Martin Lettner – Markus Diem – Robert Sablatnig – Heinz Miklas

Digital Image Analysis for the Investigation of Ancient Manuscripts

Abstract: This paper presents the preliminary results of a collaboration of philologists and computer scientists devoted to the recording, investigation and editing of two medieval Slavonic manuscripts. The goal of the project is the development of techniques and tools for the recording, restoration and analysis of such sources in order to support philological studies by automatically deriving the description and restoration of the relevant scripts. The methodology of the image analysis aspect of the project is as follows. A multi-spectral acquisition of the manuscripts will provide the basis for improving the readability of the texts. The subsequent registration process aligns one image to the other. Further steps include the extraction of "hidden information" (e.g. overwritten text on parchment) of multi-spectral images, image enhancement, and a computer-aided script description and reconstruction.

Introduction

The study of handwritten sources covers a wide field reaching from the examination of the physical body up to the text and its contents, composition and condition. Until the 1990s, this was mainly the domain of the humanities. Technical scientists were engaged predominantly in the recording and conservation of valuable objects. During recent years, however, interdisciplinary work has gained ground, no longer concentrating merely on a few special tasks, like the development of Optical Character Recognition (OCR) software, but comprising a growing amount of relevant items: the description of manuscripts, the Internet publication of texts, the imaging and restoring of watermarks, palimpsests[1] and other latent texts (e.g. Easton / Knox / Christens-Barry 2003), or a thorough description of writing systems and the identification of individual handwriting. Without overestimating the progress made so far, we can speak of a new era in the study of traditional written sources. And it may be expected that in the long run the decipherment, study and editing of such sources will be:

- achieved predominantly based on facsimile images; a method that relieves the originals and makes their investigation independent of the place of preservation,
- enabled or improved by special recording methods like multi-spectral or radiographic imaging,
- more exhaustive, precise and less time consuming through automatic image analysis, and
- executed with a set of tools obtainable in every computer shop and applicable by anyone interested in the matter with little training.

In this paper we first give a survey of recent developments in image analysis for historical handwritten documents. Furthermore we present an overview of an interdisciplinary project where philologists and computer scientists will collaborate to enable a multifold progress in their fields both in time and substance. Finally we will show some preliminary results of a multi-spectral image acquisition system of historical manuscripts and a method for a fully automatic image registration. The registration process is necessary in order to align one spectral image to the other. Until now the images of the different spectral ranges had to be registered manually. Our project is devoted to the recording, investigation and editing of two medieval Slavonic manuscripts of extraordinary importance. The objects to be edited are two Glagolitic[2] manuscripts with Cyrillic and Greek additions of the classical Old Church Slavonic (OCS) corpus, belonging to the new findings made in 1975 at St. Catherine's monastery on Mt. Sinai: *Euchologii Sinaitici pars nova* and "*Missale*" *(Sacramentarium)*

[1] A palimpsest is a manuscript page (particularly parchment) that has been written on, erased and used again.
[2] The Glagolitic alphabet is the oldest known Slavic script.

