Analysis of Scribe Ink Patterns in Medieval Manuscripts Using Hyperspectral Imaging

David Messinger
Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology
Rochester, NY, USA
dwmpci@rit.edu

Andrew Beeby
Department of Chemistry
Durham University
Durham, UK
andrew.beeby@durham.ac.uk

***Abstract*—Multispectral imaging has been used in the study of historical artifacts, particularly medieval manuscripts with faded text, for the past 20+ years. Recently, hyperspectral imaging systems are finding utility in similar studies, particularly those where subtle material differences can be leveraged to learn about the materiality of the artifacts. Previous work has largely focused on the utility of hyperspectral imaging for accurate material identification. Here we show how HSI can be used to study the process by which an artifact was created. We present a methodology using hyperspectral imagery to measure subtle variations in the ink of visible text of medieval manuscripts. We demonstrate that these differences show a semi-regular pattern that can be attributed to the process used by the medieval scribe, notably the dipping, and re-dipping, of the quill pen used to write the text. Results are shown for two manuscript folios imaged in the Durham University Library, Durham, UK.**

***Index Terms*—hyperspectral imaging, manuscript studies, heritage science**

I. Introduction

Multispectral imaging has long been used to study medieval manuscripts with a primary focus on uncovering faded or erased text [1]. More recently, multispectral and hyperspectral imaging systems have been used to study the materiality of medieval artifacts, including maps and manuscripts [2]–[4]. These approaches leverage the spectral resolution and coverage of hyperspectral system to identify materials based on their spectral signatures, allowing for pigment identification and mapping. Howell & Snijders (2020) [5] describe how hyperspectral imaging systems can be used in libraries, but with a focus equipment considerations for libraries and general data processing. Matsoukova and Pavelka (2013) [6] provide an overview of the use of hyperspectral imaging for cultural artifacts with a focus on painted works of art, while Mandal, *et al.*, (2023) [7] describes the use of hyperspectral imaging and k-means clustering to perform pigment clustering in Edvard Munch's *Self Portrait*.

However, the sensitivity of hyperspectral systems can be used to identify other material features on the artifact. Here, we demonstrate the utility of hyperspectral imaging to study a different aspect of the materiality of a medieval manuscript, related not to the identity of the materials used in its creation but to identifying patterns in *how* the manuscript was created. The process of writing the text on a particular manuscript page was laborious, typically using a quill pen and ink, most commonly iron gall ink. This required not only continuous care and management of the quill pen used, but also the regular dipping and re-dipping of that pen in the ink receptacle. We show how this process produces variability in the amount of ink deposited in particular characters, and how that variation may be visualized in hyperspectral imagery.

The images were collected with the Visible-NIR hyperspectral imaging system at Durham University, UK, described below. Two manuscripts from the Durham Library were imaged in this project and the methodology used to analyze the data, as well as the results of that analysis are shown below.

This paper is presented in the following way. Section II presents the methodology used here, including descriptions of the instrumentation, artifacts, and analysis methods developed. Section III presents the results of the analysis, and Section IV summarizes the work.

II. Methodology

A. Hyperspectral Instrumentation

The hyperspectral imaging system used is the Headwall VNIR E-Series which collects images from approximately 400 - 1000 nm with a cross-track width of 1600 pixels in a pushbroom configuration. Each image contained 374 spectral bands and all were used in this analysis.

In this setup, the imaging system is fixed on a camera stand in a nadir-pointing geometry and the artifact is placed on a translation stage such that it is scanned underneath the camera to build up the image line-by-line. The imaging setup is shown in Figure 1 with the pushbroom camera imaging a strip illuminated by a light source as the object translates through the field of view. Note that there is a second camera shown in this image, a SWIR HSI instrument, which was not used in this study.

Dark images were collected along with images of a Spectralon target, to create a "white image," and the final data cubes were directly calibrated to estimated surface reflectance.

