Schriften des Instituts für Dokumentologie und Editorik — Band 2

Kodikologie und Paläographie im digitalen Zeitalter

Codicology and Palaeography in the Digital Age

herausgegeben von | edited by

Malte Rehbein, Patrick Sahle, Torsten Schaßan

unter Mitarbeit von | in collaboration with

Bernhard Assmann, Franz Fischer, Christiane Fritze

2009

BoD, Norderstedt

Bibliografische Information der Deutschen Nationalbibliothek:
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über **http://dnb.d-nb.de/** abrufbar.

Leicht veränderte Fassung für die digitale Publikation (siehe Vorwort).

Slightly modified version to be published digitally (see preface).

Publication réalisée avec le soutien d'Apices
Association Paléographique Internationale
Culture – Écriture – Société
Dotation J.M.M. Hermans.
http://www.palaeographia.org/apices/

Herstellung und Verlag: Books on Demand GmbH, Norderstedt
ISBN: 978-3-8370-9842-6
Einbandgestaltung: Katharina Weber
Satz: XeTeX und Bernhard Assmann

The Ghost in the Manuscript: Hyperspectral Text Recovery and Segmentation

Patrick Shiel, Malte Rehbein, John Keating

Abstract

Major activities in palaeographic and manuscript studies include the recovery of illegible or deleted text, the minute analyses of scribal hands, the identification of inks, and the segmentation and dating of text. This article describes how Hyperspectral Imaging (HSI) can be used to perform quality text recovery, segmentation and dating of historical documents. It provides a comprehensive overview of HSI, and associated computational and segmentation techniques used for two experimental investigations: (i) a 16th century pastedown cover, and (ii) a multi-ink example typical of that found in, for example, late medieval administrative texts such as Göttingen's *kundige bok*.

Zusammenfassung

Paläographische Forschung und das Studium von Handschriften umfassen das Wiederherstellen unlesbaren oder getilgten Texts, die genaue Analyse von Händen, die Identifizierung von Tinten sowie die Segmentierung und Datierung von Text. Dieser Beitrag beschreibt, wie hyperspektrale Bildverarbeitung (*Hyperspectral Imaging, HSI*) für diese Aufgaben angewendet werden kann. Er liefert einen umfassenden Überblick über die Technologie und beschreibt die rechnergestützten Verfahren und Methoden der Segmentierung, die experimentell für zwei Fallstudien entwickelt wurden: (i) zur Untersuchung eines Bucheinbandes aus dem 16. Jahrhundert und (ii) zur Analyse der Gebrauchsspuren verschiedener Tinten, wie sie etwa typisch für spätmittelalterliche Amtsbücher sind.

1 Introduction

The condition of medieval manuscripts ranges from those that are fully legible to those that can only be read in part, and their legibility is determined by the manner in which they were preserved and treated throughout the ages. In some cases deterioration is due to processes such as fading or staining; in others, the text may have been interfered with in some way. For instance, in the Irish context, the oldest (12th century) surviving

Kodikologie und Paläographie im Digitalen Zeitalter – Codicology and Palaeography in the Digital Age. Hrsg. Malte Rehbein, Patrick Sahle und Torsten Schaßan, unter Mitarbeit von Bernhard Assmann, Franz Fischer und Christiane Fritze. Schriften des Instituts für Dokumentologie und Editorik 2. Norderstedt: Books on Demand, 2009. 159–174.

manuscript written entirely in Irish, *Leabhar na hUidhre* (*Book of the Dun Cow*) was subject to part-erasure and rewriting by a scribe who was active at some point between the 12th and the 14th centuries (Ó Cróinín). In the German context, we refer to a 15th century manuscript (*kundige bok*) containing a legal text that is characterised by many revisions over a period of approximately 50 years (Rehbein). It is a multi-layered text of 330 paper pages, with its different layers representing the various stages in the development of the town law over the years. Revealing its layers is crucial for understanding the text and for historical studies.

For these examples, it is not only necessary to identify the different scribes in the manuscript but also to deal with the issue of the same scribe writing at different points in time. The work on the digital edition of *kundige bok*, for example, which aims at visualising the textual evolution, has so far relied mainly on two techniques: (i) looking for dated entries and contextual (internal and external) information, and (ii) using palaeographic, codicological and linguistic methods to analyse a range of identifying factors. These would include analysis of the scribes' style of writing over the years (i.e. the writing process itself and customs in using certain words and phrases), inks (as far as the human eye can distinguish the colours), and the paper the text was written on (analysis of watermarks etc.), etc. (Bischoff).

