16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

REGISTRATION AND ENHANCING OF MULTISPECTRAL MANUSCRIPT IMAGES

Martin Lettner[†], *Markus Diem*[†], *Robert Sablatnig*[†] *and Heinz Miklas*[‡]

[†] Pattern Recognition and Image Processing Group,
Institute for Computer Aided Automation,
Vienna University of Technology, 1040 Vienna, Austria
phone: + (43)-1-58801-18353, email: lettner@prip.tuwien.ac.at

[‡] Institute for Slavonic Studies,
University of Vienna, 1010 Vienna, Austria
phone: + (43)-1-4277-42810, email: heinz.miklas@univie.ac.at

ABSTRACT

Two medieval manuscripts are recorded, investigated and analyzed by philologists in collaboration with computer scientists. Due to mold, air humidity and water the parchment is partially damaged and consequently hard to read. In order to enhance the readability of the text, the manuscript pages are imaged in different spectral bands ranging from 360 to 1000*nm*. A registration process is necessary for further image processing methods which combine the information gained by the different spectral bands. Therefore, the images are coarsely aligned using rotationally invariant features and an affine transformation. Afterwards, the similarity of the different images is computed by means of the normalized cross correlation. Finally, the images are accurately mapped to each other by the local weighted mean transformation. The algorithms used for the registration and results in enhancing the texts using Multivariate Spatial Correlation are presented in this paper.

1. INTRODUCTION

Multi- and hyperspectral imaging have been used in scientific and industrial fields including space exploration, remote sensing, medical diagnosis or food quality evaluation [13]. Recently, the techniques have become a tool for scientific analysis and documentation of old manuscripts with latent (degraded, disintegrating and overwritten) texts [3]. The main advantage of the analysis in different wavelengths extending the visible light is the additional information that the human eye cannot see [12]. Especially in analyzing palimpsests (parchment which was written on after the first text had been erased) the analysis in different spectral ranges, covering especially ultra-violet light enhances the visibility of the underwritten text [9].

While major studies in multi- and hyperspectral imaging for historical documents focused on the visualization of the underwritten text in palimpsests (e.g. [3, 14]) our goal is to enhance the readability of damaged or degraded text written on parchment. The objects to be studied are two Glagolitic (this is the oldest known Slavic script) manuscripts with Cyrillic and Greek additions belonging to the new findings made in 1975 at St. Catherine's monastery on Mt. Sinai: *Euchologii Sinaitici pars nova* and *"Missale" (Sacramentarium) Sinaiticum* [10]. Since the manuscripts are in a deplorable state (e.g. non-uniform appearance of the writing and the background, blur of the background, fading out of ink, mold, water stains and humidity) conventional imaging techniques are not gainful. So we resort to multispectral image data to improve the readability of the damaged manuscripts. The goal is to produce enhanced gray level or pseudo colored images to support philological studies on the partially undecipherable manuscripts.

The Principal Component Analysis (PCA) is a common method to produce pseudo colored images where three components are combined to one RGB image [3]. In this study we use a combination of spectral and spatial information of the multivariate image data in order to optimize the production of pseudo colored images and therefore to enhance the readability of the degraded texts. The basis for the investigation is the Multivariate Spatial Correlation Matrix proposed by D. Wartenberg [18]. Warner [17] applied the method to remotely sensed data and showed the effectiveness of the method in contrast to PCA. Consequently, we use the Multivariate Spatial Correlation (MSC) to enhance the readability of the varying appearance of the texts. To show the benefit of the method we compare the results to conventional techniques like PCA or the decorrelation stretch [5].

The paper is organized as follows: The following section shows related work and gives an overview in the fields of image acquisition and technical image analysis for historical documents. Section 3 and 4 show the multispectral image acquisition system and the registration of the individual spectral bands. In Section 5 we introduce the proposed image enhancement methodology and in Section 6 we present experiments and results. Conclusions and an outlook in Section 7 finish the paper.

2. RELATED WORK

Two representatives using multi- or hyperspectral imaging methods for historical documents are the Archimedes palimpsest [3] or Tischendorfs Codex Sinaiticus.