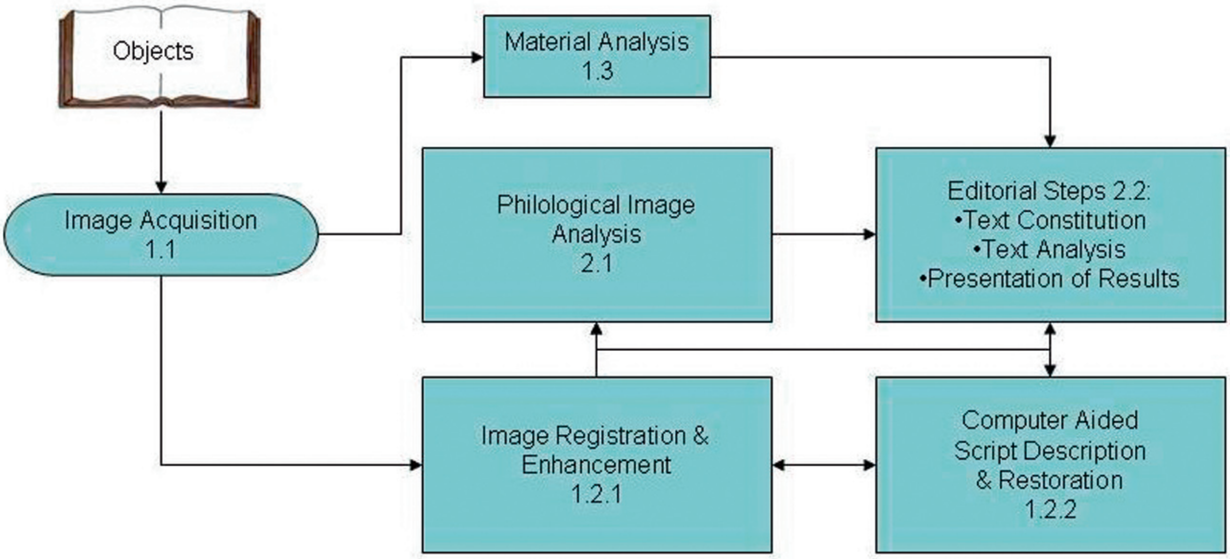

Fig. 1. Flow chart of the project subtasks.

Sinaiticum (Miklas 2003). The project is subdivided into the following subtasks:

1. Multi-spectral image acquisition
2. Image registration and image enhancement
3. Technical image analysis consisting of a computer aided script description and the reconstruction of broken character links
4. Material analysis
5. Philological image analysis and editorial steps.

First, a multi-spectral acquisition of the manuscripts shall provide the basis for improving the readability of the texts. Image registration, image enhancement, and computer-aided script description and restoration will then foster the philological work. A non-destructive material analysis will enable the definition of parchment, inks and pigments. The philological tasks include the deciphering, and the paleographic and graphemic study of the written material on the one hand, and the text constitution and text comparison, commentaries, a glossary and an introduction on the other. The most challenging part of the project from the standpoint of image analysis lies in the description and reconstruction of the relevant scripts. Yet it has to be stressed that the algorithms to be developed shall only enable the philologists to perform their tasks better and faster. *Fig. 1* shows the flow chart of the project subtasks and the collaboration between them.

This paper is organized as follows: The section that follows shows related work and gives an overview in the fields of image acquisition and technical image analysis for historical documents (manuscripts). The next section presents the methodology of the image analysis part within this project. The following sections cover the image acquisition and registration in more detail and show preliminary results. The last section gives a conclusion and an outlook.

Related Work

There have already been quite a few efforts in image analysis of historical documents. In general, differences between image analysis for ancient versus modern documents accrue particularly from the aging process of the documents, the fading out of ink, or other environmental influences which complicate the analysis. Some major and several related studies in image analysis of historical documents will be covered in this section.

Multi- and hyper-spectral imaging has been used in a wide range of scientific and industrial fields including space exploration like remote sensing for environmental mapping, geological search, medical diagnosis or food quality evaluation. Recently, the technique has been applied more and more frequently in the investigatation of old manuscripts. Two prominent representatives are the Archimedes Palimpsest (Easton / Knox / Christens-Barry 2003) and Tischendorf's Codex Sinaiticus. Easton et al. were the first to capture and enhance the erased writing of the famous Archimedes palimpsest by multi-spectral methods. In the project, it turned out that the adoption of spectral imaging produces higher and better readability of the texts than conventional thresholding methods. Balas et al. developed a computer controllable hyper-spectral imaging apparatus, capable of acquiring spectral images of 5 nm bandwidth and with 3 nm tuning steps in the spectral range of 380–1000 nm (Balas et al. 2003). Very good results in the enhancement of spectral images of palimpsests and other latent texts have also been achieved by the Italian company Fotoscientifica Re.co.rd, which for instance provided the pictures for the EC project Rinascimento Virtuale, devoted to the decipherment of Greek palimpsest manuscripts (Harlfinger 2002). The EC project IsyReaDeT developed a system for a virtual restoration and archiving of damaged manuscripts, using a multi-spectral imaging camera, advanced image enhancement and document management software (Tonazzini / Vezzosi / Bedini 2003).

