der Buchseiten im Digitalisat verzerrt, was man durch geeignete Apparaturen hätte ausgleichen können? Besitzt die Gedenktafel keine Inschrift auf der Rückseite, oder wurde sie lediglich nicht mit abgelichtet? Für den Bildwissenschaftler oder den Historiker sind dies entscheidende Fragen.

Technische Parameter der Digitalisierung: Auch technische Parameter spielen bei der Digitalisierung und bei der kritischen Beurteilung der Nutzbarkeit digitaler Bilder eine wichtige Rolle. Die in vielen Fällen wichtigsten Messgrößen sind dabei die schon besprochene Farbtiefe sowie die **Auflösung**. Der Begriff der Auflösung wird häufig mit der Bildgröße synonym verwendet (»Auflösung von 1920 × 1200«), was im Grunde falsch ist. Bei der Auflösung im eigentlichen Sinne geht es um eine Punktdichte pro Streckeneinheit. Erst wenn man die Abtastrate eines Scanners in Relation zur Größe der zu scannenden Vorlage betrachtet, erhält man ein Maß der Qualität der Digitalisierung. Je mehr digitale Pixel von einer Vorlage abgetastet werden können, d. h. je feiner die Abtastrate ist, umso höher die Qualität. Umgekehrt ist die Auflösung auch ein Maß der Schärfe eines digitalen Bildes auf einem Ausgabemedium (Bildschirm, Drucker). Je mehr Pixel des digitalen Bildes etwa auf einem DIN A4-Blatt ausgedruckt werden können, desto schärfster der Ausdruck.

Die Auflösung wird in Bildpunkten pro Streckeneinheit gemessen. In der Regel wird das Maß **pixel per inch (ppi)** oder dots per inch (dpi) verwendet, die, sieht man von Druckern ab, synonym verwendet werden können. Ein *inch* (deutsch: Zoll) sind 25,4 mm, das Maß *ppi* gibt also an, wie viele Bildpunkte auf einer Strecke von 25,4 mm abgetastet oder ausgegeben werden. Höhere Auflösungen verursachen dabei größere Bilddateien, die wiederum übermittelt und gespeichert werden müssen und einen Kosten- und Zeitfaktor darstellen.

In den Digitalisierungsprojekten der Digital Humanities werden Auflösungen daher in der Regel pragmatisch und zweckgebunden gewählt. Als Untergrenze definiert die Deutsche Forschungsgemeinschaft eine Auflösung, »bei der die Details einer Vorlage vollständig in einer gleich großen Reproduktion wiedergegeben werden können. Ein Ansatz zur Ermittlung der notwendigen Mindestauflösung eines Digitalisats ist das Auflösungsvermögen des menschlichen Auges bei deutlicher Sehweite [25 cm]« (DFG 2013, 8–9). Für Vorlagen, die für die Betrachtung mit bloßem Auge gedacht sind, ergibt sich damit eine Mindestauflösung von 300ppi. Für bestimmte Anwendungsfälle ist jedoch eine höhere Auflösung sinnvoll, etwa bei Medien, deren Bildinformationen nur in Vergrößerung vollständig erfassbar sind (Miniaturen, Dias). Hier können Auflösungen von 3000ppi und mehr notwendig sein.

Die Wahl der Auflösung bei der Digitalisierung unterliegt dem allgemeinen Prinzip der Zweckgebundenheit bei der Erzeugung von Abbildern/Modellen. Schwierig ist es jedoch, sämtliche Nutzungsszenarien der Digitalisate für die Zukunft vorherzusagen. So werden etwa im Bereich der digitalen Paläographie heute deutlich höhere Auflösungen als 300ppi bei der Digitalisierung von Handschriften gefordert, weil zukünftig nicht nur die Betrachtung des digitalen Abbildes der Handschrift mit dem bloßen Auge berücksichtigt werden sollen, sondern auch detaillierte Analyseverfahren aus dem Bereich der Computer Vision, die über die Sehkraft des menschlichen Auges hinausgehen (Hassner et al. 2014).