Since the behavior of different inks is different in multispectral bands [9] the appearance of the text depends on the spectral range and the method used for enhancing the digital images. Easton et al. used an unconstrained least squares algorithm for spectral unmixing and produced normalized and non-negative fraction maps of text and parchment to enhance the readability of the palimpsests [3]. The combination of these fraction maps can be used to highlight different classes, e.g. the underlying text and the overwriting. Salerno et al. [14] assessed the performances of the PCA and the In-

THIS WORK WAS SUPPORTED BY THE AUSTRIAN SCIENCE FOUNDATION (FWF) UNDER GRANT P19608-G12.

dependent Component Analysis (ICA) on these images in order to enhance the underwritten text. Rapantzikos and Balas developed a hyperspectral imager to improve the readability of palimpsests [12]. They separated the overwritten from the underwritten text layer in palimpsests by means of PCA and a linear spectral mixture analysis [12].

An improvement of PCA for color enhancement in multispectral images was proposed by Mitsui et al. [11]. An image fusion approach to enhance the spatial quality of the multispectral images while preserving the spectral contents is proposed by Tseng et al. [16].

In contrast to previous studies which aim particularly at the enhancement of underwritten texts in palimpsests our preferences are in the enhancement of the readability from ancient manuscripts especially written on parchment, where the fading out of ink or mold are the most challenging problems. Furthermore the appearance of the text varies on each page where the ink is partially or completely faded out. Since ICA methods are based on the assumption of mutual independence of the sources [15] these methods are not applicable for our purpose. Therefore we focus on multispectral image data and the spectral signature as well as the spatial correlation in order to enhance the readability.

3. IMAGE ACQUISITION

Since photographic techniques in the visible range have proven to be insufficient with the objects given, spectral imaging was used for the acquisition. Applied in the spectral range from Ultra-Violet (UV), Visible light (VIS) and Near Infrared (NIR), these techniques combine conventional imaging and spectrometry to acquire both spatial and spectral information from an object. It is known that especially in the UV and NIR range additional information appears which is invisible for the naked eye [9].

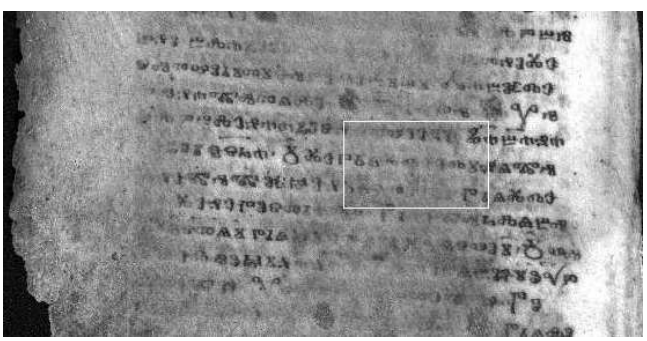

Figure 1: *"Missale" (Sacramentarium) Sinaiticum.*

For the acquisition of the manuscripts we use a Hamamatsu C9300-124 (spectral response: $330 - 1000nm$) gray scale camera and a Nikon D2Xs SLR camera. To obtain multispectral data a set of optical filters is used to select specific ranges from the spectrum. The filters are mounted within a filter wheel in front of the gray scale camera. Additionally to a conventional lighting system we use a low power UV light to gather UV reflectography and UV fluorescence images. A summary of the spectral bands can be seen in Table 1.

Figure 1 shows an example page from *"Missale" (Sacramentarium) Sinaiticum* where the text is degraded and hardly readable in the upper and lower part of the page. Selected spectral images within the white rectangle can be seen in Figure 2. The figure shows seven spectral bands as well as a color image (RGB) from the Nikon camera. The readability of miscellaneous characters differs in the spectral images. It can be seen that the contrast in image LP 800 as well as BP 780 is very low. The text in the second row can be deciphered solely in images SP 400, BP 450 and BP 550 and is not visible in the other images. SP 400 (UV reflectography) emphasizes the characters best but contains more noise than the other images due to the illumination with the low power UV light.

