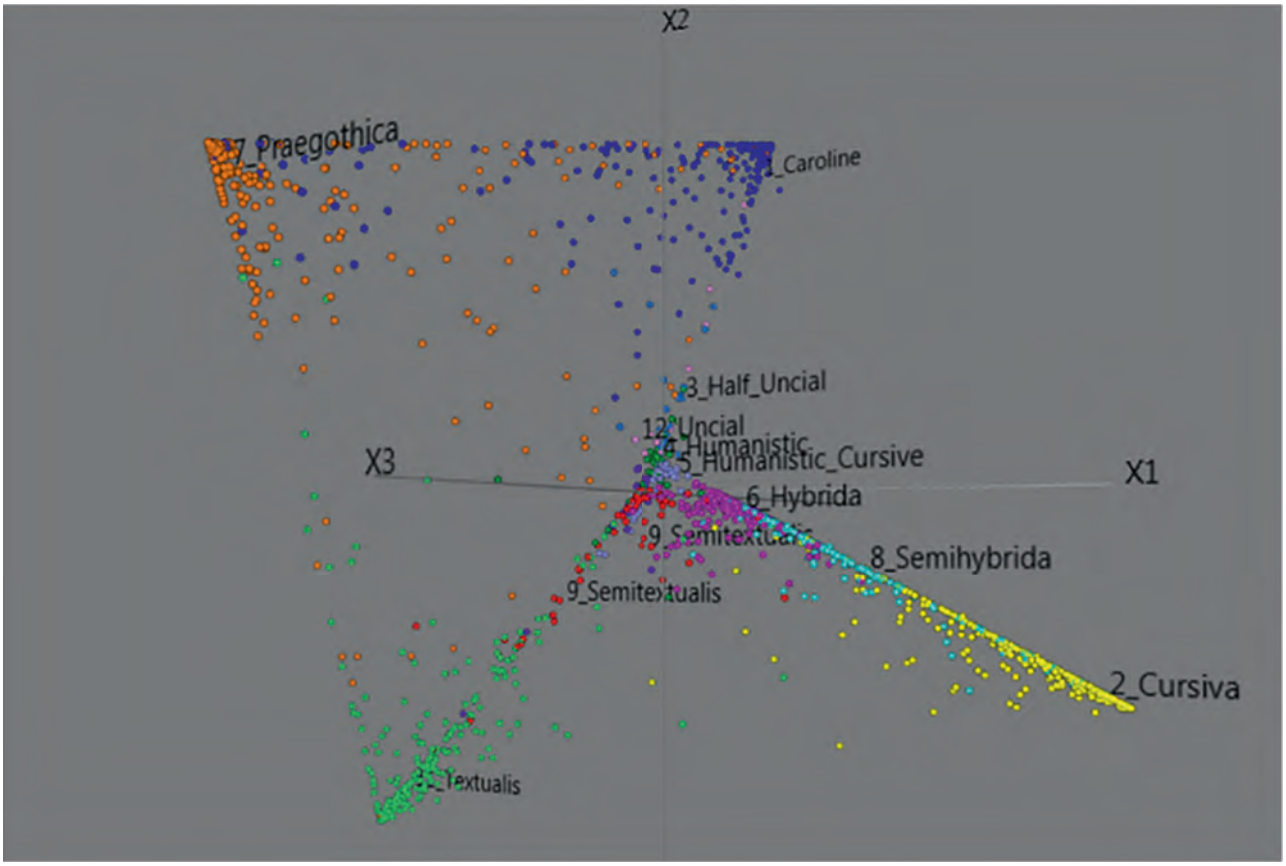
manuscript cultures

Hamburg | Centre for the Study of Manuscript Cultures ISSN 1867–9617



Publishing Information

Natural Sciences, Technology and Informatics in Manuscript Analysis
Edited by Oliver Hahn, Volker Märgner, Ira Rabin, and H. Siegfried Stiehl

Proceedings of the third *International Conference on Natural Sciences and Technology in Manuscript Analysis* and the workshop *OpenX for Interdisciplinary Computational Manuscript Research* that took place at the University of Hamburg, Centre for the Study of Manuscript Cultures, on 12–14 June 2018.

Editors
Prof. Dr Michael Friedrich
Universität Hamburg
Asien-Afrika-Institut
Edmund-Siemers-Allee 1/ Flügel Ost
D-20146 Hamburg

Tel. No.: +49 (0)40 42838 7127
Fax No.: +49 (0)40 42838 4899
michael.friedrich@uni-hamburg.de

Prof Dr Jörg Quenzer
Universität Hamburg
Asien-Afrika-Institut
Edmund-Siemers-Allee 1/ Flügel Ost
D-20146 Hamburg
Tell. No.: +49 40 42838 - 7203
Fax No.: +49 40 42838 - 6200
joerg.quenzer@uni-hamburg.de

www.csmc.uni-hamburg.de

Editorial Office
Dr Irina Wandrey
Universität Hamburg
Centre for the Study of Manuscript Cultures
Warburgstraße 26
D-20354 Hamburg
Tel. No.: +49 (0)40 42838 9420
Fax No.: +49 (0)40 42838 4899
irina.wandrey@uni-hamburg.de

Layout
Miriam Gerdes

Cover
Image of ICDAR2017 Tensmeyer's distance matrix (axes 2 and (1 and 3)), see article by Dominique Stutzmann, Christopher Tensmeyer and Vincent Christlein in this volume.

Translation and Copy-editing
Amper Translation Service, Carl Carter, Fürstenfeldbrück

Print
AZ Druck und Datentechnik GmbH, Kempten
Printed in Germany

ISSN 1867–9617

CONTENTS

ARTICLES

Editorial

Natural Sciences and Technology in Manuscript Analysis

The Centre for the Study of Manuscript Cultures (CSMC) in Hamburg strives to enhance the field of interdisciplinary manuscript studies by providing a forum for dialogue between the humanities, the natural sciences and informatics. This dialogue leads to well-defined research questions and frequently provides scientifically grounded answers to questions that could not be solved by historical and philological methods alone.

In recent years, considerable progress has been made in multi- and hyperspectral imaging to recover erased texts in palimpsests. However, digital imaging and image processing still require serious research and development, not to mention the establishment of standard procedures for imaging protocols and benchmarking, for example. In addition to conducting non-destructive material analyses of pigments and dyes, which have become standard in modern studies of illuminated manuscripts, CSMC's interdisciplinary teams are currently working on establishing guidelines for studies of writing inks. Finally, digital image processing and analysis are also gaining recognition in the fields of palaeography and codicolog .

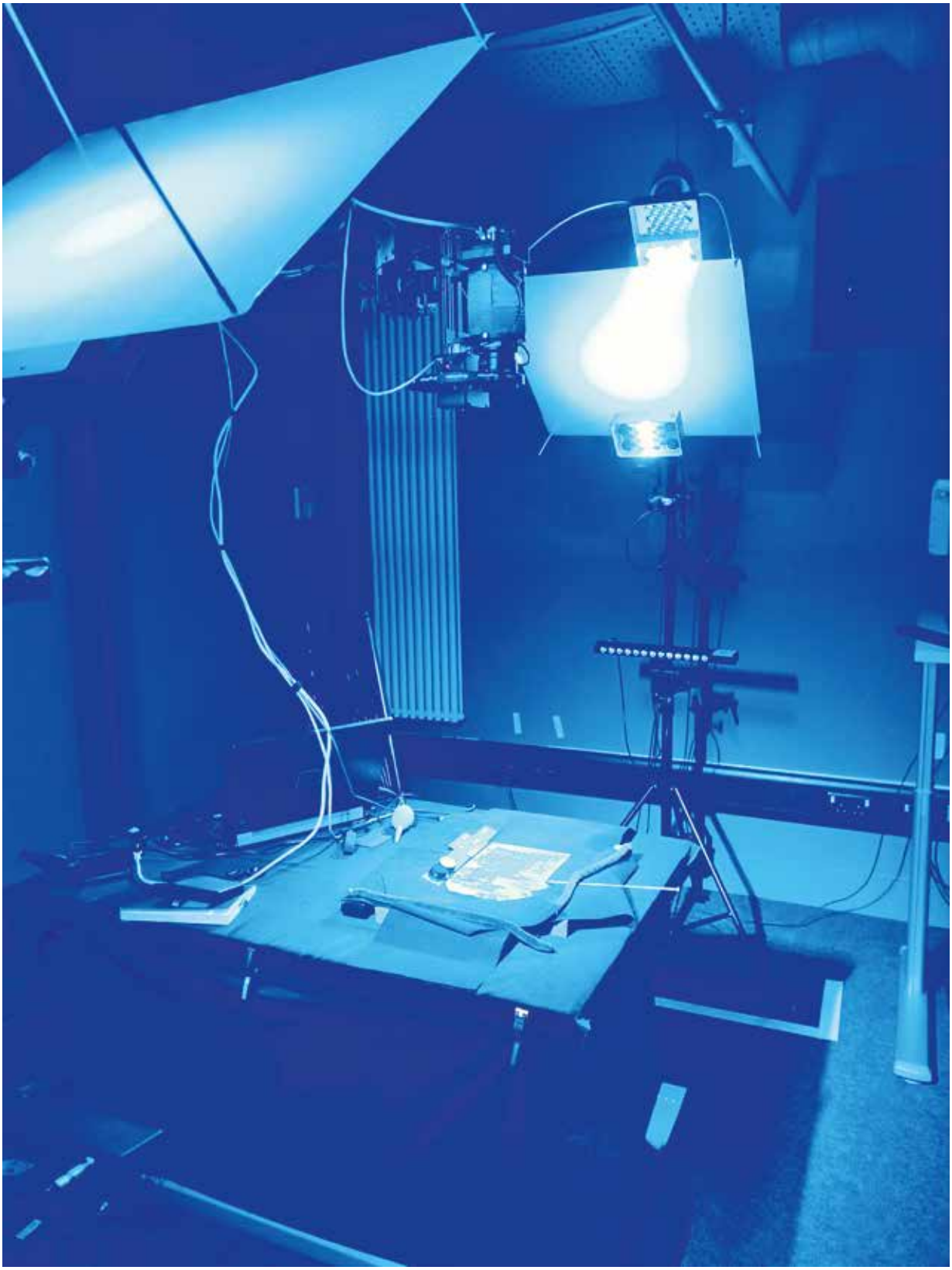
The third *International Conference on Natural Sciences and Technology in Manuscript Analysis* was held on 13–14 June 2018. Like previous conferences, it brought scientists and scholars together who were engaged in this field of interdisciplinary research and provided a forum for discussion and for presenting new methods and findings Some of the research presented in this volume was conducted at SFB 950 'Manuscript Cultures in Asia, Africa and Europe' of the University of Hamburg and sponsored by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG)

A two-day pre-conference workshop entitled *OpenX for Interdisciplinary Computational Manuscript Research* took place on 12–13 June directly before the conference. Scientists from leading European research groups presented and discussed their research results (marked by *) in the context of the ever-growing adoption of the Open Science paradigm in computational manuscript research. In this vein, OpenX not only refers to open-source software and open data in the classic sense, but to open services and platforms now on the Web and consequently involves legal implications such as copyright, licensing and data-privacy issues. The scientific, technical, and local organisation was done by the CSMC team (SFB 950, Scientific Service Project Z03 and Department of Informatics, University of Hamburg) headed by H. Siegfried Stiehl of Z03 in close cooperation with Andreas Fischer (DIVA, University of Fribourg, Switzerland), Vinodh Rajan Sampath (Department of Informatics, University of Hamburg), and Marcel Gygli (Würsch) (then DIVA, University of Fribourg, Switzerland). Funding by the German Research Foundation and support by DIVA and the Department of Informatics at the University of Hamburg are gratefully acknowledged.

This issue of manuscript cultures contains a selection of peer-reviewed papers presented at the workshop and conference along with some additional contributions presenting unique case studies. The articles were solicited for original research work illuminating the role of the natural sciences, technology and informatics in manuscript analysis. All in all, this special issue represents the state of the art, illustrating how different techniques and methodologies can be successfully applied to analytical and computational investigations in the field of manuscript analysis. We hope it will help to incorporate the natural and applied sciences and informatics into the field of manuscript studies

We would like to express our gratitude to all the authors for submitting up-to-date papers and to all the anonymous reviewers for their valuable and constructive comments.

Oliver Hahn, Volker Märgner, Ira Rabin, and H. Siegfried Stiehl



The CSMC Multispectral Imaging System in operation – recovering illegible writing of the *Manichaean Papyrus Collection* at the Chester Beatty Library in Dublin, Ireland (https://chesterbeatty.ie/conservation/revealing-the-mystery-of-mani/). This very illumination setting has proven to be key for the recovery of faded ink on these particular papyrus fragments. By projecting blue, 450nm, light and the use of a Wratten R25 long pass fil er in front of the lens – we were able to capture the fluo escence of the papyrus and thus finally see the dark t aces of the ink which now appeared in higher contrast and enabled us to read the text with ease.

Article

On Avoiding Segmentation in Handwritten Keyword Spotting: Overview and Perspectives

Marçal Rusiñol | Barcelona

Abstract

In this paper we focus on the problem of handwritten keyword spotting when we want to avoid a layout segmentation step. We strongly believe that bypassing the segmentation step is a must in the context of historical document collections where achieving a perfect word or text line segmentation is unfeasible. So, architectures that dismiss the segmentation step are a clear asset in the context of historical documents. We will provide an overview of the various state-of-the-art approaches and conclude by discussing the promising results in the field and the remaining challenges

1. Introduction

Nowadays, in order to provide access to the contents of digital document collections, their textual contents are stored in electronic format so that any search engine is able to index the data corpus and users are able to perform textual searches. When dealing with large collections, automatic transcription processes are used, since manual transcription is not a feasible solution. In the context of digital collections of historical manuscripts, handwriting recognition strategies are applied to achieve an automatic transcription of the handwritten text. However, handwriting recognition often does not perform satisfactorily enough in the context of historical documents. Documents with severe degradations or using ancient glyphs can make recognizing individual characters difficult, and the lexicon definition and language modelling steps are not straightforwardly solved in such a context. In such a scenario, the particular application of handwritten keyword spotting gained attention.

Handwritten keyword spotting is defined as the pattern recognition task aimed at locating and retrieving, from a collection of manuscript images, a particular keyword requested by the user, without explicitly transcribing the whole corpus or having the text in electronic format at hand. Such an application is thus particularly interesting in scenarios in which an automatic text transcription is deemed likely to fail by resulting in too many errors and when budget restrictions make manual transcription of the corpus' contents unfeasible.

In fact, the term *keyword spotting* was not coined in the context of historical manuscripts, but first appeared in the speech analysis community in the early seventies (Vintsyuk 1971). For its purposes, the application focused on locating the timestamps in audio signals in which a specific keyword might be uttered. Today, such technology is used by personal digital assistants such as Alexa, Siri, and Google's Home device, but, of course, national security agencies have employed keyword spotting systems to search through hours and hours of recorded conversations and isolate utterances of suspicious keywords.

In the early nineties, some preliminary works that used the term spotting in document images appeared, for instance the papers by Chen et al. 1993 and by Kuo and Agazzi 1994. Those works focused on dealing with typewritten text. At that time, however, the performance of OCR engines was already starting to be good enough to be used reliably. So, the only interest in using spotting techniques instead of recognition pipelines was computational efficienc . But, in the mid-nineties, Manmatha et al. (1996 and 1997) and Keaton et al. (1997) first used the term in the context of handwritten document images. Here, the recognition approaches were far from successful, and there was strong motivation to use spotting techniques. It was clearly an exciting research line that might have an important cultural impact. Handwritten keyword spotting systems will provide accessibility to tons of digitized manuscripts that were previously doomed to stay locked in vaults.

2. The conventional pipeline

From those early works until the end of the first decade of the 2000s, all the proposed approaches followed the same pipeline. First, a layout analysis step was needed to

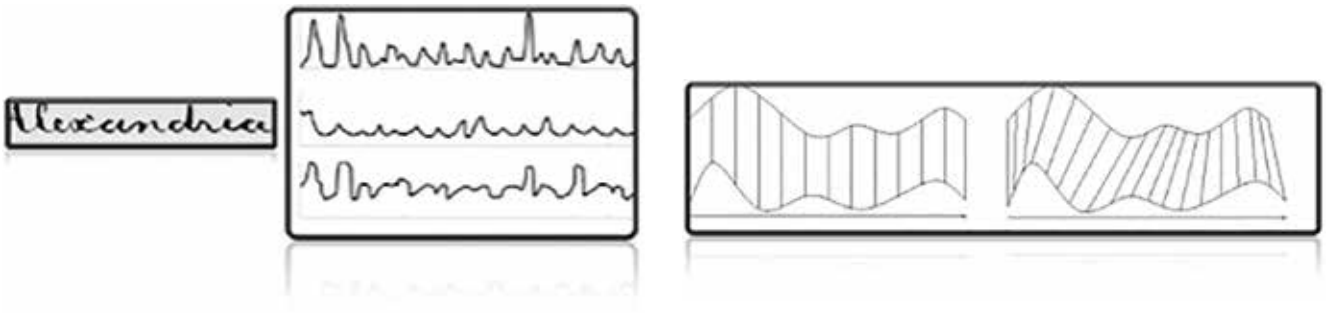


Fig. 1: Projection profiles and TW for matching individual words, presented in 2007 by Rath and Manmatha.

segment the document pages into physical elements such as paragraphs, text lines or words. At that time, most of the approaches worked at the level of the word. In a second stage, a descriptor was computed for each of the already segmented words. Such a descriptor encoded the word's shape. Finally, when the system's user made a query, giving an example of the keyword he was looking for, a distance between the query descriptor and all the descriptors from the dataset was computed to yield the final ranked list

This pipeline presents several drawbacks. On the one hand, the segmentation step is not always straightforward and is usually error-prone. In fact, although word and text line segmentation is a quite mature research topic, it is far from being a solved problem in critical scenarios dealing with handwritten text and highly degraded documents. Any segmentation errors will affect the subsequent word representations and matching steps. On the other hand, working with the query-by-example paradigm forces the user to already browse the document pages searching for an instance of the keyword he wants to query. Finally, such an approach is hardly scalable to large collections, because, at the retrieval stage, one has to compute the distances between the query and all the words in the collection.

The team of Rath, Manmatha et al. continued to work on the idea of handwritten keyword spotting, specificall on the manuscript letters of George Washington (Rath, Manmatha et al. 2002). But it was the seminal work by Kolcz et al. (2000) that achieved a breakthrough in the handwritten keyword-spotting domain by proposing the use of the Dynamic Time Warping (DTW) method (often used in speech analysis) for nonlinear sequence alignment. The use of DTW together with profile features was later popularized by the well-known works by Rath and Manmatha (2003 and 2007), and many flavours of DTW-based handwritten keyword spotting methods have appeared since their publications. However, the use of DTW-based solutions was extremely computationally demanding, and thus many other authors devoted their efforts to proposing less-demanding feature-vector representations and matching schemes, for instance, the binary representation proposed by Zhang et al. in 2003.

In this paper, we will focus on the segmentation issue, provide an overview of the various state-of-the-art approaches, and conclude by discussing the promising results in the field and the remaining challenges

3. First attempts at producing segmentation-free systems

Dependence on good word segmentation motivated the researchers of the keyword-spotting domain to move recently towards completely segmentation-free methods. One of the first groups of researchers to approach this problem, Leydier et al., in 2005, 2007 and 2009, proposed a keyword spotting methodology based on local keypoints. For a given query image, interest points were extracted and encoded by a simple descriptor based on gradient information. The keyword spotting was then performed by trying to locate zones in the document images with similar interest points. These retrieved zones were then filtered and only the ones sharing the same spatial configuration as the query model were returned.

The same idea was also used early by Rothfeder et al. in 2003 using corner features, and by Rusiñol and Lladós in

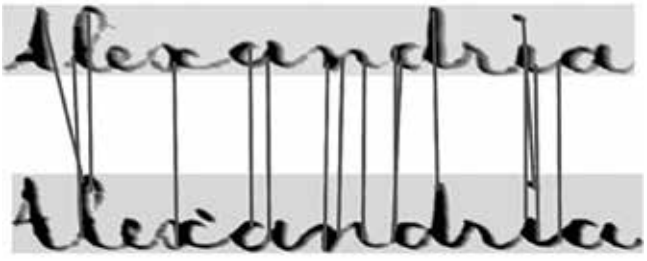


Fig. 2: Matching local features from handwritten words, from the work by Rothfeder et al. 2003.

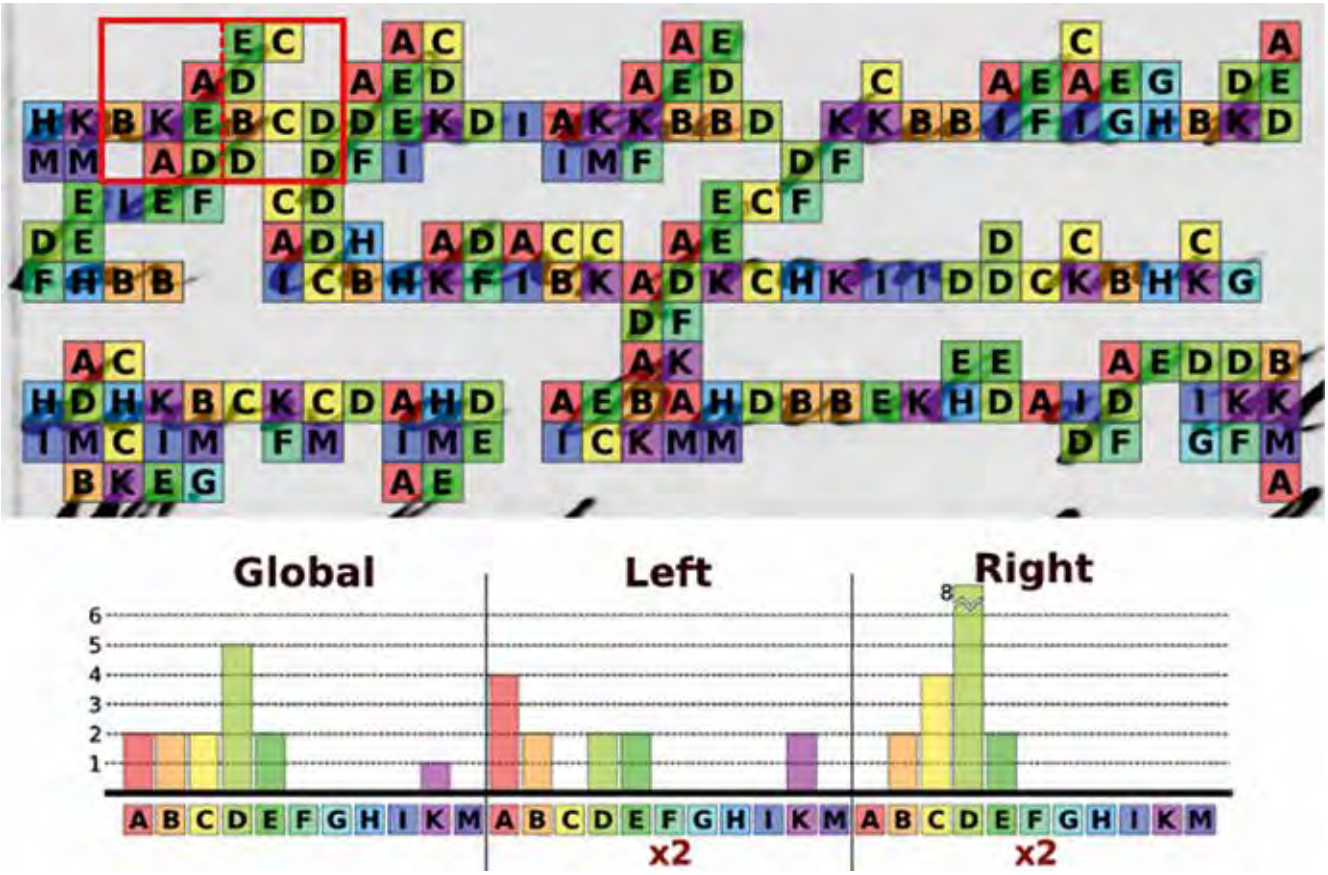


Fig. 3: Sliding window approach over a bag of visual words from Rusiñol et al. 2015.

2008 as well as by Zhang et al. in 2013 using well-known local descriptors such as SIFT or shape context to spot handwritten text.

However, directly matching local keypoints might be too computationally expensive when dealing with large datasets. A single page would easily yield tens of thousands of local descriptors that have to be matched with the query keyword snippet. Even though indexing and approximate-nearest-neighbours algorithms can help to deal with billions of such distances, it is clear that such approaches do not scale well. Thus, researchers started to apply other approaches, inspired by the latest developments in computer vision for problems such like facial and traffi sign recognition.

4. Sliding window methods

The term segmentation-free handwritten keyword spotting started to gain momentum around 2010 with the appearance of methods that used sliding window approaches over bag-of-visual-words (BoVW) descriptions. Usually, local shape features are computed densely over the whole page, and later a sliding window with a fixed geometry defines the different patches where such local descriptors are to be aggregated. The zones where the sliding window overlaps with the queried word should have a slightly more similar description than the perfectly segmented word. Examples of such an approach to handwritten keyword spotting include the works by Rusiñol et al. in 2011 and 2015, which used SIFT-powered BoVW descriptors over sliding windows; by Rothacker et al. in 2013, 2014 and 2015, which proposed to compute a HMM that sequentially analysed BoVW features; by Almazán et al. in 2014, which used a sliding window-approach based on HOG descriptors; by Ghosh and Valveny in 2015, which used a sliding window over Fisher Vector descriptors; and by Ghorbel et al., who proposed in 2015 the use of Haar-like features like those used for facial recognition in commodity cameras.

Such methods had an important advantage over the methods that tried to directly match local descriptors. By holistically describing image regions, the process of matching was drastically consolidated. The contents of the sliding windows were encoded with powerfully performant descriptors at that time, and, when combined with quantization and indexing schemes, the methods did scale up and were able to tackle large collections.

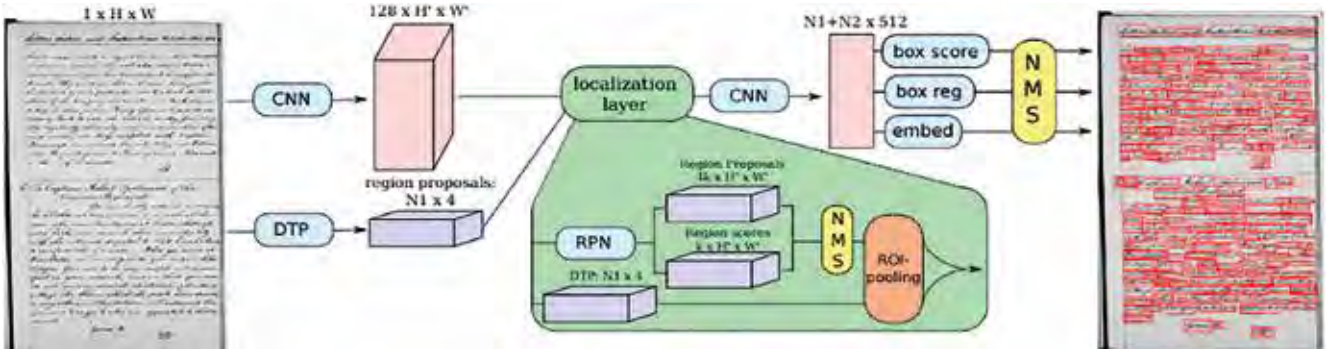


Fig. 4: End-to-end deep learning architecture for proposing and spotting word locations, by Wilkinson et al. 2017.

One might argue that using a sliding window *is* in fact some sort of very naïve word segmentation. Indeed it is, but the strength of such methods was that they were already engineered and designed from this point of view, and thus they assumed from the start that the word segmentation that they would have to process was inherently noisy.

Though in the first such proposed methods we saw that sliding windows struggled with either very long or very short words, the subsequent approaches, which either used multi-scale sliding windows (Rusiñol et al. 2015) or worked directly on an over-segmentation of the pages (Kovalchuk et al. 2014), started to be powerfully performant in such aspects.

At that point, the document image community started to address yet another important aspect that had scarcely been addressed before: performing queries by inputting typewritten electronic text, instead of querying the system by example. To deal with this problem, completely unsupervised methods were no longer fruitful, and the community started to use methods based on machine learning. At that time, deep learning re-emerged with the appearance of AlexNet, which competed in the ImageNet Large Scale Visual Recognition Challenge in 2012. The network achieved an error rate of 15%, almost 11 percentage points lower than that of the runner-up. The document image analysis community started to shift to methods based on deep learning, as well.

5. Methods based on Deep Learning

Like in other computer vision applications, the use of deep learning approaches completely superseded previous methods, notoriously improving retrieval performances on standard public datasets. Although most of the contributions in the literature still focus on segmentation-based pipelines, some works like the ones proposed by Rothacker et al. in 2017, Wilkinson et al. in 2017 and Ghosh and Valveny in 2017 have started to tackle the segmentation-free scenario. In such cases, either a neural network is trained to provide likely locations to find words – which, again, the punctilious reader might argue *is not* strictly speaking segmentation-free. Or the systems are trained end-to-end and are able to provide directly the locations of good examples of the words that the user has queried.

But, even though the results of methods based on deep learning are amazingly good, there is no such thing as a free lunch. Such methods *do* need a significant amount of annotated data to be trained properly. Even though researchers have started to apply synthetic text and data augmentation strategies (Krishnan and Jawahar 2016), they still need a significant amount of real data to fine-tune the system. This need for annotated data might be a strong deterrent to using such methods in real archival environments.

6. Discussion

First, note that most publications on handwritten keyword spotting that are based on a previous step of word segmentation are overly optimistic about their presented results. It is a common practice in our field to use the 'perfect' ground-truth segmentation provided by public datasets such as the George Washington letters or the IAM datasets (e.g. Aldavert et al. 2015). In that sense, we argue that the results provided by segmentation-free methods are usually closer to the reality than those provided by segmentation-based methods (Dey et al. 2016).

The second aspect we would like to highlight is the appropriate use of learning-based methods. Indeed, their performance is astonishing, and nowadays they have completely outperformed the rest of the state-of-the-art. But, in keyword spotting applications for libraries or archives, one can hardly assume that the users or owners of the system will be eager to provide hundreds of thousands of manually transcribed words in order to properly train the spotting system. In such

scenarios, we believe that it is best to avoid learning-based methods, and that sliding-window methods offer the important advantage that they can be used off the shelf.

But, should learning-based methods really be ruled out? So far, other aspects of handwritten keyword spotting are solved solely by learning-based methods. To provide the ability to query a system by strings instead of examples or to deal with different writers can be addressed only with machine learning systems. So the question remains.

Finally, we would like to also highlight that in recent years, public datasets such as the George Washington letters and the IAM database can be regarded as almost solved. Recent methods already claim average precision rates above ninety per cent in such datasets. In the near future, we will need larger and more challenging public annotated datasets in order to keep the handwritten keyword-spotting field alive and progressing.

