

## Automatische Mustererkennung und Historische Handschriftenanalyse

**Veranstalter:** Friedrich-Alexander Universität Erlangen-Nürnberg; Ludwig-Maximilians-Universität München; SAOT

**Datum, Ort:** 14.06.2013-15.06.2013,

**Bericht von:** Viktoria Trenkle, Erlangen

Mit den Grußworten von Joachim Hornegger, Klaus Herbers (beide Erlangen) und Irmgard Fees (München) wurde die Tagung eröffnet. VINCENT CHRISTLEIN (Erlangen) stellte anschließend das Projekt *Schrift und Zeichen. Computergestützte Analyse von hochmittelalterlichen Papsturkunden. Ein Schlüssel zur Kulturgeschichte Europas* vor. Dabei erläuterte er die Hauptanliegen dieses interdisziplinären Projektes – die automatisierte Schriftanalyse und Schreiberidentifizierung anhand eines der wichtigsten Quellenkorpora des Mittelalters, den Papsturkunden. Ziel des Projektes sei die detaillierte Untersuchung der Schriftentwicklung in den Papsturkunden des 11. und 12. Jahrhunderts, deren kulturwissenschaftliche Einbindung neue Erkenntnisse zum päpstlichen Kanzleiwesen ermöglichen. Dazu sollen Werkzeuge entwickelt werden, die sowohl die Transkription erleichtern, als auch die Schrift beschreiben und wenn möglich einem Schreiber zuweisen sollen. Teile der Tools sollen anschließend an die bereits bestehende Datenbank der Regesta Pontificorum Romanorum online angebunden werden. Diese Datenbank, die im Herbst 2013 online gehen wird, stellte THORSTEN SCHLAUWITZ (Erlangen) in seinem Vortrag ausführlich vor.

In der ersten Sektion standen Ansätze der traditionellen Paläographie im Mittelpunkt. Eröffnet wurde sie mit dem Vortrag von MARTIN WAGENDORFER (München), der mittels der Methoden der traditionellen Paläographie die Möglichkeiten der Schreiberidentifizierung aufzeigte. Die Betrachtung von Sonderzeichen, Ligaturen und Kürzungen sei besonders vielversprechend, wohingegen die Betrachtung von Buchstaben nicht immer zielführend sei.

GUDRUN BROMM (Mannheim) widmete sich in ihrem Vortrag auch der Schreiberidentifikation. Dies geschah unter dem

Blickwinkel der forensischen Schriftanalyse, welche vor allem in der Rechtsfindung bei der Prüfung der Echtheit, der Identifizierung eines Schrifturhebers sowie den Entstehungsbedingungen von großer Bedeutung sei. Frau Bromm erläuterte die für den Begutachtungsprozess verwendeten Methoden, die physikalisch-technische, die paläographische und die psychologische Schriftanalyse.

Den thematischen Schwerpunkt der zweiten Sektion bildete die Schreiberidentifizierung. Den Anfang machte dabei erneut ein Vortrag aus der mittelalterlichen Geschichtsforschung. WERNER MALECZEK (Wien) referierte über die Authentizität der Kardinalsunterschriften auf päpstlichen Privilegien des 12. und 13. Jahrhunderts. Gerade eigenhändige Unterschriften ermöglichen dem Historiker, einzelne Persönlichkeiten über einen längeren Zeitraum zu verfolgen. Die Veränderungen, die sich auch in diesen Unterschriften widerspiegeln, wurden an konkreten Beispielen gezeigt.

Anhand der Lindbergh-Entführung und Schriften von Herbert Melville stellte SARGUR SRIHARI (Büfallo) die von der Gemeinschaft für Mustererkennung und künstliche Intelligenz entwickelten automatischen Werkzeuge zur forensischen Dokumentenanalyse vor. Gleichsam stelle dieses Dokument aber auch eine Besonderheit dar, da - wie es eine neue Hypothese vermuten lässt - wohl von mehreren Schreibern verfasst wurde.