12.3 | Erschließung der Digitalisate

Metadaten: »Die Erzeugung von Metadaten, welche erst die Auffindbarkeit der Objekte gewährleisten und eine kontextualisierende Präsentation ihrer digitalen Images erlauben, ist zentraler Bestandteil der Digitalisierung« (DFG 2013, 25). Für die Nutzung von Bilddigitalisaten sind **technische, deskriptive, strukturelle und administrative Metadaten** von Bedeutung. Diese vier Typen von Metadaten unterscheiden sich in Hinblick auf ihre Anwendung.

- **Deskriptive Metadaten:** werden vorrangig zum Auffinden eines bestimmten Objektes herangezogen
- **Strukturelle Metadaten:** werden verwendet, um das aufgefundene Objekt in seinem Aufbau zu erfassen
- **Technische Metadaten:** werden verwendet, um die Qualität des Objekts zu beurteilen
- **Administrative Metadaten:** werden verwendet, um etwa Nutzungsrechte zu dokumentieren

Dies sind zwar in ihrer Erzeugung recht eindeutig festgelegte Kategorien, in der Anwendung aber keine starren Beschränkungen, denn natürlich können etwa auch technische Metadaten zum Auffinden eines Objektes herangezogen werden (z. B. »finde alle Digitalisate von mittelalterlichen Urkunden, die mit mindestens 600ppi digitalisiert wurden«).

Es folgt eine Übersicht über einige international gängige Metadaten-Standards und Spezifikationen, die im Rahmen der Digitalisierung des kulturellen Erbes eine Rolle spielen. Da sich deskriptive und strukturelle Metadaten je nach Objekttyp (z. B. Fotografie, mittelalterliche Urkunde, Roman) stark unterscheiden, kann im Rahmen dieser Einführung nicht auf Details eingegangen werden.

- **Dublin Core** ist ein internationaler Standard zur Beschreibung von digitalen und realen Dokumenten sowie anderen Objekten. Das *Dublin Core Metadata Element Set* besteht dabei im Kern lediglich aus 15 Elementen, die in verschiedenen Formaten (RDF/XML, OpenDocument, HTML) ausgedrückt werden können. Dublin Core wird von der Dublin Core Metadata Initiative entwickelt.
- **METS** (Metadata Encoding and Transmission Standard) ist ein Standard zur XML-basierten Kodierung von deskriptiven, administrativen und strukturellen Metadaten von Objekten innerhalb einer digitalen Bibliothek. METS wirkt dabei als ein Container, bei dem in einzelnen Abschnitten andere Formate (z. B. Dublin Core, MODS, MARC) untergebracht werden können. Der METS-Standard wird von der Library of Congress gepflegt und das zugehörige Datenschema von der Digital Library Federation, einem nordamerikanischen Bibliotheksverbund weiterentwickelt.
- **MODS** (Metadata Object Description Schema) ist ein XML-Format für bibliografische (deskriptive) Metadaten. Es gilt als Kompromiss zwischen der Komplexität des Bibliothekskatalogsformats MARC (Machine Readable Cataloging) und der Einfachheit von Dublin Core. MODS wird von der Library of Congress entwickelt.
- **TEI** (Text Encoding Initiative). Die TEI bietet Richtlinien und XML-Schemata vornehmlich für die Kodierung elektronischer Texte. Die Richtlinien enthalten aber auch eigene Sektionen, u. a. `<teiHeader>`, `<msDesc>` und `<bibl>`, die für deskriptive und strukturelle Metadaten herangezogen werden können. TEI wird von

einem internationalen Konsortium mit gleichem Namen gepflegt und weiterentwickelt.

- **EAD** (Encoded Archival Description) ist ein XML-Standard zur Kodierung deskriptiver Metadaten für Archivalien (Findmittel), der von Archiven, Museen und Bibliotheken genutzt wird. EAD wird von der Library of Congress gepflegt.
- **SAFT** (Standardaustauschformat). SAFT-XML ist ein XML-Austauschformat zur Kodierung deskriptiver Metadaten für Archivalien (Findmittel). SAFT wurde im Auftrag der Deutschen Forschungsgemeinschaft entwickelt.