For the enhancement of the readability Easton et al. used an unconstrained least squares algorithm for spectral unmixing and produced normalized and non-negative fraction maps of text and parchment. The combination of these fraction maps can be used to highlight different classes, such as the underlying text and the overwriting. Rapantzikos and Balas separated the overwritten from the underwritten text-layer by means of principal component analysis (PCA) and a linear spectral mixture analysis (Rapantzikos / Balas 2005). Promising results are also obtained by Tonazzini et al. who applied an independent component analysis (ICA) to the spectral components at different bands to separate bleed through and show-through texts as well as palimpsests (Tonazzini / Vezzosi / Bedini 2003).

Methodology

From a pattern recognition point of view, the problem can be subdivided into two major goals: image registration and enhancement on the one side, and the computer-aided script description and reconstruction on the other. When an image is to be utilized, it is necessary to make corrections in brightness and geometry if the accuracy of interpretation is not to be prejudiced. Geometrical and radiometrical errors (i.e. vignetting, chromatic aberration and non linear geometric distortion) have to be detected (e.g. with test patterns) and corrected to achieve the desired accuracy (Richards / Jia 1999). The following registration process aligns one image to the other by establishing a coordinate transformation that relates the pixel coordinates from one image to another. Then the resulting images will be enhanced with various image processing techniques to further increase the readability of the text and to prepare it for the following computer aided script description. To improve the legibility of the texts several methods, like contrast enhancement, independent component analysis, filters for background removal or homomorphic filtering (Pan et al. 2004) are applicable. We will combine and improve these methods also in order to prepare the text images for the computer aided description (analysis) and reconstruction of individual scripts. For specific tasks like ruling, line structure, and layout analysis, 2D projection techniques used in standard OCR will be reused, whereas primitive analysis tools must be developed and adapted. The following points summarize the specific goals of the image analysis part of the project:

1. High resolution, multi-spectral image acquisition
2. Geometrical and radiometrical corrections
3. Image registration and image enhancement to increase the readability of latent texts
4. Incorporation of OCR knowledge and philological parameters for the description of scripts and writing systems:
   - Layout analysis (shape and size of page layout)
   - Ruling analysis (position, number and angle of lines with respect to borders)
   - Line structure analysis (number of primitives per line)
   - Automatic extraction of primitives for statistical evaluation
   - Primitive analysis and classification
   - Representation of character primitives for database applications
   - Scribe analysis on the basis of variations of primitives in different parts
   - Reconstruction of missing links of primitives (characters)in enhanced images.

Image Acquisition

The first step enabling the technical and philological image analysis is the digitization of the manuscripts. Since photographic techniques in the visible range (film, digital camera) have proven to be insufficient with the objects given, spectral imaging

Fig. 2. Acquisition system.

will have to be applied. Like in the field of Art Research and Conservation, multi- and hyper-spectral imaging techniques have become a powerful tool in the scientific analysis and documentation of old manuscripts with "latent" (degraded, disintegrating and overwritten) texts, since images in different wavelengths provide information that the human eye cannot see. Applied in the spectral range from ultra-violet (UV), visible light range (VIS) up to the infrared (IR) range, these techniques combine conventional imaging and spectrometry to acquire both spatial and spectral information from an object. They produce three-dimensional images or spectral image cubes where the third dimension contains spectral information for each pixel. According to the number of spectral channels we distinguish between multi-spectral for fewer channels and hyper spectral for a large number of spectral channels.

For the acquisition of the manuscripts we use a Hamamatsu C9300-124 camera system, furnished with additional equipment. The C9300-124 is a high resolution camera with 10 Mega pixels using a high-speed interline transfer CCD chip. Its spectral response lies between 330 and 1000 nm. To obtain multi-spectral data a set of optical filters is used to select the best specific ranges from the spectrum. Easton et al. used four different spectral bands to improve the readability of the Archimedes palimpsest. A lighting system provides the required illumination (UV, VIS and IR) for the multi-spectral images, and filters fixed on the lens of the camera select specific spectral ranges. *Fig. 2* shows the acquisition system.