Am Ende dieser Sektion präsentierte LAMBERT SCHOMAKER (Groningen) einen Überblick über die Forschungen, welche zur Entstehung eines Programmes zur Schreiberidentifizierung, das unter dem Namen GIWIS bekannt ist, geführt haben. Er stellte dabei sowohl die Fortschritte der Computertechnik bei der Identifizierung von Schreiberhänden und Abfrage von Schlüsselwörtern vor und hob die Bedeutung der Stifthaltung hervor, verwies aber auch auf die methodische Problematik, dass Aussagen von computerbasierten Systemen zur Schreiberidentifizierung probabilistisch und nicht deterministisch sind.

Möglichkeiten der Digitalen Paläographie standen im Zentrum der dritten Sektion, die THOMAS FRENZ (Passau) mit einem Vortrag über die Entwicklung der Handschrift an der päpstlichen Kurie im 15. Jahrhundert eröffne-

---

te. Gerade in diesem Jahrhundert gelangte neben der gotischen Schrift auch die humanistische Schrift an die päpstliche Kurie, den größten Urkunden- und Aktenproduzenten dieser Zeit. Der Einfluss der humanistischen Schrift sei über die Sekretäre an die Kurie gelangt und fände sich hauptsächlich in den Kanzleiregistern. Andere Behörden hingegen seien gänzlich dem gotischen Schreibstil treu geblieben, wie zum Beispiel die Rota Sacra Romana, was Herr Frenz auf die dort tätigen deutschen Juristen zurückführte.

PETER STOKES (London) betrachtete anschließend das gesamte Bild der historischen Dokumentenforschung. Ausgehend von der Frage nach Forschungsfeldern, die am offensten für die automatische Analyse sind, stellte Stokes das Potenzial der Automatischen Mustererkennung für die historische Dokumentenanalyse vor. Er übte aber gleichzeitig Kritik an automatischen Methoden, da diese oft schwer verifizierbar und interpretierbar seien und somit „historical truth“ auch mit diesen Methoden allein schwierig zu erlangen sei.

Mit den Ergebnissen einer automatischen Kategorisierung von französischen Manuskripten beschäftigte sich DOMINIQUE STUTZMANN (Paris). Er stellte dabei heraus, dass es aufgrund der Uneinigkeit der Paläographen keine Eindeutigkeit und Endgültigkeit der Ergebnisgenauigkeit geben könne. Des Weiteren stellte er fest, dass die Bedeutung der Ergebnisse von der Fragestellung abhänge, man aber auch von einer Evaluierung neue Erkenntnisse erlangen könne, was anhand von Buchstabenclustern, die durch eine Hauptkomponentenanalyse (PCA) erlangt worden war, dargestellt wurde.

Abschließend stellte JOSEP LLADÓS (Barcelona) sein Projekt über Wordspotting in den Eheregistern der Kathedrale von Barcelona vor. Worspotting beschreibe dabei den Prozess, alle Instanzen eines Anfrageschlüsselwortes einer digitalen Dokumentenbibliothek zu erhalten. Herr Lladós verwies in seinem Vortrag auch auf die Schwierigkeit, bei der Abfrage große Deskriptoren zu realisieren.

Die vierte Sektion, welche sich mit der Dokumentenanalyse beschäftigte, wurde mit einem Vortrag von OTFRIED KRAFFT (Marburg) über die äußere Gestaltung von Urkunden ab Urban II. begonnen. Nach ei-

ner Vorstellung des Layouts der hochmittelalterlichen Papsturkunden, ging Herr Krafft auf die normierenden Vorstellungen, die bei der Urkundenproduktion eine wichtige Rolle spielte, ein. Dabei spiegeln Veränderungen in den Urkunden nicht immer die normierenden Vorstellungen eines institutionellen Rahmens wieder, sondern können auch von einem individuellen Schreiber abhängig sein.

Anschließend stellte ROBERT SABLATNIG (Wien) die Möglichkeiten der spektralen Bildverarbeitungsmethoden vor. Dabei komme die Verbesserung der Lesbarkeit durch die Kombination von Orts- und Spektralinformationen der multivarianten Bilddaten zustande. In seiner Präsentation zeigte Sablatnig, wie durch die Anwendung der Fisher Linear Discriminate Analysis die Dimension des multispektralen Scans reduziert und die beschädigte Schrift besser lesbar wurde.