Using this approach, it was possible to assign slightly more than half of the textual alterations to text layers (i.e. stages of the town law) and to bring them into chronological order. This required, however, a lot of experience; it took time to familiarise with scribes and scribal habits of that particular place and time, and objectivity in one's decision could not always be assured. Overall, it was a time consuming process, leaving behind a good amount of uncertainty. There are still more than 40% of the medieval scribes' changes to the text that cannot be assigned to a text layer or brought into chronological order in a satisfying way at all. The digital edition of *kundige bok* copes with this by ensuring transparency in the editorial method, making decisions and uncertainty visible to the users and allowing them to dynamically create text layers on their own using the facsimile provided alongside the transcriptions. But shall this be the end of our efforts? Uncertainty in revealing the textual evolution grows when: (i) too many entries in a particular context are not dated, (ii) the same scribe makes corrections at the same passages but (likely) at different times, (iii) ink colours are too similar for the human eye to distinguish, (iv) entries are too short to survey scriptural characteristics or do not even consist of text at all (e.g strike-throughs).

Two central research questions related to the uncertainty of textual development, of interest to all scholars, therefore, are:

- To what extent is text recovery (e.g. in the cases of palimpsest, fading, deliberate removal, etc.) in key medieval texts possible using current technologies?

- To what extent is it possible to establish irrefutable scientific evidence for interpretation of questioned documents, e.g., identify the different hands (inks)?

In this article, we illustrate how to provide answers to these questions using modern scientific techniques and emerging forensic technology, i.e. hyperspectral imaging (Chang 2003) and associated image processing techniques (Chang 2007). In particular, we are interested in the application of hyperspectral segmentation techniques to multi-layered manuscripts to help solve these problems and overcome the uncertainty of the textual development. It focuses thereby on working particularly on the dating issues of the later added entries, thus allowing us to bring them at least into a chronological order where palaeographic means alone would fail.

The availability of a hyperspectral scanner, a Forensic XP-4010, has presented the authors with opportunities to subject damaged or illegible texts to a modern scientific re-examination. The scanner has the potential to read various different layers of a manuscript in a manner not possible to the human eye and to analyse elements of its composition. As such it presents the possibility of retrieving text that has been lost through fading, staining, overwriting or other forms of erasure. In addition, it offers the prospect of distinguishing different ink-types, and furnishing us with details of the manuscript's composition, all of which are refinements that can be used to answer questions about date and provenance. This process marks a new departure for the study of manuscripts, for the authors, and may provide answers to many long-standing questions posed by palaeographers and by scholars in a variety of disciplines. Furthermore, through text retrieval, it holds out the prospect of adding considerably to the existing corpus of texts and to providing many new research opportunities for coming generations of scholars. In this introductory chapter on hyperspectral imaging, we concentrate on two key processes: text recovery and text segmentation.

2 Background and Methodology

The investigative and analytical methods described here are based on a novel and highly specialised technique called Hyperspectral Imaging (HSI), and sometimes referred to as Optical Reflectance Imaging. HSI is a non-destructive optical technique that measures reflectance (fraction of light reflected) characteristics of a document with high spatial and spectral resolution. An HSI device, operating as a reflectance spectrometer, records a sequence (typically hundreds) of digital images of the selected manuscript area (with maximum dimensions 50mm x 50mm) illuminated with monochromatic light from a tunable light source from 350nm (near-UV) through the entire visible range and up to 2400nm (infrared). The value of each image pixel in the recorded image sequence represents an accurate measurement of the reflectance curve for a tiny—13 micron square—area on the document. Analysis of all spectral curves, essentially a cube of information, provides information about the physical characteristics of questioned manuscripts.

HSI, together with modern two-dimensional spectrum software and three-dimensional image and visualisation software, provides modern researchers working in the field of historic documents analysis with opportunities for forensic examination that were heretofore unavailable. Methodologically, there are two main fields of applications of this technique: (i) the extraction of relevant historic, diplomatic and palaeographic information from documents, and (ii) the investigation of the impact of environmental conditions on document condition and of degradation effects on writing materials and substrates. In particular, reflectance curves found in different sections of the manuscripts can be compared with each other in order to determine whether different types of inks had been used during text composition or to identify modifications that occurred during the manuscripts' history. Light spectroscopy analyses may also be conducted to aid recovery and segmentation. Fluorescence occurs when an object emits a high wavelength (low energy light) following illumination by a shorter wavelength (higher energy light) due to molecular absorption of part of the incident light. Furthermore, the spectral curves may be compared with those in international databases containing typical ink spectra to determine and date the kind of ink or pigment used. The image cube recorded using the technique may be used to enhance the visibility of hidden material such as palimpsest or erased text.