B. Artifacts and Hypothesized Physical Model

Two folios from separate manuscripts in the Durham University Library Special Collection were imaged using the HSI system. BAM-SEL-6 is a liturgical psalter originally from the

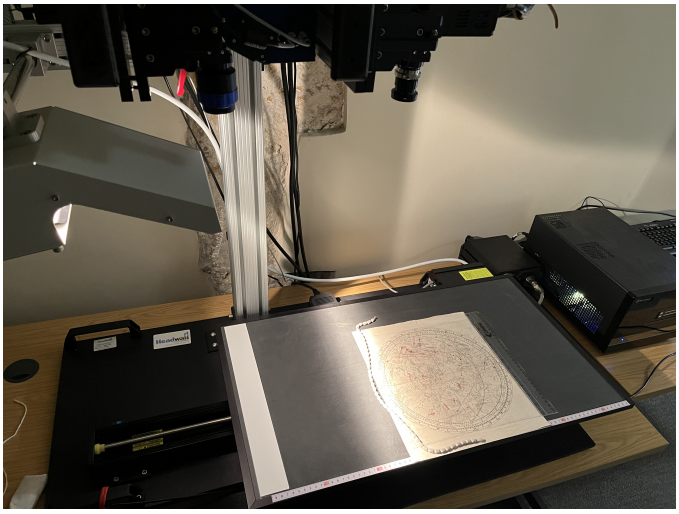


Fig. 1: Setup of the HSI imaging system and a representative artifact. The pushbroom instrument images an illuminated strip as the artifact is translated through the field of view.

Bamburgh Library, approximately 12th century. Results for folio 33 verso are shown here. USH-6 is originally from the Ushaw College Library, now housed at Durham Library. Folio 3 recto was analyzed for this work. Images of the two folios can be found in Figures 2a & 4a, respectively.

The ink primarily used in medieval texts was an iron gall based ink that becomes progressively more transparent in the NIR, depending on ink density. Scribes used quill pens to write the text, requiring not only continual care and maintenance of the pen, but also regular dipping of the pen in the ink receptacle [8]. Our hypothesis is that the rate of deposition of the ink, which is held in the quill pen largely by meniscus, or capillary, forces, varies with the amount of ink in the pen. Previous work has shown that for viscous liquids like iron gall ink, there is a flow rate dependency that can slow down the deposition of ink [9]. Here, we hypothesize that as the volume of ink in the quill decreases, the rate of deposition changes leading to variability in the amount of ink deposited onto the page as the scribe wrote the text. By identifying subtle differences in the ink density across the written text we seek to describe the spatial frequency of ink dipping used by the scribe as they were writing the text.

C. Analysis Methods

In some manuscripts, variations in ink density can be seen with simple visual inspection, indicating that the scribe was not particularly careful in applying the ink evenly. In others, simple visual inspection of images collected in the near-IR ($\approx$ 750 nm or longer) can reveal ink density differences, as the transparency of the ink in the NIR is related to the ink density. However, these visual inspection techniques do not always show the ink variation.

Here, we consider two manuscripts for which the ink darkness is visually uniform. To assess variations in ink density, which we hypothesize are related to the scribe's writing and ink dipping pattern, we propose three analysis methods: we apply a relatively simple data transformation away from the reflectance space, a novel "spatial block" Principal Components Analysis (PCA) to the data to highlight subtle variations in the signatures of the ink, and a nonlinear data transformation based on Laplacian Eigenmaps.

Based on our understanding of how the ink is deposited onto the page, we can assume that the reflectance of the ink, $\rho(\lambda)$, is non-linear with the amount of ink present. However, in the "attenuation space" the data will be more linear with the ink density. Here, we convert the data from reflectance to the attenuation space via a simple transformation given by

$$\rho(\lambda) \rightarrow \log\left(\frac{1}{\rho(\lambda)}\right). \tag{1}$$

In this space, the differences in ink density, which can be manifest in the radiance and reflectance data as nonlinear features, can sometimes be more easily distinguished by visual inspection, particularly in the NIR.