ACKNOWLEDGEMENTS

This work was partially supported by the Spanish research project TIN2014-52072-P, the CERCA Programme / Generalitat de Catalunya and the project "aBSINTHE" –Fundación BBVA, Ayudas a Equipos de Investigación Científica, 2017. We gratefully acknowledge the support of the NVIDIA Corporation with the donation of the Titan X Pascal GPU used for this research.

REFERENCES

Aldavert, David, Marçal Rusiñol, Ricardo Toledo, and Josep Lladós. (2015), 'A study of Bag-of-Visual-Words representations for handwritten keyword spotting', *International Journal on Document Analysis and Recognition*, 18.3: 223–234. doi: 10.1007/s10032-015-0245-z.

Almazán, Jon, Albert Gordo, Alicia Fornés, and Ernest Valveny (2014), 'Segmentation-free word spotting with exemplar SVMs', *Pattern Recognition*, 47.12: 3967–3978. doi: 10.1016/j.patcog.2014.06.005.

Chen, Francine R., Lynn D. Wilcox, and Dan S. Bloomberg (1993), 'Word spotting in scanned images using hidden Markov models', in *1993 International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5 (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 1–4. doi: 10.1109/ICASSP.1993.319732.

Dey, Sounak, Anguelos Nicolaou, Josep Lladós, and Umapada Pal (2016), 'Evaluation of the Effect of Improper Segmentation on Word Spotting', *arXiv.org*: arXiv:1604.06243.

Ghorbel, Adam, Jean-Marc Ogier, and Nicole Vincent (2015), 'A segmentation free Word Spotting for handwritten documents', in *13th International Conference on Document Analysis and Recognition (ICDAR)* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 346–350. doi: 10.1109/ICDAR.2015.7333781.

Ghosh, Suman K. and Ernest Valveny (2015), 'A sliding window framework for word spotting based on word attributes', in Roberto Paredes, Jaime S. Cardoso, and Xosé M. Pardo (eds), *Pattern Recognition and Image Analysis.7th Iberian Conference (IbPRIA)* (Cham: Springer), 652–661.

—— and —— (2017), 'R-PHOC: Segmentation-Free Word Spotting using CNN', in *14th International Conference on Document Analysis and Recognition (ICDAR),* vol. 1 (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 801–806. doi: 10.1109/ICDAR.2017.136.

Keaton, P., Hayit Greenspan, and Rodney Goodman (1997), 'Keyword spotting for cursive document retrieval', in *Proceedings Workshop on Document Image Analysis (DIA'97)* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 74–81. doi: 10.1109/DIA.1997.627095.

Kolcz, Aleksander, Josh Alspector, Mareijke Augusteijn, Robert Carlson, and G. Viorel Popescu (2000), 'A line-oriented approach to word spotting in handwritten documents', *Pattern Analysis & Applications*, 3.2: 153–168. doi: 10.1007/s100440070020.

Kovalchuk, Alon, Lior Wolf, and Nachum Dershowitz (2014), 'A simple and fast word spotting method', in *14th International Conference on Frontiers in Handwriting Recognition* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 3–8. doi: 10.1109/ICFHR.2014.9.

Krishnan, Praveen and C. V. Jawahar (2016), 'Generating synthetic data for text recognition', *arXiv.org*: arXiv:1608.04224.

Kuo, Shyh-Shiaw and Oscar E. Agazzi (1994), 'Keyword spotting in poorly printed documents using pseudo 2-D hidden Markov models', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16.8: 842–848. doi: 10.1109/34.308482.

Leydier, Yann, Frank Lebourgeois, and Hubert Emptoz (2005), 'Omnilingual segmentation-free word spotting for ancient manuscripts indexation', in *8th International Conference on Document Analysis and Recognition (ICDAR),* vol. 1 (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 533–537. doi: 10.1109/ICDAR.2005.171.

Leydier, Yann, Frank Lebourgeois, and Hubert Emptoz (2007), 'Text search for medieval manuscript images', *Pattern Recognition*, 40.12: 3552–3567. doi: 10.1016/j.patcog.2007.04.024.

Leydier, Yann, Asma Ouji, Frank Lebourgeois, and Hubert Emptoz (2009), 'Towards an omnilingual word retrieval system for ancient manuscripts', *Pattern Recognition*, 42.9: 2089–2105. doi: 10.1016/j.patcog.2009.01.026.

Manmatha, Raghavan, Chengfeng Han, Edward M. Riseman, and W. Bruce Croft (1996), 'Indexing handwriting using word matching', in *Proceedings of the 1st ACM international conference on Digital libraries* (New York: ACM Association for Computing Machinery), 151–159. doi: 10.1145/226931.226960.

—— and W. Bruce Croft (1997), 'Word spotting: Indexing handwritten archives', in Mark T. Maybury (ed.), *Intelligent Multimedia Information Retrieval*, 43–64.

Rath, Tony M., Shaun Kane, Andrew Lehman, Elizabeth Partridge, and Raghavan Manmatha (2002), *Indexing for a digital library of George Washington's manuscripts: a study of word matching techniques*, (Amherst: CIIR Center for Intelligent Information Retrieval), <http://ciir.cs.umass.edu/publications> (accessed 6 November 2019).

—— and Raghavan Manmatha (2003), 'Word image matching using dynamic time warping', in *2003 Conference on Computer Vision and Pattern Recognition (CCVPR)*, vol. 2 (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library). doi: 10.1109/CVPR.2003.1211511.

—— and —— (2007), 'Word spotting for historical documents', *International Journal of Document Analysis and Recognition (IJDAR)*, 9.2–4: 139–152. doi: 10.1007/s10032-006-0027-8.

Rothacker, Leonard, Marçal Rusiñol, and Gernot A. Fink (2013), 'Bag-of-features HMMs for segmentation-free word spotting in handwritten documents', in *12th International Conference on Document Analysis and Recognition (ICDAR)* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 1305–1309. doi: 10.1109/ICDAR.2013.264.

——, ——, Josep Lladós, and Gernot A. Fink (2014), 'A two-stage approach to segmentation-free query-by-example word spotting', *manuscript cultures (mc)*, 7: 47–57, <https://www.manuscript-cultures.uni-hamburg.de/mc.html> (accessed 6 November 2019).

—— and Gernot A. Fink (2015), 'Segmentation-free query-by-string word spotting with bag-of-features HMMs', in *13th International Conference on Document Analysis and Recognition (ICDAR)* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 661–665. doi: 10.1109/ICDAR.2015.7333844.

——, Sebastian Sudholt, Eugen Rusakov, Matthias Kasperidus, and Gernot A. Fink (2017), 'Word hypotheses for segmentation-free word spotting in historic document images', in *14th International Conference on Document Analysis and Recognition (ICDAR),* vol. 1 (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 1174–1179. doi: 10.1109/ICDAR.2017.194.

Rothfeder, Jamie L., Shaolei Feng, and Toni M. Rath (2003), 'Using corner feature correspondences to rank word images by similarity', in *2003 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW),* vol. 3 (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 30–30. doi: 10.1109/CVPRW.2003.10021.

Rusiñol, Marçal and Josep Lladós (2008), 'Word and symbol spotting using spatial organization of local descriptors', in *The Eighth IAPR International Workshop on Document Analysis Systems* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 489–496. doi: 10.1109/DAS.2008.24.

——, David Aldavert, Ricardo Toledo, and Josep Lladós (2011), 'Browsing heterogeneous document collections by a segmentation-free word spotting method', in *2011 International Conference on Document Analysis and Recognition (ICDAR)* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 63–67. doi: 10.1109/ICDAR.2011.22.

——, ——, ——, and —— (2015), 'Efficient segmentation-free keyword spotting in historical document collections', *Pattern Recognition*, 48.2, 545–555. doi: 10.1016/j.patcog.2014.08.021.

Vintsyuk, Taras K. (1971), 'Element-wise recognition of continuous speech recognition of words from a specified dictionary', *Cybernetics*, 7.2: 361–372 (= *Kibernetika*, 2 (1971): 133–143). doi: 10.1016/j.patcog.2014.08.021.

Wilkinson, Tomas, Jonas Lindström, and Anders Brun (2017), 'Neural Ctrl-F: segmentation-free query-by-string word spotting in handwritten manuscript collections', in *2017 International Conference on Computer Vision (ICCV)* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 4443–4452. doi: 10.1109/ICCV.2017.475.

Zhang, Bin, Sargur N. Srihari, and Chen Huang (2003), 'Word image retrieval using binary features', in *Document Recognition and Retrieval XI*, Proceedings vol. 5296 (Society of Photo-Optical Instrumentation Engineers: SPIE. Digital Library), 45–54. doi: 10.1117/12.523968.

Zhang, Xi and Chew Lim Tan (2013), 'Segmentation-Free Keyword Spotting for Handwritten Documents Based on Heat Kernel Signature', in *12th International Conference on Document Analysis and Recognition (ICDAR)* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 827–831. doi: 10.1109/ICDAR.2013.169.

Article

Writer Identification and Script Classification: Two Tasks for a Common Understanding of Cultural Heritage

Dominique Stutzmann, Christopher Tensmeyer, Vincent Christlein | Paris, Provo, UT, Erlangen

Abstract

Writer identific tion and script classification are usually considered as two separate and very different tasks, in palaeography as well as in computer science. Following the ICDAR competition on the CLAMM corpus about script classification and dating, this paper proposes to reconsider the tasks and methods of *Palaeography* and *Computer Vision* applied to *Artificial Palaeography*. We argue that, when aiming at understanding past societies and written cultural heritage in their complexity, palaeography and its core tasks may be defined as discretising the historical and social continuum at different levels of granularity. In this sense, we can consider writer identification and script classification as a single task. We then transfer this hypothesis to computer science, by running two infrastructures created for script classification on a more homogeneous dataset with a focus on writer identification. The analysis of the results confirms the uniformity of these tasks and allows us to reflect in a novel way on how to demonstrate and illustrate the historical continuum while discretising it in a non-binary way.

1. Introduction

The discipline of palaeography is quite often mistakenly defined, especially in French or Spanish, as the ability to read old documents and books. Palaeographers, however, generally acknowledge three 'levels' of palaeography: (i) reading; (ii) expertise; and (iii) historical analysis, as stated in the very first sentence of Cencetti's standard work *Lineamenti: di storia della scrittura latina:*

> La 'Paleografia è lo studio critico delle antiche scritture ed è suo scopo non solo (i) interpretare esattamente i manoscritti, (ii) ma anche datarli, localizzarli e, (iii) in generale, trarre dal loro aspetto esteriore tutti gli elementi utili allo studio del loro contenuto e in generale alla storia della cultura.[1] [numbering added].
>
> 'Palaeography' is the critical study of ancient scripts and its purpose is not only (i) to accurately interpret the manuscripts, (ii) but also to date them, identify their place of origin and, (iii) in general, extract from their external appearance all the elements useful for the study of their content and in general for the history of culture.

Computer science also has the tasks related to these three levels: (i) handwritten text recognition and word spotting; (ii) writer identification; and, tending towards (iii), script classification. Here, terms have to be defined. In computer science and in general language, 'script' means 'writing system', for instance in Wikipedia articles about Arabic and Latin scripts,[2] and 'script type' refers to what the palaeographer Malcolm Parkes – and most English-speaking palaeographers after him – would call a 'script'. Indeed, he gave the following definitions: 'A *script* is the model which the scribe has in his mind's eye when he writes, whereas a *hand* is what he actually puts down on the page.'[3] Marc Drogin wrote: '[Among] styles of writing [that] were constantly undergoing changes […], we can give names to certain segments of this changing pattern […]. There are obvious and relatively long-lived styles that we can recognize and term *scripts.*'[4] For instance, Textualis is such

[1] Cencetti 1997.

[2] 'Arabic Script' 2018; 'Latin Script' 2018.

[3] Parkes 1969, xxvi.

[4] Drogin 1989, 4.

a 'script' in palaeographical terms, 'script type' in general terms or a long-lived 'style of writing' in the words of Marc Drogin, whose characteristic features are a double bowed 'a', a loopless 'l' and a long 's' (ſ) standing on the baseline. For the sake of clarity and unless stated otherwise, this paper will consistently use the terms 'script' for the more general level of 'writing system' and 'script type' for the long-lived styles of writing.

In this paper, we argue that, when aiming at understanding past societies and the written cultural heritage in their complexity, palaeography and its core tasks may be defined as discretising the historical and social continuum. Thus, the different tasks involved in palaeography can be considered as different levels of granularity, rather than a three-stage discipline. We can list the following interconnected components: (a) distinguishing script/text from non-script/non-text, (b) distinguishing one script ('writing system') from another, (c) identifying what the text is, (d) dating or (e) localizing and pointing to a specific context of production, e.g. *scriptorium*, (f) pointing to a specific writer, (g) classifying as a script type and (h) understanding the cultural history. To perform their analytical work, palaeographers tend to agree to base their arguments on the seven aspects of a medieval hand introduced by Jean Mallon and listed here according to Mark Aussems' translation: (i) form, i.e. morphology of the letters, (ii) angle of writing, specifically pen angle, (iii) ductus, (iv) 'modulus', i.e. the dimensions of the letters, (v) contrast, i.e. the difference in thickness between hairlines and shadow lines, also called 'weight', (vi) writing support and (vii) internal characteristics.[5] The respective importance of each of these aspects, however, remains subject to debate.[6] Some of the palaeographic tasks listed above are in common with or imply a connection with other disciplines, not only those related to the internal characteristics that are connected to linguistics and history (e.g. text identification, scholarly editing, interpretation of scribal variants etc.), but also those dedicated to the 'writing support' and layout, connected with codicology, philology, diplomatics or the broader study of the '*forma mentis*'.[7] In short: there are common analytical tools for different tasks in the humanities side of palaeography, and these tasks may be defined in various manners and at different levels.

On the other hand, computer science has developed different tools for different tasks, as evidenced by the last competitions organized in ICFHR and ICDAR: layout analysis,[8] reading or indexing,[9] writer identification[10] and dating and script classific tion,[11] even if some offer a combined goal on separate 'tracks', such as layout and text recognition.[12] From a modelling perspective, however, one could argue that it is the same problem, namely to distinguish what belongs together and what does not, a question that can be addressed by the same means and in both cases entails a discretisation of a body of evidence that can be seen as a historical continuum. In the following, we will test and confirm this hypothesis. Section II discusses the fusion of script and writer identification in palaeography. Section III addresses the notion of distance in 'similarity measure' between script types, to open the discussion about historical and social continuum. Section IV applies the tools and analyses of script classification to writer identification. Finally, we conclude that the fitness of tools confirms the uniformity of the tasks.

2. Palaeography: writer identifi ation and script classifi ation

In palaeography, 'expertise'[13] is traditionally subdivided into two parts: on the one hand, identifying the writer or ascribing several written artefacts to one known or unknown scribe and, on the other hand, dating and localizing a manuscript artefact based on the form of its handwriting. Each task represents a different perspective.

Writer identific tion focuses on individual characteristics and, in its fundamental principles and goals, is related to forensics, so that, despite the divide between the fields, professors of palaeography at the École nationale des chartes and other 'chartistes' have appeared in court as experts in prominent cases such as the Dreyfus affair and the murder of Grégory Villemin. Dating and localizing, on the other

[5] Mallon 1952; Aussems and Brink 2009; Muzerelle 2013.

[6] Stutzmann 2015.

[7] Barret, Stutzmann, and Vogeler 2016.

[8] E.g. Simistira et al. 2017.

[9] Andreu Sánchez et al. 2017.

[10] E.g. Fiel et al. 2017.

[11] E.g. Cloppet et al. 2017.

[12] Clausner, Antonacopoulos, and Pletschacher 2017.

[13] Cf. Gilissen 1973.

hand, do not focus on individuals, but on the zeitgeist and the general characteristics of a given era, and searches for the place or time in which the written artefact would best fit (usually a country or region and a period of half a century or less).

At a medium level of granularity, palaeographers focus on specific areas and times. In this case, researchers try to single out what is common to a group of scribes and to understand how this common ground evolves over time.[14] The environment can be more or less closed, homogeneous, subject to external influences and develop more or less specific writing shapes and styles, e.g. monastic *scriptoria*, secular chanceries, or Florentine merchants. This kind of research is generally linked to the third level of palaeography as cultural history. Indeed, studying a *scriptorium* leads directly to the intellectual history of a social group and its individuals.[15] However, with its technical component of expertise, it highlights that there is no divide between the study of autographs and writer identification, on the one hand, and dating and localizing, on the other, but a continuous scale of times and places in which palaeographers will attempt to place a written artefact, from an individual to a *scriptorium* or workshop, to a school, to a milieu, to a country and century.

Here, we need to add that the social and palaeographical continuum is multidimensional. Most societies experience some sort of polygraphism, i.e. the use of several writing systems or several script types by different individuals or by the same individual in different situations. Observing the adoption and implementation of script types by scribes can indeed be meaningful for historians. For example, the Hybrida script type was used by specific religious circles of the *devotio moderna* ('Devotenbastarda'),[16] and particular script types were used for liturgical books and books of hours in the fourteenth to sixteenth centuries.[17] Polygraphism is a challenge not only for writer identification[18] but also for palaeographical study itself, because to assess polygraphism, one needs to define the 'script types' in question before analysing the places and times in which they may occur and their relations and reciprocal influences as evidenced by specific features. As a consequence, the notion of palaeography as cultural history cannot be a separate theoretical domain, encompassing wider historical research *in abstracto* and studying script evolutions in time and space to interpret the modes of production and reception of the written object in past societies, without getting back to the individuals and the very graphical part of the study of handwriting. Indeed, this level of study implies the combined analysis of graphic and extra-graphic factors for script evolution, and understanding the role of the individual in society and in communication processes. This task is connected to linguistics, sociology, anthropology, cognitive sciences etc., but also specifica ly to the question of milieu and the social uses of writing, including the study of script classifications.

Script classification is therefore, once again, a prerequisite for an in-depth study of a larger social and religious evolution. Indeed, any large-scale and statistical analysis needs to name its objects and to single out specific features or 'factors', because one has to be able to define features within each field before comparing fields. Indeed, as Marc Smith has put it, 'Devising a rational typology and nomenclature to encapsulate the dimensions and features of a vast number of scripts is a necessity, if experts are to agree on anything, but no easy task.'[19] Likewise, the palaeographer David Ganz appropriately quoted Jakob Burckhardt: 'It is the most serious difficult of the history of civilization that a great intellectual process must be broken up into single and often into what seem arbitrary categories, in order to be in any way intelligible.'[20] In palaeography as well, any classification aims at discretizing a complex and interconnected reality that is largely a continuum.

Discretising is a difficult task *per se*. It is difficult even to differentiate between writing systems, and which script was used in some famous artefacts, such as the Namara inscription bearing the epitaph of king Imru'l-Qays, is itself an object of debate. The genealogy and distances between writing systems are likewise a field of research.[21] Polygraphism and the use of a script within another system (e.g. of Greek characters within a document otherwise written in the Latin alphabet)

[14] Nievergelt, Gamper, and Bernasconi 2015.

[15] E.g. Ganz 1990.

[16] Oeser 1992.

[17] Stutzmann 2017, 2019.

[18] De Robertis 2013.

[19] Smith 2016.

[20] Burckhardt 1982; Ganz 1990, 121.

[21] Hosszú 2017.

also brings about situations in which it is difficult to state to which writing system a particular sign may belong. For the study of more closed milieus, such as monastic *scriptoria*, with specific styles within a script type,[22] we need to identify specific features that appear consistently in several books and ascertain a chronology in order to explain the evolution within the abbey, in relation to successive abbots or to a specific cultural change. Indeed, the scribal landscape is like a huge maelstrom in which so-called canonised scripts (as defined by Cencetti)[23] form some islands and emerge with fixed morphological features [24]

As a consequence of these difficulties and the diverging valuation of the features under scrutiny, for now there is no generally accepted classification [25] Following Albert Derolez's proposal for a general classification of Gothic book scripts, Marc Smith discussed the arbitrariness of any such classification and that it is difficult to assess the extent to which a nomenclature, however based on objective features, truly reflects historical realities.[26] According to him, scholars should therefore avoid thinking in terms of fixed, arbitrary, sometimes 'elementary' nomenclatures, rather than perceiving the complex evolutionary mechanisms.

These needs and shortcomings together show that all tasks involved in forensic examination in palaeographical research are really a single task on different levels and, depending on historical reality itself, the focus can be set on virtually any point on a continuous scale:

a. inner individual variation < individual < closed group of individuals < larger group or region in a given time < world at a given time
b. script instances, 'scribal hands' < varieties of script types < script types < writing systems < visual semiotic systems[27]

3. Image analysis and script classifi ation: introducing distance visualisation as heuristics

Parallel to the tasks of palaeography, in recent years there have been two different series of competition for the classification of historical scripts. One is on writer identification, the other on script classification. The former is by far the most prominent. It aims to identify several writing samples as having been produced by a single person,[28] even in different writing systems or a 'multi-script environment'.[29] In these competitions, the measured dissimilarity of the handwriting is generally considered to be the consequence of difference between individuals, but some research addresses the notion of script style and other biologically or sociologically defined groups.[30] Regarding the latter, it must be noted that gender, i.e. a social difference, is probably a better explanation than biological difference. Studies combining image analysis, genetics and social sciences would be welcome. In studies supposedly demonstrating the influence of hormones,[31] the difference in the expression of masculine and feminine traits is also demonstrated and probably to be linked to social behaviour, as in historical scripts, where the social division of male and female and their distinct association with script types has been observed (e.g. Cursiva for men, Textualis for women[32]).

The notion of the division of script types is rarely addressed in image analysis, even when, as in the latest competition on historical writer identification [33] the corpus covers such a large chronological span (with 720 writers) that the differences between writing samples may reflect script types rather than merely individual variety.

The other strand, started in 2016, is the competition on Latin handwriting classification [34] based on the CLAMM corpus.[35] The competitors were required to provide a normalized belonging matrix (a CSV file recording the

[22] Palaeographers sometimes call these styles 'types', sometimes using the German word 'Schrifttypen'.

[23] Cencetti 1978, 1997.

[24] Stutzmann 2018.

[25] Stutzmann 2015.

[26] Smith 2004.

[27] Here, following our vocabulary, we introduce the notion of 'variety of a script type' to render what Malcolm Parkes called 'the varieties of a single script'; Parkes 2008, 152.

[28] Fiel et al. 2017.

[29] Djeddi et al. 2015.

[30] Maadeed et al. 2012; Djeddi et al. 2015, 2016.

[31] Beech and Mackintosh 2005.

[32] Stutzmann 2012.

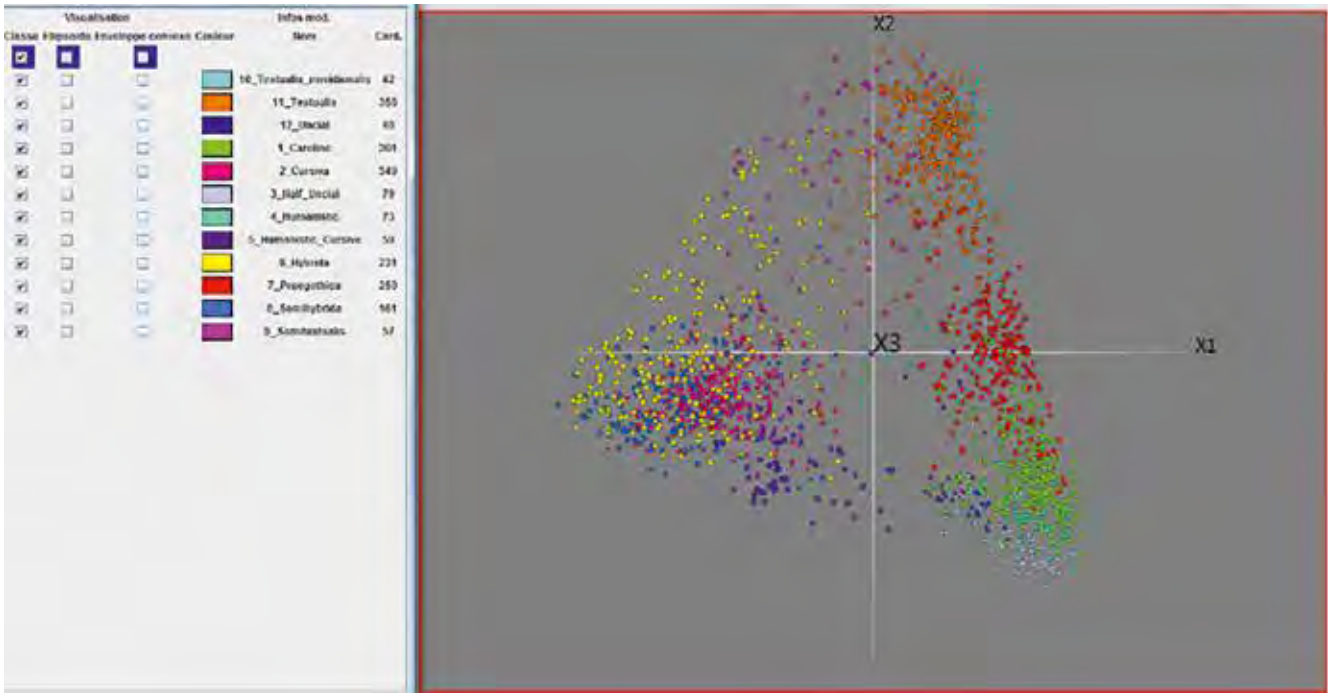
[33] Fiel et al. 2017.

[34] Cloppet et al. 2016, 2017.

[35] Stutzmann 2016.

Table 1: ICDAR2017 PCA on Christlein's belonging matrix.

Axes 1 and 2



Axes 2 and 3

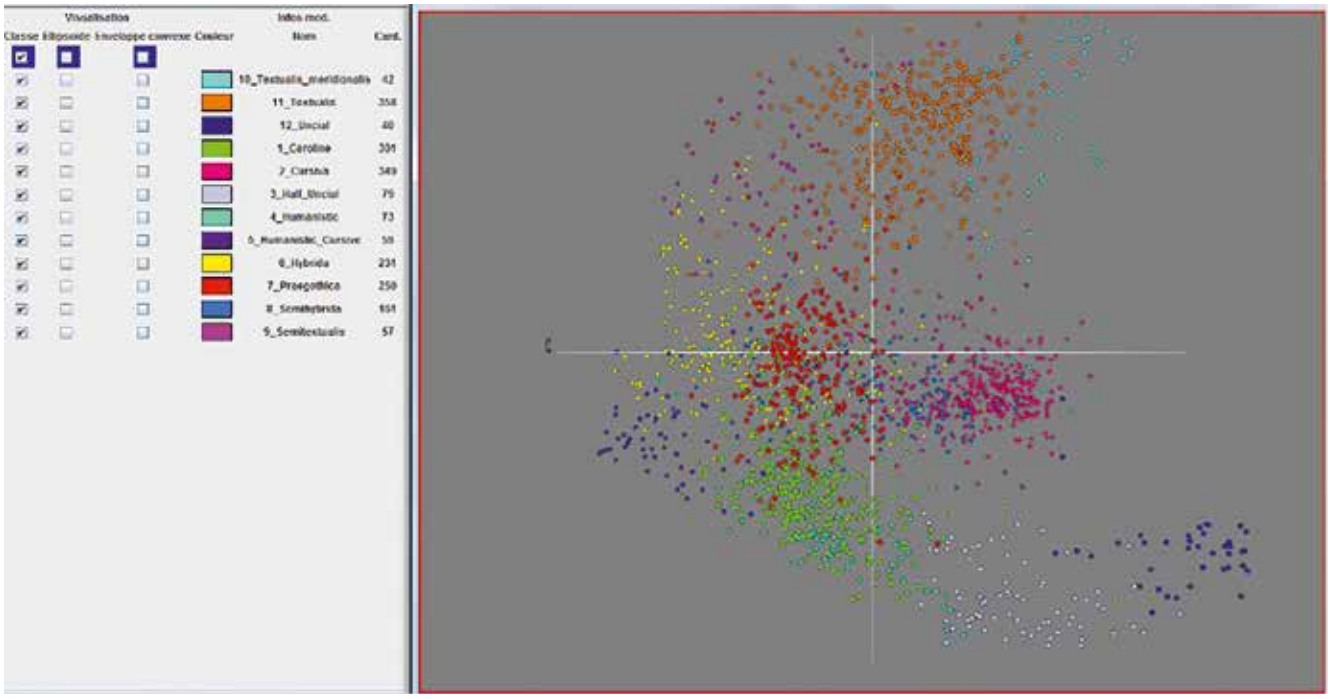
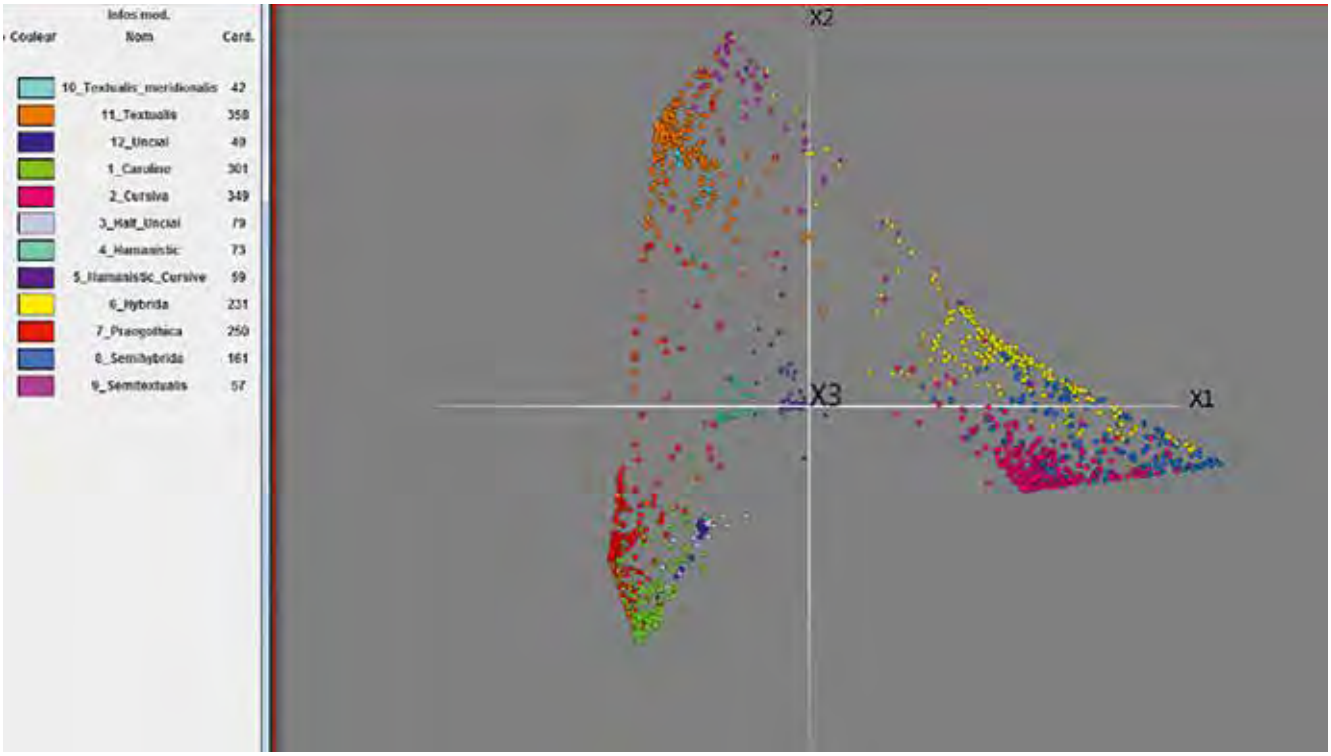
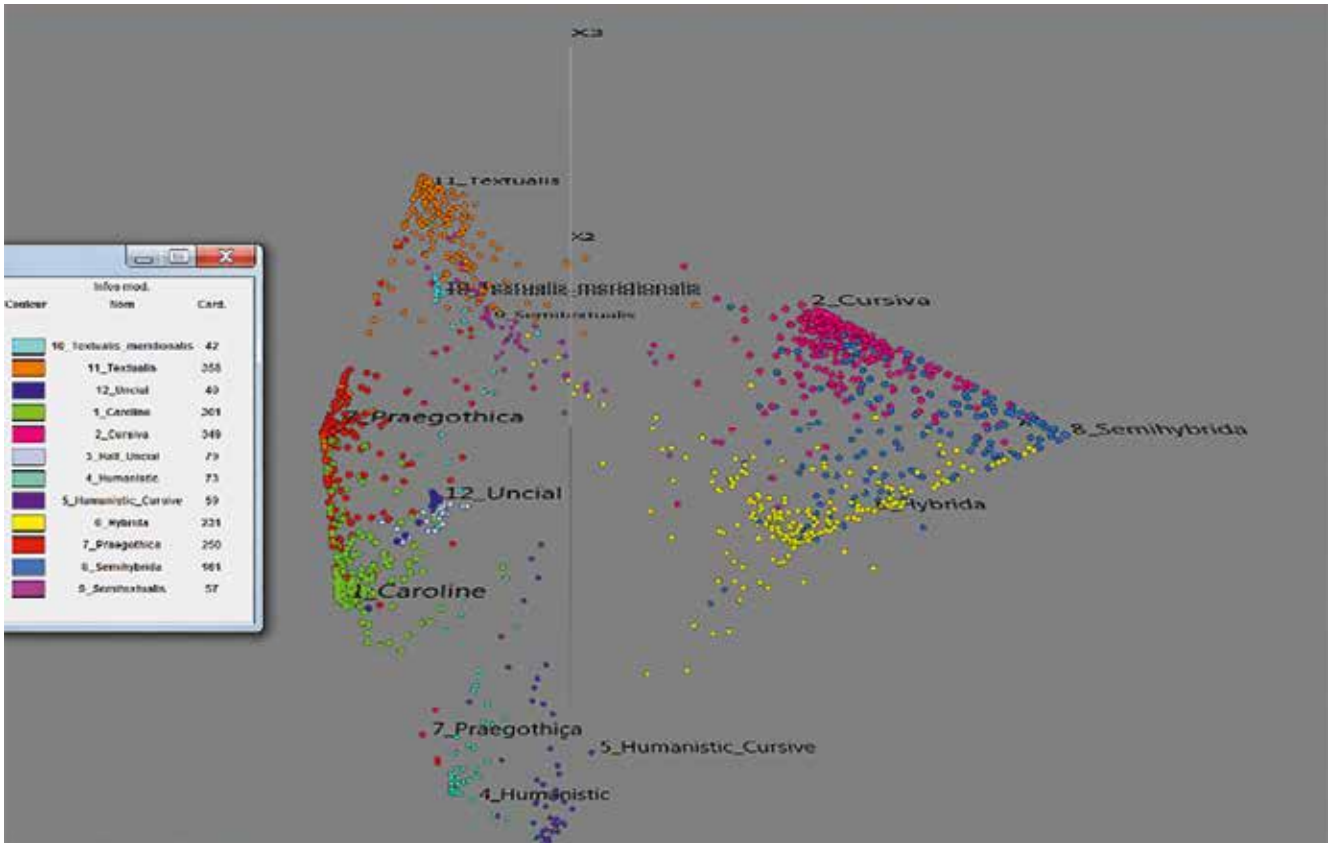