This methodology for manuscript analysis is of significant interest to archivists, conservationists, and scientists interested in non-destructive historical document analysis. Klein et al. (2008) recently provided an excellent description of the basic concepts, working principles, construction and performance of a HSI device specifically developed for the analysis of historical documents. Their custom-developed quantitative hyperspectral imager is currently used by the *Nationaal Archief* (National Archives of The Netherlands) to study degradation effects of artificial samples and original documents, exposed in their permanent exhibition area or stored in their deposit rooms. Earlier Klein et al. used their device to record the variation of spectral reflectance on a historic 17th century map, and also used the instrument to compare the local variation of the yellowness index of reference papers stored in a bound volume, and loose sheets. Using HSI they determined that yellowness occurs within 20mm border in both cases, and that the yellowness index was much higher for the bound paper than the loose paper. Ongoing HSI investigations allow researchers to detect and visualize differences in aging processes on a document, and are particularly useful when taken, for example, before and after an exhibition, whereby it is possible to investigate the effects of exhibiting and handling on document yellowing (Padoan et al.).

Padoan et al. have conducted HSI supported palaeographic investigations of the first manuscript written in Dutch (14th century), held by *Koninklijke Bibliothek* (Royal Library of the Netherlands). In particular, the investigation centred on whether a coat of arms, drawn on the lower part of the first page, could have been ascribed to the famous 15th century Flemish bibliophile *Lodewijk van Gruuthuse* or whether the element may have been added at the same time as the 17th century text surrounding it (Padoan et al.). They found that the HSI derived image shows no correlation between the coat-of-arms and the surrounding text, while strong similarities were observed within areas where corrections were made in the 15th century.

There have also been several other recent reports of promising results from the application of the HSI analysis of paintings and conservation of works of art (Fischer and Kakoulli), paper discoloration and foxing (Missori et al.), and detection of iron-gall inks (Havermans). It is our observation that the successful application of HSI in palaeographic and codicological studies emerges from equitable partnerships of researchers and scholars in humanities, computing, and natural sciences.

3 Reflectance and Fluorescence Spectroscopy

Light spectroscopy is the study of light that is emitted by or reflected from objects. When applied to hyperspectral image analysis, spectroscopy deals with capture and examination of images using a large portion of the light spectrum. Light and other forms of electromagnetic radiation are commonly expressed in terms of their wavelength; each photon of light has a wavelength determined by its energy level. When a single frequency light wave comes in contact with an object several different phenomena may be observed depending on the object's composition and wavelength of the incident radiation, for example, the incident light may be reflected, absorbed and turned to heat, or transmitted (and refracted). Single frequency incident light is uncommon, however. Normally, visible light striking an object contains many frequencies, and when this occurs the object will selectively reflect, absorb or transmit certain frequencies. Some excellent introductions to the various aspects of light spectroscopy are given by Hapke who discusses, in detail, the theory of reflectance and emittance spectroscopy, and Ashutosh and Schulman who comprehensively describe fluorescence spectroscopy in their volume.

HSI analysis, therefore, refers to the analysis of spectral images taken at a sequence of spatially aligned wavelength bands. Simple spectral imaging systems, such as a digital camera acquire intensity images in 3 bands (red, green and blue). Hyperspectral Image systems typically image a scene in hundreds of indexed bands, utilising information from a wide region of the electromagnetic spectrum. Modern hyperspectral image sensors can be used to capture the attributes of light emitted by materials, and its variation in energy with wavelength, at a series of narrow and contiguous wavelength bands. For our investigations, we used the Forensic XP-4010 forensic document examination system (from MS-Macrosystems, The Netherlands) to acquire high-resolution optical

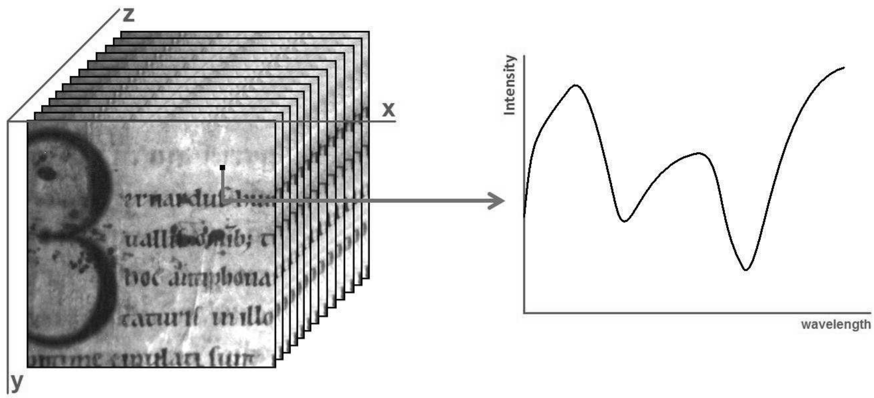

Figure 1. A Hyperspectral cube (left); hundreds of spatially recorded images acquired contiguously over a wavelength region. The reflected energy spectrum for an individual pixel is shown on the right.

and infrared images. The sensor measures reflectance, absorption, transmittance and fluorescence spectral information in the ultraviolet, and 400–1000nm region. The data obtained is in the form of a data-cube (Figure 1) which represents the image information as a data set in three dimensions; two of the data cube's axes represent the spatial data, while its third axis represents the spectral information. This image cube consists of hundreds of spatially recorded images, acquired contiguously over the wavelength region. Each pixel within a hyperspectral image-cube represents the reflected energy spectrum of materials spatially covered by the pixel.

