

## Teil 7: Konsequenzen der Digitalisierung



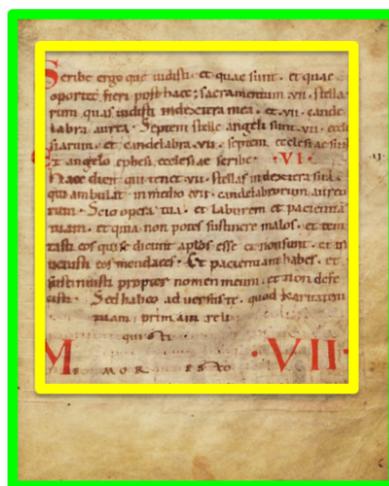
Abb. 16 Handschriften, virtuell ‚entblättert‘

Durch die weitgehende Digitalisierung historischer Dokumente tun sich neue Möglichkeiten für die Forschung auf. Ohne die Digitalisierung wäre es zum Beispiel nicht möglich, eine Komplettübersicht über alle Seiten eines Kodex zu erhalten. Mit der aufgeschlagenen Originalhandschrift könnte dieser Effekt nur dann erzielt werden, wenn man die einzelnen Blätter aus ihrer Bindung lösen würde (siehe Abb. 16).

Um zum Beispiel herauszufinden, ob das Layout eines Buchbestandes einem bestimmten Muster folgte, lassen sich an den Digitalisaten Verfahren der Bildverarbeitung anwenden. Diese Verfahren erlauben es, einfache Layoutmerkmale auf vielen gescannten Handschriftenseiten zu extrahieren. Im Projekt „eCodicology“ wird ausgetestet, wie sowohl die Bilddigitalisate als auch die Metadaten aus dem „Virtuellen Skriptorium St. Matthias“ wissenschaftlich ausgewertet werden können. Hierbei kommen auch Verfahren der Bildverarbeitung zum Einsatz. Diese Methoden liefern einen Eindruck davon, auf welche Weise mit den digitalisierten Beständen weiter gearbeitet werden kann.

Um zum Beispiel herauszufinden, ob das Layout eines Buchbestandes einem bestimmten Muster folgte, lassen sich an den

Das Projekt bezieht seine Ausgangsdaten aus dem mittelalterlichen Bibliotheksbestand der Benediktinerabtei St. Matthias in Trier, der im „Virtuellen Skriptorium St. Matthias“ (insg. ca. 170.000 digitalisierte Seiten) vorliegt. Es sollen Algorithmen entwickelt werden, die prägnante Gestaltungsmerkmale auf den gescannten Seiten automatisiert erkennen und extrahieren können. Für die Untersuchung wurden zunächst grundlegende Parameter identifiziert: *Seitengröße*, *Schriftraum*, *Bildraum* (siehe Abb. 17); weitere Merkmale sind der freigelassene Raum sowie die Spalten- und Zeilenzahl. Für diese Merkmale soll ihre Ausdehnung (Höhe x Breite), Position (Koordinaten) und Anzahl (Spalten, Zeilen) auf jeder Seite in einem zuvor erarbeiteten Metadatenschema verzeichnet werden.



Legend:  page size,  written size,  pictorial space

Abb. 17 Layoutmerkmale auf Kodexseiten aus St. Matthias

Damit der Computer bestimmte Bildmerkmale erkennen kann, müssen die Scans in mehreren Schritten „vorbehandelt“ werden. Am Anfang steht die Farb- und Größenkalibrierung der Bilder. Durch diese Bearbeitungsschritte werden Ungenauigkeiten in Auflösung und Farbechtheit, beispielsweise durch die Verwendung unterschiedlicher Scanner, ausgeglichen. Das ‚Bildrauschen‘ wird durch verschiedene Filter minimiert oder sogar ganz beseitigt.

Um die Verarbeitungszeit zu reduzieren, werden die Bilder in geringeren Auflösungen gespeichert. Die Segmentierung ist schließlich der erste Schritt der Bildanalyse und der letzte Schritt vor der eigentlichen Merkmalsextraktion. Bei der Segmentierung wird das Bild der Handschriftenseite in zusammenhängende Bereiche unterteilt, indem anhand seiner Pixel bestimmte Areale vom Hintergrund abgetrennt (segmentiert) werden. Diese Areale können anschließend extrahiert und vermessen werden.

Die Messdaten aus der Extraktion der Layoutbestandteile können anschließend automatisiert in bereits bestehende Metadaten einfließen und diese sinnvoll ergänzen. Im Projekt „eCodicology“ stammen die bereits bestehenden Metadaten aus den Beschreibungen in den älteren Handschriftenkatalogen und sind als XML in der Datenbank des „Virtuellen Skriptoriums St. Matthias“ hinterlegt. Um mit allen Metadaten arbeiten zu können (d.h. den bestehenden beschreibenden Daten und den neu hinzugewonnenen Layoutdaten), ist ein Webportal sinnvoll, über das bestimmte Suchanfragen an den Bestand gerichtet werden können. So wäre es zum Beispiel denkbar, einen, mehrere oder alle Kodizes nach einem bestimmten Gestaltungsmerkmal zu durchsuchen. Eine statistische Auswertung aller festgehaltenen Werte kann Regelmäßigkeiten oder Veränderungen innerhalb des Bestandes von St. Matthias aufzeigen. Auf diese Weise können zum Beispiel zeitliche Entwicklungen in der Seitengestaltung ermittelt werden.

Nähere Informationen zum Projekt entnehmen Sie der zugehörigen Website

[www.ecodicology.org](http://www.ecodicology.org)