Table 2: ICDAR2017 PCA on Tensmeyer's belonging matrix.

Axes 1 and 2



Axes 1 and 3



relative score of all twelve proposed script type labels for each image) and a distance matrix describing all images. The evaluation was based on the belonging matrix, in which the script type with the highest score had to match the ground-truth label. A first palaeographical interpretation was given based on the confusion matrices, showing similarities and recurring confusions between some of the script types.[36]

Based on the belonging and distance matrices, we can establish a new heuristic and interpretation mode, as we suggested in 2015 and then implemented and commented with Mike Kestemont.[37] Principal component analysis (PCA) of the belonging matrix easily reduces the 12 dimensions to a three-dimensional space. Here, we use the Explorer3D software.[38] For the following interpretation, we compare the data from the two best-performing systems in the ICDAR 2017 competition. One was developed by Vincent Christlein,[39] the other by Christopher Tensmeyer.[40] The results are quite different.

In Christlein's results (Table 1), the first two axes of the PCA analysis can be represented as a triangle, and the points gather: (a) Uncial, Half-Uncial, Caroline and Humanistic, (b) Textualis, Textualis meridionalis and Semitextualis; (c) Cursiva, Hybrida and Semihybrida. One script type, namely Praegothica, forms the edge between poles (a) and (b), as does Humanistic Cursive between (a) and (c), while Semitextualis connects with Hybrida. A view along axes 2 and 3 highlights the fact that the script types in the three poles are also well defined

In Tensmeyer's results (Table 2), the best visualisation combines the first three axes and gives a different view of the relations between scripts, with the position of the Humanistic group and that of Praegothica being greatly changed within a space with four main zones: (a) Uncial, Half-Uncial, Caroline and most Praegothica samples belong in one corner; (b) Cursiva, Semihybrida and Hybrida in another corner; (c) Textualis, Semitextualis, and Textualis meridionalis in a compact zone with two corners; and (d) Humanistic minuscule and Humanistic Cursive in a distinct zone that appears to be almost completely autonomous (except for the fact that, in this zone, Humanistic minuscule is, as expected, on the side of Caroline and Textualis, while Humanistic Cursive is closer to Cursiva).

These results, despite or because of the differences, are very valuable for freshening our perspective on the history of the Latin writing system. The training data set contained only labels and no statement about the distance between scripts in history. Despite this, not only do the results correspond to our historical knowledge, but the discrepancies between the two sets of results also correspond to on-going palaeographical debates. They force historians to reconsider the respective positions of each script. They also illustrate how entangled the different script types are. At the same time, coeval script families may be either very clearly separated (e.g. Textualis and Cursiva) or completely integrated (e.g. Cursiva and Semihybrida). This is a first illustration of the historical continuum, in which some script families emerge and separate from a common stem.

Another visualisation of this continuum is to represent the 2000 × 2000 distance matrix of all images. This is also done with the Explorer3D software and uses classical multidimensional scaling (MDS). The first observation is that Christlein's and Tensmeyer's distance matrices are surprisingly different. In Christlein's method, a major step is the definition of the features of the learnt classes,[41] and the distances between the images themselves do not initially appear to be very different (Table 3). Indeed, a 3D projection shows a compact, populated kernel and some secluded groups. The images of each script are not distributed in coherent classes with a small intra-class distance, but rather along axes either starting at the centre or traversing it and developing mostly in one direction each. This happens in such a way that two otherwise clearly separate scripts, such as Cursiva and Textualis, are distributed in axes pointing almost in the same direction, even if the subsequent classification distinguishes them without mistake. This is an intriguing view of the history of script. Each family comprises samples that are closer to the other families than to its siblings. And, here, 'intriguing' does not mean 'false': on the contrary, it corresponds perfectly to the experience of palaeographers who can always find script images for which the so called 'first impression' is misleading or may correspond to other features than letter shapes.

[36] Cloppet et al. 2016, 2017.

[37] Stutzmann 2015; Kestemont, Christlein, and Stutzmann 2017.

[38] Exbrayat and Martin 2017.

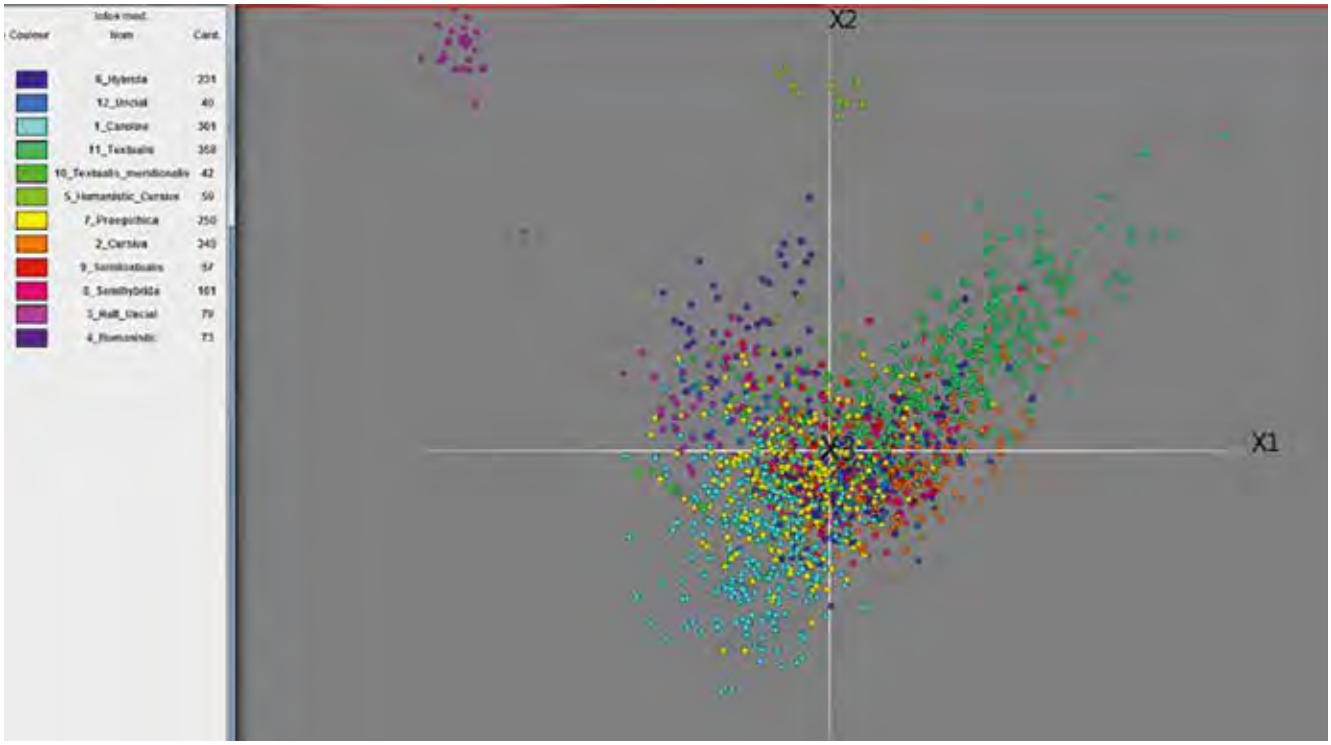
[39] Christlein et al. 2017; Christlein [2017] 2018.

[40] Tensmeyer, Saunders and Martinez 2017; Christopher Tensmeyer 2017.

[41] Kestemont, Christlein, and Stutzmann 2017.

Table 3: ICDAR2017 Christlein's distance matrix.

Axes 1 and 2



Axes 2 and 3

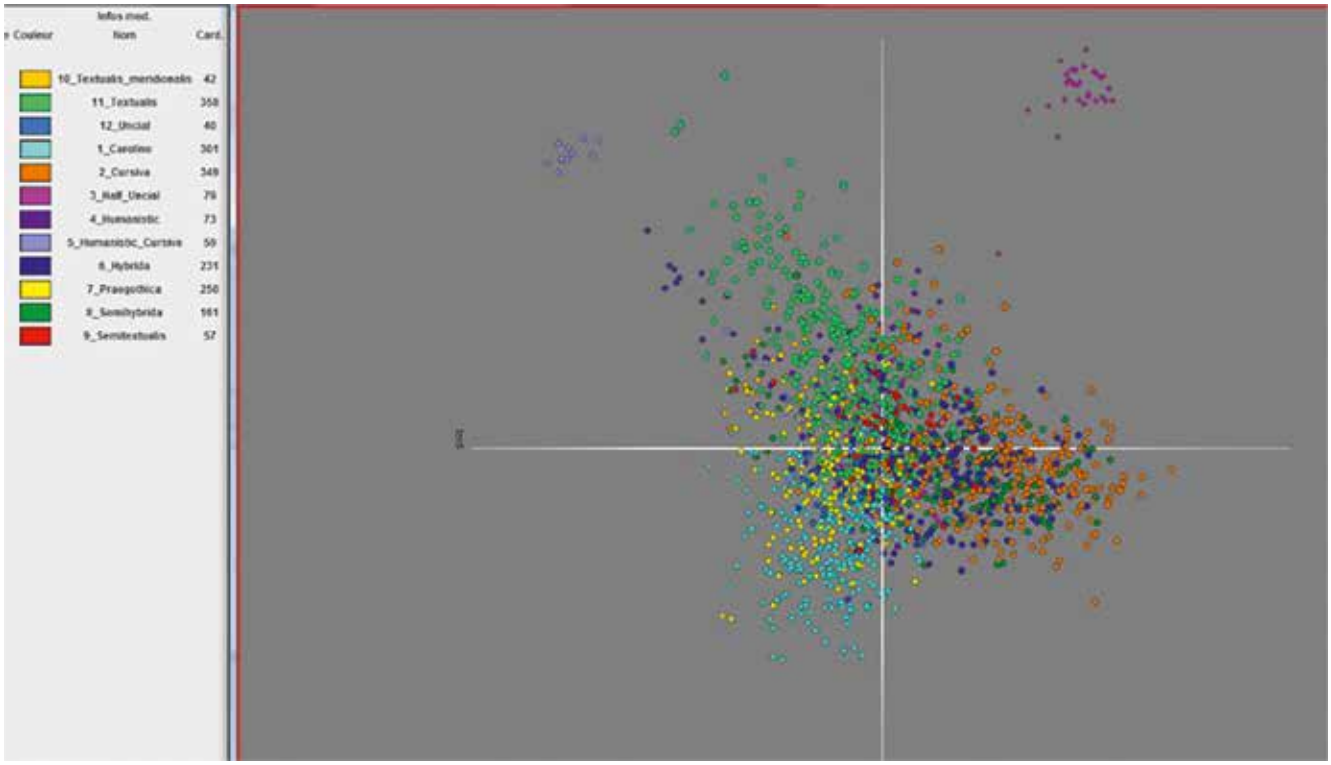
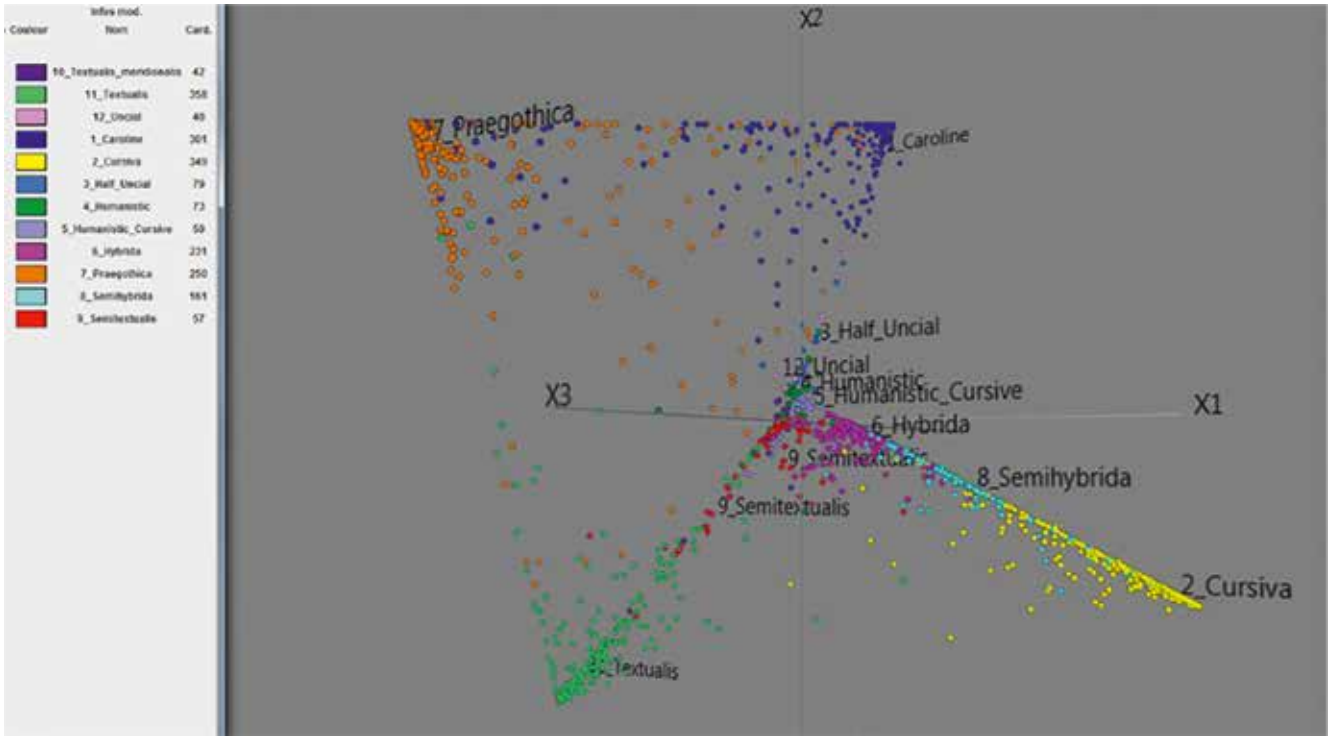


Table 4: ICDAR2017 Tensmeyer's distance matrix .

Axes 2 and (1 and 3)



Tensmeyer's distance matrix (Table 4), however, can be represented in a shape that is much more similar to our view of the belonging matrix. The shape is a triangular pyramid, whose edges are populated very differently, as well as diagonals to the centre of gravity. The corners are Caroline, Praegothica, Textualis and Cursiva. The centre of gravity or central node joins, in an unexpected way, Humanistic and Humanistic Cursive to (a) Uncial and Half-Uncial (very close, but rather on the diagonal to Caroline), (b) Hybrida and Semihybrida on the diagonal to Cursiva or slightly dispersed on the plane between Cursiva and Textualis and (c) Semitextualis and Textualis meridionalis on the diagonal to Textualis. The whole representation is striking, because one can make some sense of it, but only in a very fuzzy way. The centrality of Humanistic, Humanistic Cursive, Uncial, Half-Uncial and, to some extent, Semitextualis may correspond to what students perceive as the readability or legibility of a script, but it clearly does not correspond to any historical evolution, nor to morphological features at the level of single letters (according to which Humanistic should be located between Caroline and Praegothica). In this sense, we are quite uncertain how to interpret the distance matrix. It is interesting to note that this system also attempts to analyse the non-textual components of the manuscript pages to classify the script type of the page and that there is a significant bias about which scripts could be identified from non-textual content (e.g. Praegothica, Humanistic).[42] It is not clear how much this implementation influences the present output.

Nevertheless, this kind of visualisation is very closely connected to what one palaeographer envisioned in the 1970s as being a key model for a 'Cartesian nomenclature',[43] that is, a 3D space in which each zone would correspond to a specific script – with the only exception that axes would be defined according to a limited set of allographs.

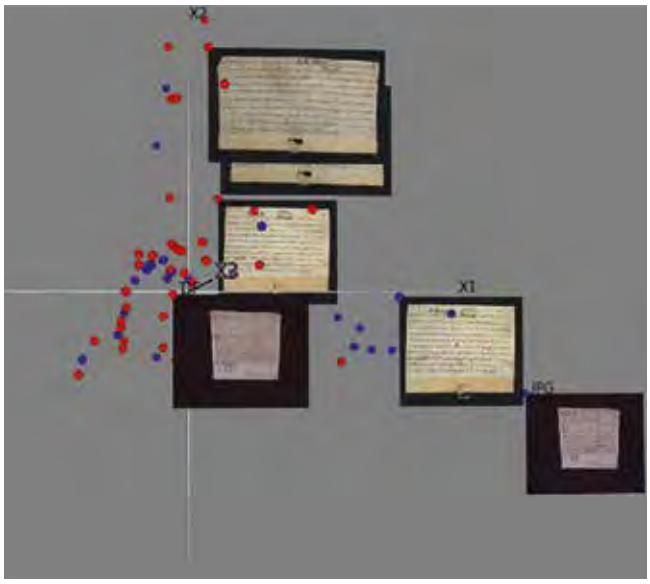
Distance visualisation on the results of script classification is a way to gain access to the notion of a historical continuum. It does not abolish the need for or the validity of existing classifications, but it allows us to go back to the most confusing artefacts and then either reshape our criteria or understand the features that influence our perception and, at the same time, better understand what the system has learnt.

[42] Tensmeyer, Saunders, and Martinez 2017.

[43] Gumbert 1976.

Table 5: Tensmeyer's belonging matrix for duplicate images (red = tif ; blue = jpg).

For three documents, the thumbnails show that jpg and tif are not classified identically.



4. Image analysis and writer identifi ation

Section III has demonstrated that systems designed to identify script classes may produce belonging and distance matrices in which a palaeographer may try not only to validate the labelling, but also to explore the distances between script classes and the distribution of script samples within the class and in respect to other classes. Closeness may be interpreted in morphological and phylogenetic-historical terms.

In this section, we examine the same systems applied to a new data set built for writer identification. The systems have not been retrained with new labels, but only applied to new images as if they were part of the original test set. The new data set encompasses 377 images of written documents and books produced or received by the Cistercian abbey of Fontenay during the twelfth century. Among them are thirty-seven pairs of duplicate images, present as tif and jpg files, to test the impact of format on the systems. For the documents, we created a ground truth based on our study

Table 6: Christlein's belonging matrix.

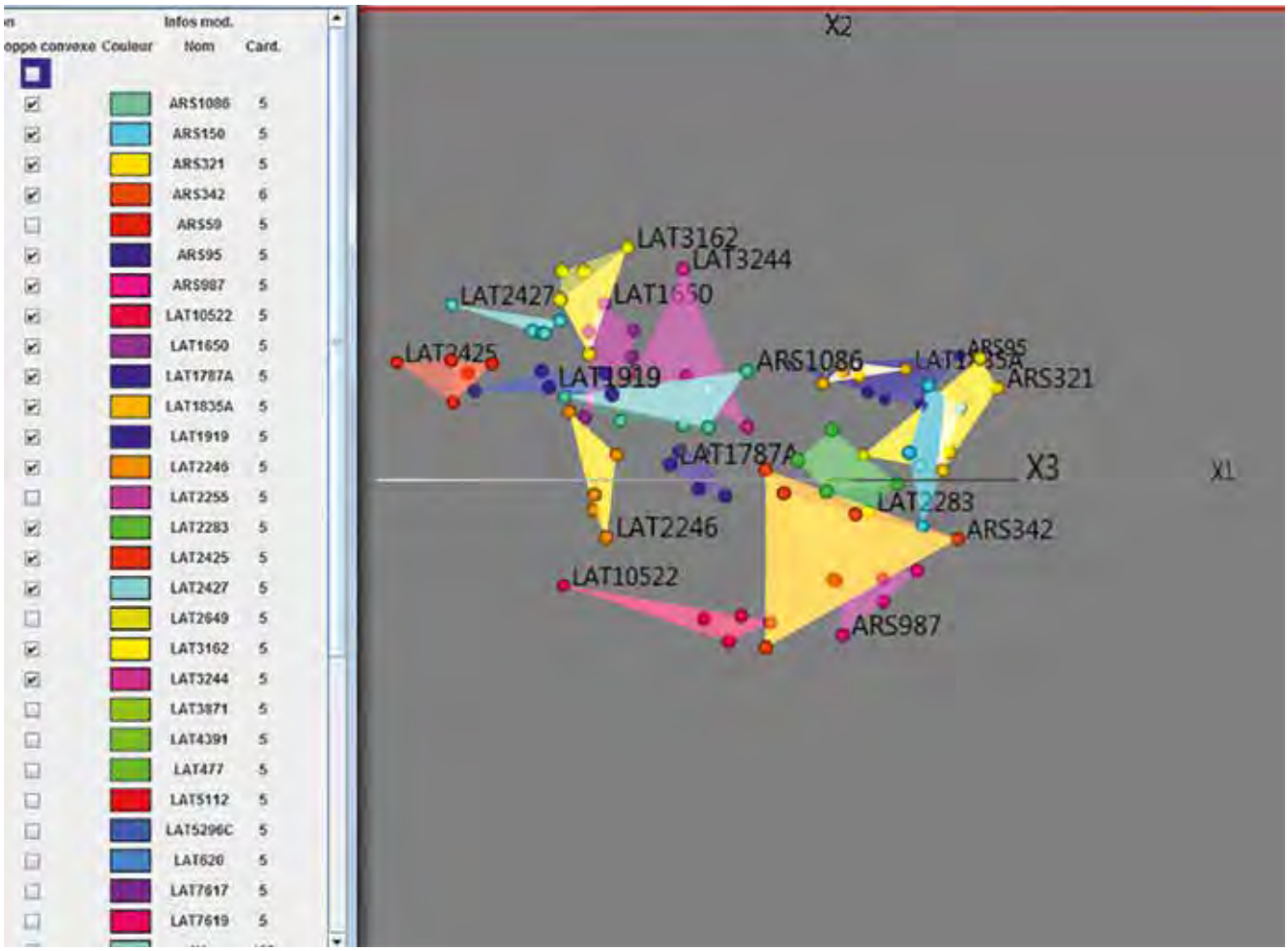
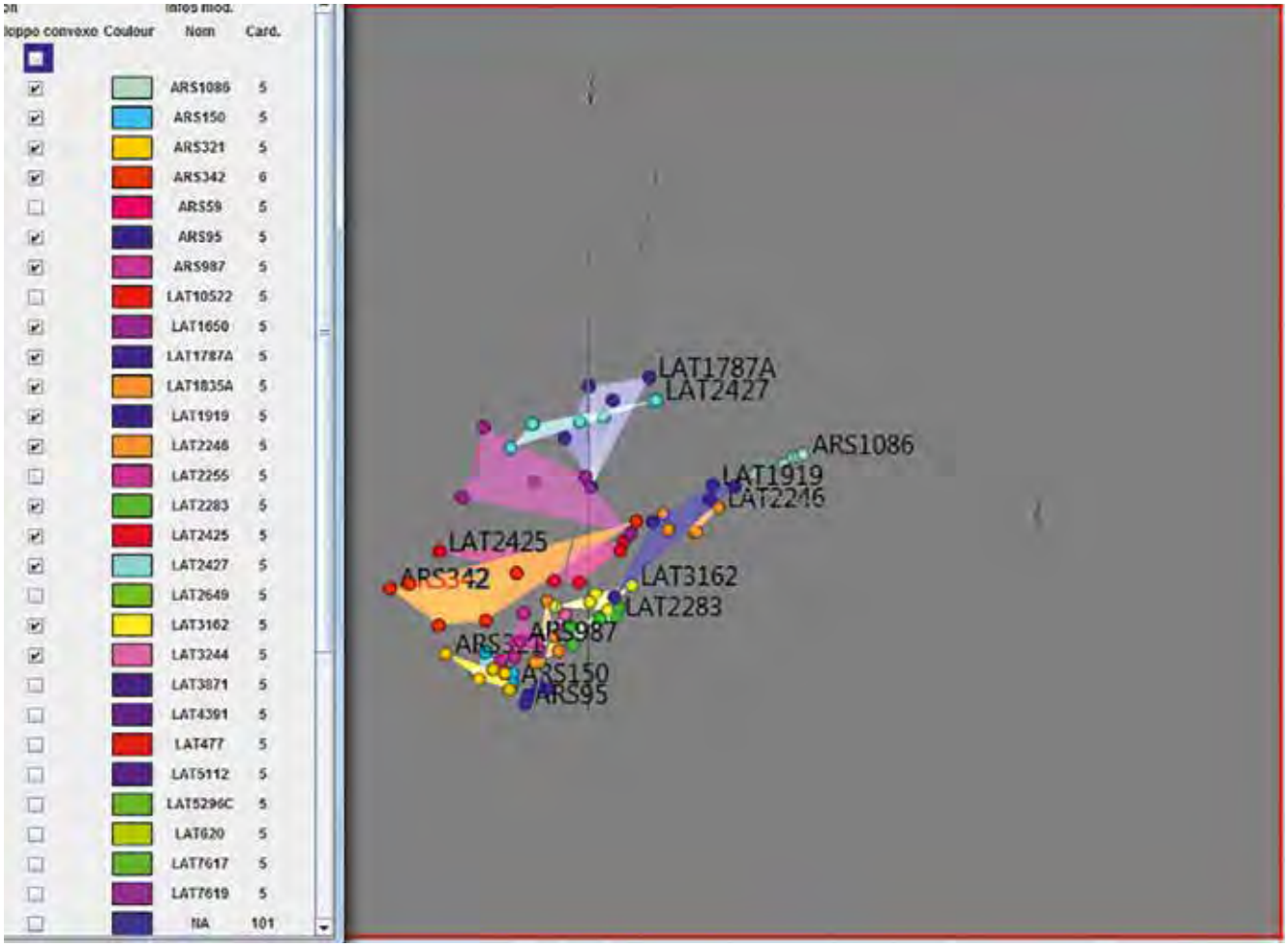


Table 7: Tensmeyer's belonging matrix.



of the scriptorium.[44] For the books, we used five pages of each volume and considered that each of these samples of five pages could represent one anonymous 'hand' – that is, a single writer in a single context. The relative distance/closeness between the different sets of anonymous hands can be a starting point for suggesting the identity of separate writer sets, for which there is, however, no formalized ground truth.