12.4 | Textdigitalisierung

Die zuvor besprochenen Verfahren zur Digitalisierung bilden eine Vorlage als zweidimensionale Tabelle von Farbwerten in Form einer digitalen Rastergrafik ab. Dabei werden lediglich die geometrischen Positionen der einzelnen Pixel gespeichert, nicht aber ihre Bedeutung im Bild oder gar der Bedeutungszusammenhang einzelner Pixel untereinander. Eine Information zum Beispiel dargestellt, dass eine Gruppe von schwarzen Pixeln die Tinte repräsentiert, mit der ein bestimmter Buchstabe eines Textes gedruckt wurde, liegt in der Rastergrafik nicht vor. Handelt es sich bei der Vorlage also um einen Träger von Text (etwa ein Buch), ist folglich ein weiterer Bearbeitungsschritt notwendig, um diesen Text zu erkennen, aus der Rastergrafik zu extrahieren und daraus maschinenlesbare Zeichencodes zu generieren. Diesen Vorgang nennen wir **Texterkennung** oder **Transkription** (von lat. *transcribere*: abschreiben, übertragen). Wir können zwischen manuellen und automatischen Verfahren unterscheiden.

Die **manuelle Texterkennung**, auch *keying* genannt, erfolgt durch Abtippen der Vorlage per Hand. Hierbei erfolgt implizit durch den Erfasser des Texts eine Kodierung in ein maschinenlesbares Zeichensystem (Unicode). Da bei der manuellen Texterkennung Lese- und Flüchtigkeitsfehler auftreten bzw. bei schlechter Vorlage oder unbekannter Schrift Interpretationsmöglichkeiten bestehen, wird häufig das Verfahren des *double-keying* angewandt, bei dem zwei Erfasser unabhängig voneinander den gleichen Text transkribieren. Beide Transkriptionen werden anschließend verglichen und Differenzen durch einen Dritten korrigiert. In großen Digitalisierungsprojekten wird das *keying* oft mittels **Outsourcing** durch externe Dienstleister durchgeführt, die eine Texterkennungsqualität von bis zu 99,997 % korrekt erfasster Buchstaben anstreben (DFG 2013, 33).

Bei der **automatischen Texterkennung** kommen Verfahren der **Optical Character Recognition** (Optische Zeichenerkennung, OCR) zum Einsatz. OCR-Verfahren wurden erstmalig bereits gegen Ende des 19. Jahrhunderts beschrieben und kamen mit der Erfindung der Digitalcomputer in den 1940er und 1950er Jahren zum kommerziellen Einsatz (Eikvil 1993). Da für automatische Texterkennung auch außerhalb der Wissenschaften ein großer Bedarf besteht, werden die Verfahren seitdem kontinuierlich weiterentwickelt. Für OCR kommen kommerzielle und frei verfügbare Software-Produkte in Frage, die entweder generisch für den breitgefächerten Einsatz oder für spezielle Dokumentarten und Schrifttypen entwickelt werden. In den Digital Humanities werden zumeist generische Software-Produkte eingesetzt, die gegebenenfalls für besondere Anwendungsfälle speziell trainiert werden.

Der Prozess der Texterkennung gliedert sich in der Regel in mehrere Phasen: einer Vorverarbeitung der grundlegenden Rastergrafiken, der eigentlichen Zeichenerkennung und einer Nachbereitung des erkannten Textes. Was im Folgenden lediglich modellhaft

dargestellt werden kann, variiert dabei in den verschiedenen Software-Implementierungen. Auch werden zahlreiche Optimierungsverfahren diskutiert (Mühlberger).

In der **Vorverarbeitung** (*pre-processing*) wird zunächst versucht, die Rastergrafik als Vorlage für die OCR zu optimieren. Hierzu zählt die Ausgleichung eventueller Verzerrungen oder Rotationen der Vorlage, wobei eine gute Bilddigitalisierung diesen Schritt vorwegnehmen kann und idealerweise auch sollte. Mit der anschließenden **Binarisierung** wird jeder Pixel der Rastergrafik entweder Text oder Hintergrund zugeordnet, was die OCR-Software auf Basis eines Algorithmus entscheidet. Hierbei können fließende Übergänge, Verunreinigungen oder Zerfall der Vorlage (bei historischen Dokumenten) ebenso Probleme verursachen wie eine zu geringe Auflösung des Digitalisats.