It is known that especially in the UV range additional information appears which is invisible to the naked eye (Mairinger 2003). The near UV band (400–320 nm) excites visible fluorescent light in conjunction with specific inorganic and organic substances. For our preliminary studies, we have recorded historical manuscripts with UV fluorescence and UV reflectography, showing palimpsests of ancient handwritings:

UV reflectography to visualize retouching, damages and changes through luminescence, for example. Therefore the visible range of light has to be excluded in order to concentrate on the long wave UV light. This is only achievable with difficult technical effort.

UV fluorescence shows only changes in the upper script (paint) layer. In principle, taking the visible fluorescence of objects is possible with every camera.

Thus we have two different images of one and the same manuscript: the UV reflectography image and the UV fluorescence image. These are only preliminary test images (in order to develop the registration process) but we are going to expand the number of the spectral images with a couple of spectral filters.

Image Registration

Following the acquisition of the manuscript images, the images have to be registered. Image registration is a fundamental task in image processing used to match two or more pictures taken under different conditions. For instance, during the acquisition of images in different bands, the objects or the camera may be affected as a result of filter changes. Until now, the images had to be registered manually using commercial image processing software. Image registration is the process of estimating the ideal transformation between two different images of the same scene taken at different times and/or different viewpoints. It geometrically aligns images so that they can be overlaid. There is a wide variety of different methods (especially in remote sensing and medical imaging applications). An overview of image registration methods is given elsewhere (Zitova / Flusser 2003). The image registration process proposed in this paper and in the majority of registration methods can be separated into four steps:

- Feature detection: Features are detected automatically in the so-called reference image. For further processing, these features are represented by a point which is placed in their center and denominated as control point.
- Feature matching: The correspondence between the feature of the reference image and an area in the sensed image is established. The similarity measurement is calculated by the normalized cross correlation.
- Transform model estimation: The parameters of the mapping function, aligning the sensed image with the reference image, are estimated using the previously computed corresponding control points.
- Image resampling and transformation: The sensed image is transformed by means of the local weighted mean transformation which is a local mapping function. Non-integer coordinates are interpolated.

Feature Extraction, Matching and Transformation

Initially the input images are prepared for registration by means of low level image processing. These steps include homomorphic filtering (Pan et al. 2004), non-linear diffusion (Mrazek 2001) and thresholding in order to detect regions (e.g. characters) that allow the best possible cross correlation. The homomorphic filtering and the non-linear diffusion are intended to reduce noise so that the cross correlation compares characters exclusively. The binary image is used to find the "best" regions for the cross correlation. Characters, as well as punctuations and so on, recur within lines and paragraphs. If their visual information is sparse (e.g. an "i" or a point in the Latin font), the mistake rate of the cross correlation will be high.

The cross correlation calculates the difference of two windows by means of a modified Euclidean distance. The so-called template image is an image detail of the reference image. Its image size corresponds to the bounding box of the feature found afore. The template image is shifted over the entire detail of the sensed image. For each shift the correlation between the template and the display window (the detail in the sensed image) is computed. The resultant function indicates the strongest correspondence of the template image and the sensed image by the absolute maximum. Hence the control point of the sensed image is seated at the coordinates of the absolute maximum. The cross correlation is variant to changes in the image amplitude caused, for example, by changing lighting conditions. Consequently, the correlation coefficient normalizing the template as well as the display window of the sensed image is computed.

Having determined the control points, the parameters of the mapping function are computed. Images which possess only global distortions (e.g. rotation) may be registered with a global mapping function. The global rigid, affine and projective transformations are most frequently used (Likar / Pernuš 2001).