A first comparison is based on the duplicate images in jpg and tif and between grey-scale and colour images (Table 5). Both systems deal accurately with both grey-scale and colour images. Tensmeyer's system uses two separate analysis mechanisms for tif and jpg, so that the two spaces overlap but are not identical, and duplicate images in different formats may be somewhat separated

However, for both Tensmeyer and Christlein, if we compare images in the same format, we see very coherent and compact spaces for each manuscript (same scribe, same moment). If we add the script classes used in documents, written either by different scribes or by a single scribe at different moments and in different sizes and layouts, the spaces are slightly less homogeneous, but still offer a very neat classification. The following figures (Tables 6 and 7), representing the results of PCA on the belonging matrix, are reduced to the distribution of all five images from only sixteen different manuscripts. The full data is too compact to be easily illustrated, but each group of images pertaining to one writer is sufficiently compact to be distinguished from the other groups.

Here, we will not go into the details of the analysis. However, we can stress that the belonging matrix keeps the labels of the original task, which are totally irrelevant for the present task. The classification of the images in the present data set into different groups was based mainly on

[44] Stutzmann 2009.

morphological features,[45] which are not at all in line with those used for describing the long-term history of scripts. Christlein and Tensmeyer's systems, however, were trained to classify writing samples produced from the sixth to the sixteenth centuries. It is therefore striking and unexpected that they are able to 'zoom in' onto a very homogeneous *scriptorium* production and adequately separate different manuscripts and slightly differing styles within the closed milieu of a monastic *scriptorium*, when they rely on learnt features that we would not deem applicable to the present data set.

Once again, the visualisations based on the data generated by each system do not render exactly the same landscape, but both measures are in line with palaeographical expertise and may shed light on the debate about respective dates of production (given that none of the manuscripts is dated or signed).

5. Conclusion

This paper reunites writer identification and script classification as a single task, from a theoretical perspective not only in palaeography, but also in computer science. The hypothesis of unity is corroborated by the experiment made for this research: applying programs trained for script classification for the medieval millennium to writer identification in the closed and homogeneous production of a monastic *scriptorium*. Moreover, we also propose to use and analyse distance metrics as a way to show the historical and social continuum of written production. This opens a non-binary solution to the task of script or writer identificatio and a more balanced view of model (what is in the 'mind's eye'), influence, imitation and evolution. It also opens the discussion about the features used in image analysis, the meaning of measured distance and visualisation. Further experiments may now be organized. The first one should measure how well a unique scribe can be ascertained across different scripts and script types. The corpus under scrutiny would be composed of the autographs of known writers using several script types (*di-* or *polygraphism*). A second one should extend the notion of *continuum* that we used in this paper and test the measures of distance between different script systems (e.g. Latin, Greek, Hebrew, Arabic, Chinese etc.).

[45] Stutzmann 2014.

ACKNOWLEDGMENT

This research was initiated with the support of the French National Research Agency for the project ORIFLAMMS (ANR-12-CORP-0010) and continued in the project HIMANIS, funded by the European Commission in JPI Cultural Heritage (ANR-15-EPAT-003). The software Explorer3D used for the visualisation was developed by Matthieu Exbrayat and Lionel Martin (Université d'Orléans) and is freely available at http://www.univ-orleans.fr/lifo/software/Explorer3D. The authors wish to thank Marc H. Smith and Jacob Currie for their comments.

REFERENCES

Andreu Sánchez, Joan, Verónica Romero, Alejandro Héctor Toselli, Mauricio Villegas, and Enrique Vidal (2017), 'Competition on Handwritten Text Recognition on the READ Dataset', in *14th International Conference on Document Analysis and Recognition (ICDAR)*, vol. 1 (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 1383–1388. doi: <10.1109/ICDAR.2017.226>.

'Arabic Script' (2018), in *Wikipedia*. <https://en.wikipedia.org/w/index.php?title=Arabic_script&oldid=857390673.8.2.2019> (accessed 19 November 2019).

Aussems, Mark, and Axel Brink (2009), 'Digital Palaeography', in Malte Rehbein, Patrick Sahle, and Torsten Schaßan (eds), *Kodikologie und Paläographie im digitalen Zeitalter – Codicology and Palaeography in the Digital Age* (Norderstedt: Books on Demand; Schriften des Instituts für Dokumentologie und Editorik 2), 293–308.

Barret, Sébastien, Dominique Stutzmann, and Georg Vogeler (eds) (2016), 'Introduction', in *Ruling the Script in the Middle Ages: Formal Aspects of Written Communication (Books, Charters, and Inscriptions)* (Turnhout: Brepols; Utrecht Studies in Medieval Literacy 35), 1–24.

Beech, John R., and Isla C. Mackintosh (2005), 'Do Differences in Sex Hormones Affect Handwriting Style? Evidence from Digit Ratio and Sex Role Identity as Determinants of the Sex of Handwriting', *Personality and Individual Differences,* 39.2: 459–468. doi: <10.1016/j.paid.2005.01.024>.

Burckhardt, Jacob (1982), *Über das Studium der Geschichte: Der Text der "Weltgeschichtlichen Betrachtungen" auf Grund d. Vorarbeiten von Ernst Ziegler nach d. Hs. hrsg. von Peter Ganz* (München: C.H. Beck).

Cencetti, Giorgio (1978), *Paleografia latina* (Roma: Jouvence; Guide allo studio della civiltà romana 10,3).

—— (1997), *Lineamenti di storia della scrittura latina, dalle lezioni di paleografia*, 2nd rev. edn, ed. by Gemma Guerrini Ferri (Bologna: Pàtron).

Christlein, Vincent, Martin Gropp, Stefan Fiel, and Andreas Maier (2017), 'Unsupervised Feature Learning for Writer Identification and Writer Retrieval', in *14th International Conference on Document Analysis and Recognition (ICDAR)*, vol. 1 (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 991–997. doi: <10.1109/ICDAR.2017.165>.

Christlein, Vincent [2017] (2018), *Unsupervised Feature Learning for Writer Identification and Writer Retrieval: VChristlein/Icdar17code*, Python. <https://github.com/VChristlein/icdar17code> (accessed 19 November 2019).

Clausner, Christian, Apostolos Antonacopoulos, and Stefan Pletschacher (2017), 'ICDAR2017 Competition on Recognition of Documents with Complex Layouts – RDCL2017', in *14th International Conference on Document Analysis and Recognition (ICDAR)*, vol. 1 (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 1404–1410. doi: <10.1109/ICDAR.2017.229>.

Cloppet, Florence, Véronique Églin, Marlène Helias-Baron, Cuong Kieu, Nicole Vincent, and Dominique Stutzmann (2017), 'ICDAR2017 Competition on the Classification of Medieval Handwritings in Latin Script', in *14th International Conference on Document Analysis and Recognition. (ICDAR)*, vol. 1 (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 1371–1376. doi: <10.1109/ICDAR.2017.224>.

Cloppet, Florence, Véronique Églin, Van Cuong Kieu, Dominique Stutzmann, and Nicole Vincent (2016), 'ICFHR2016 Competition on the Classification of Medieval Handwritings in Latin Script', in *15th International Conference on Frontiers in Handwriting Recognition (ICFHR)* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 590–595. doi: <10.1109/ICFHR.2016.0113>.

De Robertis, Teresa (2013), 'Una mano tante scritture. Problemi di metodo nell'identificazione degli autografi', in Nataša Golob (ed.), *Medieval Autograph Manuscripts. Proceedings of the XVIIth Colloquium of the Comité International de Paléographie Latine, held in Ljubljana, 7–10 September 2010* (Turnhout: Brepols; Bibliologia, 36), 17–38. doi: <10.1484/M.BIB.1.101466>.

Djeddi, Chawki, Somaya Al-Maadeed, Abdeljalil Gattal, Imran Siddiqi, Abdellatif Ennaji, and Haikal El Abed (2016), 'ICFHR2016 Competition on Multi-Script Writer Demographics Classification Using 'QUWI' Database', in *15th International Conference on Frontiers in Handwriting Recognition (ICFHR)* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 602–606. doi: <10.1109/ICFHR.2016.0115>.

Djeddi, Chawki, Somaya Al-Maadeed, Abdeljalil Gattal, Imran Siddiqi, Labiba Souici-Meslati, and Haikal El Abed (2015), 'ICDAR2015 Competition on Multi-Script Writer Identification and Gender Classification Using 'QUWI' Database', in *13th International Conference on Document Analysis and Recognition. (ICDAR)* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 1191–1195. doi: <10.1109/ICDAR.2015.7333949>.

Drogin, Marc (1989), *Medieval Calligraphy: Its History and Technique* (New York: Dover Publications).

Exbrayat, Matthieu, and Lionel Martin (2017), 'Explorer3D', Université d'Orléans, Laboratoire d'informatique Fondamentale d'Orléans. <https://www.univ-orleans.fr/lifo/software/Explorer3D/> (accessed 19 November 2019).

Fiel, Stefan, Florian Kleber, Markus Diem, Vincent Christlein, Georgios Louloudis, Nikos Stamatopoulos, and Basilis Gatos (2017), 'ICDAR 2017 Competition on Historical Document Writer Identifica ion (Historical-WI)', in *14th International Conference on Document Analysis and Recognition (ICDAR)*, vol. 1 (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 1377–1382. doi: <10.1109/ICDAR.2017.225>.

Ganz, David (1990), *Corbie in the Carolingian Renaissance* (Sigmaringen: Thorbecke; Beihefte der Francia, 20). <http://www.perspectivia.net/content/publikationen/bdf/ganz_corbie> (accessed 19 November 2019).

Gilissen, Léon (1973), *L'expertise des écritures médiévales. Recherche d'une méthode avec application à un manuscrit du XIe siècle: le lectionnaire de Lobbes (Codex Bruxellensis 18018)* (Gand: E. Story-Scientia; Les publications de "Scriptorium", 6).

Gumbert, Johan Peter (1976), 'A Proposal for a Cartesian Nomenclature', in Johan Peter Gumbert and Max J.M. Haan (eds), *Essays Presented to G.I. Lieftinck*, vol. 4: *Miniatures, Scripts, Collections* (Amsterdam: van Gendt; Litterae Textuales, 4), 45–52.

Hosszú, Gábor (2017), 'Phenetic Approach to Script Evolution', in Hannah Busch, Franz Fischer, and Patrick Sahle (eds), *Kodikologie und Paläographie im digitalen Zeitalter – Codicology and Palaeography in the Digital Age 4* (Norderstedt: Book on Demand; Schriften des Instituts für Dokumentologie und Editorik 11), 179–252, <http://kups.ub.uni-koeln.de/7787/> (accessed 19 November 2019).

Kestemont, Mike, Vincent Christlein, and Dominique Stutzmann (2017), 'Artificial Paleography: Computational Approaches to Identifying Script Types in Medieval Manuscripts', *Speculum,* 92.S1: S86–S109. doi: <10.1086/694112>.

'Latin Script' (2018), in *Wikipedia*. <https://en.wikipedia.org/w/index.php?title=Latin_script&oldid=858277087. 8.2.2019> (accessed 19 November 2019).

Maadeed, Somaya Al, Wael Ayouby, Abdelâali Hassaïne, and Jihad Mohamad Aljaam (2012), 'QUWI: An Arabic and English Handwriting Dataset for Offline Writer Identification', in *2012 International Conference on Frontiers in Handwriting Recognition (ICFHR)* (Institute of Electrical and Electronics Engineers: IEEE Xplore Digital Library), 746–751. doi: <10.1109/ICFHR.2012.256>.

Mallon, Jean (1952), *Paléographie Romaine* (Madrid: Consejo superior de investigaciones cientificas; Scripturae. Monumenta et Studia, 3).

Muzerelle, Denis (2013), 'Jeux d'angles et jeux de plume. I. Retour sur l'hypothèse du biseautage de la plume', *Gazette du livre médiéval,* 60.1: 1–27. doi: <10.3406/galim.2013.2032>.

Nievergelt, Andreas, Rudolf Gamper, Marina Bernasconi Reusser, Birgit Ebersperger, and Ernst Tremp (eds) (2015), *Scriptorium: Wesen, Funktion, Eigenheiten, Comité International de Paléographie Latine, XVIII. Kolloquium, St. Gallen, 11.–14. September 2013* (München: C.H. Beck; Veröffentlichungen der Kommission für die Herausgabe der mittelalterlichen Bibliothekskataloge Deutschlands und der Schweiz).

Oeser, Wolfgang (1992), 'Beobachtungen zur Entstehung und Verbreitung schlaufenloser Bastarden. Eine Studie zur Geschichte der Buchschrift im ausgehenden Mittelalter', *Archiv für Diplomatik, Schriftgeschichte, Siegel- und Wappenkunde,* 38: 235–343.

Parkes, Malcolm B. (1969), *English Cursive Book Hands 1250–1500* (Oxford: Clarendon Press; Oxford Palaeographical Handbooks).

—— (2008), *Their Hands Before Our Eyes: A Closer Look at Scribes. The Lyell Lectures Delivered in the University of Oxford 1999* (Aldershot: Ashgate).

Simistira, Fotini, Manuel Bouillon, Mathias Seuret, Marcel Würsch, Michele Alberti, Rolf Ingold, and Marcus Liwicki (2017), 'ICDAR2017 Competition on Layout Analysis for Challenging Medieval Manuscripts', in *14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 1361–70. doi: <10.1109/ICDAR.2017.223>.

Smith, Marc H. (2004), 'Review of Albert Derolez (2003), *The Palaeography of Gothic Manuscript Books: From the Twelfth to the Early Sixteenth Century* (Cambridge: Cambridge University Press)', *Scriptorium*, 58: 274–279.

——— (2016), 'Learning to Look at Texts', in Marc H. Smith and Laura Light (eds), *Primer 9*: *Script* (New York: Les Enluminures), 2–7.

Stutzmann, Dominique (2009), *Écrire à Fontenay: Esprit cistercien et pratiques de l'écrit en Bourgogne (XIIe-XIIIe siècles)* (Thèse de doctorat Université Paris 1 Panthéon-Sorbonne).

Stutzmann, Dominique (2012), 'Paléographie latine et vernaculaire (livres et documents)', *Annuaire de l'École pratique des hautes études (EPHE), Section des sciences historiques et philologique, Résumés des conférences et travaux*, 143: 152–57.

——— (2014), 'Conjuguer diplomatique, paléographie et édition électronique: Les mutations du XIIe siècle et la datation des écritures par le profil scribal collectif', in Antonella Ambrosio, Sébastien Barret, and Georg Vogeler, *Digital Diplomatics: The Computer as a Tool for the Diplomatist* (Cologne: Böhlau; Archiv für Diplomatik, Beiheft ,14), 271–290.

——— (2015), 'Clustering of Medieval Scripts through Computer Image Analysis: Towards an Evaluation Protocol', *Digital Medievalist*, 10.<https://journal.digitalmedievalist.org/articles/10.16995/dm.61/> (accessed 19 November 2019).

——— (2016), *Competition on the Classification of Medieval Handwritings in Latin Script* (Paris: Institut de Recherche et d'Histoire des Textes). <http://clamm.irht.cnrs.fr/>. (accessed 19 November 2019)

——— (2017), 'Les "manuscrits datés", base de données sur l'écriture', in Teresa De Robertis and Nicoletta Giovè Marchioli (eds), *Catalogazione, storia della scrittura, storia del libro: I Manoscritti datati d'Italia vent'anni dopo* (Florence: SISMEL – Edizioni del Galluzzo), 155–207.

——— (2018), 'Variability as a Key Factor For Understanding Medieval Scripts: The ORIFLAMMS Project (ANR-12-CORP-0010)', in Stewart J. Brookes, Malte Rehbein, and Peter A. Stokes (eds), *Digital Palaeography* (London: Routledge, Taylor & Francis Group; Digital Research in the Arts and Humanities). <https://halshs.archives-ouvertes.fr/halshs-01778620> (accessed 19 November 2019).

——— (2019), 'Résistance au changement? Les écritures des livres d'heures dans l'espace français (1290–1550)', in Eef Overgaauw and Martin J. Schubert, *Change in Medieval and Renaissance Scripts and Manuscript: Proceedings of the 19th Colloquium of the Comité international de paléographie latine (Berlin, September 16-18, 2015)* (Turnhout: Brepols; Bibliologia, 50), 97–116.

Tensmeyer, Chris, Daniel Saunder, and Tony Martinez (2017), 'Convolutional Neural Networks for Font Classification', in *14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 985–990. doi: <10.1109/ICDAR.2017.164>.

——— (2017), *Submission for CLaMM Competition as Part of ICDAR 2017: Ctensmeyer/Clamm_2017*, Python. <https://github.com/ctensmeyer/clamm_2017>.(accessed 19 November 2019).

Article

Z-Profile: Holistic Preprocessing Applied to Hebrew Manuscripts for HTR with Ocropy and Kraken

Daniel Stökl Ben Ezra und Hayim Lapin | Paris

While considerable high quality textual and lexical data is openly available for many languages, such as Greek or Latin (although there is room for improvement here as well), researchers of classical Hebrew and Aramaic together with many other important European languages such as Armenian or Georgian are groping in the dark.[1] Most of the texts available online are vulgate editions, not scholarly reliable texts.[2]

Among the most important classical Hebrew texts are those redacted during the tannaitic Rabbinic period, around the third century CE: the Mishnah and the Tosefta, two juridical works, and the 'Halakhic' or 'Tannaitic' Midrashim, commentaries to the Bible (Exodus–Deuteronomy).[3] The Tosefta is a text closely related to the Mishnah that follows the same overall structure and clearly 'knows' the Mishnah, but it incorporates legal traditions with a complex intertextual relationship to those included in the Mishnah.

All of these sources have a strong interest in legal matters (the first two are in fact juridical texts) that illuminate Jewish life in Palestine, an Eastern province of the Roman Empire. Better known to the world outside of Jewish studies is the Babylonian Talmud, a later text, which is in fact a commentary to the Mishnah.[4] (An over-simplified analogy might be the relationship between two synoptic Gospels: the works share a recognizable outline and common materials but also are also distinct from one another.) The length of these texts is substantial, e.g. about two hundred thousand words for the Mishnah, about three hundred thousand words for the Tosefta, and they probably represent the most extensive sources from the pre-Christian Roman Empire that are still extant and not written in Greek or Latin. Their importance for our understanding of the development not only of classic rabbinic Judaism, but also of Roman provincial law and social history cannot be overstated.

Despite the importance of these texts there are neither full critical editions in either digital or print format. While there are several low quality online open source texts, there are no high-quality transcriptions that are openly available. An excellent linguistically annotated transcription of one manuscript of the Hebrew part of almost all texts can be accessed via the website of the Israel Academy of the Hebrew Language.[5] While access to its resources is free of charge, there are significant restrictions put on the use of the transcriptions by the Academy. Other projects put their transcriptions behind a substantial pay wall.[6]

Three years ago, Hayim Lapin from the University of Maryland and Daniel Stökl Ben Ezra from the École pratique des hautes études / Paris Sciences & Lettres in Paris have joined their respective projects on these texts in the eRabbinica project to start closing this gap. They secured funding from different sources for different subprojects. A pilot edition of three treatises of the

[1] E.g. Perseus Digital Library at <https://github.com/PerseusDL>. The ERC Project Lila at <https://lila-erc.eu/>. McGillivray 2013. The Classical Language Toolkit (cltk.org). Bizzoni et al. 2014. The HIMANIS (HIstorical MANuscript Indexing for user-controlled Search) and ORIFLAMMS (Ontology Research, Image Features, Letterform Analysis on Multilingual Medieval Scripts) projects have produced some Latin script data sets, see at <https://github.com/oriflamms/>. Dominique Stutzmann's most recent project, HORAE (The HORAE project: A textual exploration of Books of Hours, https://horae.digital/) will produce an even vaster data set.

[2] See e.g. the texts in Sefaria: A Living Library of Jewish Texts <https://sefaria.org> or on <https://en.wikisource.org>. This is certainly not due to disdain for such texts. Yet, OCR of out of copyright print editions are easier to acquire.

[3] Good introductions are Stemberger 2011, Ben-Eliyahu, Cohn and Millar 2013.

[4] 'The Talmud' usually refers to the Babylonian Talmud. In fact, there is a separate somewhat earlier Talmud from Byzantine Galilee called the Palestinian or Jerusalem Talmud.

[5] *Ma'agarim* (The Academy of the Hebrew Language, The Historical Dictionary Project), <http://maagarim.hebrew-academy.org.il/>.

[6] *Cooperative Development Initiative* <www.lieberman-institute.com>.

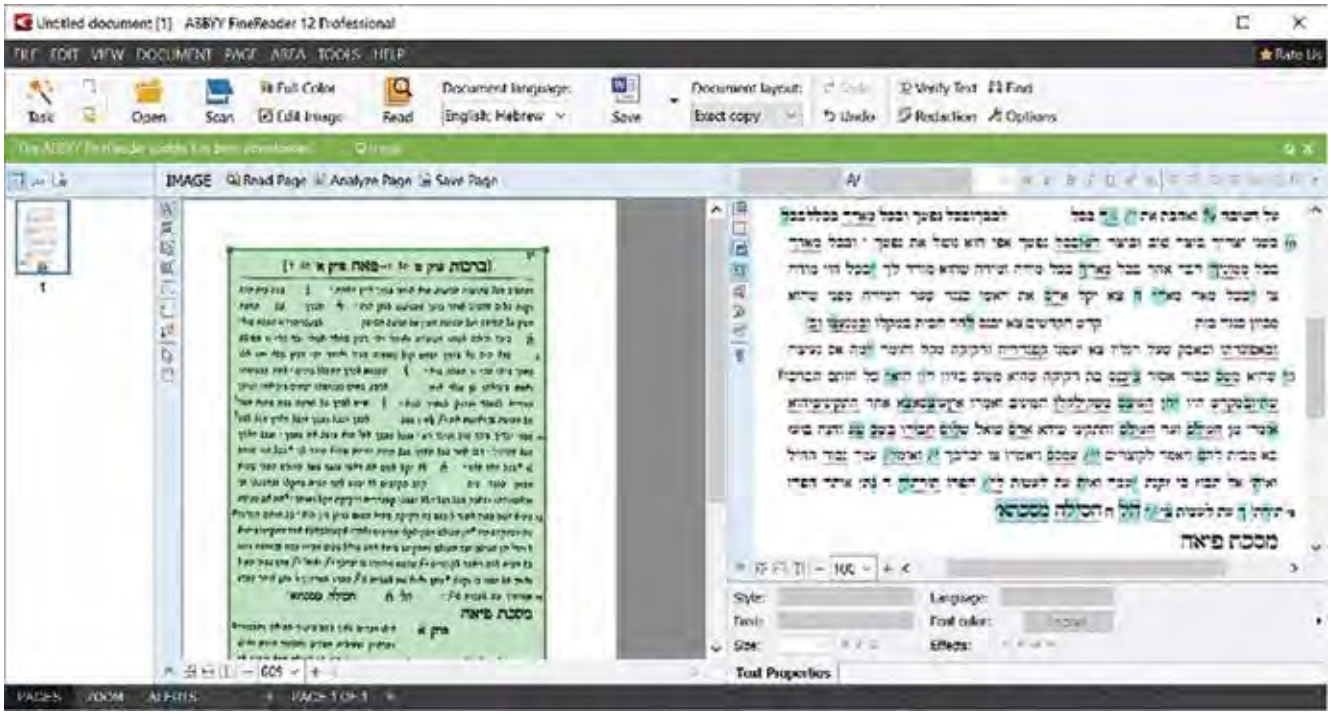


Fig. 1: Screen capture of a random page of Lowe OCRed with ABBYY.

Mishnah with transcription, automatic textual criticism, French and English translation and linguistic annotation, based on TEI/XML and the open source edition software TEI-Publisher by eXist[7] has recently been published.[8] This paper briefly introduces a selection of the computer vision and machine learning algorithms applied in our project following a chronological sequence to perhaps encourage others with similar projects, especially those coming from the humanities. If a given infrastructure constructed for a general purpose achieves bad results, there are means to arrive at good local solutions with open source code.

1. The original motivation: a tailor-made OCR

One of the most important manuscripts of the Mishnah is the Cambridge Ms. Add 470.1 from fifteenth century Byzantium.[9] In 1883, William H. Lowe published an extremely precise transcription that represents faithfully not only the text of the manuscript but also changes in writing style using various fonts and special placement of characters above the line for interlinear additions and at the end of lines or of paragraphs in the margins for marginal additions. Dots above letters indicate abbreviations and corrections.

At first, we tried to train a commercial OCR of this nineteenth century transcription, yet the multiplicity of fonts and special characters and ligatures and the use of the less common font (commonly, but imprecisely called *Rashi*) did not give very good results (Fig. 1).

Furthermore, it would have been difficult to preserve all the precious semantic information conveyed in the letter positions. Therefore, in 2016 Stökl Ben Ezra developed a tailor-made simple but very effective OCR engine. In a first step, horizontal projections were used to locate headers and footers, both of which were subsequently excluded from further analysis. The next step was the creation of a huge database of all connected components on the main part of all pages. The central architecture consists of a k-means clustering of 335 classes based on HOG-features (Histogram of Gradients) of connected components (letters or, in the case of ligatures, letter groups).[10] The vector for the Euclidean distance k-means clustering consisted of a concatenation

[7] TEI-Publisher <https://teipublisher.com/index.html>, accessed 26 May 2019.

[8] For the life interface, see *Digital Mishnah* <editions.erabbinica.org>. For the web interface code, see: <https://gitlab.existsolutions.com/mishnah/mishnah>. For the data see: <https://gitlab.existsolutions.com/mishnah/mishnah-data>.

[9] Accessible at <https://cudl.lib.cam.ac.uk/view/MS-ADD-00470-00001>.

[10] More precisely, letters consisting of several unconnected strokes, such as *he* or *qof* or all letters with a dot or stroke above or inside where combined into a single connected component via an intermediate vertical morphological transformation.

Fig. 2: Sample of Letter Clusters.

of HOG-features of 3 resized representations of each single connected component: 64×64 square and a flat 32×128 and a tall 128×32 rectangle with a cell size of 4×4 and 8×8, plus the height, width and the height-width proportion of each connected component. The 335 clusters were identified with characters manually. Clusters representing the same glyph with only visual but not semantic differences (e.g., the result of broken type) were grouped into one cluster, while clusters representing glyphs with semantically important differences (i.e. letters of different typefaces and sizes or with or without diacritical dots) were kept separate (Fig. 2).

Paragraph segmentation and identification of marginal additions were done with vertical profiles. Row segmentation was based on horizontal profiles. All connected components could then be assigned to rows. A combination of the clustering result and the centroid position plus the top and bottom boundaries vis-à-vis the row base-line served to evaluate whether a letter was superscripted and where it was to be placed on the horizontal axis. All this semantic layout position and font information was inserted into the transcription of the letters via tags. These tags where translated into Microsoft Word styles for visualization. Subsequently, the automatic transcription was corrected manually. We estimate that the precision of the automatic transcription was higher than 99.5% (Fig. 3).

2. Holistic manuscript layout analysis for writing block detection

Originally, we had anticipated applying the above-mentioned system to manuscripts for automatic alignment and transcription. After some manual clustering, the system indeed attained a transcription precision of about 85%. This was not high enough to replace hand transcription or even to provide a searchable 'background' transcription for the publication of images. The main challenge consisted in the letter segmentation of the connected script. With the help of morphological transformations we made some progress in dissecting connected letters, but the process was completely manuscript and scribe dependent and too labor intensive. Transcription-glyph alignment based on synthetic 'manu-scriptization' of the transcription was more successful, but still not precise enough for production.[11]

In a lecture in the e-philologie lecture series at PSL Université Paris, Marcus Liwicki mentioned Ocropy.[12] Despite the statement of Thomas Breuel that Ocropy was not suited for transcription of handwritten documents, the biLSTM of Ocropy is in fact quite powerful at least for certain medieval manuscripts.[13] Its central problem

[11] Some of these steps were presented as a poster in the CSMC conference in February/March 2016 by Stökl Ben Ezra.

[12] Breuel 2014.

[13] We made our first attempts in autumn 2016. Jean-Baptiste Camps reported useful results in 2017, too.

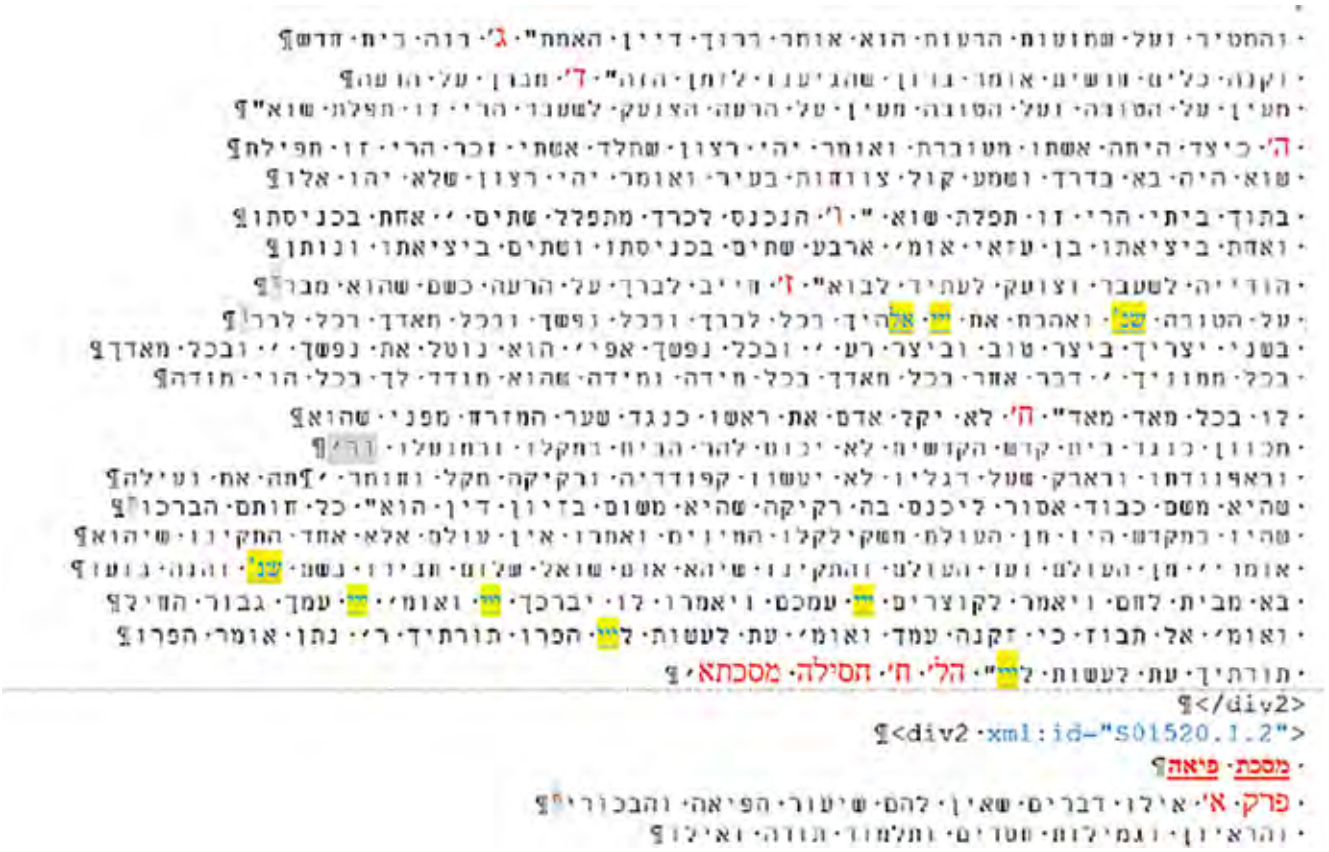


Fig. 3: Same page from Lowe in our OCR in Microsoft WORD output.

is the layout analysis, which suits the needs of printed documents but not the small and larger irregularities of manuscripts as well as binarization. The solution was to develop the binarization and column/writing block and line segmentation ourselves and to subsequently feed the preprocessing results into Ocropy. We should note, that we have since then moved from Ocropy to the more advanced OCR-engine Kraken by Benjamin Kiessling, because it is natively RTL (right-to-left), bidirectional and unicode enabled (but allows also for non-unicode codecs) and has a superior recognizer.[14] In the frame of the Scripta-PSL project, we are in the final stages of creating an open-source web-based infrastructure that integrates Kraken and will in the future also permit deep annotation of philological (additions, deletions etc.), historical (e.g. named entities), linguistic and palaeographical nature.[15] It is to this system that we will turn in the near future.