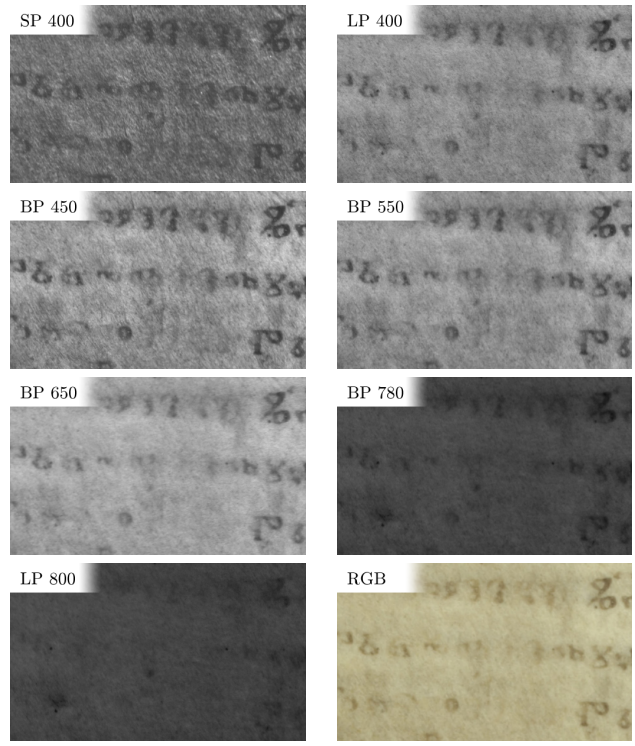

Figure 2: Multispectral images showing the marked detail from Figure 1.

Table 1: Description of the multispectral images containing the channel number (H for Hamamatsu and N for Nikon), the filter type and the methodology of the image acquisition. LP depicts a long pass filter, SP a short pass filter and BP is a band pass filter.

Channel	Filter type	Annotation
H1	SP 400	UV reflectography
H2	LP 400	VIS-IR
H3	BP 450	VIS-IR
H4	BP 550	VIS-IR
H5	BP 650	VIS-IR
H6	BP 780	VIS-IR
H7	LP 800	IR reflectography
H8	no filter	VIS-IR
N1	RGB	VIS-IR
N2	RGB	UV fluorescence

Since the manuscript pages are repositioned between the two cameras and the use of filters in different wavelengths a registration process is necessary in order to combine the spectral images.

4. IMAGE REGISTRATION

The process of registering the Nikon images to the multi-spectral Hamamatsu images is illustrated in Figure 3. Both Nikon images are coarsely registered to the reference image (H8) by an affine transformation. This compensates the rotations caused by repositioning the manuscript pages. The feature matching is done using rotationally invariant local descriptors of the Scale-Invariant Feature Transform (SIFT) [8]. Since the computation of the scale-space is computationally expensive and the size of the manuscript pages is similar in different images, the scale-space is not computed in our approach. Thus, each control point detected has the same scale.

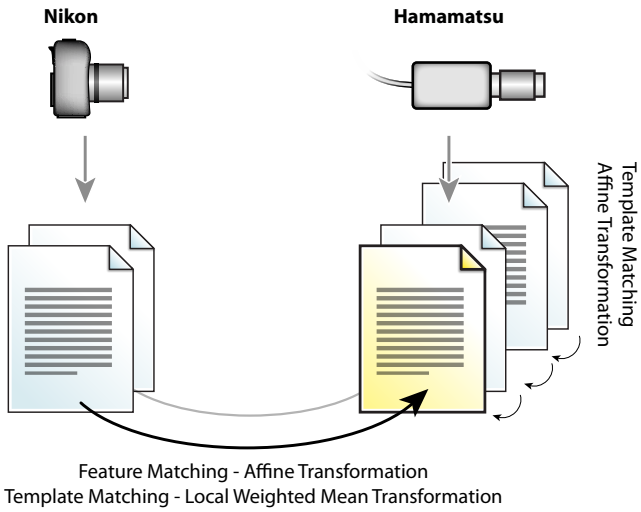

Figure 3: Diagram showing the registration process. The bold border indicates the reference image (H8).

The Difference of Gaussians (DoG) detects too many control points for a registration task. Thus, the control points are localized using the Harris Corner Detector [7] which detects less control points with the same scale parameter σ and is robust against rotational changes too. The orientation assigned to each control point is computed similar to Lowe's implementation [8]. First the image gradient magnitude $m(x,y)$ and the orientation $\theta(x,y)$ are computed for each pixel of the smoothed image $L(x,y)$. An orientation histogram with 36 bins corresponding to 360° is created. Each sample added to the histogram is weighted by its gradient magnitude and a Gaussian weight. Afterwards, the histogram is smoothed with a Gaussian kernel. The maximum of the histogram indicates the dominant direction of local gradients.

In order to compute a local descriptor that characterizes each control point the image gradients $m(x,y)$ and the orientations $\theta(x,y)$ in a $16 \times 16px$ window around each control point are considered. The coordinates of the descriptor and the gradient orientations are rotated relative to the orientation of the control point so that the features are rotationally invariant. Each gradient region is weighted by a Gaussian window of $\sigma = 8$ so that the descriptor does not change significantly with small changes in the position of the window. The control point descriptor consists of eight 4×4 planes where each plane represents the spatial distribution of the gradients for eight different directions. The location of a gradient within the local descriptor depends on the rotated coordinates and its orientation. Each gradient is interpolated to its eight neighbors of the descriptor.