Spectral reflectance, the fractional amount of incident energy that is reflected from a surface with respect to wavelength, is one of the fundamental attributes obtained when performing hyperspectral analysis of documents. In general, an object's reflectance varies with wavelength, as incident energy at particular wavelengths is absorbed or scattered in different directions. Materials that have a similar reflectance under natural light may have vastly different reflectance under light of a specific wavelength from other regions of the light spectrum. The importance of this, in the context of analysis of documents, is that this difference in spectral reflectance allows for classification of different features of the document which, under natural light, look identical. Furthermore, the spectral reflectance also depends on the orientation of the object surface which can be a problem with items that are deformable, such as manuscripts. This may be overcome, however, by implementing diffuse illumination within a light-proof observation chamber. In an experimental investigation on text recovery, detailed below, we demonstrate how spectral reflectance may be used to segment text written with three different pens that appear similar to the human eye.

Luminescence is a phenomenon where an object emits light due to chemical reaction, electrical energy, subatomic motions, or stress. A particular type of luminescence, called fluorescence, is of particular interest when performing hyperspectral analysis of documents. Fluorescence occurs when an object emits a high wavelength (low energy light) following illumination by a shorter wavelength (higher energy light) due to molecular absorption of part of the incident light. The object absorbs part of the incident light which causes excitement of the molecules within the object and then emits energy in the form of lower energy light usually within the visible range of the spectrum. In an example, later in this chapter, we show how the application fluorescence spectroscopy can be useful in recovering unreadable text from historical documents in our sample 16th century book cover.

Tilley provides an excellent discussion on the relationship between light, the optical properties of materials and colour, and is particularly useful for those interested in using HSI to distinguish different materials, for example, inks on paper. HSI observations of historical documents written on paper, parchment, etc. are complex spectral combinations of the reflectance of a collection of materials that have different temporal degradation properties. Vaarasalo provides a useful discussion on the optical properties of paper, which loses its optical properties as time passes (van der Reyden). More recently, de la Rosa and Bautista have discovered that to find the presence and concentration of different colorants or components in the paper, it is only necessary to know the spectra and fluorescence lifetimes (at 337.1nm). They indicate that these kinds of measurements could be useful for studying the paper's long-term stability and how aging affects it, and is particularly important in the preservation of paper-based historical records (Committee on Preservation of Historical Records). Light spectroscopy has also been particularly useful in the examination of paper aging, for example, in accelerated aging experiments (Bansa) and on the evaluation of whiteness and yellowness (Smith).

4 Experimental Investigation: Recovery of Hidden Text

We are especially interested in the use of hyperspectral analysis to support the recovery of hidden text, and in particular, the kind of recovery that requires substantial conservation efforts, or the disassembly of manuscripts and their bindings. We believe that it is possible to employ non-invasive and non-destructive investigative techniques to search for hidden text prior to embarking on physical analyses and treatment of historical texts.

A particularly interesting example of this high-precision, skilled and manual work is reported by Quandt who conducted a detailed physical analysis and treatment of a late