Our binarization is based on a sequence of well-known morphological transformations (Fig. 4): 1. closing with a structuring element in the form of a disk of a size depending on resolution and script size to calculate background, usually 30 or 50 gave excellent results; 2. deducing background from image to create foreground; 3. adjusting image intensity values of the foreground; 4. Otsu binarization of the resulting image. Despite its simplicity, the results were good enough on our material.[16]

With regard to layout analysis, Stökl Ben Ezra's approach was to better exploit the regularity of literary manuscripts. Strangely, documents are frequently considered as two-dimensional objects (even though pages are warped) in automatic document analysis and as a collection of individual pages. However, in particular for our corpus of Hebrew manuscripts, without illuminations, it seemed absurd to deal with pages of a manuscript one by one as if each one was completely new and unrelated to the others. Even if lines can be slightly oblique or curved, or paragraphs can be oblique, or there are frequent marginal additions, the page layout of these manuscripts of literary texts tends to be highly regular in plan. Columns, too, have a relatively constant position,

[14] Kiessling 2019.

[15] Stokes et al. forthcoming.

[16] Stökl Ben Ezra 2018.

Fig. 4: Binarization process: a) input; b) closing for background calculation; c) deducing b; d) image intensity values adjustment; e) Otsu binarization and inversion of d; f) direct Otsu binarization of a for comparison.

width and height. They can be interrupted by intermediary titles or empty space, but in principle their position on the page is quite regular.

Our 'holistic approach' takes into serious consideration the overlooked third dimension of manuscripts, the z-axis in addition to y and x.[17] Instead of a horizontal or a vertical profile of single pages, the method consists of calculating a z-profile from the vertical superposition of the two-dimensional images, as if an X-ray was looking through the manuscript and then applying horizontal and vertical profiles [18]

In a first step, a cube is calculated from all images of the manuscript.[19] They are transformed to grayscale and padded in order to fit the width and height of the largest picture. The z-profile is simply the sum of all images divided by their number. If the manuscript has been photographed as single pages, one can calculate the z-profile for even and odd pages apart. Average columns are calculated by setting all values below the median value of the average image to 0 and all those equal to or above the median value to 255. In a final step all connected components beyond the expected number of columns are deleted.

The resulting image is applied as a mask to each binarized individual page of the cube. It defines the area suspected to contain the centroid of the effective column(s) of each individual page. The median height of the remaining connected components is used for a vertical dilatation. Only connected components inside the masked area(s) are kept. All connected components inside an area become linked to each other via their centroids and holes are filled. The coordinates of the resulting connected component(s) are expected to agree to the coordinate of the major column(s) (Fig. 5).

The z-profile provides excellent information about the regularity of writing block disposition in a manuscript or printed book. Areas that are more frequently part of a writing block have higher z-profile values than those that are not. While this idea may appear extremely simple, it has proven very efficient in practice (Fig. 6).

This approach also makes it possible to calculate the variability of writing block width and height and the distance to marginal additions. Manuscripts with a very regular

[17] This absence is probably not by chance. There is no English adjective that would express the z-axis of an object in the analogue way as the terms horizontal and vertical. In fact, our current usage of the terminology in document analysis presumes a sheet hold vertically in the hand, not lying on a table, because vertical actually refers to the axis going up and down. In the Cartesian system the technical terms are abscissa axis and ordinate axis for x and y and applicate axis for z.

[18] Code available on GitHub at <https://github.com/ephenum/z-profile column_segmentation>.

[19] They may have to be reduced in size in order to fit the memory of the CPU.

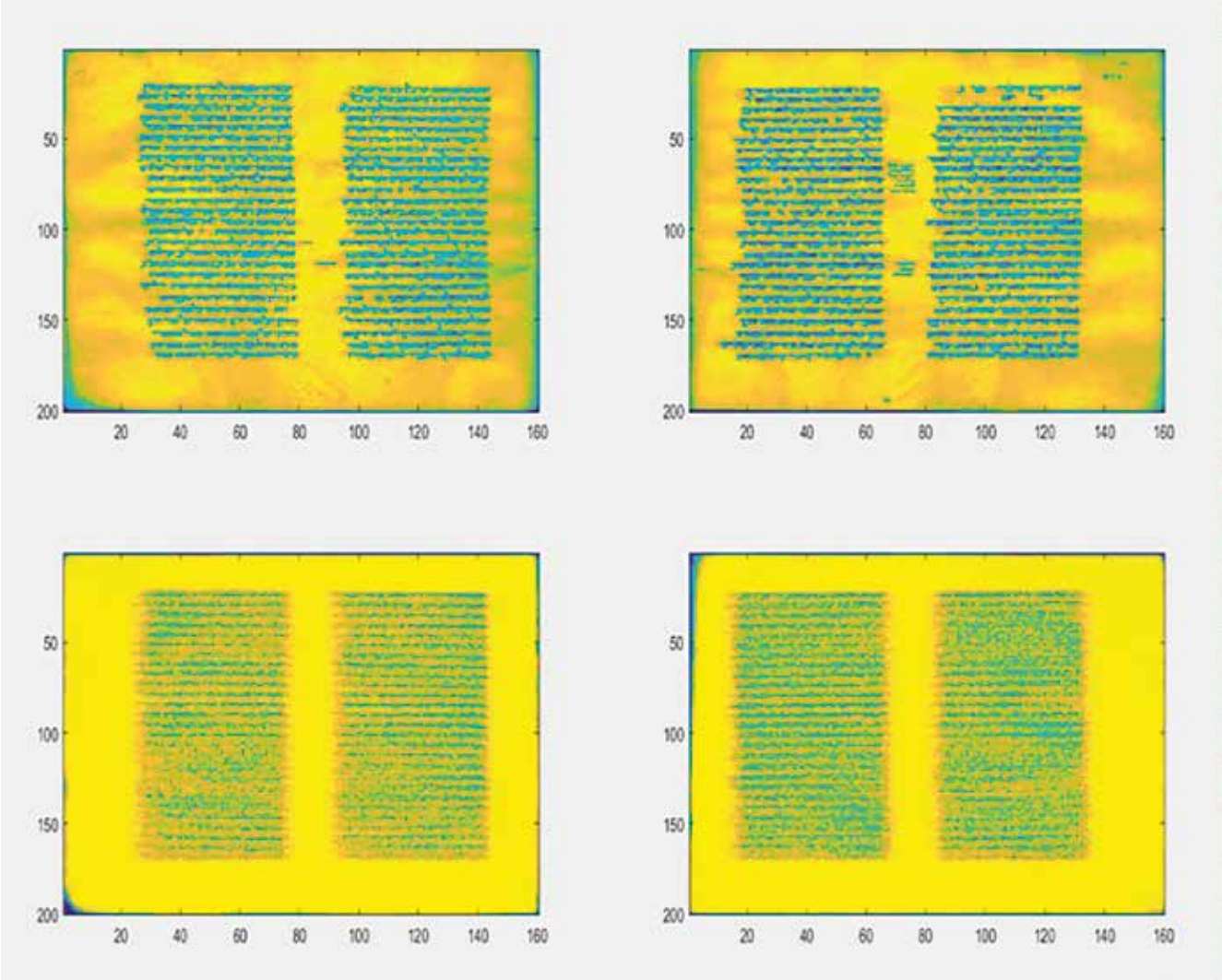


Fig. 5: The upper row shows two manuscript images of a manuscript written in two columns, where one can well discern the marginal additions particular to each page. The lower row shows the z-profile of the even and odd pages of the manuscript where only the main columns remain visible (Ms. Kaufmann A50 from the Library of the Hungarian Academy of the Sciences, Budapest).

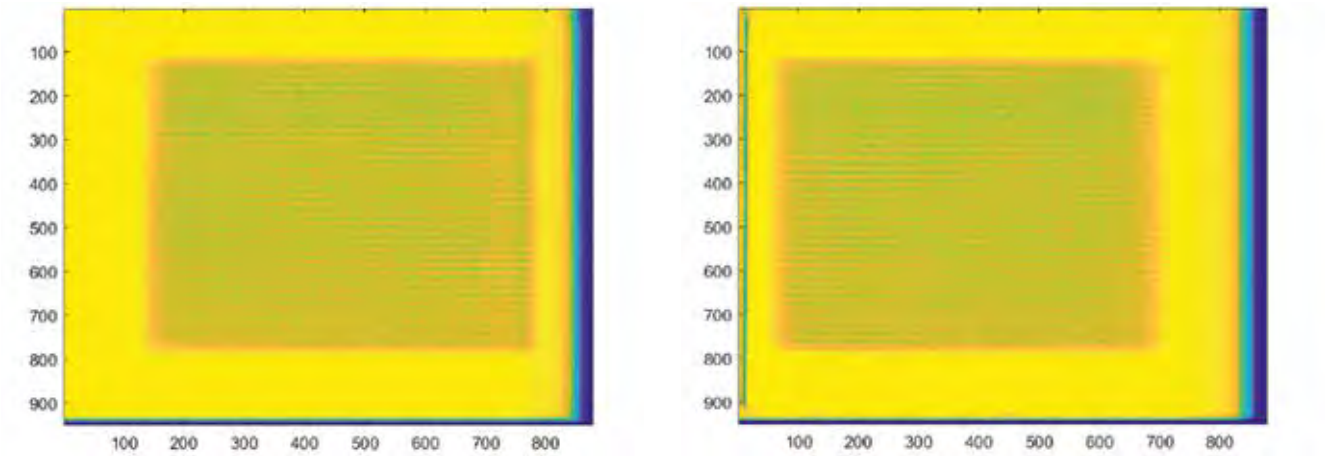


Fig. 6: Erfurt Tosefta z-profil . Even (left picture) and odd (right picture) pages.

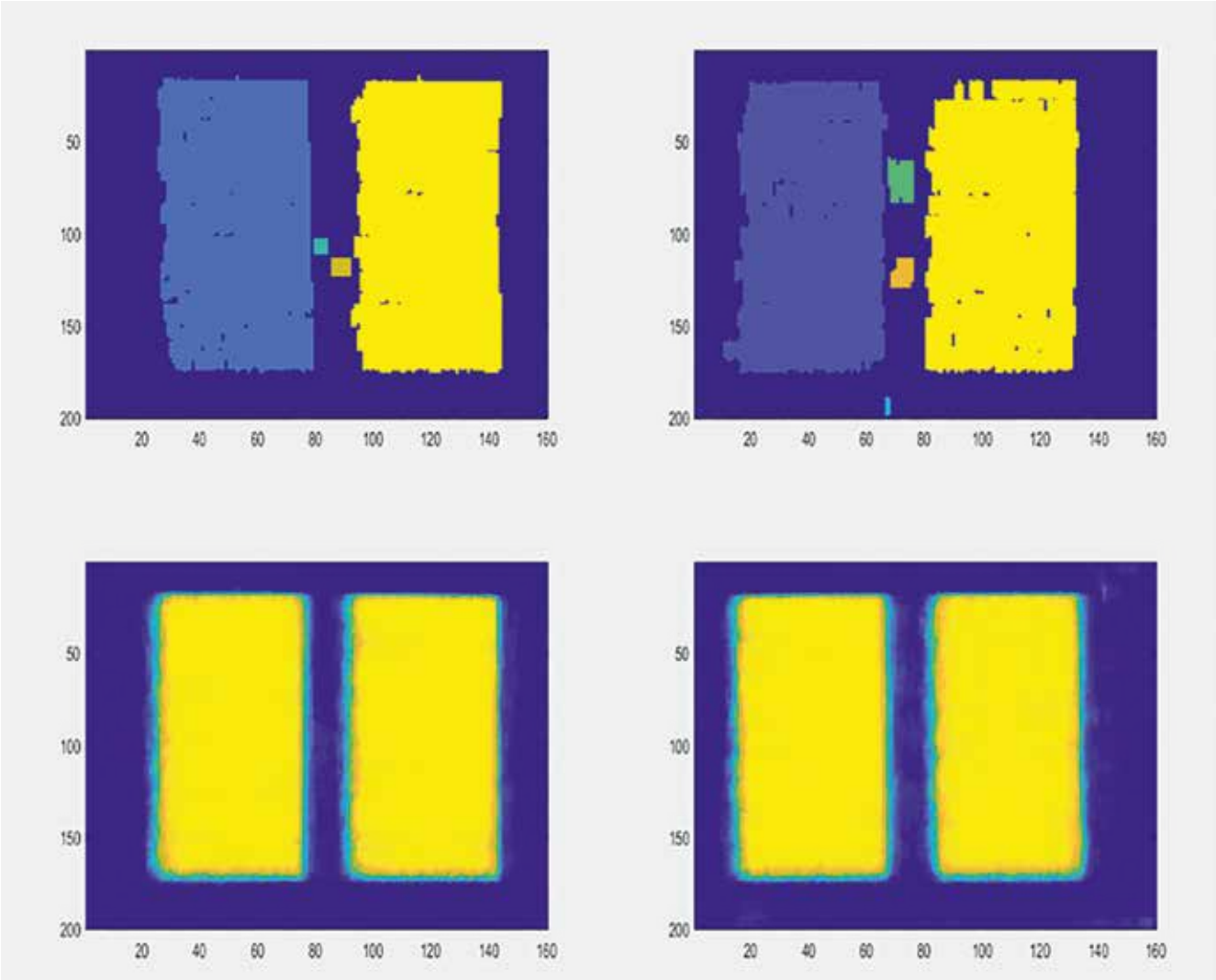


Fig. 7: In the top row an even and an odd page of the two column Kaufmann A50 manuscript after a morphological transformation. Main columns and marginal additions have been discerned with connected components into separate entities. In the bottom the z-profile a ter a morphological transformation.

layout have a very sharp z-profile, while manuscripts with a less regular layout have a more blurred z-profile. Marginal additions that are difficult to detect in a one-page-a-time approach, become discernible with the z-profile that distinguishes the normative basis from the addition. We should note that the system depends on a good binarization in order to distinguish between marginal additions and other dark areas, e.g. through deterioration of the manuscript or shadows. For this reason, we have started training an off-the-shelf Convolutional Neural Networks (CNNs) to distinguish between marginal areas with and without ink. However, as the main aim is the main text, this distinction is the cherry on the cake (Fig. 7).

Distinguishing between the z-profile of even and odd pages further sharpens the z-profile since many if not most manuscripts have a mirrored layout. Calculating the distance from the z-profile for each page can subsequently help to establish different z-profiles for different parts of the manuscript, e.g. for the material in the beginning and the end of the book, or for pages that commence a new chapter, pages with illustrations or tables etc. Exploiting this information might also improve existing algorithms, probably even in the age of convolutional neural network layout analysis.

3. Line segmentation with the heartbeat-seamcarve algorithm

Most recent line-segmentation approaches strive to find a general solution for the ultimate problem of finding any line in any orientation and on any position of any document

Fig. 8: Ms. Kaufmann A50 164[ra] with automatic line segmentation of seemingly easy material. On the left side Transkribus. On the right side heartbeat seamcarve.

image. CNNs have achieved excellent results on such cases.[20] However, on the main texts of regular manuscripts standard approaches may be just as good or even superior because current CNNs have a problem with large blank spaces (*vacat*s) inside a semantic line, especially, e.g. if one considers poetic manuscripts. Nevertheless, the standard approaches can be unsatisfactory even on a completely regular, quite simple manuscript as the following screenshots show (Fig. 8). The left image was produced using the state-of-the-art infrastructure Transkribus, while the image on the right uses the heartbeat seamcarve method. Both show the same random page of the test manuscript (Ms. Kaufmann A50).

The Transkribus looks clean from afar, yet, it has eight errors (lines 1, 2 (2×), 9 (2×), 10 (2×), 13), while the heartbeat seamcarve has no mistake in line recognition.[21] Tests with the CNN algorithm resulted in significant errors that did not arise with the heartbeat seamcarve (Fig. 9).

In the Middle Ages, preparing parchment and paper for writing was an expert task. In literary manuscripts, pages very frequently have been carefully ruled before the inscription, indicating lines, columns and/or writing blocks. In literary manuscripts, the distance between these lines is mostly very regular. While the lines can be empty, end early or start in the middle of the column or be interrupted by a large *vacat*, the vertical distance is mostly as constant as the width of the columns (but not the length of each individual line). Very often, the z-profile picture reveals not only the columns but even the number of lines (see Fig. 5, above).

One of the line segmentation algorithms is seam-carving.[22] The success of the seam-carving algorithm depends to a large extent on single column segmentation and on the correctness of the detection of median lines. If, however, a line ends early and the subsequent line starts late, the simple median line approach will consider the second a continuation of the first which will result in the two lines being seam-carved as a single line.

[20] As shown, e.g. in Diem et al. 2017 and Kiessling et al. (submitted).

[21] It cuts the head of two *lameds* in line 9 and 13 but transkribus seems to cut all ascenders and descenders (at least in its visualization).

[22] Saabni and El-Sana 2011.

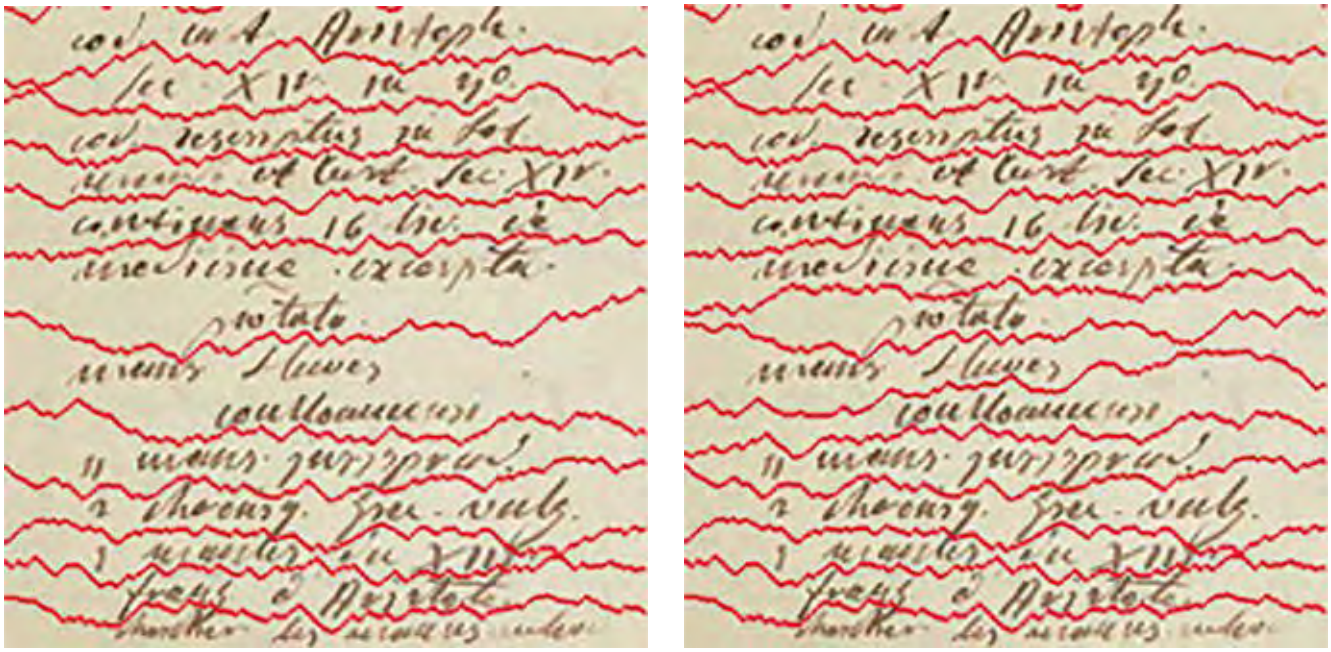


Fig. 9: Seamcarve on a sample from the Washingon[25] dataset. Left without, right with heartbeat.

Mathias Seuret, Marcus Liwicki and Stökl Ben Ezra started to improve the seam-carving algorithm by the assumption of regularity in the analysis of the manuscript.[23] Based on a Fourier transformation of the horizontal projection of each of *n* slices of a writing block, the procedure calculates the median line length. Wherever the line is too short or empty and therefore the horizontal profile misses a peak, the algorithm adds one or several artificial peaks according to the regular line distance with regard to the lines above and below. The algorithm is now implemented in the DIVAServices.[24]

4. Manuscript transcription and transcription-glyph alignment

Once a pipeline for the production of relatively clean manuscript-line-image and transcription was established, we were able to train models with Ocropy showing useful results. A preliminary step was data augmentation. We used the well-known methods of salt and pepper as well as shearing of the manuscript line image in different dosages, angles and combinations to multiply input pairs by a factor of nine.

A challenging stage was the production of transcription text lines that correspond to the visible signs in the main text block. All marginal or interlinear additions had to be deleted. On the other hand, all deletions of the main text by simple strikethrough had to be kept. Numbers and paratext such as *eschatocols* of chapters or treatises had to be kept. Letters functioning as simple line fillers without importance for the linguistic text, a frequent practice in Hebrew manuscripts, had to be kept, too. Abbreviations had to remain unresolved. Ligatures had to be represented by special marks. Our transcription markup distinguished between the various forms of addition and deletion and it was mainly a question of the order of transformation steps.

For the preparation of the most complex manuscript, we used Microsoft Word with numerous styles to emulate XML tagging because XML editors like Oxygen are still difficult to manage with RTL scripts whose writing direction counters that of the tags. Hayim Lapin wrote a series of conversion scripts to convert the Word documents (docx) to XML/TEI that used Visual Basic to prepare the transcription for transformation to raw XML, and subsequent conversion to a TEI schema-conformant document using XSLT. The forthcoming eScriptorium platform will permit us to combine automatic layout analysis and HTR with deep annotation.

So far, we have applied our pipeline to the following Hebrew manuscripts:

- K: Mishnah: Ms. Kaufmann A50 in the Library of the Hungarian Academy of the Sciences, Budapest. Written in Italian script from the eleventh or twelfth century. 256 folios.

[23] Seuret, Stökl Ben Ezra, and Liwicki 2017.

[24] Würsch, Ingold, and Liwicki 2017.

[25] Fischer et al. 2012.

- C: Mishnah: Ms. Add. 470.1, Cambridge University Library. Written in Byzantine script from the fifteenth century. 250 folios.
- M: Mishnah: Cod. Ebr. 95 of the Bayerische Staatsbibliothek, Munich (on the part of the Mishnah only because the resolution is very low for the tiny script of the Talmud itself). The manuscript was written in 1342 probably in France. 576 folios. Semi-manual manuscript layout segmentation of the complex Talmudic layout was very kindly provided with by the Larex team around Christian Reul.
- V: Tosefta: Cod. Hebr. 20, Austrian National Library, Vienna. Written around the fourteenth century in square Sephardic script. 327 folios.
- L: Tosefta: Add. 27296, British Library, London. Written in fifteenth century Sephardic script. 73 folios

This makes altogether ca. 1,400 pages. For most manuscripts, we could achieve a character error rate (CER) <5%, sometimes <3%. We tried different sizes for the hidden layer, different learning rates, different sizes of training data. We also mixed training data from different manuscripts with encouraging results. 5 columns (171 lines, 200 neurons) achieved a CER <10% for the Vienna manuscript, while 19 columns (645 lines) sufficed for 2.1% CER (43k iterations). Due to limited manpower and calculation power (Ocropy runs on CPUs only and demands the transformation of all RTL scripts into LTR), we did not apply all tests on all materials. The main aim was not to improve the LSTM but to arrive at exploitable results.[26] Of course, in a manuscript of 1,000,000 characters, 3% CER still means 30,000 errors to spot and to correct. However, where a vulgate text is available or one manuscript of a text is already transcribed, we can automatically align both versions with CollateX or with Shmidman-Koppel-Porat algorithm.[27]

5. Outlook to the Future

Since June respectively September 2018, two new projects have started around eRabbinica, both with the National Library of Israel and their manuscript portal Ktiv.[28] In Tikkoun Sofrim ('scribal error correction') with Haifa University we have worked on correction of automatic transcriptions of post-classical Midrashim of the Tanhuma-Yelamdenu genre via crowdsourcing.[29] We have so far transcribed four manuscripts with CER's of 2.8%, 2.9%, 6.9% and 8.9%.[30] The manuscript with 8.9% CER has been the first to be submitted to the crowdsourcing process and we have been able to reconstruct a complete text with the help of CollateX and a majority vote on the word level.[31] The plan is to link coordinates for words, and where possible glyphs via IIIF to the manuscript images and to integrate them with the help of the Mirador viewer at Ktiv. The National Library will serve as repository for long-term preservation.

In the second project, Sofer Mahir (*tachygraph*, or 'rapid [i.e., skilled] scribe') with the University of Maryland and Dicta[32] we collaborate on the creation of a pipeline to produce open source manuscript transcriptions of all major manuscripts of the principal tannaitic compositions: approximately twenty substantial manuscripts with about 6,000 pages. In collaboration with Dicta, the texts will be automatically analyzed linguistically. We hope to be able to integrate the linguistic analysis directly into the transcription pipeline to further reduce the error ratio. In the LAKME project, we have already annotated 25,000 words lexically and morphologically and created the corresponding lexicon in French, English and German in order to apply a neural network architecture developed by Dicta on all of our transcriptions.[33] In a related project, Lapin is creating a database of shared text (identified by string matching) among the corpora that will be part of the infrastructure of future editions.

The resulting text will be presented according to the developing Distributed Text Service (DTS) API (Application Programming Interface).[34] An extension of the Canonical Text Service (CTS) specification, first developed for the

[26] Results will be published on the eRabbinica.org website and on GitHub after full transition from Ocropy to Kraken.

[27] Shmidman, Koppel and Porat 2018.

[28] <http://web.nli.org.il/sites/nlis/en/manuscript>.

[29] Wecker et al. 2019. Tikkoun Sofrim website: <https://tikkoun-sofrim.firebaseapp.com/>. Tikkoun Sofrim GitHub: <https://github.com/drore/tikkoun>.

[30] Kuflik et al. 2019

[31] Dekker et al. 2015, Decker and Midell 2011.

[32] <http://dicta.org.il/>.

[33] Stökl Ben Ezra et al. 2018.

[34] <https://distributed-text-services.github.io/specifications/> (accessed 26 May 2019).

Homer Multitext Project,[35] the DTS specification provides a common, predictable protocol for sharing texts at various levels of granularity, as well as information about texts and the collections they appear in. A number of significant projects have begun to use the CTS/DTS specification, among them the Perseus, Kitab (Knowledge, Information Technology, and the Arabic Book), and ARTFL Encyclopédie.[36] This approach has internal advantages for the project (for instance, it allows us to build applications around texts without constructing a purpose-built querying system). As the adoption by multiple projects working in diverse languages and types of texts suggests, the CTS/DTS model provides a standardized way of sharing information between projects, so that (as we have learned all too well in our own work) each project does not have to reproduce the work of every other.

The amount of work put into the pipeline, the preparation of the training data and the correction is very substantial. Our preliminary measurements agree with results of previous teams that it is more time-consuming to correct a text with a CER > 10% than to transcribe it from scratch. For the London Tosefta, e.g. the correction of a page created with the first model, trained with only 15 pages (376 lines), took about 30 minutes per page, about the same time as a transcription from scratch. However, after the correction of 20 pages, we trained a new model whose correction only takes ca. 15 minutes per page on the average and we hope to further reduce this amount of time with the next model. The amount of human labor involved in extracting the data from the current interface and launching a retraining is about 2 hours. In the very near future, however, this procedure will just demand a series of button clicks inside the eScriptorium platform. While for short manuscripts the effort to automatize the transcription and then correct it manually may be greater than the time needed to prepare a completely manual transcription, the immediate humanist gain is the direct link between image and transcription. Doing this by hand for manuscripts would usually involve much more human labor. Secondly, the automatization of the transcription will improve with the quantity of ground truth acquired and with future improvement of algorithms for layout segmentation and transcription as well as language modelling. Ergonomics for ground-truth-creation and for post-correction tools will be further improved. We will attain a stage soon, where we can transcribe most of the historical manuscripts automatically with a grade of precision sufficiently high for human reading and machine exploitation. For Medieval Hebrew, we are on the brink of doing it.

[35] <https://www.homermultitext.org/> (accessed 26 May 2019).

[36] Perseus: <http://www.perseus.tufts.edu/hopper>; Kitab: <http://kitab-project.org; Encyclopédie: http://encyclopedie.uchicago.edu>.

REFERENCES

Ben-Eliyahu, Eyal, Yehudah Cohn, and Fergus Millar (2013), *Handbook of Jewish Literature from Late Antiquity: 135–700 CE* (Oxford: Oxford University Press).

Bizzoni, Yuri, Federico Boschetti, Riccardo Del Gratta, Harry Diakoff, Monica Monachini, and Gregory Crane (2014), 'The Making of Ancient Greek WordNet', in Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis (eds), *Proceedings of the Ninth International Conference on Language Resources and Evaluation* (European Language Resources Association), 1140–1147.

Breuel, Thomas (2014), *Ocropy: Python-based tools for document analysis and OCR.* <https://github.com/tmbdev/ocropy> (accessed 6 November 2019).

Camps, Jean-Baptiste (2017), 'Homemade Manuscript OCR (1): Ocropy'. <https://graal.hypotheses.org/786> (accessed 6 November 2019).

Dekker, Ronald H., and Gregor Middell (2011), 'Computer-Supported Collation with CollateX: Managing Textual Variance in an Environment with Varying Requirements', *Supporting Digital Humanities*, University of Copenhagen, Denmark, 17–18 November 2011. <*https://doc.anet.be/docman/docman.phtml?file=.irua.4ec3d9.935894d4.pdf*> (accessed 6 November 2019).