In addition to converting the data to the attenuation space, as mentioned above, two processing methods were developed to derive this spatial block approach to PCA. In both, we seek to highlight the subtle variations *within the text itself* to drive the PCA transformation. In the first case, a single block of pixels containing only black text on parchment was identified as a Region of Interest (ROI), and the mean and covariance matrix for just those pixels were computed and used to derive the PCA transformation. That transformation was saved and then applied sequentially to other blocks of text. For the pixels in this ROI, the primary sources of variability will be the differences between the text and the parchment and the variability within the text. These variations will dominate the covariance matrix and so the derived PCA eigenvectors will be aligned with these variations and the subtle differences in the text can be highlighted with a simple false color image.

A second approach for even more subtle text variations creates a text mask, extracting pixels that are only on the text and do not contain measurements of the background parchment, to estimate the mean and covariance matrix used to develop the PCA transformation. In this case, the variability in the data will be dominated by variations in just the text that may have been suppressed by the spectral differences between the text and the parchment. Once this transformation is derived, it can be applied to an entire block of text and the results visualized.

Here we show two visualizations of the PCA results, along with a simple grayscale visualization of the attenuation space approach. The first PCA visualization uses a simple false color combination of three bands in the transformed image blocks layered on top of the RGB of the original manuscript. In this case, pixels of similar visual color represent materials with similar spectral response, while pixels of different visual color represent materials with distinct spectral differences.

The second visualization uses a single PCA band which highlights the spectral differences solely in the ink. This

grayscale image is then placed "on top" of the original RGB as a transparent layer (here shown as 50% transparent) again highlighting the subtle differences in the ink optical density on the text.

The third approach applies the Laplacian Eigenmaps (LE) non-linear, manifold learning transformation to the data. LE uses a graph-based model of the data to derive an estimate of the non-linear manifold on which the data lie in the spectral domain [10], [11]. Once this manifold is determined, the original data can be transformed into the manifold coordinates where the data will be more linear.

Functionally, to create the LE transformation one creates a nearest neighbor similarity graph of the data representing the connected structures in the spectral space. Then, through solving an eigenproblem, the embedding is derived and the data can be transformed into the new, linear space. Visualizations of the data in that space are analogous to those in Principal Components Analysis. While this manifold learning algorithm is computationally complex, it does provide an inherently non-linear approach to analysis of this dataset.[1]

III. Results

Results for this approach to analyzing subtle differences in ink density are shown in Figures 2 & 4. For the Bamburgh Library manuscript (BAM-SEL-6 f33v), the RGB image (Figure 2a) shows no visual differentiation in ink density across the black text. For this artifact, we show the band at 850 nm after being converted to the attenuation space as described above (Figure 2b). Here, the brightness variations in the text are indicative of the variation in the ink density. In this case, for the PCA analysis, a single block of text on parchment was identified and used to develop the PCA transformation, which was then applied sequentially to the other blocks of black text. The result is shown in Figure 2c as blocks of false colored images overlaid onto the RGB image. Here, the background parchment is generally "blue" while variations between "green" and "red" in the false color blocks indicate variations in the ink spectral features, representing subtle variations in the ink density. We note that these color variations are also correlated to brightness differences in the attenuation space image. One can see how the ink changes across the written words, with sometimes sharp jumps at individual characters, indicating positions where the scribe likely re-dipped the quill pen being used.

Figure 3 shows a comparison of all three analysis methods on a portion of the text in high resolution. Figure 3b shows the second band after application of the PCA transformation to the image, while Figure 3c shows the second band after application of the LE transformation. The same data set converted to the $\log 1/\rho(\lambda)$ data is shown in Figure 3d. Interestingly, the LE transformation highlights subtle differences around the edges of the characters, a phenomenon that requires further investigation. However, we can see that across all three analysis methods similar patterns in the bright and dark cycles of the ink are apparent, further supporting our interpretation of the results.