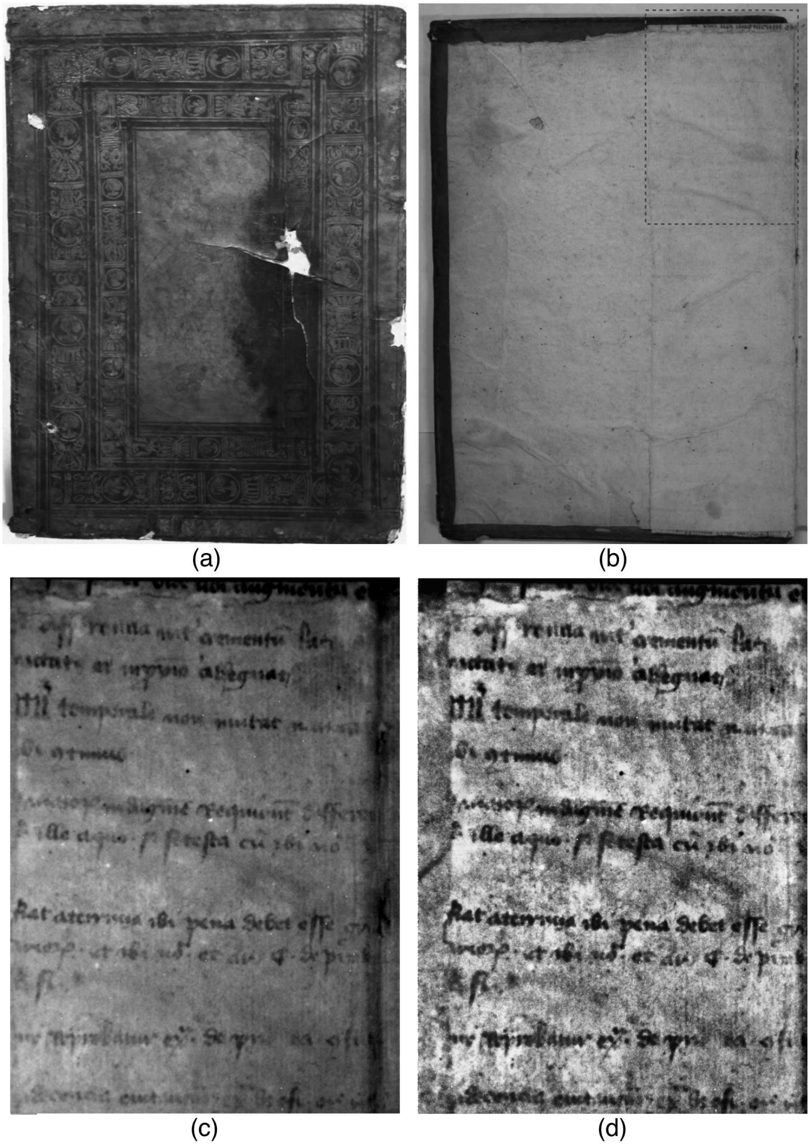

Figure 2. Hyperspectral text recovery of a 16th century book cover showing (a) the exterior, (b) the interior pastedown, (c) recovered text, and (d) thresholded text. Source: Russell Library, Maynooth, Co. Kildare.

13th century copy of the *Etymologies* of Isidore of Seville. The manuscript, although damaged over the centuries, retained most of its original medieval binding structure, and it is reported that part of the project included the compilation of technical evidence that would lead to accurate localisation of the text and a reconstruction of the binding history of the volume. One aspect of this work included the removal of apparently blank pastedown from the upper board; however, disassembly revealed that the pastedown contained writing (cursive Latin text) on the verve. Quandt concludes that manuscript fragments, such as those used as pastedowns, are potentially important, as they may serve to document the origin of the manuscript and its medieval binding.

In order to investigate non-invasive hyperspectral techniques for text recovery, we obtained an exterior 16th century book cover (located in the Russell Library at National University of Ireland Maynooth). This had become degraded with time and suffered from mould in places (Figure 2(a)). The exterior cover, which is intact for the most part, is not of interest here. The interior cover's structure, shown in Figure 2(b), consists of an underlying text which has been pasted over with a clean blank faced sheet of paper. Using fluorescence spectroscopy, i.e, light induced fluorescence in the page, it is possible to reveal the underlying text as shown in Figure 2(c) and 2(d). The underlying text is assumed to be degraded with time but is unreadable due to the presence of the overlaying, pasted down sheet.

When the cover was illuminated with high energy light (505nm) it was found that this light produced fluorescence in the pasted down and underlying pages. Both pages absorb part of this incident light, and in turn emit a low energy (red) light with wavelength 720nm. Using an appropriate filter, our camera records a greyscale image of the resultant fluorescence intensities. As a result of the ink having very low fluorescence relative to the paper, any portion of the interior cover that contains the ink will not fluoresce at 720nm, and is represented by black pixels in our greyscale image, thereby revealing the underlying text.

Following initial light spectroscopy images, we apply hyperspectral analysis techniques and algorithms to further enhance the contrast between page and text and increase the legibility of the recovered text. As described by Salerno et al. techniques have been developed to model the book cover and page as a series of different layers. These layers represent different patterns in the hyperspectral image such as the clean text, the mould pattern and the original parchment or paper pattern. The digital restoration of this book cover involves differentiating the text pattern from interfering patterns and the parchment pattern. This is typically achieved using statistical and image processing algorithms such as Principal Component Analysis (Jolliffe) and Independent Component Analysis (ICA) (Hyvärinen).

Principal component analysis (PCA), for example, is a statistical analysis technique that is generally used to condense high dimensional data into data of a lower dimension. PCA can also be used to identify the most representative elements of the data, called Principal Components, which is of use in hyperspectral image analysis. These principal components are selected and ranked by their relative variance, judging the most important representative element to be the most variant. Selectively removing certain principal components, and emphasising others, allows the isolation of each of the layers in a hyperspectral image.