—— , Dirk van Hulle, Gregor Middell, Vincent Neyt, and Joris van Zundert (2015), 'Computer-supported Collation of Modern Manuscripts: CollateX and the Beckett Digital Manuscript Project', *Literary and Linguistic Computing*, 30: 452–470.

Diem, Markus, Florian Kleber, Stefan Fiel, Tobias Grüning, and Basilios Gatos (2017), 'Cbad: ICDAR 2017 Competition on Baseline Detection', *International Conference of Document Analysis and Recognition*, 1355–1360. doi: 10.1109/ICDAR.2017.222.

Fischer, Andreas, Andreas Keller, Volkmar Frinken, and Horst Bunke (2012), 'Lexicon-Free Handwritten Word Spotting Using Character HMMs', *Pattern Recognition Letters*, 33.7: 934–942.

Kiessling, Benjamin, Daniel Stökl Ben-Ezra, and Matthew T. Miller (2019), 'BADAM: A Public Dataset for Baseline Detection in Arabic-script Manuscripts', *Historical Document Imaging and Processing, International Conference of Document Analysis and Recognition*. *arXiv.org*: <arXiv:1907.04041> (accessed 6 November 2019).

—— (2019), 'Kraken – an Universal Text Recognizer for the Humanities', *DH2019*, Utrecht University, 8–12 July 2019. <https://dev.clariah.nl/files/dh2019/boa/0673.html> (accessed 6 November 2019).

Kuflik, Tsvi, Moshe Lavee, Daniel Stökl Ben Ezra, Avigail Ohali, Vered Raziel-Kretzmer, Uri Schor, Alan Wecker, Elena Lolli, and Signoret, Pauline (2019), 'Tikkoun Sofrim – Combining HTR and Crowdsourcing for Automated Transcription of Hebrew Medieval Manuscripts', *DH2019*, Utrecht University, 8–12 July 2019. <https://dev.clariah.nl/files/dh2019/boa/0568.html> (accessed 6 November 2019).

Lowe, William H. (1883), *The Mishnah on Which the Palestinian Talmud Rests*, (Cambridge: Cambridge University Press).

McGillivray, Barbara (2013), *Methods in Latin Computational Linguistics* (Leiden: Brill).

Saabni, Raid, and Jihad El-Sana (2011), 'Language-Independent Text Lines Extraction Using Seam Carving', *International Conference on Document Analysis and Recognition,* 563–568. doi: 10.1109/ICDAR.2011.119.

Seuret, Mathias, Daniel Stökl Ben Ezra, and Marcus Liwicki (2017), 'Robust Heartbeat-based Line Segmentation Methods for Regular Texts and Paratextual Elements', *Historical Document Imaging and Processing, International Conference of Document Analysis and Recognition,* 71–76. doi: 10.1145/3151509.3151521 (accessed 6 November 2019).

——, Daniel Stökl Ben Ezra, and Marcus Liwicki (2018), 'Heartbeat Seamcarve Line Segmentation'. <https://github.com/ephenum/heartbeat_seamcarve_line_segmentation> (accessed 6 November 2019).

Shmidman, Avi, Moshe Koppel, and Ely Porat (2018), 'Identification of Parallel Passages Across a Large Hebrew/Aramaic Corpus', *Journal of Data Mining & Digital Humanities, Special Issue on Computer-Aided Processing of Intertextuality in Ancient Languages, Towards a Digital Ecosystem: NLP. Corpus infrastructure. Methods for Retrieving Texts and Computing Text Similarities*. *arXiv.org*: <arXiv:1602.08715v2>.

Stemberger, Günter (2011), *Einleitung in Talmud und Midrasch* (Munich: C.H. Beck).

Stokes, Peter Anthony, Daniel Stökl Ben Ezra, Benjamin Kiessling, and Robin Tissot (2019, forthcoming), 'EScripta: A New Digital Platform for the Study of Historical Texts and Writing', *DH2019*, Utrecht University, 8–12 July 2019.

Stökl Ben Ezra, Daniel (2016), *Why Should Philologists Learn Computer Vision*, CSMC Conference Poster, February/March 2016. <https://www.academia.edu/22887575/Why_should_philologists_learn_computer_vision> (accessed 6 November 2019).

—— (2018), 'Simple Binarization for Manuscript Images'. <https://github.com/ephenum/binarization> (accessed 6 November 2019).

——, Avigail Ohali, Emmanuelle Main, Hayim Lapin, Avi Shmidman, Shaltiel Shmidman, Meni Adler, Yael Netzer, and Thierry Poibeau (2018), 'The Creation of a Lemmatized and Morphologically Annotated Open Source Corpus for Rabbinic Hebrew', in Jean-Baptiste Camps, Thierry Poibeau, and Daniel Stökl Ben Ezra, *Creating Richly Annotated Linguistic Corpora for Languages with Few Linguistic Resources*, Multiple Paper Panels, Conference of the European Association for Digital Humanities (EADH), Oral presentation, Galway, 2018. <https://eadh2018.exordo.com/programme/session/41> (accessed 6 November 2019).

—— (2019), 'z-Profile Column Segmentation'. <https://github.com/ephenum/z-profile_column_segmentation> (accessed 6 November 2019).

Wecker, Alan, Daniel Stökl Ben Ezra, Vered Raziel-Kretzmer, Uri Schor, Tsvi Kuflik, Avigail Ohali, Dror Elovits, Moshe Lavee, and Pauline Stevenson (2019), 'Tikkoun Sofrim: A WebApp for Personalization and Adaptation of Crowdsourcing Transcriptions', *27th Conference on User Modeling, Adaptation and Personalization UMAP'19*, Adjunct Publication (Larnaca. New York: ACM Press). doi: 10.1145/3314183.3324972.

Würsch, Marcel, Rolf Ingold, and Marcus Liwicki (2017), 'DIVAServices—A RESTful Web Service for Document Image Analysis Methods', *Digital Scholarship in the Humanities*, 32.suppl.1: i150–i156. doi: 10.1093/llc/fqw051.

Article

On Digital and Computational Approaches to Palaeography: Where Have we Been, Where Are we Going?

Peter A. Stokes | Paris

Although palaeographers have been using technological developments since the discipline's beginning, nevertheless the last decade has seen an enormous increase of and very rapid advances in computational approaches to the field. This seems likely to prove highly transformative and disruptive and has not gone unnoticed by palaeographers. Indeed, debate has been ongoing for many years now about the nature of palaeography as an 'art' or a 'science', as well as the role (if any) of quantitative or other 'hard' evidence in palaeographical studies.[1] Originating over a century ago, if not longer, this debate has resurfaced with the ready availability of digitised images of manuscript pages, along with relatively powerful personal computers and developments in machine vision, artificial intelligence, and related topics. At the same time, recent years have also seen increasing interest from computer science and related fields; this is demonstrated by increasing recognition of historical materials at conferences such as the International Conference for Handwriting Recognition (ICFHR) and the International Conference for Document Analysis and Recognition (ICDAR), by centres like the CSMC, and by the recent creation of a chair in digital and computational humanities applied to the study of historical writing in France. Although fears of a split into 'digital' and 'non-digital' palaeography now seem thankfully unwarranted,[2] nevertheless real questions remain about how the many different approaches can and should best be integrated, ensuring that all parties gain through fruitful dialog and that none simply supplies services or 'answers' to the other. The objective of this paper is therefore to review these developments, focussing not on technical advances *per se* but rather on the 'view from the (Digital) Humanities', making comparison with previous reviews to suggest some points around where we have been and where we might be going.[3]

1. A (very) brief overview of digital and computational work to date

Significant work in the last decade or more has been done on topics such as automatic analysis of handwriting for dating, localising and (especially) writer identification, script classification, and layout analysis. This work has typically drawn on the very large volume of digitised material and has developed and applied methods in machine vision and artificial intelligence to their analysis. Recently, good progress has been made in fields such as line detection, Handwritten Text Recognition (HTR), wordspotting of handwritten documents, as well as automatic layout analysis and the dating of script.[4] More specifically aligned to palaeographical research is work on the characterisation of script.[5] Success has also been achieved in aligning images of text to pre-existing transcripts, as well as the identification of fragments of manuscripts from the same original document, and the identification of specimens of script likely to have been written by the same individual.[6] Also relevant here are methods in Machine Vision such as automatic feature extraction, whereby the software determines automatically what aspects of the image are significant. This raises the possibility that such features could be meaningful and useful

[1] The discussion has been summarised briefly by Stokes 2015; relatively recent contributions include Costamagna et al. 1995 and 1996; Gumbert 1998; Derolez 2003, 1–9; Ciula 2005; Sirat 2006, 309ff and 493; Canart 2006; Davis 2007; Schomaker 2007; Stokes 2009; Stokes 2012; and Stutzmann 2011, among others.

[2] For an expression of these concerns see Stokes 2009.

[3] Previous reviews include Hassner et al. 2013 and Hassner et al. 2015.

[4] Examples of such work are given by Sudholt and Fink 2016, applied to Cuneiform by Rothacker et al. 2015, and recent work for the project HIMANIS; Chen and Seuret 2017, Seuret, Stökl ben Ezra and Liwicki 2017, and Xue Yu's recent success at ICDAR 2017; Christlein, Gropp and Maier 2017; and Kestemont, Christlein and Stutzmann 2017, respectively.

[5] See, for example, Kestemont, Christlein, and Stutzmann 2017.

[6] See, for example, Fischer et al. 2011a and 2011b; Stutzmann et al. 2015; Sadeh et al. 2015; Wolf et al. 2011a; Wolf et al. 2011b; and work by Schomaker et al. such as He, Wiering and Schomaker 2015, respectively.

to palaeographers, perhaps even highlighting new areas on which to focus. Indeed, there are standard techniques in machine vision which can be helpful here, such as the so-called 'Bag of Words' which automatically builds a set of visual elements which make up the image; these visual elements are established by the machine and so may not be meaningful to scholars, but in principle they are very similar to the graphic elements referred to by palaeographers such as 'wedge', 'foot-serif", 'hook' and so on, and some work has been done on making these automatic 'visual words' meaningful to palaeographers as well.[7] Similarly, Deep Learning has recently been applied to the classification of medieval script (among many other things), and not only was this technique successful in itself, but it was also possible to 'look inside' the system to see what elements of the script it found distinctive.[8]

Another important development is the degree to which these methods and techniques are becoming freely available for use by other projects, through free or Open Source software but also through web APIs, a noteworthy example of which is the suite provided by the DIVA group at the University of Fribourg.[9] This is potentially a significant boon to palaeographers with some understanding of digital methods but without the resources or expertise to implement their own code, and it may also go some way towards addressing the need for benchmarking and standard algorithms.[10] This refers to the need in computer science to have standard sets of data which different groups can use in order to establish meaningful comparisons between algorithms. However, the very existence of these benchmarking datasets tends to encourage people to produce algorithms that are suited to that data, and so it is important that the datasets themselves give a good representation of the range of manuscript material. For this reason increasing numbers of such datasets are becoming available to cover a wider range of historical material, as for instance in competition datasets produced by the IRHT for use at ICDAR.[11] This is badly needed, as the 'standard' datasets that are typically used for benchmarking are those such as the George Washington Papers which comprise images of documents written by the famous eighteenth-century general.[12] Although a useful example, his writing bears very little resemblance at all to medieval manuscripts from five hundred or a thousand years earlie .

In addition to this, important work is also continuing on how to better present and interact with palaeographical knowledge in tractable ways through symbolic representation and modeling, interface design and UI/UX, visualisation, and so on. An example of this is DigiPal and its successors, particularly Archetype, which focussed on structured descriptions of handwriting which researchers could enter into software. This involves manually drawing annotations on images of handwriting and entering descriptions of the letters. Because the database already contains information about the components or essential elements of letters (for instance that **b**, **h** and **l** all contain ascenders), it therefore becomes very easy for researchers to find and compare forms in ways that are palaeographically meaningful (for instance searching for examples of ascenders by a particular scribe or from a particular region). The emphasis is firmly based on knowledge creation through experimentation, exploration and visualisation, as well as the communication of evidence to support the resulting argument.[13] This approach has proven successful insofar as the methods and results are now increasingly used by palaeographers and others, and have been successfully applied to many different writing systems.[14] This aspect of discovery, analysis and communication relates directly to larger questions in Digital Humanities and beyond about how one represents expert knowledge in systems that are tractable to the computer, connecting to areas and technologies such as ontologies, formal modelling, Linked Open Data and the Semantic Web. In this respect other important projects on modelling manuscripts and documents are relevant here, particularly ORIFLAMMS but also

[7] Examples here include Wolf et al. 2011; Hassner et al. 2015; and Kestemont, Christlein and Stutzmann 2017.

[8] Kestemont, Christlein and Stutzmann 2017; compare also Rajan Sampath 2016 for a different approach.

[9] Eichenberger et al. 2015; Garz et al. 2015; Würsch, Ingold and Liwicki 2016; other examples include Sudholt and Fink 2016 and Kestemont, Christlein and Stutzmann 2017.

[10] Hassner et al. 2013, and see also discussion below.

[11] Kestemont, Christlein and Stutzmann 2017.

[12] Washington Database 2016, and compare those in Kestemont, Christlein and Stutzmann 2017. For this need see also Hassner et al. 2015, 118, among others.

[13] Brookes et al. 2015.

[14] Brookes et al. 2015; Stokes 2017.

Europeana, Biblissima, IIIF, CRMtex and others.[15] These tend to focus on the document more than the script, but some formal ontologies are starting to emerge for describing writing itself. ORIFLAMMS, for instance, established important groundwork towards a substantial ontology for script, work which is still continuing.[16] CRMtex adds a small extension to the CIDOC-CRM model for museums and objects: it comprises entities TX1 Written Text, TX2 Writing, TX3 Writing System and TX4 Written Field, as well as properties such as TXP1 Used Writing System.[17] More complex are the ontologies developed for the IDIOM project to model Mayan hieroglyphics which include not only the distinction between graphs and signs but also represent graphs that are derived from other graphs, allographic relationships between signs, reading hypotheses through sign functions, confidence levels of different readings, and so on.[18] Although not (yet) expressed as a formal ontology, a further model which focuses more on the graphical aspect of writing and written communication is that developed for the DigiPal project which includes structural relationships between allographs and graphs, relations between different graphs written by the same scribe, and so on.[19] Similarly, models and ontologies for manuscripts as objects have been developed as part of several projects. One of these is an event-based model for Hebrew manuscripts which seeks to model the object's history and its cultural context.[20] Another ontology has been published by the Biblissima team which extends FRBRoo with entities such as Binding, Annotation, Digital Surrogate, Electronic Edition and more, with relationships such as Used as Source, Used as Copy, Used Style of Script, and so on.[21] IIIF, on the other hand, is less an explicit model and more a standard for interoperability of images, including not only the images themselves but also regions, basic manipulation such as rotating, but also relationships between images such as sequences of images representing consecutive pages in a book.[22]

This is of course a small overview of just some of the work that is underway towards modelling books and writing, but the promise of such developments is starting to be seen. Although by no means using the most advanced computational methods, IIIF in particular is nevertheless transforming the way in which scholars in the Humanities are working with manuscripts today. The projects together provide (among other things) stable protocols for addressing and manipulating images of manuscripts and other cultural heritage over the Web, in a system which is becoming increasingly widely used by libraries, archives and other cultural heritage institutions. This means that we are now able to refer unambiguously to images of manuscript pages and to regions in those images, and to access the images directly from many different repositories. For the Humanities researcher this means access to material and – really for the first time – the ability to easily compare images from different institutions in the same software. It also responds very directly to the need for ready and open access to data which has been noted in the Dagstuhl events and elsewhere.[23]

2. Some Continuing or Future Directions?

The necessarily very brief survey above points to some of the main directions in recent research, at least as seen by the present author. It also alludes to some issues and gaps, many of which have been recognised for some time but which still require further work. Without claiming completeness or indeed originality, the following issues are relevant to the Humanities and seem to this author likely to become increasingly important in the near future

2.1 Algorithmic Accountability

In terms of future developments, an important question that has often been raised with regards to computational methods and has long been discussed in Digital Humanities is the need to be able to interpret and understand algorithms and their approaches.[24] This approach formed the premise of DigiPal

[15] For examples see Stutzmann 2013 (as well as Stutzmann 2012); Biblissima [n.d.]; IIIF [n.d.]; Murano and Felicetti 2017; Felicetti and Murano 2017; and Zhitomirsky-Geffet and Prebor 2016.

[16] Stutsmann 2013, 85 and 89–90. For the the ontology by the end of the project see Stutzmann 2017, 17.

[17] Murano and Felicetti 2017; Felicetti and Murano 2017.

[18] Gronemeyer and Diehr 2018.

[19] Brookes et al. 2015. 'Allograph' and 'graph' are here used in the palaeographic/forensic sense, where 'allograph' refers to different ways of writing the same letter (such as roman **a** and italic *a*), and 'graph' refers to a single instance of a letter on the page. See further Davis 2007, 254–55.

[20] Zhitormsky-Geffet and Prebor 2016.

[21] Biblissima [n.d.].

[22] IIIF [n.d.].

[23] Hassner et al. 2013; Hassner et al. 2015.

[24] For further discussion see Ciula 2005; Davis 2007; Sculley and Pasanek 2008; Stokes 2009; Clement, Steger and Unsworth 2009; among others.

(now Archetype), the basis of which was the insight that although techniques in machine vision and machine learning are increasingly successful, they typically are opaque to researchers, and are therefore difficult to interpret or verify.[25] However, if palaeographers cannot interpret the statistical results or understand their limitations then it becomes difficult or impossible to use the software. Furthermore, many of the computational systems listed above must be trained on a 'Ground Truth', namely a relatively large corpus of material that is taken to be known and without doubt, but the very assumption that such a ground truth exists is problematic in palaeography, where the quantity of known material is relatively small and where so much doubt often surrounds this content.[26] Rather than going away, the problem of understanding, inherent bias and therefore trust in algorithmic approaches has instead become increasingly prominent and is now widely recognised, particularly under the rubric of 'algorithmic accountability'. The Association for Computing Machinery (ACM) US Public Policy Council and Europe Policy Committee, for instance, published a joint 'Statement on Algorithmic Transparency and Accountability' which notes in words similar to those of Davis from a decade earlier that 'there is … growing evidence that some algorithms and analytics can be opaque, making it impossible to determine when their outputs may be biased or erroneous'.[27] The ACM Council goes on to observe that '[d]ecisions made by predictive algorithms can be opaque because of many factors, including technical (the algorithm may not lend itself to easy explanation), economic (the cost of providing transparency may be excessive, including the compromise of trade secrets), and social (revealing input may violate privacy expectations). Even well-engineered computer systems can result in unexplained outcomes or errors, either because they contain bugs or because the conditions of their use changes, invalidating assumptions on which the original analytics were based.' Although work on historical documents does not have the societal implications of the cases discussed by the ACM, nevertheless machine-generated features and highly computational methods are often not meaningful to humans and particularly to those in the Humanities. This difficulty is further compounded by the financial pressures that surround this work, meaning that the underlying software and algorithms must often be proprietary and so become commercial secrets rather than being fully Open Source. This is less of a problem if the results are easily verifiable by a human observer: in the case of HTR, for instance, it is normally very easy to see if the machine's results are valid or not, and in this context how the results are obtained is relatively unimportant. However, the problem is significantly greater in cases where judgement is required, such as whether two samples are by the same scribe or not, or indeed in a broader context what sentence should be delivered to an offender, and in this case it is essential that one can 'look inside the machine' in order to understand and judge the validity of underlying assumptions and the impact of the biases that inevitably underlie them, without which these results cannot sustain the necessary scepticism and critical questioning that is the foundation of the academic and scientific method. Important progress is already being made here but more work remains to be done.[28] The problem of 'looking inside the machine' is famously difficult, particularly for very deep neural networks. Even if we can see the final layer(s) this is only a small part of the whole computational process, and it remains very unclear how we can go about understanding the process more fully. As always, there is also still the risk of misunderstanding, particularly for those without expertise in the algorithms: statistics, visualisations and other information can be misleading, for instance by misunderstanding or misusing the confidence levels in a classifie . Granted this opacity is nothing new, as human palaeographical expertise has long proven very difficult to articulate, with some even arguing that it cannot be taught but can only be 'acquired' through years or decades of exposure to the materials.[29] It does, however, raise the question to what extent these new computational methods are simply reproducing the problems of the old manual ones and how, if at all, we can get beyond this.

[25] Davis 2007 n. 27; Stokes 2009; Hassner et al. 2013; Stokes 2015; and Hassner et al. 2015, 113–4, where the question is usefully rephrased as one of trust rather than the 'black box'

[26] Hassner et al. 2013.

[27] ACM 2017.

[28] See especially Kestemont, Christlein and Stutzmann 2017 for recent work applied directly to handwriting. A useful discussion of these systems from the point of view of a palaeographer is also given by Smit 2011.

[29] For this argument see, for example, Pfaff 1977, 104, discussed by Derolez 2003, 1.

2.2 Interdisciplinarity and combining approaches

Closely related to accountability in manuscript studies is the field s inherent interdisciplinarity. This has already been discussed at length elsewhere[30] but always needs emphasising: that manuscripts are extremely complex objects that can be approached along infinitely many dimensions, whether textual (philological, literary, linguistic, historical), cultural (as art, as status-object, as reflection of learning and world-view), visual (palaeographical, art-historical) or material (codicological, archaeological), from the 'hard' sciences (chemistry, physics, biology), as images (machine vision, pattern-matching), as three-dimensional objects (3D scanning, Reflectance Transformation Imaging), and many, many more. This – thankfully – has led to many interdisciplinary studies of manuscripts, many of which have been highly successful. Such work in turn inevitably leads to challenges in communication and, more significantl , in finding true collaboration such as research questions that are relevant to all parties and finding an appropriate balance of power and trust.[31] For instance, in practice 'collaboration' can in fact be one group providing 'solutions' to the other group's 'problems'. This carries the implication that this is a one-way process, with one group holding the power and having nothing to learn from the other, a risk that is not far from one identified by Gillian Beer, namely 'becoming merely disciples because not in control of a sufficient range of knowledge'.[32] The converse is also a risk, and is also observed in practice, namely that 'the problems preoccupying those working in another discipline may sometimes (initially, arrogantly) seem quite simple – because we are not familiar with the build-up of arguments across time that has reached this moment of dilemma'.[33] Fortunately the experience of this author has been that both of these risks have largely been avoided in the case of manuscript studies, not least because so many in the field have committed to the hard work that is required for true collaboration, and so fears raised by the author in 2009 of the possible split into two disciplines – 'digital' and 'non-digital' palaeography – seem happily misguided, but nevertheless this work must necessarily continue. In either case, though, is the question of trust once again, which Beer phrased as 'the matter of competence … Is this a raiding party? Is there time to question and to learn? *How much must be taken on trust?*'[34] Although perhaps not an issue of trust or the 'raiding-party', nevertheless even different computer-based approaches rely on different types of expertise, each of which (usefully) has different strengths and weaknesses and is therefore better adapted to different questions. For instance, research in machine vision, clustering and large-scale statistical approaches have produced very good results with relatively large amounts of data and more-or-less established Ground Truth, giving a type of palaeographical 'distant reading' or helping one find a way into a large and complex corpus. Fields like knowledge representation and knowledge elicitation seek to model and represent palaeographical content in 'symbolic' ways that are closer to those of the domain experts, tending towards more of a 'close' reading of a defined and relatively manageable corpus. User Interface and User Experience (UI/UX) is another relevant domain, looking more at visualisation and communication with the user, and so on. Relatively little attention has been paid to interface design in the field of document analysis, but there are obvious areas of potential difficulty here, for instance an interface reflecting the model from one discipline being unfamiliar, difficult to understand or even misleading from those of other disciplines. Such difficulties are often observed in complex search forms which might reflect the underlying data structure but which do not necessarily reflect the users' understanding of their materials, for instance. In this sense, then, the interface risks perpetuating disciplinary and conceptual differences, rather than helping to bring disciplines together. Furthermore, it is very often the case that off-putting, unfamiliar or inefficient interfaces are not tolerated by target users: rather than learning to use seemingly difficult systems, people tend to simply abandon the software and go back to their previous practices. Once again, then, combining these different domains of research is an important step that needs to continue, even if the disciplinary nature of academic research and its need for specialised research and publications may sometimes make this difficult in practice

[30] See especially Hassner et al. 2013 and Hassner et al. 2015 as two examples among many.

[31] This has been discussed particularly by Hassner et al. 2013, and Hassner et al. 2015, among others.

[32] Beer 2006.

[33] Beer 2006.

[34] Beer 2006; my emphasis.

2.3 Multigraphism and a Conceptual Reference Model

Another recurring question is the need for a precise, transversal ontology and conceptual reference model (CRM) for handwriting. The need for this has been recognized for some time,[35] but its importance is becoming increasingly evident as research towards digital resources for the field is largely resulting in different and incompatible models. The advantages of sharing and interconnecting palaeographical data are evident, but at the same time the likelihood of a single universal terminology is remote and not necessarily desirable: as Petrucci argued, every palaeographical terminology is based on a particular view of scribal practice and all such views are potentially valid,[36] and this argument suggests we should encourage rather than limit the proliferation of analytic frameworks. Nevertheless, the need for a consistent and standardised way of at least referring to manuscripts is evident, and indeed significant progress has already been made here.[37] However, as work to date is increasingly demonstrating, a successful model needs clarification of more fundamental concepts. What *is* writing, after all, and how can we accurately distinguish between its many aspects: visual, stylistic, phonological, functional, chemical, mechanical, semantic, kinetic, graphematic, allographic, and so on? What *exactly* is a script, and how does one define the boundaries between one script or another, or decide if two samples of writing are in the same script or not? As discussed in Section 1 above, some work is beginning to appear in this area but much more is needed. A true CRM should also account for different writing systems, and indeed interest in transversal and multigraphic approaches is growing rapidly: how, for instance, can one compare Arabic and Hebrew written by the same scribe? All this suggests the need for a formal CRM that can provide formal precision while functioning meaningfully across all the world's writing systems. Clearly it is not possible to create a complete model that will fully capture all aspects of all the world's written communication: indeed this is not even necessarily desirable, as one of the values of modelling is that it simplifies, allowing one to focus on particular aspects which are intractable in the domain being modelled.[38] The question, rather, is whether it is possible to create a model that is sufficiently broad to be useful to palaeographers and others and to allow genuine transversal study of scripts. The answer to this question is unclear at present. To raise just some of the challenges that it presents, researchers in both palaeography and computer science are now recognising the difficulties of multigraphism, namely cases where individuals or cultures simultaneously use entirely different scripts, alphabets, or even writing systems. Most computational methods have been developed at least theoretically independently of any given script or writing system, and examples in practice include the same software working successfully across Hebrew, Tibetan, Old English and Greek, for instance, or Word Spotting in Latin and Arabic.[39] In these cases, however, the work has generally been applied to only one script at a time, rather than to different scripts simultaneously. However, many – perhaps almost all – cultures wrote and still write in different scripts and alphabets or writing systems at the same time. This problem is beginning to be recognised from informatics and also from palaeography.[40] Indeed, this is becoming so much a 'hot topic' that a competition on the subject will be run at ICFHR 2018, and at least one large grant on the subject has recently been submitted for funding by the European Research Council. However, the challenges are significant, particularly in modelling, as (for instance) it requires a clear distinction between a letter's normal physical form (the allograph), its given instance on the page (the graph), but also its function whether graphematic, phonetic, semantic and so on. For instance, two given forms that look identical such as **H** and **H** would normally be assumed to represent the same grapheme but may not: if we write **H**ABEAM and I**H**ΣOYΣ then it becomes clear that the first **H** is the Latin letter (Unicode U+0048) and the second **H** a Greek Eta (Unicode U+0397).[41] A scholar interested in palaeographically comparable forms in a Greek-Latin multigraphic context would presumably want to find examples of both allographs, and this requires the capacity to search by form rather than graphematic function.

[35] See, for example, Stutzmann 2011 and Hassner et al. 2013.

[36] Petrucci 2001, 70–71.

[37] Examples of relevant projects addressing this question include MESA, Biblissima, Pinakes, TRAME, SNAP DRGN and Medium, among others.

[38] McCarty 2004.

[39] Hassner, Wolf and Dershowitz 2013, and Sudholt and Fink 2016, respectively.

[40] Examples from informatics include Singh et al. 2014; He, Wiering and Schomaker 2015; Bertolini, Oliviera and Sabourin 2016; Ahmed et al. 2017; Ubul et al. 2017. Discussions from palaeography include Petrucci 2005, p. 53; cf. also Cavallo 1990 and Smith 2004, p. 429; de Robertis 2012, p. 223; de Robertis 2013, p. 17; cf. also Ceccherini 2012, 237 (but contrast that of De Gregorio 2000, pp. 19–2); Dewartes 2013; Radiciotti 2006; Stokes 2017; and Stokes 2018.

[41] Stokes 2018, drawing on Burgarski 1993.

In other cases, however, scribes might use a Greek Φ in an otherwise Latin context for the phonetically equivalent **F**, or the Latin symbol **&** (meaning *et*) for the semantically equivalent *and* in an English context or *ocus* in Old Irish. These complexities multiply significantly in other more complex writing systems, for instance logosyllabographies such as Sumerian or Mayan where signs can function semantically or phonetically depending on context, and where signs can take significantly different forms, even incorporating elements of one sign within another, and so on.[42]

3. Concluding notes

As mentioned above, it must be emphasised once again that the preceding discussion is just one of many views on the field, from one of many possible directions. Certainly there is much to do, and many challenges still to face; indeed, it still seems largely the case that digital and computational approaches to palaeography have yet to reach their full potential in terms of having tangible impact on new results in palaeographical scholarship. Nevertheless the feeling of this author is one of optimism. There is much evidence of genuine and productive collaboration, with many examples of people making a real effort to listen, understand, learn from and gain from each other. The genuinely collaborative workshops to date have shown that the combination of digital/computational with palaeography is not one of providing tools for use or answers to questions in a one-way and uncritical flo , and neither is it providing ready access to texts or visual reproductions of manuscript pages. Rather, it is profoundly changing the way that people think in palaeography and – I hope – in computer science and informatics as well. What does it mean to work with an image instead of the object, and how do the results of our research change as a result? How does a 'distant' view of a manuscript or corpus change our view of it, and how much does this matter? What does the instability of text and writing mean for computer classification? Once again, what *exactly* is a letter anyway, and what is its relation to the texts, images, and physical and digital objects that transmit it? Asking new questions, seeing new points of view, and reflecting on these and how they change our very world-view: these are the real gains of Digital Humanities, and this is where we need to focus in future.

[42] Daniels 1990; Gronemeyer and Diehr 2018.

REFERENCES

ACM US Public Policy Council and Europe Policy Council (2017), *Statement on Algorithmic Transparency and Accountability*. <https://www.acm.org/binaries/content/assets/public-policy/2017_joint_statement_algorithms.pdf>.