Figure 4 shows the application of the spatial block PCA technique on the folio from Ushaw College (USH-6 f3r). Again, the RGB image of the folio is shown in Figure 4a. Here we see significant "cockling" on the page, spatial wrinkling of the parchment that creates brighter and darker variations in the image due to the bidirectional reflectance distribution function (BRDF) effects of the parchment. Fortunately, regions at the top and bottom right corner of the page are sufficiently flat to remove this source of variability from the analysis.

The result here (Figure 4b) is shown as a transparent layer over the RGB. After transformation, PCA band 3 was used as a 50% transparent layer to highlight differences in the ink density of the analyzed regions. Again, subtle differences in the ink density pattern across the written text are now easily viewed showing the pattern of dipping, writing, and re-dipping the pen.

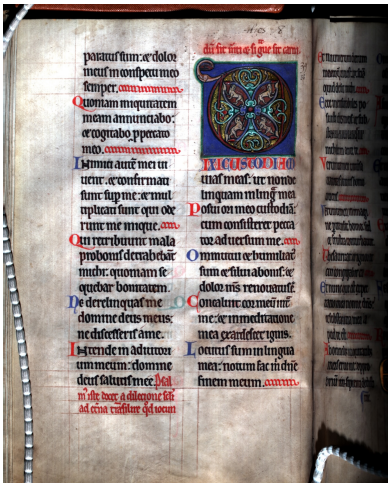
These two results, across two significantly different manuscript folios, demonstrate the ability of HSI to discriminate these subtle differences in ink density on the page, which in turn leads to better understanding of the process by which the manuscript was created.

IV. Summary and Conclusions

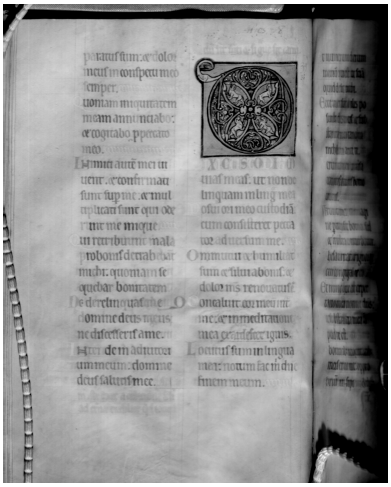
Here we demonstrate a novel approach to using hyperspectral imagery to analyze medieval manuscripts. While previous work has used HSI to enhanced faded or erased text or perform pigment analysis, here we use the enhanced spectral resolution of HSI systems over MSI systems to study ink density on the written text and infer the scribe pattern of writing and ink dipping used.

This paper presents multiple contributions. The first is a novel method for using the standard Principal Components Analysis transformation on hyperspectral imagery. Instead of the common practice of developing the transformation on the entire image, or commonly the entire image after cropping out items such as calibration targets, here we use two approaches to selectively develop and apply the transformation on specific pixels of interest, highlighting subtle variations in their spectral signatures. When using a "global" approach to developing the transformation, these subtle variations often contribute very little to the covariance matrix of the data, given the much more significant variability between materials of different classes with large spectral differences between them. Here, we are interested in developing the transformation to highlight *within material differences*, and so need to derive the associated statistics for only those pixels of interest. We present two approaches to this problem, one using specific spatial blocks as regions of interest that only contain pixels of ink and parchment, and one that isolates only text pixels using a spatial masking approach. These pixels are used to derive the transformation which can then be applied to large spatial regions of the image. Additionally, we show how a simple conversion of the data to the attenuation space (eq. 1) can

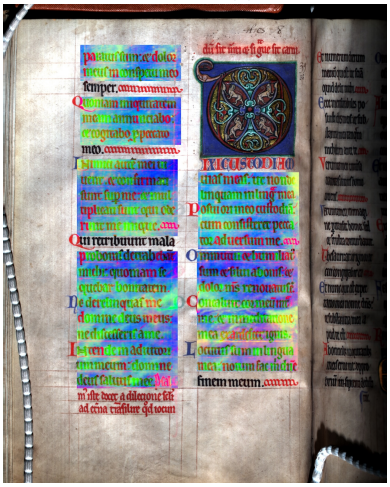
[1]Here, the scikit-learn SpectralEmbedding python package was used for the LE results.