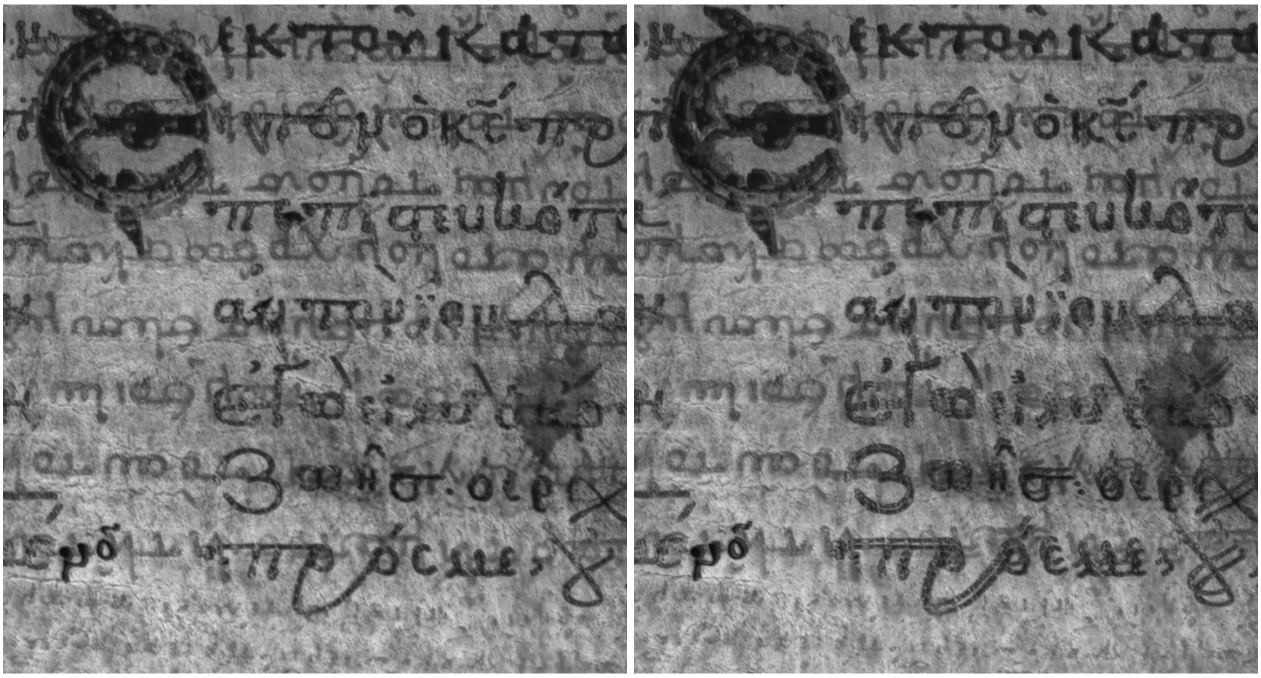

Fig. 3. UV fluorescence image registered to the UV reflectography image using an affine mapping function (left). Registration with the same control points using a local weighted mean transformation (right).

As a consequence of non-rigid distortions such as the changing lenses, illumination or curvature of a single page, the images have to be registered using a curved transformation. To overcome this problem a local mapping function is applied. The local weighted mean method (Ghostasby 1988) is a local sensitive interpolation method. It requires at least six control points which should be spread uniformly over the entire image. Polynoms are computed by means of the control points. Thus, the transformation of an arbitrary point is computed by the weighted mean of all passing polynomials. Besides this, a weighting function is defined which guarantees that solely polynomials near an arbitrary point influence its transformation.

Results

Resulting images of the image registration are presented in *Fig. 3*. The local weighted mean method was compared to a global interpolation method. As a consequence of the global mapping function the registered images correspond only to certain parts of the images. Hence, the farther the points are away from the corresponding area, the more they differ. That is why a local mapping function was used to compute the transformation.