In the hyperspectral analysis of historical documents it is seen that individual pixels can be a combination of various substances. In our example, a mould pattern has formed spatially covering different pixels. Pixels spatially covered by this mould pattern will have a spectral signature different to that of pure ink pixels, and pure mould pattern pixels. This type of pixel is termed to be a spectrally mixed pixel. The spectral (pixel) un-mixing of the signals leads to the determining of the contribution of each material in the mix. Un-mixing hyperspectral image data can be seen as an unsupervised method for blind source separation where the objective is to determine the contribution of each component in the mixed signal without prior knowledge of the sub-components. Independent component analysis (ICA), sometimes referred to as *blind-source separation*, is an unsupervised source separation process which provides one method to un-mix the different components of a mixed pixel. ICA can be applied to hyperspectral images where the data consists of linearly mixed signals; for this example, we can use ICA to separate the book cover into the aforementioned layers of the cover model. This enhances the readability of the recovered text by separating it from any interference patterns.

In the case of the treatment of the *Etymologies of Isidore of Seville* copy mentioned above, the actual binding structure was not harmed in any way, and the recovered evidence proved to be extremely useful in reconstructing the history of the medieval manuscript. The authors admit that the uncovering and removal of binding fragments seemed at times to be too invasive a procedure. We have shown that the employment of hyperspectral methods for text recovery is a viable option during times when treatment may be too invasive a procedure. The procedures, albeit non-invasive, are complex, non-automated, and require further investigation into appropriate humanities research questions in this domain prior to the production of fitting computer software.

5 Experimental Investigation: Text Segmentation of Different Inks

The general objectives of cursive text segmentation include tasks such as word spotting, text/image alignment, authentication and extraction of specific fields (Likforman-Sulem). An important step associated with all of these tasks is the segmentation of the document into logical units, for instance text lines, words or letters. In general, this is difficult due to the low quality and complexity of these documents, and automatic text

Figure 3. Hyperspectral segmentation of simulated textual evolution using three different inks; (a) the test sample, (b) hyperspectral spectrogram, (c) original text (white), and (d) later annotations (in black and grey).

segmentation of such kind is an open research field (see also in this volume: Aussems – Brink, Ciula and Stokes). Sophisticated image processing of single-image documents is hence the norm so far (Likforman-Sulem). Here we describe our recent approach towards segmentation of a different kind, which we refer to as hyperspectral segmentation; the technique is based on the separation and segmentation of different inks by recording and analysing their reflectance properties. This technique is particularly useful for the segmentation of texts that have been edited by various authors over a long time period. Thus, it helps in answering basic palaeographic questions and allows the dating of text by comparing the segments with known dates, or using repositories containing hyperspectral properties of different materials (e.g. inks, paper, etc.).

Figure 3 shows a simulation of the textual evolution of *kundige bok* (Rehbein). It is a sentence from this medieval town law, though re-written on modern paper with modern inks for demonstration purposes. As can be seen in the sample (Figure 3(a)), it is hardly possible for the human eye (if at all) to detect whether the changes of the text were made with a different ink or not, thus failing to give an indication whether the two changes within the sentence originate from different points in time in the writing process or not.

In this recently conducted experiment, the hyperspectral scans revealed the different inks surprisingly easily. Figure 3(b) shows the spectrogram of the simulation, measuring the reflectance on three different pixels on the manuscripts, marked green, blue and red. While the spectrograms for two pixels (green and blue) are very close to each other, the third one (red) shows a significant difference. The change, made at the position indicated by the red pixel ('3' substituted by '2') was likely made with a different ink than the change indicated by the blue pixel ('100' substituted by '150'), while the latter was likely made with the same ink with which the original text was written (green dot). Taking into account that medieval scribes produced their ink individually (Wattenbach, 240; Hoheisel 102), using traditional recipes, it can be concluded—if the simulation was based on a real medieval manuscript—that the sample was original text (green) with a simultaneous correction (blue) and a later revision (red). In the *kundige bok* case study we refer to here, this would be an important step towards a complete revelation of the textual evolution and with it the development of medieval town law in late 15th century—information that was not known before.

These preliminary and experimental results thus give hope that work on the original manuscript of *kundige bok* or a similar text, intended to be undertaken in co-operation with An Foras Feasa, National University of Ireland, Maynooth and the Stadtarchiv Göttingen, in Summer 2009, will lead to promising results also. Furthermore, we are also encouraged by the recent results of Klein et al. (2006; 2008) who successfully combine HSI, mathematical feature extraction and classification techniques to analyse documents where several types of ink have been applied, and documents where one ink displayed various degrees of degradation.

Identifying segments of the text written with the same or different ink is only the first step, however, and does not solve the dating issues by itself. It must be accompanied by the expert's view on the manuscript. Dating requires one more piece of information. Consider, for instance, an entry that is undated but from which hyperspectral analysis reveals the same ink signature as a dated entry elsewhere in the book, or even a different source. It can then be dated by inference. However, building up a database of historical ink signatures in a certain (chronological and/or local) context, could establish the basis for an (semi-) automatically created facsimile edition of a manuscript (be it medieval or modern) by visualisation of the different stages (see Figure 3(c)) of the text and could also lead to automated markup preparation—catering, for example, as a tool for the creation of genetic editions.