Es folgt eine Analyse des Layouts und die Segmentierung der Vorlage. Ziel ist es, die Bereiche zu identifizieren, die überhaupt Text enthalten. Andere Elemente wie Bilder oder Grafiken werden als solche markiert, fließen aber nicht in die weitere Texterkennung ein. In den Textbereichen werden anschließend Zeilen und häufig auch Wörter voneinander separiert. Moderne OCR-Software speichert die Layout-Information (Koordinaten) von Textblöcken, einzelner Zeilen, Wörter oder auch Zeichen zusammen mit dem erfassten Text, so dass bei einer späteren Volltextsuche auch die Position des Treffers im Bilddigitalat visuell hervorgehoben werden kann.

Moderne OCR-Software bemüht sich außerdem darum, das Dokument in seiner logischen Struktur zu erfassen, etwa Kopfzeilen, Fußzeilen, Marginalien und andere Paratexte vom eigentlichen Text zu trennen, Überschriften und Absätze zu erkennen, eingebettete Bilder und Grafiken als solche zu markieren oder eine Tabellenstruktur zu rekonstruieren. Dies geschieht zunächst rein auf Basis typographischer Kriterien, wie Abstandsmessungen zwischen Textblöcken oder dem Wechsel von Schriftarten (*fonts*) und kann daher letztendlich nur als Indikator für eine logisch-inhaltliche Struktur des Textes dienen, die in der Regel nachkorrigiert werden muss.

Für die eigentliche **Zeichenerkennung** sind mehrere Verfahren beschrieben, die von den Software-Produkten in verschiedenen Konstellationen und Konfigurationen eingesetzt werden. Ausgehend von der Extraktion der einzelnen Glyphen wird diese beim Musterabgleichsverfahren auf eine Matrix reduziert (s. Abb. 48) und berech-

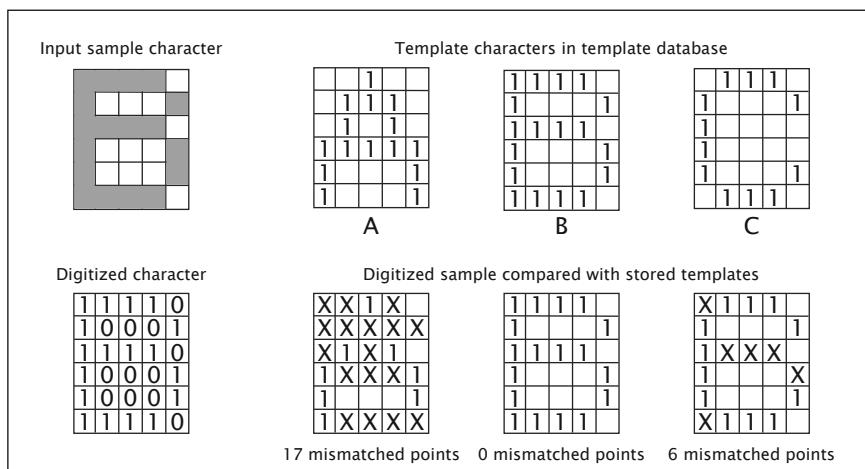


Abb. 48 OCR-Musterabgleich nach Li (1991)

net, welchem Zeichen eines modellhaften Zeichensatzes (Klassifikatoren) diese Glyphe am wahrscheinlichsten entspricht (Li 1991). Einen anderen Ansatz verfolgt das fortgeschrittenere Verfahren der *feature extraction*, bei der die Glyphe als Komposition von Einzelbestandteilen (z. B. Abstriche, Bögen) betrachtet, in diese zerlegt (Vektorisierung) und mit Prototypen verglichen wird.

In beiden Verfahren basiert die Zeichenerkennung auf Datenbanken mit bereits bekannten Glyphen-Mustern. Moderne OCR-Software kennt dabei die gängigen (Druck-)Schriftarten und kann zudem bei unbekannten und stark abweichenden Schriftarten ihre Datenbank durch Training entsprechend erweitern.