After the features for the reference and the sensed image are computed, they are matched using the nearest-neighbor algorithm. The correspondence of two control points is indicated by the minimal Euclidean distance. Since a control point may exist solely in one of the two images, corresponding control points are rejected if their distance to the nearest-neighbor is less than 0.8 times the distance to the second-nearest neighbor. Control points which have more than one correspondence are discarded too. Having discarded the control points according to this scheme approximately 200 corresponding control points are left for a $391 \times 493px$ image.

Since false matches can exist after discarding the previously mentioned control points and a single outlier changes the transformation estimation of the Least Squares Solution dramatically, the RANSAC [4] method is used to discard all remaining outliers. This approach computes the affine transformation using three randomly selected matching points. Having tested all remaining control point pairs, the model is reestimated from the entire set of hypothetical inliers. These steps are repeated until the distances between points and the model meet a given threshold. This method discards in our approach $\approx 8.3\%$ of previously matched control points. Matched and discarded control points of two manuscript images can be seen in Figure 4. Afterwards, an affine transformation matrix is computed by means of the Least Squares Solution with all remaining corresponding control points.

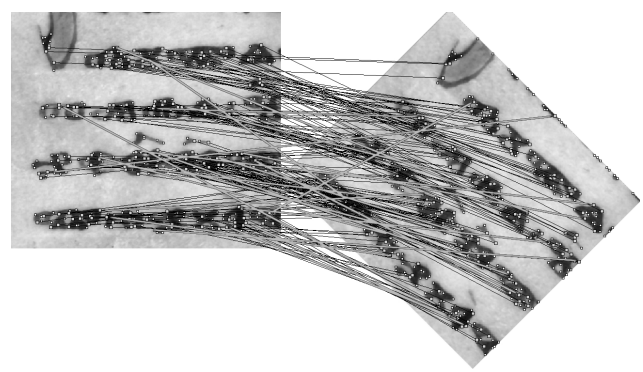

Figure 4: Matched control points. False correspondences are eliminated by means of RANSAC (bold gray lines).

Having aligned both Nikon images coarsely to the reference image using adapted SIFT features and a global affine transformation, a template matching and a subsequent local transformation is performed in order to correct non-rigid distortions caused by changing page curvatures. Due to feature matching and RANSAC it is not possible to spread corresponding control points uniformly across both pages. Therefore the normalized cross correlation is computed at the locations of the previously found control points. Each template contains one character. Having determined the control points, the parameters for the transformation matrix are computed.

Transformations using polynomials of order n are defined by at least $n+1$ parameters, which results in a complex similarity functional that has many local optima. To overcome this problem a local mapping function is applied. The local weighted mean method [6] is a local sensitive interpolation method. It requires at least 6 control points which should

be spread uniformly over the entire image. Polynomials 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, a weighting function is defined which guarantees that solely polynomials near an arbitrary point influence its transformation.

Since the spectral images from the gray level camera are misaligned (minor rotations) template matching combined with an affine transformation is applied, see Figure 3. A more detailed explanation of the registration method is presented in [1].

5. MULTIVARIATE SPATIAL CORRELATION

Multispectral image data is often highly correlated, i.e. they are visually similar [13]. The correlation arises through sensor band overlap and material spectral correlation. The Principal Component Analysis (PCA) is a common transformation which removes this redundancy:

$$x' = A^t(x-\mu), \tag{1}$$

where A denotes the transformation matrix, x denotes the multivariate data and μ is the d-dimensional mean vector [2]. A^t is the transpose of A. The columns of A consist of the k most valuable eigenvectors which are computed from the $d \times d$ covariance matrix.

Multivariate Spatial Correlation is a method for quantifying spatial autocorrelation in multiband data [17]. Wartenberg extended a common univariate method of spatial correlation analysis for multivariate data. The method called Multivariate Spatial Correlation (MSC) was primarily used for geographical analysis, e.g. geographical distribution of blood groups [18]. Warner [17] took up this method for the analysis of remotely sensed data and showed the robustness in the presence of noise. Compared to the results of synthetic data, the MSC matrix explained 99% of the multivariate spatial correlation, whereas the first three components of PCA explained only 75%.