6 Conclusion

In conclusion, our initial experimental investigations demonstrate the advantages of high-spatial reflectance and fluorescence spectroscopy measurement for the non-invasive examination of historical documents. In particular, it can support codicology research by revealing binding structure of a codex or creating a database of ink signatures, and palaeographic research by making visible hidden text or by giving support to identify scribes and to solve dating issues. We believe that the inclusion of hyperspectral imaging devices as standard research equipment for usable non-destructive analysis of historic documents is both affordable and attainable and would encourage humanities research institutes, libraries and archives to invest in the technologies, methodologies and proficient personnel to maximise their potential. Furthermore, we believe that equitable humanities computing partnerships are an essential component in hyperspectral imaging projects in order to provide realistic *use cases* for the development of the necessary software tools to support disruptive codicology and palaeography research.

Bibliography

Ashutosh, Sharma and Stephen G. Schulman. *An introduction to Fluorescence spectroscopy.* New York: John Wiley & Sons, 1999.

Bansa, Helmut. "Accelerated Aging of Paper: Some Ideas on its Practical Benefit." *Restaurator* 23.2 (2002): 106–117.

Barrow, William J. *Manuscripts and documents. Their deterioration and restoration.* 2. ed. Charlottesville: Univ. Press of Virginia, 1976.

Bischoff, Bernhard. *Paläographie des römischen Altertums und des abendländischen Mittelalters.* Berlin: Schmidt, 1979.

Brannahl, Günther and Malte Grause. "Untersuchungen an Tinten." *Archivalische Zeitschrift* 70 (1974): 79–98.

Brown, K. L. and R. J. H. Clark,. "The Lindisfarne Gospels and two other 8th century Anglo-Saxon/ Insular manuscripts: pigment identification by Raman microscopy." *Journal of Raman Spectroscopy* 35 (2004): 4–12.

Cahill, Thomas. A., Bruce H. Kusko, and Richard N. Schwab. "Analyses of inks and papers in historical documents through external beam PIXE techniques." *Nuclear Instruments and Methods* 181 (1981): 205–208.

Chang, Chein-I. *Hyperspectral Imaging: Techniques for Spectral Detection and Classification.* New York: Kluwer Academic, 2003.

Chang, Chein-I. *Hyperspectral Imaging: Signal Processing Algorithm Design and Analysis.* New York: John Wiley and Sons, 2007.

Clarke, Mark and Marieke Mejers. "Simplification of near-infrared visualisation techniques for identifying blue pigments in situ on manuscripts." *Care and Conservation of Manuscripts* 6 (2000): 242–249.

Erastov, Dimitri. P. "Optico-photographic methods in research in manuscripts and in the visualisation of invisible texts." *Care and Conservation of Manuscripts* 3 (1997): 52–62.

De la Rosa, J. and F. J. Bautista. "Optical properties of paper at 337.1nm." *Revista Mexicana de Física* 51.1 (2005): 110–113.

Fischer, Christian and Ioanna Kakoulli. "Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications." *Reviews in Conservation*, 7 (2006): 3–16.

Fuchs, Robert. "The history of chemical reinforcement of texts in manuscripts – What should we do now?" *Care and Conservation of Manuscripts* 7 (2002): 159–170.

Hapke, Bruce. *Introduction to the Theory of Reflectance and Emittance Spectroscopy.* Cambridge: Cambridge University Press, 2001.

Havermans, John B. G. A., Abdul Aziz, Hadeel, and Scholten, Hans. "Non destructive detection of iron-gall inks by means of multispectral imaging, Part 2: Application on original objects affected with iron-gall-ink corrosion." *Restaurator*, 24 (2003): 88–94.

Hoheisel, Peter: *Die Göttinger Stadtschreiber bis zur Reformation. Einfluß, Sozialprofil, Amtsaufgaben.* Göttingen: Vandenhoeck & Ruprecht (Studien zur Geschichte der Stadt Göttingen, 21), 1998.

Hyvärinen, Aapo, Juha Karhunen, and Erkki Oja. *Independent Component Analysis*, New York: John Wiley and Sons, 2001.

Jolliffe, Ian T. *Principal Component Analysis.* New York: Springer-Verlag, 2002.

Klein, M. E., et al. "The Quantitative Hyperspectral Imager – A Novel Non-destructive

Optical Instrument for monitoring Historic Documents." *International Preservation News* 40 (2006): 4–9. <http://www.art-innovation.nl/fckfiles/file/Downloads/Articles/2006/2006_The_Quantitative_Hyperspectral_Imager.pdf>.