Conclusion

This paper introduced an interdisciplinary project of computer scientists and philologists. The main goals of the project are the preparation of an edition of two medieval Slavonic manuscripts of extraordinary importance where the technical image analysis supports the philological studies. The methodology of the image processing part is as follows. Firstly, multi-spectral image acquisition will improve the readability of the manuscripts. Secondly, a registration process aligns the images of the different spectral bands. Preliminary results of the registration process are presented in this paper. We registered an UV reflectography image and an UV fluorescence image in order to avoid a manual registration which causes errors. The next steps of the project will cover the enhancement of the readability of the text. For this we will use PCA or comparable methods in order to find the most effective method for our purposes. Then page segmentation and the segmentation of the primitives (characters) will be exhaustively analyzed as a basis for the philological studies. The major goal of the project from the standpoint of image analysis lies in the description and reconstruction of the relevant scripts which might also be the most challenging one.

Acknowledgments

This work was supported by the Austrian Science Fund (FWF) under grant P19608-G12.

References

Balas et al. 2003
C. Balas / V. Papadakis / N. Papadakis / A. Papadakis / E. Vazgiouraki / G. Themelis, A novel hyper-spectral imaging apparatus for the non-destructive analysis of objects of artistic and historic value. Journal of Cultural Heritage 4,1, 2003, 330–337.

Easton / Knox / Christens-Barry 2003
R. L. Easton / K. T. Knox / W. A. Christens-Barry, Multispectral Imaging of the Archimedes Palimpsest. In: Proceedings of the 32nd Applied Imagery Pattern Recognition Workshop, Washington DC, USA, October 15–17, 2003 (Washington 2003) 111–118.

Ghostasby 1988
A. Ghostasby, Image registration by local approximation methods. Image and Vision Computing 6, 1988, 255–261.

Harlfinger 2002
D. Harlfinger, Rinascimento virtuale. Digitale Palimpsestforschung. Rediscovering written records of a hidden European cultural heritage. In: D. Harlfinger (ed.), Berichtband der Konferenz des Netzwerks Rinascimento virtuale zur digitalen Palimpsestforschung (Wien 2002) 28–29.

Likar / Pernuš 2001
B. Likar / F. Pernuš, A hierarchical approach to elastic registration based on mutual information. Image and Vision Computing 19, 2001, 33–34.

Mairinger 2003
F. Mairinger, Strahlenuntersuchung an Kunstwerken. Bücherei des Restaurators 7 (Berlin 2003).

Miklas 2003
H. Miklas, Die slavischen Schriften: Glagolica und Kyrillica. In: W. Seipel (ed.), Der Turmbau zu Babel: Ursprung und Vielfalt von Sprache und Schrift. Ausstellungskatalog des Kunsthistorischen Museums 3a (Wien 2003) 243–249.

Mrazek 2001
P. Mrazek, Nonlinear Diffusion for Image Filtering and Monotonicity Enhancement (Prague 2001).

Pan et al. 2004
X. Pan / M. Brady / A. K. Bowman / C. Crowther / R. S. O. Tomlin, Enhancement and feature extraction for images of incised and ink texts. Image and Vision Computing 22,6, 2004, 443–451.

Rapantzikos / Balas 2005
K. Rapantzikos / C. Balas, Hyperspectral imaging: potential in non-destructive analysis of palimpsests. In: Proceedings of the IEEE International Conference on Image Processing 2, Genoa, Italy, September 11–14, 2005 (Pisa 2005) 618–621.

Richards / Jia 1999
J. A. Richards / X. Jia, Remote Sensing Digital Image Analysis: An Introduction (Berlin 1999).

Tonazzini / Vezzosi / Bedini 2003
A. Tonazzini / S. Vezzosi / L. Bedini, Analysis and recognition of highly degraded printed characters. International Journal on Document Analysis and Recognition 6,4, 2003, 236–247.

Zitova / Flusser 2003
B. Zitova / J. Flusser, Image registration methods: a survey. Image and Vision Computing 21,11, 2003, 977–1000.

Martin Lettner
Markus Diem
Robert Sablatnig

Vienna University of Technology
Pattern Recognition and Image Processing Group
Favoritenstr. 9–11 // 183
1040 Vienna
Austria
lettner@prip.tuwien.ac.at

Heinz Miklas

University of Vienna
Institute for Slavonic Studies
Spitalgasse 2, Hof 3
1090 Vienna
Austria