Ahmed, Ahmed Abdulla, Harith Raad Hasan, Fariaa Abdalmajeed Hammed, and Omar Ismael Al-Sanjary (2017), 'Writer Identification on Multi-Script Handwritten using Optimum Features', *Kurdistan Journal of Applied Research*, 2: 178–185. doi: 10.24017/science.2017.3.64.

Archetype (2018), *Archetype: An Accessible and Evolving Toolset to Empower Digital Research* (London: King's College). <https://www.archetype.ink>.

Beer, Gillian (2006), 'Dame Gillian Beer's Speech on the Challenges of Interdisciplinarity' (Durham: University of Durham). <https://www.dur.ac.uk/ias/news/annual_research_dinner/>.

Bertolini, Diego, Luiz S. Oliveira, and Robert Sabourin (2016), 'Multi-script Writer Identification using Dissimilarity', *23rd International Conference on Pattern Recognition (ICPR)*. <https://www.etsmtl.ca/getattachment/Unites-de-recherche/LIVIA/Recherche-et-innovation/Publications/Publications-2016/Bertolini_icpr_2016.pdf>.

Biblissima [n.d.], 'Ontologie', *Biblissima Documentation*. <http://doc.biblissima-condorcet.fr/ontologie>.

Brookes, Stewart, Peter A. Stokes, Matilda Watson, and Debora Marques de Matos (2015), 'The DigiPal Project for European Scripts and Decorations', in Aidan Conti, Orietta da Rold, and Philip Shaw (eds), *Writing Europe 500–1450: Texts and Contexts* (Cambridge: D. S. Brewer; Essays and Studies n.s., 68), 25–58.

Bugarski, Ranko (1993), 'Graphic Relativity and Linguistic Constructs', in Robert J. Scholes (ed.), *Literacy and Linguistic Analysis* (Hillsdale, NJ: Erlbaum), 5–18.

Canart, Paul (2006), 'La Paléographie est-elle un art ou une science?', *Scriptorium*, 60: 159–185.

Cavallo, Guiglielmo (1990), 'Écriture grecque et écriture latine en situation de "multigrafismo assoluto"', in Colette Sirat, Jean Irgoin, and Emmanuel Poulle, *L'Écriture: le cerveau, l'œil et la main* (Turnhout: Brepols), 349–362.

Ceccherini, Irene (2012), 'Poligrafia nel Quattrocento: Sozomeno da Pistoia', *Medioevo e Rinascimento*, 26: 237–251.

Chen, Kai, and Matthias Seuret (2017), 'Convolutional Neural Networks for Page Segmentation of Historical Document Images', *Proceedings of the 4th International Workshop on Historical Document Imaging and Processing* (New York: Association for Computing Machinery), 101–106.

Christlein, Vincent, Martin Gropp, and Andreas Maier (2017), 'Automatic Dating of Historical Documents', in Hannah Busch, Franz Fischer, and Patrick Sahle (eds), *Kodikologie und Paläographie im digitalen Zeitalter 4—Codicology and Palaeography in the Digital Age 4* (Norderstedt: Books on Demand), 151–164.

Ciula, Arianna (2005), 'Digital Palaeography: Using the Digital Representation of Medieval Script to Support Palaeographic Analysis', *Digital Medievalist*, 1. <http://www.digitalmedievalist.org/journal/1.1/ciula/>.

Clement, Tanya, Sara Steger, and John Unsworth (2008), 'How Not to Read a Million Books', *Seminar on the History of the Book* (Rutgers University, New Brunswick NJ, 5 March). <http://www.people.virginia.edu/~jmu2m/hownot2read.html>.

Costamagna, Giorgio *et al*. (1995), 'Commentare Bischoff', *Scrittura e Civiltà*, 19: 325–348.

—— *et al*. (1996), 'Commentare Bischoff', *Scrittura e Civiltà*, 20: 401–407.

Daniels, Peter T. (1990), 'Fundamentals of Grammatology', *Journal of the American Oriental Society*, 119: 727–731. doi: 10.2307/602899.

Davis, Tom (2007), 'The Practice of Handwriting Identification', *The Library*, 7th series, 8: 251–276.

De Gregorio, Giuseppe (2002), 'Tardo medioevo greco-latino: manoscritti bilingui d'Oriente e d'Occidente', in Francesco Magistrale, Corinna Drago, and Paolo Fioretti (eds), *Libri, documenti, epigrafi medievali: possibilità di studi comparativi. Atti del convegno internazionale di studio dell'associazione italiana dei paleografi e diplomatisti, Bari 2–5 ottobre 2000* (Spoleto: Centro italiano di studi sull'alto Medioevo), 17–135.

De Robertis, Teresa (2012), 'Digrafia nel Trecento: Andrea Lancia e Francesco di ser Nardo da Barberino', *Medioevo e Rinascimento*, 26: 221–235.

—— (2013), 'Una mano tante scritture: Problemi di metodo nell'identificazione degli autografi', in Nataša Golob (ed.), *Medieval Autograph Manuscripts: Proceedings of the XVIIth Colloquium of the Comité International de Paléographie Latine* (Turnhout: Brepols), 17–38.

Derolez, Albert (2003), *The Palaeography of Gothic Manuscript Books from the Twelfth to the Early Sixteenth Century* (Cambridge: Cambridge University Press).

Dewartes, Thomas (2013), 'Polygraphisme et mixité graphique: Note sur les additions d'Arias (1060–1070) dans l'Antiphonaire de León', *Territorio, Sociedad y Poder: Revista de Estudios Medievales*, 8: 67–84.

DigiPal (2010–14), *Digital Resource and Database of Palaeography, Manuscripts and Diplomatic* (London: King's College). <http://www.digipal.eu>.

Eichenberger, Nicole, Angelika Garz, Kai Chen, Hao Wei, Rolf Ingold, and Marcus Liwicki (2015), 'DIVADesk: A Holistic Digital Workspace for Analyzing Historical Document Images', *manuscript cultures*, 7: 69–82.

Felicetti, Achille, and Francesca Murano (2017), 'Scripta manent: a CIDOC CRM Semiotic Reading of Ancient Texts', *International Journal of Digital Libraries*, 18: 263–270. doi: 10.1007/s007.

Fischer, Andreas, Emanuel Indermühle, Volkmar Frinken, and Horst Bunke (2011a), 'HMM-Based Alignment of Inaccurate Transcriptions for Historical Documents', *International Conference on Document Analysis and Recognition.* doi: 10.1109/ICDAR.2011.20.

—— Volkmar Frinken, Alicia Fornés, and Horst Bunke (2011b), 'Transcription Alignment of Latin Manuscripts Using Hidden Markov Models', *Proceedings of the 2011 Workshop on Historical Document Imaging and Processing* (New York: Association for Computing Machinery), 29–36.

Garz, Angelika, Nicole Eichenberger, Marcus Liwicki, and Rolf Ingold (2015), 'HisDoc 2.0: Toward Computer-Assisted Paleography', *manuscript cultures*, 7: 19–28.

Gronemeyer, Sven, and Franziska Diehr (2018), 'Organising the Unknown: A Concept for the Sign Classification of not yet (Fully) Deciphered Writing Systems Exemplified by a Digital Sign Catalogue for Maya Hieroglyphs', in J. Girón Palau and I. Gallina Russell (eds), *Digital Humanities Book of Abstracts* (Mexico City: La Red de Humanidades Digitales). <https://dh2018.adho.org/en/organising-the-unknown-a-concept-for-the-sign-classification-of-not-yet-fully-deciphered-writing-systems-exemplified-by-a-digital-sign-catalogue-for-maya-hieroglyphs/>.

Gumbert, J. Peter (1998), 'Commentare "Commentare Bischoff"', *Scrittura e Civiltà*, 22: 397–404.

Hassner, Tal, Malte Rehbein, Peter Stokes, and Lior Wolf (eds) (2013a), 'Computation and Palaeography: Potentials and Limits', *Dagstuhl Manifestos*, 2: 14–35. doi: 10.4230/DagMan.2.1.14.

———, Lior Wolf, and Nachum Derschowitz (2013b), 'OCR-Free Transcript Alignment', *12th International Conference on Document Analysis and Recognition (ICDAR)*, 1310–1314. doi: 10.1109/ICDAR.2013.265.

—— Robert Sablatnig, Dominique Stutzmann, and Ségolène Tarte (2015), 'Digital Palaeography: New Machines and Old Texts', *Dagstuhl Reports*, 4: 127–128. doi: <10.4230/DagRep.4.7.112>.

He, Sheng, Marco Wiering, and Lambert Schomaker (2015), 'Junction Detection in Handwritten Documents and its Application to Writer Identification', *Pattern Recognition*, 48: 4036–4048.

HIMANIS (2018), *The HIMANIS Project: Indexing the Trésor des Chartes Registers*. <http://www.himanis.org>.

IIIF [n.d.], *IIIF: International Image Interoperability Framework*. <http://iiif.io/>.

Kestemont, Mike, Vincent Christlein, and Dominique Stutzmann (2017), 'Artificia Palaeography: Computational Approaches to Identifying Script Types in Medieval Manuscripts', *Speculum: A Journal of Medieval Studies*, 92: S86–S109. doi: 10.1086/694112.

McCarty, Willard (2004), 'Modeling: A Study in Words and Meanings', in Susan Schreibman, Ray Siemens, and John Unsworth (eds), *A Companion to Digital Humanities* (Oxford: Blackwell), 254–270.

Medium (2016), *Medium: Répertoire des manuscrits reproduits ou recensés* (Paris: IRHT). <http://medium.irht.cnrs.fr>.

MESA [n.d.], *MESA: Medieval Electronic Scholarly Alliance*. <http://www.mesa-medieval.org>.

Murano, Francesca, and Achille Felicetti (2017), *Definition of the CRMtex: An Extension of CIDOC-CRM to Model Ancient Textual Entities*, version 0.8 (CIDOC-CRM). <http://www.cidoc-crm.org/crmtex/sites/default/files/CRMtex_v0.8.pdf>

Petrucci, Armando (2001), *La descrizione del manoscritto: storia, problemi, modelli* (Rome: Carocci).

—— (2005), 'Digrafismo e bilettrismo nella storia del libro', *Syntagma: Revista del Instituto de Historia del Libro y de la Lectura*, 1: 53–75.

Pfaff, Richard W. (1977), 'M. R. James on the Cataloguing of Manuscripts: A Draft Essay of 1906', *Scriptorium*, 31: 103–118

Pinakes (2016), *Pinakes: Textes et manuscrits grecs* (Paris: Institut de recherche et d'histoire des textes). <http://pinakes.irht.cnrs.fr>.

Radiciotti, Paolo (2006), 'Il problema del digrafismo nei rapporti fra scrittura latina e greca nel medioevo', *Nea Rômê*, 3: 5–55.

Rajan Sampath, Vinodh (2016), *Quantifying Scribal Behaviour: A Novel Approach to Digital Paleography* (PhD Dissertation, St Andrews: University of St Andrews). <http://hdl.handle.net/10023/9429>.

Rothacker, Leonard, Denis Fisseler, Gerfrid G. W. Müller, Frank Weichert, and Gernot A. Fink (2015), 'Retrieving Cuneiform Structures in a Segmentation-free Word Spotting Framework', *Proceedings of the International Workshop on Historical Document Imaging and Processing* (New York: Association for Computing Machinery). <http://patrec.cs.tu-dortmund.de/pubs/papers/Rothacker2015-RCS.pdf>.

Sadeh, Gil, Lior Wolf, Tal Hassner, Nachum Dershowitz, and Daniel Stökl Ben-Ezra (2015), 'Viral transcript alignment', *13th International Conference on Document Analysis and Recognition (ICDAR)*, 711–715. doi: 10.1109/ICDAR.2015.7333854.

Schomaker, Lambert (2007), 'Advances in Writer Identification and Verification', *Proceedings of the 9th International Conference on Document Analysis and Recognition (ICDAR)*, 2: 769–773.

Sculley, D., and Bradley M. Pasanek (2008), 'Meaning and Mining: The Impact of Implicit Assumptions in Data Mining for the Humanities', *Literary and Linguistic Computing*, 23: 409–424.

Seuret, Matthias, Daniel Stökl ben Ezra, and Markus Liwiki (2017), 'Robust Heartbeat-based Line Segmentation Methods for Regular Texts and Paratextual Elements', *Proceedings of the 4th International Workshop on Historical Document Imaging and Processing* (New York: Association for Computing Machinery), 71–76.

Singh, Pawan Kumar, Arafat Mondal, Showmik Bhowmik, Ram Sarkar, and Mita Nasipuri (2014), 'Word-Level Script Identification from Handwritten Multi-script Documents', *Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA)*, 551–558. doi: 10.1007/978-3-319-11933-5_62.

Sirat, Colette (2006), *Writing as Handwork: A History of Handwriting in Mediterranean and Western Culture*, (Turnhout: Brepols).

SNAP DRGN [n.d.], *Standards for Networking Ancient Prosopographies* (London: King's College). <http://snapdrgn.net>.

Smit, Ginna (2011), 'The Death of the Palaeographer? Experiences with the Groningen Intelligent Writer Identification System (GIWIS)', *Archiv für Diplomatik, Schriftgeschichte, Siegel- und Wappenkunde*, 57: 413–425.

Smith, Marc H. (2004), 'Les «gothiques documentaires»: un carrefour dans l'histoire de l'écriture latine', *Archiv für Diplomatik*, 50: 417–465.

Stokes, Peter A. (2009), 'Computer-Aided Palaeography: Present and Future', in Malte Rehbein, Patrick Sahle, and Torsten Schaßan (eds), *Kodikologie und Palãographie im Digitalen Zeitalter—Codicology and Palaeography in the Digital Age* (Norderstedt: Books on Demand), 309–338. <http://kups.ub.uni-koeln.de/2978/>.

Stokes, Peter A. (2012), 'Palaeography and the "Virtual Library"', in Brent Nelson and Melissa Terras (eds), *Digitizing Medieval and Early Modern Material Culture* (Arizona: Center for Medieval and Renaissance Studies), 137–169.

—— (2015), 'Digital Approaches to Paleography and Book History: Some Challenges, Present and Future', *Frontiers in Digital Humanities*, 2:5. doi: <10.3389/fdigh.2015.00005>.

—— (2017), 'Scribal Attribution across Multiple Scripts: A Digitally-Aided Approach', *Speculum: A Journal of Medieval Studies*, 92: S65–S85.

—— (2018), 'Modelling Multigraphism: The Digital Representation of Multiple Scripts and Alphabets', in J. Girón Palau and I. Gallina Russell (eds), *Digital Humanities Book of Abstracts* (Mexico City: La Red de Humanidades Digitales), 292–296. <https://dh2018.adho.org/modelling-multigraphism-the-digital-representation-of-multiple-scripts-and-alphabets/>.

Stutzmann, Dominique (2011), 'Paléographie statistique pour décrire, identifie , dater...: Normaliser pour coopérer et aller plus loin?', in Franz Fischer, Christiane Fritze, and Georg Vogeler (eds), *Kodikologie und Paläographie im Digitalen Zeitalter 2—Codicology and Palaeography in the Digital Age 2* (Norderstedt: Books on Demand), 247–277. <https://kups.ub.uni-koeln.de/4353/>.

—— (2012), 'Modélisation des signes graphiques (1)', *Écriture médiévale et numérique* [project blog]. <http://oriflamms hypotheses.org/921>.

—— (2013), 'Ontologie des formes et encodage des textes manuscrits médiévaux: Le projet ORIFLAMMS', *Document numérique*, 16: 81–95. doi: <10.3166/DN.16.3.69-79>.

—— Théodore Bluche, Alexei Lavrentev, and Yann Leydier (2015), 'From Text and Image to Historical Resource: Text-Image Alignment for Digital Humanists', in *Digital Humanities 2015: Book of Abstracts* (Sydney: University of Western Sydney). <http://dh2015.org/abstracts/xml/STUTZMANN_Dominique_From_Text_and_Image_to_Histor/STUTZMANN_Dominique_From_Text_and_Image_to_Historical_R.html>.

—— (2017), *Projet ANR-12-CORP-0010 ORIFLAMMS, Programme Corpus données et outils 2012 – Compte-rendu de fin de projet*. <https://oriflamms. ypotheses.org/files/2017/04 Oriflamms-Compte-rendu-final.pdf

Sudholt, Sebastian, and Gernot A. Fink (2016), 'PHOCNet: A Deep Convolutional Neural Network for Word Spotting in Handwritten Documents', in *Proceedings of the International Conference on Frontiers in Handwriting Recognition.* <http://patrec.cs.tu-dortmund.de/pubs/papers/Sudholt2016-PAD>.

TRAME (2013), *TRAME: Text and Manuscript Transmission of the Middle Ages* (Florence: Fondazione Ezio Franceschini). <http://trame.fefonlus.it/>.

Transkribus [n.d.], *Transkribus: Transcribe, Collaborate, Share* (Innsbruck: University of Innsbruck). <https://transkribus.eu/>.

Ubul, Kurban, Gulzira Tursun, Alimjan Aysa, Donato Impedovo, Giuseppe Pirlo, and Tuergen Yibulayin (2017), 'Script Identification of Multi-Script Documents: A Survey', *IEEE Access*, 5: 6546–6559. doi: <10.1109/ACCESS.2017.2689159>.

Washington Database (2016), *Washington Database* (Bern: University of Bern). <http://www.fki.inf.unibe.ch/databases/iam-historical-document-database/washington-database>.

Wolf, Lior, Rotem Littman, Naama Mayer, Tanya German, Nachum Dershowitz, Roni Shweka, and Yaacov Choueka (2011a), 'Identifying Join Candidates in the Cairo Genizah', *International Journal of Computer Vision*, 94: 118–135. <doi:10.1007/s11263-010-0389-8>.

—— Nachum Dershowitz, Liza Potikha, Tanya German, Roni Shweka, and Yaacov Choueka (2011b), 'Automatic Paleographic Exploration of Genizah Manuscripts', in Franz Fischer, Christiane Fritze, and Georg Vogeler (eds), *Kodikologie und Paläographie im Digitalen Zeitalter 2—Codicology and Palaeography in the Digital Age 2* (Norderstedt: Books on Demand), 157–179.

Würsch, Marcel, Rolf Ingold, and Marcus Liwicki (2016), 'SDK Reinvented: Document Image Analysis Methods as RESTful Web Services', *12th IAPR Workshop on Document Analysis Systems (DAS)*, 90–95.

Zhitomirsky-Geffet, Maayan, and Gita Prebor (2016). 'Towards an Ontopedia for Historical Hebrew Manuscripts', *Frontiers in Digital Humanities*, 3:3. doi: <10.3389/fdigh.2016.00003>.

Article

Creating Workflows with a Human in the Loop for Document Image Analysis

Marcel Gygli (Würsch), Mathias Seuret, Lukas Imstepf, Andreas Fischer, Rolf Ingold | Windisch, Fribourg

Abstract

Workflows in document image analysis have special requirements compared with workflows in most other scientific fields. As many methods are semi-automatic, the user needs to be kept in the loop and interact with the workflo , change parameters and investigate results in order to generate the best results. However, most traditional workflow management systems do not allow for this and can execute only a fully automatic workflow from beginning to end

With DIVA-DIP, we introduce a new workflow management system that focuses on the special needs of document image analysis workflows. DIVA-DIP makes it possible to design workflows visually and for processes to have a specific state in the execution phase, in which they wait for user interaction, thus keeping the user integrated in the process.

To test the application on existing document image analysis methods, we integrated methods that are offered through DIVAServices as Web Services. This enables DIVA-DIP to execute methods without the need to install them on the local machine.

Introduction

Document image analysis (DIA) has a long history in computer science research. The goal of DIA is to process digital document images and transform the information contained, such that computers can use it. Methods developed by researchers are used to analyse the layout elements on a document page or to recover the written text by means of either optical character recognition (OCR) if it is printed text or handwritten text recognition (HTR) if it is handwritten text.

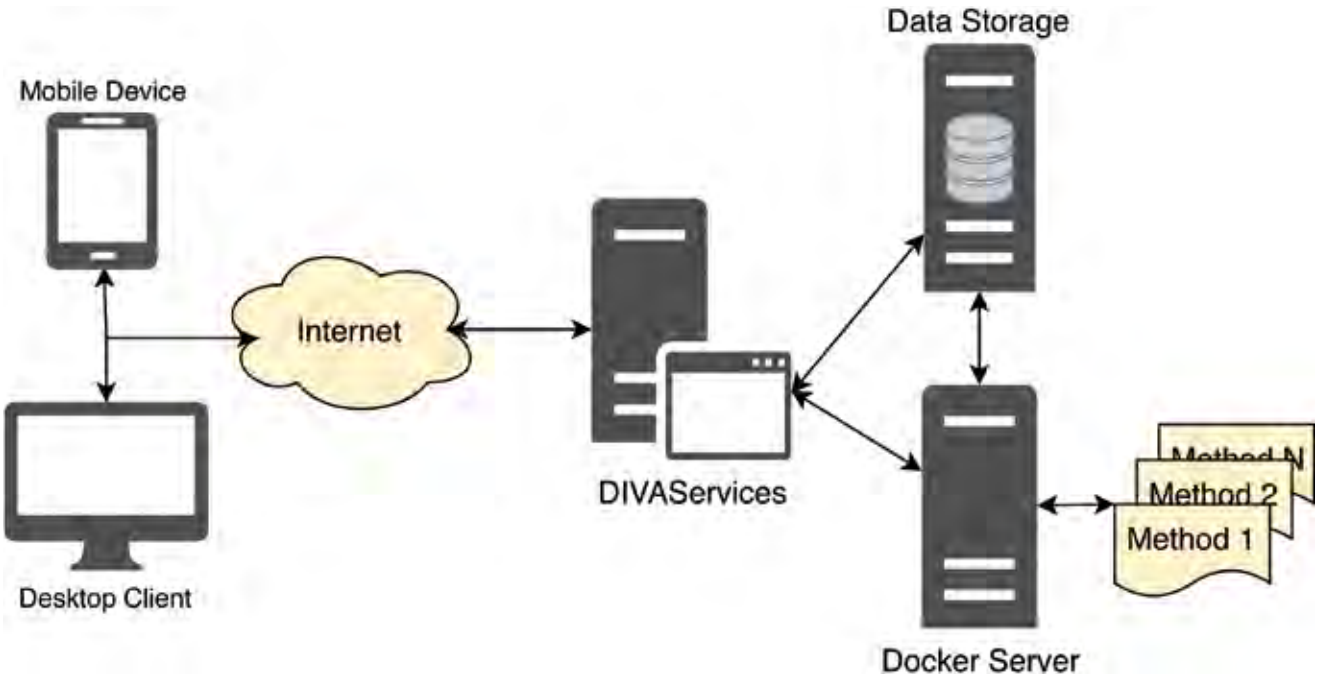


Fig. 1: Architecture overview of DIVAServices. At DIVAServices, we use three different servers, one for the general handling of requests (DIVAServices), one for storage of the data (Data Storage) and one on which the actual methods are installed and executed (Docker Server). This architecture setup allows DIVAServices to scale in case of usage growth. (Figure taken from Würsch, Liwicki, and Ingold 2018.)

Creating and executing workflows for document image analysis is not a simple task. One has to find suitable DIA methods, typically fin -tune the parameters of such methods, adjust them to work together and then execute them on a specific dataset. With DIVAServices, Würsch et al. introduced a solution for one part of this problem, by introducing a framework for providing access to DIA methods as RESTful Web Services (Würsch, Ingold, and Liwicki 2016).

In this paper, we introduce the DIVA Document Image Processor (DIVA-DIP), a workflow management system (WMS) that can be used to design and execute complex workflows using methods provided by DIVAServices. We take special care to address the special needs of researchers in DIA. This means that, in comparison with most traditional WMS, DIVA-DIP allows the user to interact and adapt the workflows while they are using it

DIVA-DIP makes use of the Human-in-the-Loop paradigm (Karwowski 2001). This paradigm allows users of a system to directly interact with it, change its behaviour and therefore have more control over the behaviour of the application. Additionally, we also make use of visual programming (Myers 1986). This will allow users to create their workflows by connecting very simple building blocks. Users will therefore not require programming knowledge, but can interact with the user interface that allows them to build their workflows

The remainder of this article is structured as follows. In the next section we introduce relevant related work, focusing on other WMS from the literature. In Section 3, we provide an overview of DIVAServices, since DIVA-DIP will make use of methods provided by it. This is followed by the explanation of DIVA-DIP and how it works. The article closes with a summary and a look at possible research areas for future work.

Related work

In recent years, various WMS have been introduced: Pegasus (Deelman et al. 2015) workflow solution for scientific experiments with a focus on exploiting distributed computing infrastructures, and Taverna (Wolstencroft et al. 2013), a domain-independent WMS that is used mainly in the life sciences. Neither of these tools found relevant adaptation in the DIA community. We believe that this is due to the special nature of the domain. Most of the WMS are designed to execute workflows with zero interaction. In DIA, however, it is often the aim to keep the human in the loop, meaning that the user has the possibility to interact with the workflow as it is running

State-of-the-art, fully automatic DIA systems are still prone to errors when applied to difficult cases, most notably historical documents. This applies across various tasks such as optical character recognition (OCR), handwriting recognition, writer identificati n and manuscript dating. Considering the sheer amount of different historical scripts and languages, the goal instead should be to provide human experts with interactive DIA tools that support them in their work.

Examples of such tools include the CATTI system for computer-assisted transcription of historical documents (Romero, Toselli, Rodríguez, and Vidal 2007), Aletheia for annotating historical prints (Clausner, Pletschacher, and Antonacopoulos 2011), PhaseGT for the binarization of historical manuscripts (Nafchi, Ayatollahi, Moghaddam, and Cheriet 2013) and GraphManuscribble for intuitive interaction with digital facsimiles (Garz, Seuret, Fischer, and Ingold 2016), to name just a few. Over time, annotations provided by humans can be used to train methods based on machine learning, which in turn can improve the suggestions made by DIA tools, thus closing the loop between human experts and the semi-automatic systems.

DIVAServices – Making DIA methods available as Web Services

Our goal is to build a WMS that can be used as a workflow designer for methods demonstrated in the DIVAServices framework (Würsch, Ingold, and Liwicki 2015). This framework allows DIA researchers to provide access to their own methods through a unified API and for DIA methods to be executed over the Internet without the need to perform any local installation. For a thorough technical explanation of how DIVAServices works and how to interact with it, we refer the reader to our publications on the topic (Würsch et al. 2016, 2018), as well to as our online documentation.[1]

Figure 1 provides an architectural overview of DIVAServices and all the components it interacts with. DIVAServices handles all incoming requests and performs the necessary actions. This includes storing data on our own data storage server and the execution of methods on our own Docker server, where all methods are installed. A good introduction to Docker for reproducible research is

[1] See *DIVAServices: A RESTful Web Service Framework for Document Image Analysis Methods* <https://diva-dia.github.io/DIVAServicesweb/>.

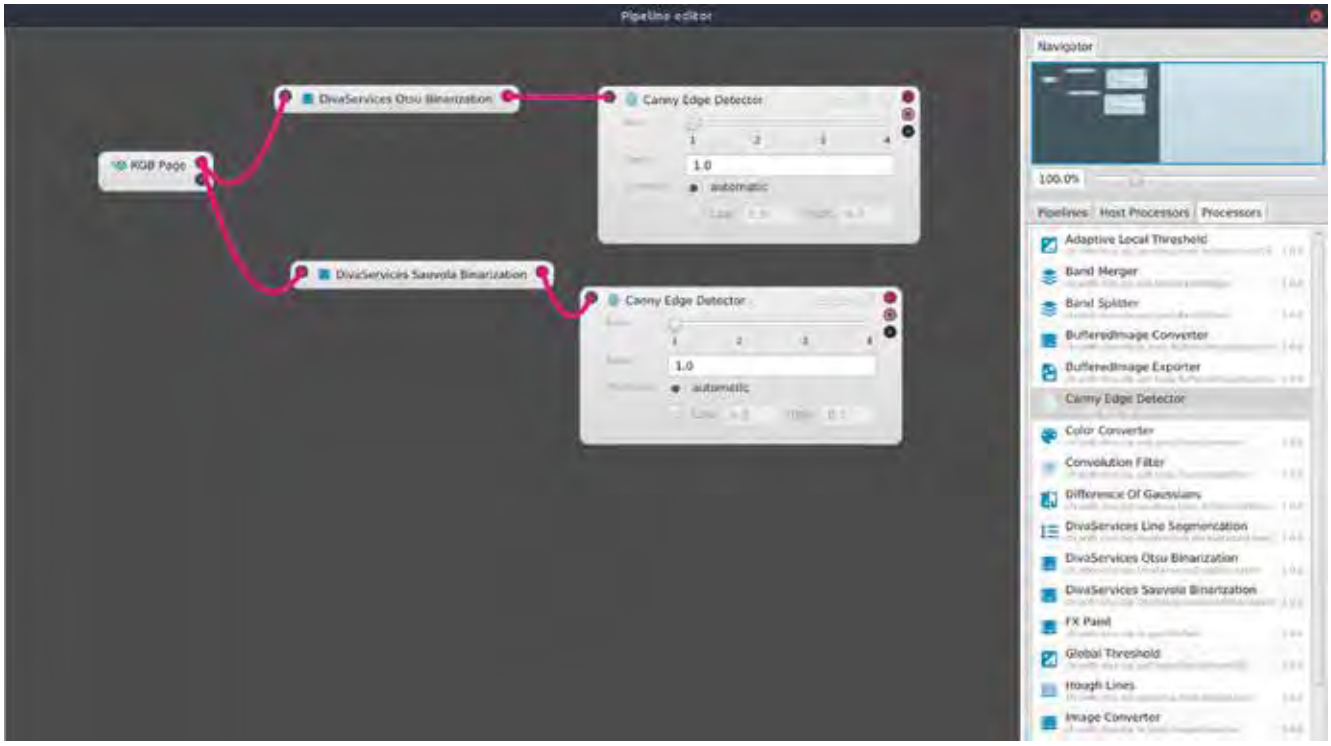


Fig. 2: Workfl w designer of DIVA-DIP. The various inputs and outputs are colour-coded for simplicity. Users can change available parameters for each method.

available in Boettiger 2015.When using DIVAServices, the data must be transferred onto its servers to be processed. For this, we require all data that are processed with DIVAServices to be released under a Creative Commons[2] license. We are aware that this is a problem in many humanities institutions where data is locked behind very restrictive licenses. For this reason, we provide the complete DIVAServices framework under an Open Source LGPL v2.1[3] license. This allows everyone to install their own instance of DIVAServices on their local infrastructure, making use of the same principles without losing access to their data.

Currently, when working with DIVAServices, the only way to design and execute complex workflows is by interacting and by programming it. This requires a developer who writes the necessary code that brings together the inputs and outputs of the various methods and chains them together. Some of the users of DIVAServices, like scholars from the humanities, sometimes lack these skills. In this work, we aim at building a WMS that allows them to design their workflows without the need to write a single line of code

The DIVA Document Image Processor (DIVA-DIP)

DIVA-DIP is a novel WMS for designing DIA workflows of methods offered through DIVAServices. The goal is to provide a tool that allows non-technical experts to design and execute workflows without the need to programme them. With the direct integration of DIVAServices, DIVA-DIP is able to offer state-of-the-art DIA methods without the need to install any of them on the local computer.