Klein, M. E., et al. "Quantitative Hyperspectral Reflectance Imaging." *Sensors* 8 (2008): 5576–5618. <http://www.art-innovation.nl/fckfiles/file/Downloads/Articles/2008/sensors-38-02-original.pdf>.

Likforman-Sulem, Laurence, Aderrazak Zahour, and Bruno Taconet. "Text line segmentation of historical documents: a survey." *International Journal on Document Analysis and Recognition* 9.2-4 (2006): 123–138.

Missori, M., M. Righini, and S. Selci. "Optical reflectance spectroscopy of ancient papers with discoloration or foxing." *Optics Communications* 231 (2004): 99–106.

Mitchell, Charles Ainsworth. "Inks. Their composition and manufacture, including methods of examination and a full list of British patents." 4. ed. London: Griffin & Co, 1937.

The National Academy of Sciences. *Preservation of Historical Records*, Chapter 4: "Paper", 1986. <http://www.nap.edu/openbook/030903681X/html/33.html>.

Ó Cróinín, Dáibhí, ed. *A New History of Ireland: Prehistoric and early Ireland*, Oxford: Oxford University Press, 2005.

Padoan, R., et al. "Quantitative Hyperspectral Imaging of Historical Documents: Technique and Application." *ART Proceedings* (2008). <http://www.art-innovation.nl/fckfiles/file/Downloads/Articles/2008/padoan2008.pdf>.

Penders, Nathalie J. M. C. and John B. G. A. Havermans. "Preventive conservation related to iron-gall ink deterioration." *Care and Conservation of Manuscripts.* 6 (2000): 18–32.

Quandt, Abigail B. "The Documentation and Treatment of a Late 13th century Copy of Isidore of Seville's Etymologies" *The Book and Paper Group Annual.* Ed. Book and Paper Group (BPG) of the American Institute for Conservation of Historic and Artistic Works, Volume 10, Washington, D.C., 1991. <http://aic.stanford.edu/sg/bpg/annual/v10/bp10-15.html>.

Rehbein, Malte. "Reconstruction the Textual Evolution of A Medieval Manuscript." *TEI@20. Special Issue. LLC. The Journal of Digital Scholarship in the humanities.* (forthcoming).

Roselieb, Hans. "Die Chemie alter und neuer Tinten." *Archivalische Zeitschrift* 70 (1974): 74–78.

Salerno, Emanuele, Anna Tonazzini, and Luigi Bedini. "Digital image analysis to enhance underwritten text in the archimedes palimpsest." *International Journal on Document Analysis and Recognition* 9.2–4 (2007): 78–87.

Scholten, J. H., M. E. Klein, and Th. A. G. Steemers. "Hyperspectral Imaging – Concepts and Potential in Paper and Writing Durability Research." *Proceedings Durability of*

Paper and Writing. Ljubljana, 2004. <http://www.art-innovation.nl/fckfiles/file/Downloads/Articles/2004/2004_Hyperspectral_Imaging.pdf>.

Scholten, J. H., M. E. Klein, Th. A. G. Steemers, and G. de Bruin. "Hyperspectral imaging - a novel nondestructive analytical tool in paper and writing durability research." *Proceedings Art 05*. Lecce, 2005. <http://www.art-innovation.nl/fckfiles/file/Downloads/Articles/2005/2005_Hyperspectral_Imaging.pdf>.

Smith, K. J. *Evaluation of Whiteness and Yellowness in Color Physics for Industry*. 2nd Edition, Society of Dyers and Colourists, 1997.

Tilley, Richard. *Colour and the Optical Properties of Materials: An Exploration of the Relationship Between Light, the Optical Properties of Materials and Colour*. New York: John Wiley & Sons, 2000.

Vaarasalo, J. "Optical Properties of Paper in Papermaking Science and Technology." 17. *Pulp and Paper testing*. Eds. J. Gullichsen and H. Paulapuro. Helsinki: Fapet Oy, 1999: 162–181.

Van der Reyden, Dianne. "Recent Scientific Research in Paper Conservation." *The Journal of the American Institute for Conservation* 31 (1992): 117–138. <http://www.si.edu/MCI/downloads/RELACT/paper_properties_degrade.pdf>.

Wattenbach, Wilhelm. *Das Schriftwesen im Mittelalter*. Zweite, verm. Auflage. Leipzig: Hirzel, 1875.

Wunderlich, Christian-Heinrich. "Geschichte und Chemie der Eisengallustinte. Rezepte, Reaktionen und Schadwirkung." *Restauro* 100 (1994): 414–421.

Zerdoun Bat-Yehouda, Monique. *Les encres noires au Moyen Âge. Jusqu'á 1600*. Paris: Éd. du Centre Nat. de la Recherche Scientifique (Documents, études et répertoires publ. par l'Institut de recherche et d'histoire des textes), 1983.