(a) RGB image of the folio.

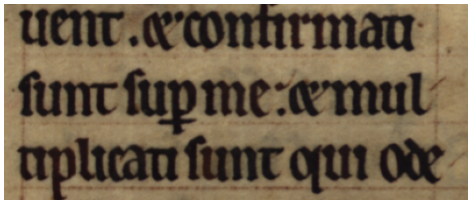


(b) 850 nm band shown in the attenuation space.



(c) Spatial block PCA of the folio.

Fig. 2: (a) RGB image of BAM-SEL-6 f33v (b) 850 nm band shown in grayscale after conversion to the attenuation space. (c) Spatial block PCA results shown in false color highlighting subtle differences in the ink thickness. Images courtesy of Lord Crewe's Charity.



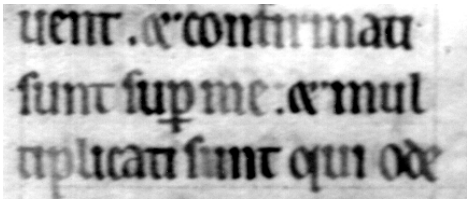
(a) RGB image of the text subset used to for comparison.



(b) PCA band 2.

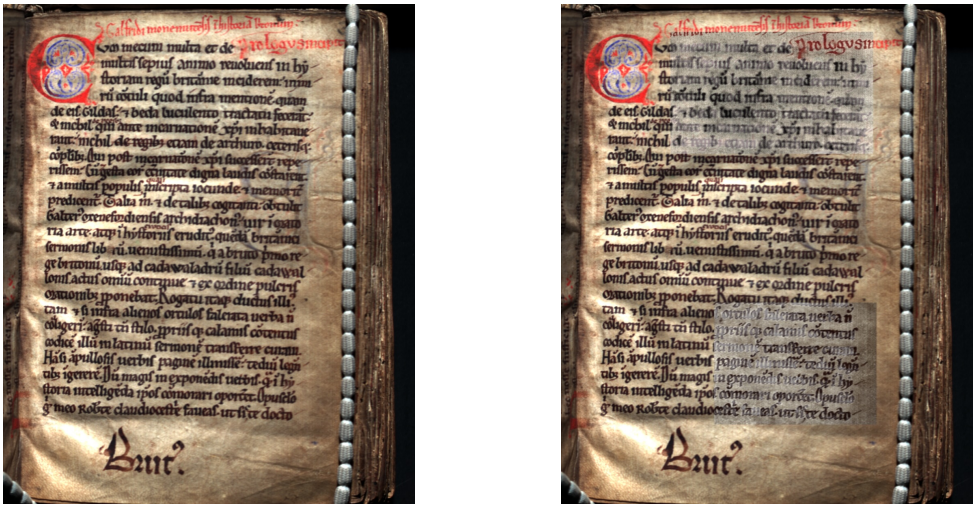


(c) LE band 2.



(d) $\log 1/\rho(\lambda)$ (inverted for comparison).

Fig. 3: (a) RGB image of a subset BAM-SEL-6 f33v for comparison (b) PCA band 2. (c) Laplacian Eiganmaps band 2. (d) $\log 1/\rho(\lambda)$ at 850 nm. Images courtesy of Lord Crewe's Charity.



(a) RGB image of the folio. (b) Spatial block PCA of the folio.