The spatial correlation methodology can be thought of as a part of a generalized principal component analysis, for details see the Appendix in [18]. The MSC matrix of a d band $n \times mpx$ image is defined as follows:

$$M = ZWZ'. \tag{2}$$

Z is a $d \times nm$ matrix containing the multivariate image data, W is a $nm \times nm$ weight matrix and Z' denotes the transpose of Z. The $nm \times nm$ weight matrix W is in the simplest case an adjacency matrix with $w_{ij} = 1$ if i and j are adjacent (e.g. $d_{ij} < c$ where c denotes a critical distance), otherwise $w_{ij} = 0$. The matrix is standardized so that its sum is equal to 1. In the optimal case the weight matrix should be a binary mask containing the characters - which is exactly what we want to obtain. As a starting point we generate a weight matrix containing the text lines as a mask. The mask is generated manually for this time due to the fact that the lining has a certain scheme and does not vary between the pages.

The spatial correlation matrix M, which is in quadratic form, can again be decomposed into orthogonal components using eigenvector analysis [18]. The components reflect the distribution of variations, comparable to PCA. In this case the result is spatially weighted throughout the multivariate field. Similar to the computation of the principal components we calculate

$$x'' = B^t(x-\mu), \tag{3}$$

where the columns of B consist of the eigenvectors which are computed from the $d \times d$ MSC matrix M.

6. EXPERIMENTS AND RESULTS

We applied the MSC to the image data obtained with the multispectral image acquisition system as presented in Section 3. Since the weight matrix achieves a size of $nm \times nm$ were n and m depict the original image size we use only fractions of the image which are additionally reduced in resolution.

The results shown in this section are based on the input images in Section 3 which size depicts $300 \times 500px$. Figure 5 shows the distribution of the eigenvalues from MSC compared to PCA. It can be seen that the first eigenvalue from

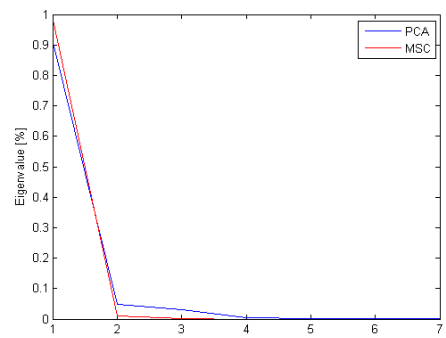

Figure 5: Eigenvalues.

the MSC matrix includes more than 99%. Compared to the eigenvalues from the PCA which extend only 91.3% in the first and 4.8 % in the second component.

The resulting images from MSC and PCA can be seen in Figure 6 and 7. The images are shown without any post processing. Regarding the input images in Figure 2 it can be seen that the characters especially in the upper left corner and in the middle of the second and third row are hardly visible. Regarding the second band obtained from the MSC transformation the visibility of the characters is clearly enhanced. The readability is feasible to the philologists who additionally use the context to read the texts. The third and fourth band of the MSC contain no useful information as the distribution of the eigenvalues shows.

Figure 7 shows the results obtained with PCA. Here the first three components contain some visible characters which is again adequate to the distribution of the eigenvalues from the covariance matrix. In this case the readability showed no useful improvements. Also the decorrelation stretch [5] which is an extension of PCA could not improve the readability significant.

7. CONCLUSION AND OUTLOOK

Multispectral imaging supports the investigation of ancient manuscripts where the text is hardly visible in conventional RGB images or for the human eye. A drawback of the method is the assemblage of highly correlated image data and the need for registration. The Principal Component Analysis

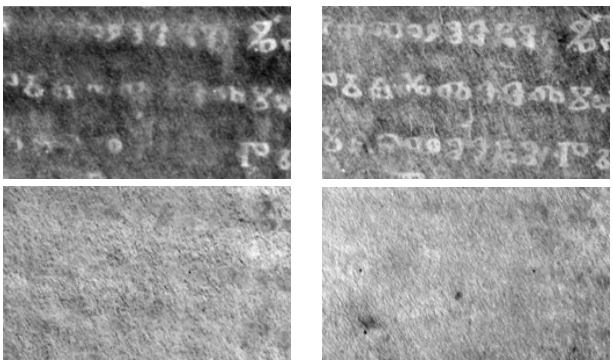

Figure 6: MSC Results.

