

Aktuelle Studentenarbeiten

9.1 Hochskalierung von Bildern für FM-Raster

Durch die Umstellung auf einen digitalen Workflow sind Grossdruckereien heutzutage imstande, hochgenaue Druckplatten herzustellen. Zudem wurde der physikalische Druckprozess insgesamt stabiler, was eine Reproduktion mit höheren Auflösungen und geringeren Schwankungen ermöglicht. Ein aktuelles KTI-Projekt in unserer Gruppe hat nun die Reproduktion mit hochfeinen FM-Raster (10 Mikrometer-Dots, entspricht einer Dot-Auflösung von 2400 dpi) im Offsetdruck zum Gegenstand.

Es soll zusammen mit der Gutenberg-Druckerei in Liechtenstein durchgeführt werden, welche mit dem Druck von NP-Rastern und Sicherheitsdruck (Beispiel: schweizer Briefmarken) grosse Erfahrung hat. Die hohe konzeptionell erreichbare Auflösung lässt sich in der Praxis jedoch selten erzielen, da normalerweise die Qualität der Kundendaten damit nicht überstimmt. Bilddaten von Kunden werden in typischen Auflösungen von 250 bis 450 Pixel per inch geliefert, ideal wären 2400. Bei den dadurch erforderlichen Skalierungsfaktoren von 5 bis 10 sind die Standardverfahren (grundsätzlich Photoshop bikubisch) überfordert und müssen durch neue, «intelligente» Skalierungsverfahren ersetzt werden.

Die primäre Anforderung an ein solches Skalierungsverfahren ist die Beibehaltung sowohl der Kantenschärfe als auch von weichen Übergängen, jedoch möglichst ohne den originalen Bildeindruck selbst zu verändern. Als zusätzliche Anforderung muss die Methode sehr schnell und speicherplatzeffizient sein, da normalerweise Bilddaten von mehreren Gigabytes verarbeitet werden müssen und die Software auf einem RIP-Server laufen soll. Mit dieser

Hochskalieren von Bildern für FM-Raster

- ❖ **Stand der Technik:** 2400 dpi bei FM-Raster
- ❖ **Kundendaten:** 250–450 Pixel pro inch
 - ideal wären 2400 Pixel pro inch
- ❖ **Ziel:** «intelligentes» Hochskalieren des Bildes
- ❖ **Rahmenbedingungen**
 - gute Kantenschärfe
 - glatte Übergänge sollen glatt bleiben
 - effiziente Algorithmen, da grosse Datenmengen
 - die visuelle Güte soll mit einem psychovisuellen Test überprüft werden

Scannen und Komprimierung bei Massendigitalisierungen

- ◇ **gegenwärtig:** Massendigitalisierungen für Kulturerbe-Projekte
 - eher schlechte Scan-Qualität, kaum lesbar durch Menschen
- ◇ **Problem:** Scan-Auflösung ⚡ Speicherplatz
- ◇ **Ziel:** bessere Scan-Qualität und einfache Komprimierung
 - optimierter Scan-Workflow durch ein Target-Konzept
 - Entwicklung von Bildqualitätsmassen zur Target-Auswertung
 - Entwicklung eines math. einfachen Komprimierungsformats
 - Analyse der visuellen Performance des Konzepts

hohe Scan-Auflösung + hohe Komprimierung

2

klaus simon farbe im digitalen publizieren farbmeterik

Arbeit sollen erste Schritte in Richtung des angestrebten Verfahrens gemacht werden, beginnend mit einer Literaturrecherche und mit dem Ziel, eine erste funktionierende Methode mit den erwähnten Anforderungen zu implementieren, mit Potential zu Erweiterungen.

Diese Arbeit ist geeignet für Semester- oder Master-Arbeiten für Informatiker mit Interesse in Bildanalyse (Kantendetektion) sowie einem Geschick für Rasterbilder. Gute bis sehr gute Programmierkenntnisse in C/C++ sind erforderlich.