Fig. 4: (a) RGB image of USH-6 f3r (b) Spatial block PCA results shown shown as a transparent layer highlighting subtle differences in the ink thickness. Images courtesy of Ushaw: Historic House, Chapels, and Gardens.

highlight spatial variability in the ink density. Finally, we apply a graph-based manifold learning model to hyperspectral data used in Heritage Science. While these methods have been used in remote sensing in the past, the application here to very high resolution imagery of a medieval manuscript is new and provide a pathway to future studies of complex radiometric process in this type of imagery.

The second contribution is in the use of HSI to study the materiality of the manuscript, not just through analysis of the specific inks or pigments, but in inferring the *process* by which the scribe used specific tools to create the manuscript. These results leverage the sensitivity of the imaging system to detect subtle variations in the ink density which can be inferred to be related to the physical process of dipping the pen in the ink bottle, writing a few characters or words of text, and then re-dipping the pen and repeating. The physical process of ink flowing through the pen produces subtle variations in the ink density which can be measured by the hyperspectral system, as shown here.

These results will hopefully provide insights to medieval scholars about not just the *content* of the manuscripts, but also the *context* of how the manuscript was created. We also note that given the appropriate sensitivity, spectral coverage, and spectral resolution, these approaches can be applied to multispectral imagery as well, producing similar results with an appropriate imaging system.

Acknowledgment

The authors would like to thank Prof. Richard Gameson for his assistance in identifying and accessing the manuscripts used in this study. The authors also thank the UK Arts and Heritage Research Council for CapCo funding to acquire the hyperspectral imaging system. Images reproduced with permission from Lord Crewe's Charity and Ushaw: Historic House, Chapels and Gardens.

References

[1] R. Easton, K. Knox, W. Christens-Barry, K. Boydston, M. Toth, D. Emery, and W. Noel, "Standardized system for multispectral imaging of palimpsests," in *Computer Vision and Image Analysis of Art*, vol. 7531. SPIE, 2010.

[2] D. Bai, D. W. Messinger, and D. Howell, "Hyperspectral analysis of cultural heritage artifacts: pigment material diversity in the Gough map of Britain," *Optical Engineering*, vol. 56, no. 8, pp. 081 805–081 805, 2017.

[3] ——, "A hyperspectral imaging spectral unmixing and classification approach to pigment mapping in the Gough & Selden maps," *Journal of the American Institute for Conservation*, vol. 58, no. 1-2, pp. 69–89, 2019.

[4] R. Gameson, A. Beeby, F. Fiorillo, C. Nicholson, P. Ricciardi, and S. Reynolds, *The Pigments of British Medieval Illuminators: A Scientific and Cultural Study*. Archetype Books, 2023.

[5] D. Howell and L. Snijders, *Conservation Research in Libraries*. De Gruyter Saur, 2020.

[6] E. Matouskova and K. Pavelka, "Hyperspectral approach to a non-invasive survey of cultural heritage objects," *IEEE WHISPERS 5th Workshop*, vol. 5, 2013.

[7] D. Mandal, H. Deborah, S. George, and J. Hardeberg, "Unsupervised clustering for works of art using hyperspectral imaging: A case study on edvard munch's self-portrait (1905)," *IEEE WHISPERS 13th Workshop*, vol. 13, 2023.

[8] R. Clemens and T. Graham, *Introduction to Manuscript Studies*. Cornell University Press, 2008.

[9] S. Cai and B. Bhushan, "Meniscus and viscous forces during normal separation of liquid-mediated contacts," *Nanotechnology*, vol. 18, 2007.

[10] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," *Advances in Neural Information Processing Systems*, vol. 14, 2001.

[11] J. Benedetto, W. Czaja, J. Dobrosotskaya, T. Doster, and K. Duke, "Spatial-spectral operator theoretic methods for hyperspectral image classification," *International Journal on Geomathematics*, vol. 7, 2016.