In der **Nachbereitung** (*post-processing*) wird schließlich versucht, die Qualität des erkannten Textes zu verbessern. Hier werden häufig fachspezifische Wörterbücher zum Abgleich der Texterkennung eingesetzt, in dem in einem Rückkopplungsprozess zweifelhafte Transkriptionen gegen Wörter im Lexikon ausgetauscht werden, zum Beispiel bei Verwechslungen von Buchstaben und Ziffern. Auch computerlinguistische Verfahren werden in der Nachbereitung angewendet. Aufwändiger werden die Nachkorrekturen vor allem bei historischen Texten und bei Texten mit nicht normierten Schreibweisen oder hohem Anteil von Eigennamen, die schwer in Wörterbüchern zu finden sind.

Qualitätskriterien für die Texterkennung: Das Maß für die Qualität der Texterkennung ist die **Erkennungsgenauigkeit**. Wir können zunächst zwischen Wort- und Buchstabengenauigkeit unterscheiden, die im Wesentlichen aussagen, wie hoch der Prozentsatz der korrekt repräsentierten Wörter bzw. Zeichen in Bezug auf die Vorlage ist. Sie wird in der Regel durch Stichproben ermittelt. Auch die Varianz in der Verteilung der Erkennungsfehler spielt eine Rolle, kann hier aber nicht vertieft werden. Die Erkennungsgenauigkeit entscheidet darüber, ob und wie brauchbar ein erfasster Text für verschiedene Anwendungszwecke ist. Als Richtgröße gibt die DFG eine Buchstabengenauigkeit von 99,95 % als »wissenschaftlich zuverlässig« an, da sie auch Negativsuchen erlaubt (d. h. liefert eine Volltextsuche nach einem bestimmten Wort keinen Treffer, so kann davon ausgegangen werden, dass dieses Wort nicht im Text vorkommt). Deutliche schlechtere Genauigkeiten im Bereich von 80–90 % können je nach Betrachtungsperspektive noch für positive Suchen im Retrieval verwendet werden, wobei jedoch vor allem ein deutlich geringerer Recall zu erwarten ist. Eine Buchstabengenauigkeit von unter 80 % gilt als »fragwürdig« (DFG 2013, 32). Wichtig ist in jedem Fall, sich vor Benutzung digitalisierter Texte über deren Qualität zu informieren.

Sehr hohe Erkennungsgenauigkeiten werden bei modernen Buchdruck- und Computerschriften erreicht. Auch Typoskripte (Schreibmaschine) werden in der Regel gut erkannt, solange das Schriftbild sauber ist. OCR-Software stößt jedoch häufig an die Grenze, wenn es sich um historische Texte handelt. Als Faustregel gilt: je älter der Druck, desto geringer die Wahrscheinlichkeit einer ausreichend hohen Qualität in der Texterkennung. Während etwa ein Frakturdruck des 19. oder frühen 20. Jahrhunderts mittlerweile oft verhältnismäßig gut erkannt werden kann, stellen frühere Drucke häufig noch ein Problem dar und weisen eine deutlich geringere Erkennungsgenauigkeit auf.

Im Bereich der **Handschriftenerkennung**, die vor allem für historische Texte relevant ist, werden verschiedene Verfahren diskutiert. Ein allgemeingültiger Ansatz, der eine gedruckten Vorlagen entsprechende hohe Erkennungsgenauigkeit bietet, ist auf Grund der Individualität und starken Varianz der menschlichen Schrift auch in nächster Zeit nicht zu erwarten (Hassner et al. 2014). Hier bleibt weiterhin die ma-

nuelle Transkription der einzig gangbare Weg der Texterfassung für wissenschaftliche Zwecke, auch wenn diese durch verschiedene computerbasierte paläographische Hilfsmittel unterstützt werden kann. Von dieser retrospektiven (offline) Handschriftenerkennung zu unterscheiden ist die sogenannte online-Handschriftenerkennung, wie sie etwa auf Tablet-Computern oder Smartphones erfolgt. Da hierbei der Duktus, d. h. die Stiftführung beim Schreiben, mit aufgezeichnet wird und somit ausgewertet und die Erkennungs-Software zudem auf den spezifischen Schreiber trainiert werden kann, sind sehr hohe Erkennungsgenauigkeiten erzielbar.