Die Arbeit wird von TOBIAS STAMM, dem Leiter des Gutenberg-Projektes, betreut.

Literatur

- Henry R. Kang, *Digital Color Halftoning*, IEEE Press, «0-7803-4741-2».
- Daniel L. Lau and Gonzalo R. Arce, *Modern Digital Halftoning*, CRC Press, «978-1-4200-4753-0».
- E. Reinhard, E. Khan, A. Akyüz and G. Johnson, «*Color Imaging*», AK Press, «978-1-56881-344-8».

9.2 Scannen und Komprimieren bei Massendigitalisierungen

Weltweit werden gegenwärtig im Kontext von Bibliotheken Digitalisierungen der Massenbestände durchgeführt. Da es sich um Grossaufträge handelt, werden Scanner eingesetzt, die auf leichte Bedienbarkeit ausgelegt sind und nicht auf eine gute fotografische Qualität. Die damit erreichte Scan-Qualität ist denn auch im Allgemeinen eher schlecht. Die Proben aus laufenden Digitalisierungsprojekten sind durchwegs an der Grenze des für Menschen Lesbaren. Es ist deshalb nicht verwunderlich, dass die handelsübliche OCR-Software bei der Rekonstruktion der Texte überfordert ist und nicht

nur, weil es sich bei den Schriften um Fraktur oder Ähnliches handelt. Das Problem wird zusätzlich durch den hohen Speicherplatzbedarf für die eingescannten Bilder verschärft. Andererseits sollen die erzeugten Daten das Kulturerbe langfristig sichern, was komplexe proprietäre Datenformate eigentlich ausschliesst.

Um die Situation zu verbessern, muss zum einen die effektive Scan-Qualität verbessert werden und zum zweiten das gescannte Bild so komprimiert werden, dass kein erweiterter Speicherplatzbedarf resultiert.

Der erste Schritt zur Verbesserung bzw. zur Kontrolle der Scan-Qualität sollte in dem Mitscannen eines handelsüblichen Targets bestehen. Dabei versteht man unter einem Target z.B. ein schwarzes Muster auf einer transparenten Folie. Die zu entwickelnde Auswertesoftware muss dann das Target im eingescannten Bild identifizieren und auf Grund der Eigenschaften des Target-Scans ein Qualitätsmass für den Scann-Vorgang liefern. Insbesondere sollte eine Aussage über die erreichte räumliche Auflösung bzw. der nutzbaren Bittiefe erfolgen. Die letztere Aussage kann dann zur Steuerung der Bildkomprimierung dienen. Zur Implementierung der Bildkomprimierung sollte ein einfaches Fileformat entworfen werden, das z.B. auf Haar-Wavelets basieren könnte. Der Qualitätsfortschritt gegenüber dem Stand der Dinge sollte in einem psychovisuellen Vergleichstest nachgewiesen werden.

Literatur

- Michael Stelmach and Don Williams, *When Good Scannings Goes Bad: A Case for Enabling Statistical Process Control in Image Digitizing Workflows*, Proc. IS&T 2006 Archiving Conference, Ottawa, Canada.
- E. Reinhard, E. Khan, A. Akyüz and G. Johnson, «*Color Imaging*», AK Press, «978-1-56881-344-8».
- Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Third Edition, Pearson Education, «978-0-13-505267-9».

OCR für PDF/A und Massendigitalisierungen

- ❖ vier verschiedene Themen, einzel oder gemeinsam zu bearbeiten
- ❖ gegeben ist ein gescanntes Bild
 1. Layout-Rekonstruktion (mit Industriepartner)
 - Identifizierung von Textteilen, Bildern, Logos, ...
 2. Aufbau eines einfachen OCR-Systems
 - Zeichenerkennung der Textteile
 3. Fontgenerierung on-the-fly (mit Industriepartner)
 - Erkennung bzw. Rekonstruktion der Schriften
 4. Semantische Rechtschreibkontrolle (Uni München, Uni Zürich)
 - nachträgliche Verbesserung der OCR

3

klaus simon farbe im digitalen publizieren farbmetrik

- Peter G. Engeldrum, *Psychometric Scaling*, IMCOTEK Press, «0-9678706-0-7».