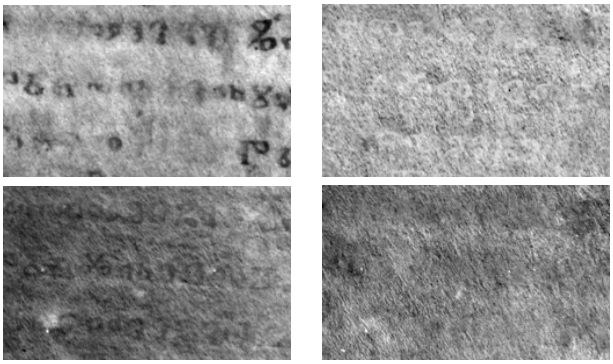

Figure 7: PCA Results.

is a conventional method to reduce the spectral image data and to produce pseudo colored images.

In this study we first presented a two stage registration process which enables the coarse alignment of images from different acquisition systems and additionally aligns locally distortions caused through the multispectral image acquisition. In order to improve the readability we use the Multivariate Spatial Correlation matrix which includes the spatial and the spectral image data to remove spectral correlation. The results show that already in one of the first three components the text appears clearly enhanced which outperforms the PCA.

Future work will investigate the automatic generation of the weight matrix. In order to approximate the spatial correlation we are going to investigate the ruling scheme of the manuscripts. An automatic segmentation of the text lines will be a first step. Another alternative is the masking of separate characters. This can be done for instance recursively which means that a new weight matrix is generated after each calculation of the MSC components until a stopping criterion is reached.

REFERENCES

[1] M. Diem, M. Lettner, and R. Sablatnig. Registration of multi-spectral manuscript images. In *8th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*, pages 133–140, Brighton, UK, 2007.

[2] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern Classification*. Wiley-Interscience, New York, 2 edition, 2001.

[3] R.L. Easton, K.T. Knox, and W.A. Christens-Barry. Multispectral Imaging of the Archimedes Palimpsest. In *32nd Applied Image Pattern Recognition Workshop*, pages 111–118, Washington, DC, October 2003.

[4] M.A. Fischler and R.C.Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, June 1981.

[5] A.R. Gillespie, A.B. Kahle, and R.E. Walker. Color enhancement of highly correlated images. i. decorrelation and hsi contrast stretches. *Remote Sens. Environ.*, 20(3):209–235, 1986.

[6] A. Goshtasby. Image registration by local approximation methods. *Image and Vision Computing*, 6:255–261, 1988.

[7] C.G. Harris and M. Stephens. A Combined Corner and Edge Detector. In *4th Alvey Vision Conference*, pages 147–151, 1987.

[8] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[9] F. Mairinger. *Strahlenuntersuchung an Kunstwerken*. Bücherei des Restaurators, Band 7. E.A. Seemann, 2003.

[10] H. Miklas. Zur editorischen Vorbereitung des sog. Missale Sinaiticum (Sin. slav. 5/N). In H. Miklas, V. Sadovski, and S. Richter, editors, *Glagolitica - Zum Ursprung der slavischen Schriftkultur*, volume XV-XVI, pages 117–129. ÖAW, 2000.

[11] M. Mitsui, Y. Murakami, T. Obi, M. Yamaguchi, and N. Ohyama. Color Enhancement in Multispectral Image Using the Karhunen-Loeve Transform. *Optical Review*, 12(2):69–75, 2005.

[12] K. Rapantzikos and C. Balas. Hyperspectral Imaging: Potential in Non-Destructive Analysis of Palimpsests. In *IEEE International Conference on Image Processing*, volume 2, pages 618–621, 2005.

[13] J.A. Richards and X. Jia. *Remote Sensing Digital Image Analysis : An Introduction*. Springer, 1999.

[14] E. Salerno, A. Tonazzini, and L. Bedini. Digital image analysis to enhance underwritten text in the Archimedes palimpsest. *International Journal on Document Analysis and Recognition*, 9(2-4):79–87, 2007.

[15] A. Tonazzini, L. Bedini, and E. Salerno. Independent component analysis for document restoration. *International Journal on Document Analysis and Recognition*, 7(1):17–27, March 2004.

[16] D.-C. Tseng, Y.-L. Chen, and M. Liu. Wavelet-based multispectral image fusion. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 1956–1958, Sydney, Australia, 2001.

[17] T.A. Warner. Analysis of Spatial Patterns in Remotely Sensed Data Using Multivariate Spatial Correlation. *Geocarta International*, 14(1):59–65, 1999.

[18] D. Wartenberg. Multivariate Spatial Correlation: A Method for Exploratory Geographical Analysis. *Geographical Analysis*, 17(4):263–283, 1985.