12.5 | Weitere Digitalisierungsverfahren

In diesem Kapitel haben wir uns vorrangig mit zwei Facetten der Digitalisierung beschäftigt: der (zweidimensionalen) Bilddigitalisierung und der daraus abgeleiteten Erfassung von Text. So vielseitig wie die geisteswissenschaftlichen Disziplinen und so divers ihre Gegenstände sind, so weit gestreut ist aber auch das Feld der Digitalisierung. Abschließend sollen nun weitere Bereiche thematisiert werden. Dies kann jedoch nur exemplarisch und skizzenhaft erfolgen.

Schon bei der Digitalisierung von Schriftgut fällt auf, dass bei einer zweidimensionalen (flachen) Abbildung – wie sie eine Kamera oder ein Buchscanner liefert – manche Eigenschaften, wie etwa die Wellung eines mittelalterlichen Pergaments, die Beschaffenheit eines Bucheinbandes oder die Tiefe von Gravuren, nicht oder nicht vollständig erfasst werden. Bei anderen Gegenständen der Digital Humanities, etwa historische oder archäologische Artefakte, ist die Dreidimensionalität der Objekte noch offensichtlicher, wie man sich leicht an einer Münze vergegenwärtigen kann: Vorderseite, Rückseite, Tiefe der Prägung und der umlaufende Münzrand lassen sich durch eine zweidimensionale Bilddigitalisierung nicht vollständig abbilden.

Multispektralfotografie

Zu den spezialisierten und fortgeschrittenen Verfahren zählt auch die Digitalisierung durch Multispektralkameras. Multispektralfotografie ist eine Methode für spezielle Anwendungen wie die geographische Fernerkundung (Luftbilder), deren Möglichkeiten seit einiger Zeit auch verstärkt in den Geisteswissenschaften genutzt wird.

Im Unterschied zur herkömmlichen Farbfotografie wird das Lichtspektrum bei der Multispektralfotografie nicht in nur drei Bereiche (rot, grün, blau), sondern in eine größere Zahl variabel abgrenzbarer Spektralbänder unterteilt, die separat von digitalen Sensoren erfasst werden. Dies gewährleistet eine große Flexibilität bei der Erfassung der Farben, weshalb eine Einsatzmöglichkeit die hochgradig farbgetreue Reproduktion von Gemälden ist.

Über den Bereich des Sichtbaren hinaus werden auch die angrenzenden Bereiche des für das menschliche Auge nicht unmittelbar sichtbaren ultravioletten und infraroten Lichtes erfasst. Diese Lichtbereiche können wissenschaftlich wertvolle Informationen über Oberfläche oder tieferliegende Schichten der Objekte sowie ihre chemischen Bestandteile liefern. In Gemälden können bestimmte Farbpigmente (z. B. Bleiweiß) nachgewiesen oder kunsthistorisch bedeutsame Unterzeichnungen (Entwurfszeichnungen) wie auch konservatorisch relevante Schäden am Objekt aufgedeckt und Originale von Fälschungen unterschieden werden. Die unterliegenden älteren Textschichten mehrfach beschriebener Pergamente

(Palimpseste wie das berühmte »Archimedes-Palimpsest«) können ebenso sichtbar gemacht werden wie Text auf ausgebliebenen oder beschädigten Manuskripten. Multispektralfotografie ist eine nichtinvasive Methode, d. h. das untersuchte Objekt wird dabei nicht beschädigt oder zerstört. Sie ist daher auf fragile und kostbare Objekte des kulturellen Erbes (und damit für dessen Erforschung) anwendbar.

In diesem Kontext gewinnt die dreidimensionale **Objektdigitalisierung** zunehmend an Bedeutung. Vorreiter sind hier die Archäologie und die Museen. Es sind unterschiedliche Verfahren beschrieben worden, wobei die Spannbreite von multiperspektivischen 2D-Aufnahmen mit anschließender softwarebasierter Berechnung eines 3D-Modells über Laserscan-Verfahren bis hin zur Computertomographie reicht. Im Gegensatz zur Bilddigitalisierung von »Flachware« erfordert die Objektdigitalisierung gegenwärtig jedoch noch einen deutlich höheren Aufwand und setzt häufig die Anwendung mathematischer Verfahren der Modellbildung voraus, um wissenschaftlich brauchbare Ergebnisse zu erzielen.