9.3 OCR für PDF/A und Massendigitalisierungen

Die folgenden Themen können je nach Ausgestaltung als Semester-, Bachelor- oder Masterarbeit bearbeitet werden. Auch können sie einzeln oder parallel von mehreren Studenten ausgeführt werden.

Das Generalthema ist die Interpretation einer gescannten Textseite, wie sie bei Massendigitalisierungen bzw. deren Speicherung als PDF/A-Dokument anfällt. Im Gegensatz zu den Anfangszeiten der OCR in den 80er Jahren geht es nicht nur um die Erkennung von Schreibmaschinenmanuscripten, sondern um die Erkennung bzw. die Rekonstruktion des gesamten Layouts.

1. Der erste Schritt besteht in der Erkennung der einzelnen Layoutelemente der gescannten Textseite. Es muss zwischen Überschriften, Textteilen, Logos, Graphiken, Bildern usw. unterschieden werden. Hier handelt es sich um ein klassisches Problem der Bildanalyse.
2. Im zweiten Teil sollte ein einfaches OCR-System realisiert werden. Dabei sollte spezifisch auf die Umstände im Bibliothekswesen eingegangen werden, d.h. es sollte nicht nur der zu Grunde liegende Ascii-Kode rekonstruiert werden, sondern auch der entsprechende Font, z.B. Fraktur. Andererseits sollte die Qualität des OCR-Resultats als Funktion der Scan-Qualität analysiert werden.
3. Die dritte Teilaufgabe bezieht sich auf die Speicherung des OCR-Resultats in einem PDF/A-Dokument. Ist ein in Teil 2 identifizierter Font dem System nicht bekannt, d.h. er liegt nicht als Schriftdatensatz (Type 1, true type, OpenType) vor, dann muss eine entsprechende Fontbeschreibung on-the-fly, z.B. in true type, generiert werden. Hier kann man auf die Erfahrungen des Fotosatzes zwischen 1970–90 zurückgreifen, wo

die klassischen Bleisatzfonts eingescannt und in digitale Formate überführt wurden.

4. Im Allgemeinen sind die Ergebnisse einer OCR-Analyse nicht perfekt. Besonders bei älteren Texten haben sich Korrekturverfahren aus dem Bereich der Computerlinguistik bzw. der Wahrnehmungspsychologie bewährt. In dieser Arbeit soll zunächst eine Übersicht über die verschiedenen Ansätze erstellt werden. Danach sollen etwa 3–4 Konzepte ausgewählt, implementiert und auf ihre Wirksamkeit im Kontext des Kulturerbes hin untersucht werden.

Literatur

- M. Cheriet, N. Kharama, C. Liu and C. Suen, *Character Recognition Systems*, Wiley, «978-0-471-41570-1».
- Peter Karow, *Digital Typefaces*, Springer, «3-540-56509-4».
- Yannis Haralambous, *Fonts & Encodings*, O'Reilly, «978-0-596-10242-5».
- K.-U. Carstensen, CH. Ebert, C. Ebert, S. Jekat, R. Klabunde und H. Lange (Hrsg.), *Computerlinguistik und Sprachtechnologie*, Spektrum, «978-3-8274-2023-7».
- E. Reinhard, E. Khan, A. Akyüz and G. Johnson, «*Color Imaging*», AK Press, «978-1-56881-344-8».
- Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Third Edition, Pearson Education, «978-0-13-505267-9».
- Peter G. Engeldrum, *Psychometric Scaling*, IMCOTEK Press, «0-9678706-0-7».