Zu Beginn der abschließenden Sektion stellten BJÖRN GOTTFRIED (Bremen) und MATTHIAS LAWOW (Berlin) das Projekt Diptychon vor. Dieses unterstütze den Editor bei der Transkription und dem Vergleich mittelalterlicher Handschriften. Dieses standardisierte Verfahren erleichtere durch den automatischen Vorschlag zur Segmentierung von Zeilen und Glyphen den Analyseprozess. Bislang wurde das Programm an zwei mittelalterlichen Schriften des Hugo von Flavigny und des Kanzlers von Kaiser Karl IV. mit vielversprechenden Ergebnissen getestet.

Dem Thema der Handschriftenerkennung widmete sich ebenso R. MANMATHA (Massachusetts). Dabei stellte er zunächst einen Ansatz vor, der Texterkennung und eine anschließende Textsuche kombiniere. Anhand der Georg Washington Manuskripte erklärte Herr Manmatha die Probleme bei der Erstellung eines Handschriftenabfragesystems, wie die Segmentierung, Vorverarbeitung und das Erstellen eines statistischen Modells zur Abfrage.

Die Werkzeuge zur Untersuchung von handgeschriebenen historischen Manuskripten stellte LIOR WOLF (Tel Aviv) in den Mittelpunkt seiner Überlegungen. Dabei präsentierte er jüngste Forschungsergebnisse, die er hauptsächlich an den Genizah-Manuskripten entwickelt hat, sowie Zukunftspläne. Neben einem Verbindungs- und einem Cluste-

ringwerkzeug, welches verbundene Paare zu Mehrseiten-Verbunden zusammensetzt, wurde auch ein Paläographie- und ein Übersetzungswerkzeug vorgestellt.

Die Abschlussdiskussion, die von KURT GÄRTNER (Trier) eröffnet und geleitet wurde, griff noch einmal die Schwerpunkte der Tagung auf. Erst der direkte Dialog und die enge Zusammenarbeit der einzelnen Teildisziplinen ermöglichten ein erfolgreiches Gelingen dieser Projekte und die Entwicklung neuer Ideen, wie mehrfach unterstrichen wurde. Dazu hat die Tagung einen wichtigen Beitrag geleistet.

### **Konferenzübersicht**

Grußworte von Joachim Hornegger, Klaus Herbers und Irmgard Fees

Vincent Christlein: Script and Signs. A Computer-based Analysis of High Medieval Papal Charters. A Key to Europe's Cultural History

Thorsten Schlawitz: Presentation of the Papal Documents Database

#### *Section I. Traditional Palaeography*

Martin Wagendorfer: Considerations of the Identification of Scribes: Aims and Methods of Traditional Palaeography

Gudrun Bromm: Forensic Handwriting Analysis

#### *Section II. Writer Identification*

Werner Maleczek: The Right Hand of the Pope: on the Authenticity of the Cardinal Signatures on Papal Privileges from the 12th and 13th Centuries

Sargur Srihari: Role of Automation in the Examination of Handwritten Items: the Lindbergh Case

Lambert Schomaker: The Necessity of Simultaneous Multiple Perspectives in Digital Identification of the Hand

#### *Section III. Digital Palaeography*

Thomas Frenz: The Evolution of Handwriting in the Papal Curia of the 15th Century (lecture in German)

Peter Stokes: In Meaning versus Mining, and

Putting the Palaeographer in Charge

Dominique Stutzmann: Image Analysis and Clustering of Medieval Scripts: an Evaluation Protocol

Josep Lladós: Handwritten Word Spotting in Historical Documents: the Project Five Centuries of Marriages

#### *Section IV. General Document Analysis*

Otfried Krafft: Layout of the page and scribe identification in papal privileges

Robert Sablatnik: Multispectral Image Acquisition and Analysis for Manuscript Research

*Section V. Automatic Handwriting Recognition and Analysis*

Björn Gottfried, Matthias Lawo: Diptychon: a Transcription Assistant System for the Separation of Glyphs in Medieval Handwritings

R. Manmatha: Searching Handwritten Manuscripts

Lior Wolf: Automatic Tools for Historical Manuscript Analysis

Kurt Gärtner: Concluding Discussion

Tagungsbericht *Automatische Mustererkennung und Historische Handschriftenanalyse*. 14.06.2013-15.06.2013, , in: H-Soz-Kult 07.09.2013.