Die bislang hier vorgestellten Digitalisierungsarten beruhen auf optischen Verfahren. Auf Akustik beruht hingegen die **Audiodigitalisierung**. Wir verstehen hierunter eine Retrodigitalisierung, also die Umwandlung bereits vorliegender analoger Tonaufzeichnungen, von der die unmittelbare digitale Aufnahme zu unterscheiden ist. Da Töne transient (flüchtig) sind, besteht in ihrer Aufzeichnung die einzige Möglichkeit, sie möglichst authentisch zu dokumentieren. Handelt es sich dabei um gesprochene Sprache, ist zwar die Transkription als geschriebener Text ebenfalls möglich, doch gehen dabei in der Regel wichtige Eigenschaften (z. B. Intonation, Sprechpausen, Dialekt) verloren. Tondokumente sind seit dem Ende des 19. Jahrhunderts in analoger Form bekannt, gängige Medien waren Wachswalzen, Schallplatten und Tonbänder. Ihre Digitalisierung erfolgt mittels Analog-Digital-Wandlung. Auf die technischen Verfahren im Einzelnen kann in dieser Einführung nicht eingegangen werden.

Digitalisierung ist für die Digital Humanities ebenso zentral wie facettenreich. Für manche Aufgaben wie die Buchdigitalisierung in Bibliotheken oder die Schriftgutdigitalisierung in Archiven sind die Verfahren inzwischen wohl bekannt und skalierbar. Eine gute Orientierung bieten die Praxisregeln der DFG 2013. Bei jeder Digitalisierung sind jedoch die besonderen Eigenschaften des Objekts und der Digitalisierungszweck zu berücksichtigen. Hieraus sind die Digitalisierungsparameter abzuleiten, konsequent anzuwenden und in den Metadaten zu dokumentieren.

Literatur

- Cohen, Daniel/Rosenzweig, Roy: *Digital History. A Guide to Gathering, Preserving, and Presenting the Past on the Web*. Philadelphia 2005.
- DFG: »Praxisregeln Digitalisierung« (2013), http://www.dfg.de/formulare/12_151/12_151_de.pdf (19.10.2015).
- Eikvil, Line: »OCR. Optical Character Recognition« (1993), <https://www.nr.no/~eikvil/OCR.pdf> (19.10.2015).
- Hassner, Tal/Rehbein, Malte/Stokes, Peter A./Wolf, Lior: »Computation and palaeography: potentials and limits«. In: *Dagstuhl Manifesto* 2/1 (2014), 14–35.
- Li, Ning: *An Implementation of OCR System Based on Skeleton Matching*. Canterbury 1991.

- Mühlberger, Günter: »Digitalisierung Historischer Zeitungen. Aus dem Blickwinkel der Automatisierten Text- und Strukturerkennung (OCR)«. In: *Zeitschrift für Bibliothekswesen und Bibliographie* 58/1 (2011), 10–18.
- Mori, Shunji et al.: »Historical review of OCR research and development«. In: *Proceedings of the IEEE* 80/ 7 (1992), 1029–1058.
- Saffady, William: *Micrographics: Technology for the 21st Century*. Prairie Village, Kan. 2000.
- Tanner, Simon et al.: »Measuring mass text digitization quality and usefulness. Lessons learned from assessing the OCR accuracy of the British Library's 19th century online newspaper archive«. In: *D-Lib Magazine* 15/78 (2009).
- Terras, Melissa: *Digital Images for the Information Professional*. Farnham 2008.
- The Library of Congress: »Technical standards for digital conversion of text and graphic materials« (2006), <http://memory.loc.gov/ammem/about/techStandards.pdf> (19.10.2015).

Internetquellen

Spezifikationen von Metadaten-Standards (Auswahl):

- Dublin Core: <http://dublincore.org/specifications/>.
- EAD: <http://www.loc.gov/ead/>.
- METS: <http://www.loc.gov/standards/mets/>.
- MODS: <http://www.loc.gov/standards/mods/>.
- NISO: »Understanding Metadata« (2004), <http://www.niso.org/publications/press/UnderstandingMetadata.pdf> (19.10.2015).
- SAFT: <http://www.archivschule.de/DE/forschung/retrokonversion/vorstudien-und-saft-xml/>.
- TEI: <http://www.tei-c.org>.

Malte Rehbein