Figure 2 shows the most important interface of DIVA-DIP, the workflow designe . In the designer, the user can see all currently available methods (called processors). To start, each workflow needs to have at least one input processor that represents the input page (digital image). Processors can be added to the workspace and the current workflow by dragging them onto the workspace. DIVA-DIP supports the user in connecting processors together using a colour-coding schema. Each input and output has a specific colour and connections can be made only between colours that match. This should help non-technical users in particular to find out what inputs and outputs work together. Additionally, if a method takes additional arguments, they can be provided for the method but can also be changed later when executing the workflo .

[2] See *Creative Commons* <https://creativecommons.org/choose/> for the different type of Creative Commons license. We suggest using CC BY-NC 4.0, which allows DIVAServices to create adaptations of the original work without any problems.

[3] See *GitHub* <https://github.com/DIVA-DIA/DIVAServices>.

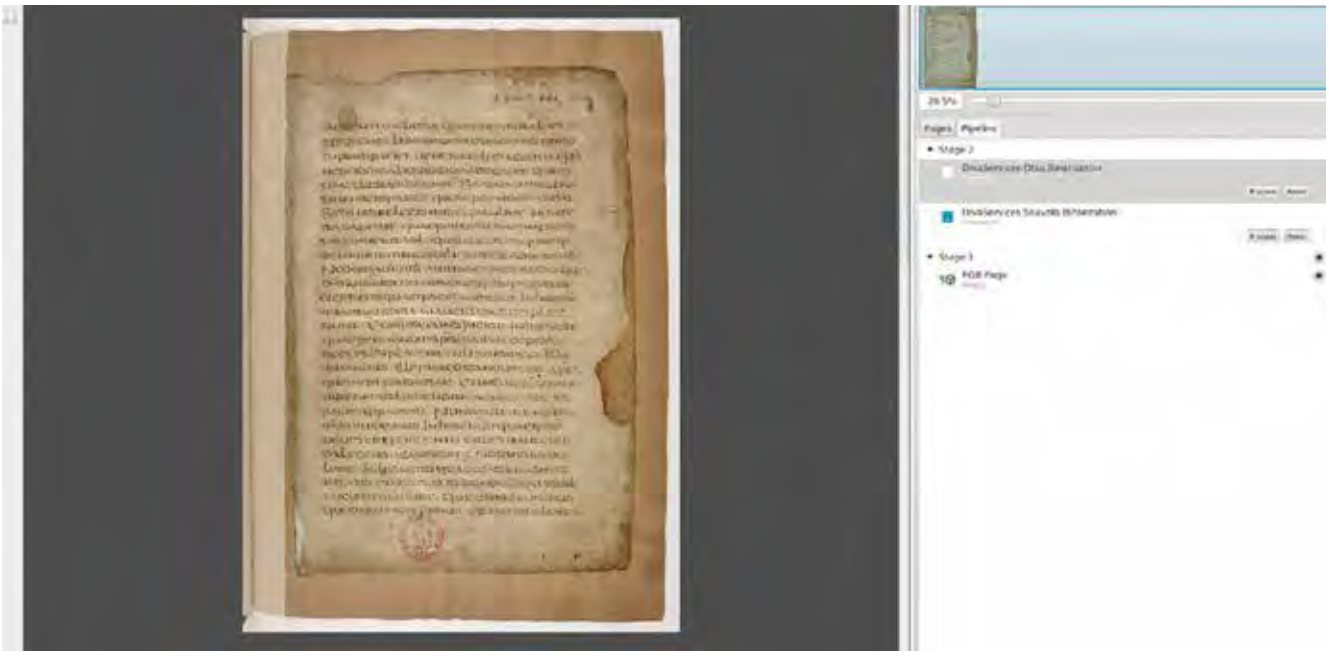


Fig. 3: The User Interface of DIVA-DIP. The focus is on providing much space to the actual document. All workfl w related information is available on the right.

Fig. 4: Visualization of a computed result. The user can switch between the various results using the radio buttons on the right side.

Unlike traditional WMS that only execute the workflo , DIVA-DIP offers additional features. The most important distinction is that, in DIVA-DIP, processors can enter a specific state where they wait for user input. This makes it possible to design interactions with semi-automatic methods without having to build specific applications. An example of where this could be used is in correcting recognized text line bounding boxes when performing layout analysis. A second difference of DIVA-DIP is that it allows for inspecting the results computed at each step. This enables a user to inspect each result individually and make changes to the workflow based on this visual feedback. Based on changes made after

this visual inspection, the user can then re-run only the necessary parts of the workflow without having to re-run the complete workflo .

Figure 3 provides an overview of the regular user interface. This interface focuses on providing as much space as possible to the document that the user is currently working on, pushing all information related to the current workflow to the right of the screen. There the user can see all the individual steps of the current workflow and can execute them

After a computation successfully finishes, the user can visualize the result of this computation, as shown in Figure 4. The status of the workflow step (the small square box next to its name) will turn green when the computation, in this case a binarization, is successfully executed. Using the radio buttons enables the user to select the result of the step they would like to have visualized on the workbench.

In our continued effort to open science and reproducibility, DIVA-DIP is fully Open Source released under LGPL v2.1.[4] This makes it possible for everyone to add processors to the application, offering methods that do not necessarily need to be hosted on DIVAServices but could be served either as local applications or through other Web Service frameworks.

Future Work and Conclusions

DIVA-DIP is a new WMS that focuses intensely on the specific differences of workflow in DIA, by keeping the user in the loop. For this, we introduced a new state that processes can enter when they wait for user input. Additionally, DIVA-DIP offers access to state-of-the-art DIA methods through DIVAServices and also eliminates the need to perform any local installation for the methods.

We believe that these additional features make it a great tool for scholars in the humanities, as it keeps them engaged in the process and enables them to investigate the impact of each step involved in the workflo . DIVA-DIP also completely removes the need for any programming knowledge to design the workflows, but offers a visual interface that is easy to understand.

Our hope is to evolve DVA-DIP into the main workflow design tool for DIVAServices. In order to get there, we need to add some more features in the future, namely: develop a way that automatically provides access to all methods currently offered on DIVAServices. Currently, a developer needs to programme and add each method as a processor in DIVA-DIP. We would like to change this, such that at start-up DIVA-DIP would automatically check which methods are available on DIVAServices and provide access to these.

We would also like to engage with the community at large and invite its members to provide access to their own methods through DIVA-DIP as well, so that it could become a WMS providing access to as many DIA methods as possible.

[4] See *Github* <https://github.com/DIVA-DIA/dip>.

REFERENCES

Boettiger, Carl (2015), 'An Introduction to Docker for Reproducible Research', *ACM SIGOPS Operating Systems Review*, 49.1: 71–79. doi: <10.1145/2723872.2723882>.

Clausner, Christian, Stefan Pletschacher, and Apostolos Antonacopoulos (2011), 'Aletheia: An Advanced Document Layout and Text Ground-Truthing System for Production Environments, *Proceedings of the International Conference on Document Analysis and Recognition*, 48–52. doi: <10.1109/ICDAR.2011.19>.

Deelman, Ewa, Karan Vahi, Gideon Juve, Mats Rynge, Scott Callaghan, Philip J. Maechling, Rajiv Mayani, Weiwei Chen, Rafael Ferreira da Silva, Miron Livny, and Kent Wenger (2015), 'Pegasus: A Workflow Management System for Science Automation', *Future Generation Computer Systems*, 46: 17–35. doi: <10.1016/j.future.2014.10.008>.

Garz, Angelika, Mathias Seuret, Andreas Fischer, and Rolf Ingold (2016), 'GraphManuscribble: Interact Intuitively with Digital Facsimiles', *Second International Conference on Natural Sciences and Technology in Manuscript Analysis, CSMC, Book of Abstracts*, 61–63.

Karwowski, Waldemar (ed.) (2001), *International Encyclopedia of Ergonomics and Human Factors* (London: Crc Press).

Myers, Brad A. (1986), 'Visual Programming, Programming by Example, and Program Visualization: A Taxonomy, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 59–66. doi: <10.1145/22627.22349>.

Nafchi, Hossein Ziaei, Seyed Morteza Ayatollahi, Reza Farrahi Moghaddam, and Mohamed Cheriet (2013), 'An Efficient Ground Truthing Tool for Binarization of Historical Manuscripts', *12th International Conference on Document Analysis and Recognition*, 807–811. doi: <10.1109/ICDAR.2013.165>.

Romero, Verónica, Alejandro H. Toselli, Luis Rodríguez, and Enrique Vidal (2007), 'Computer Assisted Transcription for Ancient Text Images', *Image Analysis and Recognition* (Berlin, Heidelberg: Springer), 1182–1193. doi: <10.1007/978-3-540-74260-9_105>.

Wolstencroft, Katherine, Robert Haines, Donal Fellows, Alan Williams, David Withers, Stuart Owen, Stian Soiland-Reyes, Ian Dunlop, Aleksandra Nenadic, Paul Fisher, Jiten Bhagat, Khalid Belhajjame, Finn Bacall, Alex Hardisty, Abraham Nieva de la Hidalga, Maria P. Balcazar Vargas, Shoaib Sufi, and Carole Goble (2013), 'The Taverna Workflow Suite: Designing and Executing Workflows of Web Services on the Desktop, Web or in the Cloud, *Nucleic Acids Research*, 41: W557–W561. doi: <10.1093/nar/gkt328>.

Würsch, Marcel, Rolf Ingold, and Marcus Liwicki (2015), 'DIVAServices: A RESTful Web Service for Document Image Analysis Methods, *Digital Scholarship in the Humanities*, 32.suppl.1: i150–i156. doi: <10.1093/llc/fqw051>.

Würsch, Marcel, Rolf Ingold, and Marcus Liwicki (2016), 'SDK Reinvented: Document Image Analysis Methods as RESTful Web Services', *12th IAPR Workshop on Document Analysis Systems*, 90–95. doi: <10.1109/DAS.2016.56>.

Würsch, Marcel, Marcus Liwicki, and Rolf Ingold (2018), 'Web Services in Document Image Analysis: Recent Developments and the Importance of Building an Ecosystem', *13th IAPR Workshop on Document Analysis Systems*.

Article

Building an Evaluation Framework Researchers Will (Want to) Use

Joseph Chazalon | Le Kremlin-Bicêtre

Abstract

Researchers often face difficult es regarding the availability and the quality of public evaluation frameworks when attempting to compare individual contributions with the state of the art. Firstly, they are often limited to a poorly documented raw dataset that is hard to obtain. Secondly, they seldom provide consistent specifications of the tasks they are designed to evaluate – the dataset and the evaluation methods. Lastly, they often are complex to use, require the use of centralized platforms or may not be available in the long term. Based on the experience we have in building and using evaluation frameworks, we present here a proof of concept for a modified version of the smartDoc competition (challenge 1). This prototype takes the form of a fully open Python package that offers great ease of use: it can be installed with a single command; loading the dataset and its associated ground truth can be reduced to one line of code; and benchmarking a given method takes only 8 lines of code.

1. Introduction

To cope with the ever-increasing pace of research work, researchers need to be able to re-use research products that can be trusted and for which results can be easily reproduced. Open initiatives about methods, tools, services, datasets etc. lay the foundations we can build on. We believe *Evaluation Frameworks* play a particularly critical role here, as the comparison between methods is of paramount importance to allow researchers to make the right choice at the right moment.

By *Evaluation Frameworks (EFs)*, we mean a set composed of:

1. the clear definiti n of a processing task (input and outputs, goal) — see Fig. 1 for an example;
2. a set of representative input and expected output data;
3. an evaluation protocol based on proven procedures (along with tools implementing them) that make it possible to measure how well predicted results match expected ones.

Input image

Ideal output segmentation

Fig. 1: Overview of the segmentation task ('Task 1' – Original task of the SmartDoc competition – challenge 1): from a raw video frame, locate the coordinates of the 4 corners of the document outline.

Building such evaluation frameworks is often considered a complex task, and few researchers invest time in this direction. However, we believe that using the right approach makes this goal easier to achieve than expected and rewards the creators of *EFs* with many benefits

What we present here is a proof of concept for an evaluation framework focused on a particular set of tasks for Document Analysis and Recognition (DAR). One of these tasks (a segmentation task) is illustrated in Fig. 1. We have

tried to make these tools as simple as possible to re-use for anyone with basic experience in the DAR field. They take the form of a Python package, which can be installed with one single command:

```
$ pip install smartdoc15_ch1
```

Given a function *you defined* named detect_object(image) → contour that takes an image as input and returns the outline of the document found in the image, the *complete* evaluation code of *your* method can be reduced to:

```
from smartdoc15_ch1 import (
  Dataset,
  evaluate_segmentation)
from my_module import detect_object
d = Dataset()
pred_seg = []
for frame in d:
  seg = detect_object(frame.read_image())
  pred_seg.append(s)
evaluate_segmentation(
  pred_seg,
  d.segmentation_targets,
  d.model_shapes,
  print_summary=True)
```

The source code for all the material presented here is available at the following URLs:
<https://github.com/jchazalon/smartdoc15-ch1-dataset>
<https://github.com/jchazalon/smartdoc15-ch1-pywrapper>.

After a brief review of a selection of notable initiatives that support open or reproducible research, with a focus on DAR (Sec. 2), we present our contribution (Sec. 3): a proof-of-concept evaluation framework based on a task-oriented definition of the SmartDoc 2015 dataset (Sec. 4.1), a new distribution scheme for the raw dataset (Sec. 4.2) and an easy-to-use Python wrapper for data loading and method evaluation (Sec. 4.3).

2. Related work

As mentioned in the introduction, we take into consideration three aspects of evaluation that constitute a consistent evaluation framework: task definitions, datasets and evaluation methods.

Within the Document Analysis and Recognition (DAR) community, the Technical Committees 10 and 11 maintain a list of publicly available datasets for research use.[1] While the datasets listed by the curators are generally free to obtain and have research-friendly licenses, it is not uncommon that the download links get broken, causing long-term availability issues because of unreliable hosting solutions. Furthermore, there are even fewer options for evaluation tools and it is very rare that free and open tools are released. A notable exception is the excellent UNLV-ISRI set of OCR evaluation tools,[2] which is still used and maintained, even 20 years after its release.

To our knowledge, the first initiative that really embraced the three aspects of an evaluation framework is the Robust Reading Competition (RRC) series.[3] RRC provides a clear definition of the different problems (or tasks) that researchers can evaluate their methods against. Datasets are hosted on the platform; upon submission of method results, an evaluation and a ranking of the methods are automatically performed. This approach now faces the issue of rising hosting and maintenance costs. In response, the platform is now being progressively opened to avoid interrupting the service. This solution could also solve the issue of having closed-source evaluation procedures and secret ground truth for some datasets.

The Document Analysis and Evaluation (DAE) platform[4] was another important initiative. It was more general in the sense that it aimed at offering an orchestration platform that could be used to compose data processing units (exposed as web services) and data sources. The complexity of the packaging of processing units made the dataset distribution aspect much more successful than the processing one.[5] To avoid availability issues for datasets, a new version of the platform (DAE-NG) now focuses on building a federation of synchronized dataset repositories. However, this still requires specific deployment and maintenance skills

[1] Available at http://iapr-tc10.univ-lr.fr and http://tc11.cvc.uab.es.

[2] Rice and Nartker 1996.

[3] Karatzas et al. 2011.

[4] Lamiroy and Lopresti 2012.

[5] Lamiroy 2017.

A last initiative from the DAR community we would like to mention is the DIVAServices platform.[6] It enables the easy packaging of document processing tools using Docker containers and their testing or combination using a REST API. Hosting evaluation methods on this platform is promising, but it requires hosting datasets on a separate platform and can lower the consistency between data and evaluation. Finally, while DIVAServices creators build the platform with openness in mind, deploying and maintaining a complex platform is necessary to run experiments, penalizing long-term availability unless a consortium eventually takes over the project.

It is interesting to note that the DAE, DAE-NG and DIVA platforms were inspired by other communities like bioinformatics, which successfully built massive platforms[7] where researchers can download and upload massive datasets (genomics for instance) and even launch computations on public clouds against standard datasets. We fully support the initiatives of our colleagues from the DAR community, as these efforts are federating the community around important problems and encourage collaboration, and we do not want to minimize their impact. Our goal here, however, is slightly different and fully compatible: we believe datasets and associated evaluation frameworks should be as open as possible and available over the long term. Such evaluation frameworks should then be hosted on those platforms, but we believe that if users can easily reproduce results by themselves, this will make things a bit more durable. Furthermore, even in communities like bioinformatics or medical imaging, evaluation methods and tools are often used for competition, but not always released in an open way. This reduces the impact of such work, as other researchers may not be able to reproduce the results by themselves.

An approach that reconciles long-term availability, ease of use and a strong consistency between dataset content and evaluation procedures is the Scikit-learn Dataset API.[8] This free and open Python library can be installed with a single command, and Scientifi Python is a new standard for computer vision research. It often makes evaluation very easy to implement, thanks to the many standard functions provided. The Dataset API makes it possible to load many datasets in a computable format with a single line of code. We believe such an approach is an interesting alternative to centralized evaluation platforms.

3. Building an open evaluation framework

Based on the original version of challenge 1 of the smartDoc 2015 competition, we revised its task definition to enable a wider use of the dataset and the evaluation tools we created (Sec. 3.1). To facilitate the dissemination and evolution of this new evaluation framework, we separated the implementation of the raw dataset distribution (Sec. 3.2) from the Python library ('wrapper'), which enables the manipulation of its content as *computable* objects, as well as the straightforward evaluation of any method that complies with the previous task definition (Sec 3.3).

As previously mentioned, all the sources and products of this proof of concept are freely available online. Our main goals were to build an evaluation framework that would be:

1. as easy to use as possible;
2. available in the long term;
3. reliable and trustworthy.

3.1 New task definitions

As previously mentioned, the original dataset we used as the basis of this proof of concept is the SmartDoc 2015 database for document capture (challenge 1).[9] This dataset was initially created to evaluate the performance of smartphone applications for document image acquisition, focusing on one of the first stages of the pipeline: the segmentation of the document outline in video frames or pictures, in order to allow the correction of perspective distortion.

The dataset was built by capturing 30 document models (5 for each of the 6 different types as shown in Fig. 2) under 5 different background scenarios (as visible in Fig. 3). Some small noise and margins from the original document images were removed and finally the images were rescaled to have the same size and fit an A4 paper format, resulting in several variants of the 30 model images. In addition to the video clips, a picture of each of the documents was captured to be used as another set of models variants.

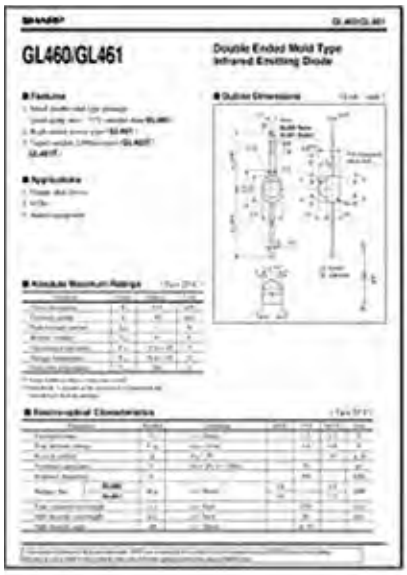
Each of these documents was printed using a colour laser jet and captured using a Google Nexus 7 tablet. The dataset consists of 150 video clips, comprising nearly 25,000 frames, captured by hand while holding and moving the tablet. The

[6] Würsch, Ingold, and Liwicki 2016.

[7] For an example, see <https://www.genouest.org/hosted-resources-and-tools/>.

[8] Buitinck et al. 2013.

[9] Burie et al. 2015.



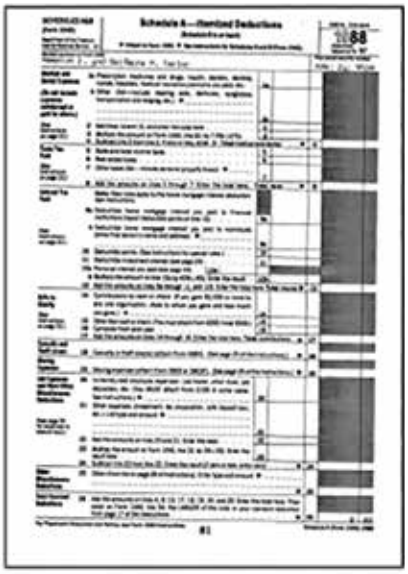
Datasheet 2

Letter 4

Magazine 2

Paper 5

Patent 4



Tax 3

Fig. 2: Sample documents used in our dataset.

video frames present realistic distortions, such as focus and motion blur, perspective, change of illumination and even partial occlusions of the document pages. The ground truth of segmentation data was created by semi-automatically annotating the quadrilateral coordinates of the document position for each frame in the collection.

The new version of the dataset makes use of the model images we created and captured. Each of the new tasks we are about to introduce can have two variants: 1) a *model-agnostic* variant with no knowledge of the original document models; 2) a *model-aware* variant based on the knowledge of the complete set of document model images. This second type of scenario allows researchers to test applications like augmented reality or form digitization.

Researchers can test their methods against three tasks using this new dataset. For each of these tasks, the *model-aware* variant is obtained by adding one extra input: the set of model images (or the result of the indexation of the latter).

3.1.1 Task 1: Segmentation (original task)

Inputs are video frames and expected output comprises the coordinates of the four corners of the document image in each frame (top left, bottom left, bottom right and top right). The evaluation is performed by computing the intersection

Background 1

Background 2

Background 3

Background 4

Background 5

Fig. 3: Sample backgrounds used in our dataset.

over union ('IoU' or also 'Jaccard index') of the expected document region and the detected region. The frame coordinates are projected onto the document referential to allow comparisons between different frames and different document models. The original evaluation code is available online,[10] and the Python wrapper also contains an implementation using the new data format.

3.1.2 Task 2: Model classification (new task

Inputs are video frames and expected output is the identifier of the document model represented in each frame. There are 30 models named 'datasheet001' to 'tax005'. The evaluation is performed using usual multi-class classification metrics: mean accuracy, precision, recall etc.

3.1.3 Task 3: Model type classification (new task

Inputs are video frames and expected output is the identifier of the document model *type* represented in each frame. There are 6 model types, each having 5 members, named 'datasheet', 'letter', 'magazine', 'paper', 'patent' and 'tax'. The evaluation is performed using usual multi-class classification metrics: mean accuracy, precision, recall etc.

3.2 Raw dataset distribution

We changed the dataset format to adapt it to the new task specifications. We did so with the objective of supporting long-term availability, which also has consequences for the hosting strategy we followed. The resulting product has a dedicated GitHub project available online.[11] GitHub hosting

[10] <https://github.com/jchazalon/smartdoc15-ch1-eval>.

[11] <https://github.com/jchazalon/smartdoc15-ch1-dataset>.

has the immediate benefit of making the project easy to find and self-documented, thanks to the embedded documentation viewer.

To design a dataset still usable in ten years, we reviewed and changed the data format used for the original version of the dataset. The original distribution was based on a set of video files, along with XML files in a custom format for the ground truth. The files were available at some secure file server using a procedure sent by email to users after they registered and accepted the dataset license on the main website. This process is sustainable, thanks to email automation, but it does not support automated download from a Python script. Furthermore, the explicit license agreement is an unnecessary burden for a standard Creative Commons license.

We started by producing a format that makes reading data as simple as possible. We first extracted all the frames from the video files and saved them as JPEG images to minimize decoding issues. We then created a simple CSV file for storing all metadata about each video frame: each line represents a frame observation, the columns store either information about file location or the ground truth for each task. The format is documented to remove any ambiguity about content types. The resulting files (images, metadata, documentation, license etc.) are packaged into a gzipped TAR archive for maximal compatibility. The model images were packaged using the same process.

Regarding data hosting and distribution, we considered several options with the constraint of being widely accessible and durably available. At the time we started working on this project, we chose to use GitHub[12] *releases*, as they feature very simple HTTPS upload and download while supporting version numbers. Since then, we started using Zenodo,[13] an EU-funded platform that specifically targets the archiving and distribution of research datasets. Zenodo and GitHub work well together, as Zenodo can automatically archive GitHub files (sources and binaries in *releases*). We now strongly encourage researchers to consider Zenodo, as it is entirely dedicated to the archiving and dissemination of research products and has many interesting features.

The resulting solution (either using GitHub or Zenodo) makes data archives directly downloadable (and verifiable using SHA256 checksums) with an implicit license agreement. Users will be able to download specific versions of the dataset based on the version number they carry. We hope versioning will encourage collaboration and the improvement of datasets, while allowing researchers to keep track of which sets of version numbers produce comparable evaluation results.

3.3 Python wrapper for data loading and evaluation

Using this reliable dataset distribution, we built a Python wrapper with three goals in mind: 1) making the dataset '*computable*' thanks to the automated loading of the raw archive as ready-to-compute objects; 2) making evaluation so easy that researchers will want to work with it; and 3) leveraging openness to maximize trust in the correctness of the method, as well as enabling long-term availability.

The resulting code has a very simple API organized around two kinds of functions: loading functions (one for the frames, one for the models), which have options to load images, pre-process them and load the associated ground truth for each task; evaluation functions (one for each of the three tasks). We tried to comply as much as possible with Python and SciPy philosophies in order to provide researchers with a plug-and-play library, as illustrated in the listings of the first page. In particular, every data series is a Numpy array that supports all the useful operations researchers are familiar with. For instance, is it possible to directly pass the result of the dataset loading to the automated function of Scikit-learn to separate training and test sets to perform cross-validation. The evaluation therefore complies with the usual evaluation pattern for estimators, comparing the target results with actual ones.

To maximize the usability of this proposed evaluation framework, we packaged this Python wrapper as a PIP package listed on the central Python Package Index (PyPI) and installable with a single `pip install` command. Finally, the source of these tools can be inspected, forked, maintained, improved and contributed to at: <https://github.com/jchazalon/smartdoc15-ch1-pywrapper>. Thanks to the online documentation viewer, this address also features a simple and nice entry point for any researcher willing to experiment with this prototype.

4. Conclusion

We presented a proof of concept for the distribution of evaluation frameworks within the document analysis and recognition community. Greatly inspired by initiatives

[12] <https://github.com>.

[13] <https://zenodo.org>.

from other communities like bioinformatics, we proposed leveraging the Python Package Index to distribute code, as well as public storage offered by GitHub or Zenodo to distribute data and archive code. This facilitates the distribution of complete evaluation frameworks to researchers. We believe that distributing data, documentation and code all together is a key to engaging researchers in using such evaluation frameworks and to encouraging reproducible research.

Of course, we still rely on external platforms like GitHub (now owned by Microsoft) and its release feature (a core element of their product), Zenodo (backed by the CERN and part of the infrastructure that stores data for the Large Hadron Collider) or software frameworks like SciPy and Python (which receive massive attention and funding these days). These choices are, in our opinion, the most reliable options today, but this will obviously change. However, we believe that providing researchers with tools they can use easily, with few dependencies and great benefits, are a key enabler for long-term support of software (hence the implemented methods) and massive dissemination of datasets. We hope our simple proof of concept will encourage researchers to look for ready-to-use *evaluation platforms* and maybe even start distributing some.

REFERENCES

Buitinck, Lars, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, and Jaques Grobler (2013), 'API Design for Machine Learning Software: Experiences from the Scikit-Learn Project', *arXiv.org*: <arXiv: 1309.0238>.

Burie, Jean-Christophe, Joseph Chazalon, Mickaël Coustaty, Sébastien Eskenazi, Muhammad Muzzamil Luqman, Maroua Mehri, Nibal Nayef, Jean-Marc Ogier, Sophea Prum, and Marçal Rusiñol (2015), 'SmartDoc: Smartphone Document Capture and OCR Competition', *13th International Conference on Document Analysis and Recognition*, 1161–65. doi: <10.1109/ICDAR.2015.7333943>.

Karatzas, Dimosthenis, S. Robles Mestre, Joan Mas, Farshad Nourbakhsh, and P. Pratim Roy (2011), 'Robust Reading Competition-Challenge 1: Reading Text in Born-Digital Images (Web and Email)', *International Conference On Document Analysis and Recognition (ICDAR)*, 1485–1490. doi: <10.1109/ICDAR.2011.295>.

Lamiroy, Bart (2017), 'DAE-NG: A Shareable and Open Document Image Annotation Data Framework', *14th IAPR International Conference On Document Analysis and Recognition (ICDAR)*, vol. 4, 31–34. doi: <10.1109/ICDAR.2017.309>.

Lamiroy, Bart, and Daniel Lopresti (2012), 'The Non-Geek's Guide to the DAE Platform', *10th IAPR International Workshop on Document Analysis Systems*, 27–32. doi: <10.1109/DAS.2012.87>.

Rice, Stephen V. and Thomas A. Nartker (1996), *The ISRI Analytic Tools for OCR Evaluation* (Las Vegas: University of Nevada, Information Science Research Institute; Technical Report TR-96-02). <https://isri-ocr-evaluation-tools.googlecode.com/svn-history/r2/trunk/analytic-tools/user-guide.pdf.>.

Würsch, Marcel, Rolf Ingold, and Marcus Liwicki (2016), 'SDK Reinvented: Document Image Analysis Methods as RESTful Web Services', *12th IAPR Workshop on Document Analysis Systems (DAS)*, 90–95. doi: <10.1109/DAS.2016.56>.

Article

Turning Black into White through Visual Programming: Peeking into the Black Box of Computational Manuscript Analysis

Vinodh Rajan Sampath and H. Siegfried Stiehl | Hamburg

Abstract

In the recent past, an ever-increasing plethora of computational methods and tools for computational manuscript research, and by extension digital palaeography, have been developed and provided by the scientific community of digital image processing and analysis (or, in a wider sense, computational vision). Invariably, however, many of these methods and tools suffer from low usability from the point of view of humanities scholars, the actual end users. The black box problem is the most commonly cited reason for this sheer fact. While it may not be possible to completely eliminate the problem, we can alleviate it by dismantling the computational black box into smaller boxes and eventually turn them white. We discuss how visual programming as a paradigm will make peeking into the computational black box possible. In this context, we introduce the iXMan_Lab and AMAP, a collaborative web-based platform that facilitates the development and experimental validation of tools and workflows using an innovative visual programming paradigm. We will also briefly relate our approach to design thinking and open science.

1. Introduction

1.1 Computing in the context of interdisciplinary manuscript research

With the recent advances in the theories, methods and applications of various computational methods (pertaining to digital image processing, analysis and archiving – among other things) in the humanities and the consequent emergence of digital humanities as a scientific discipline, we are witnessing a burst of tools enabling digital palaeography and the thorough multi-facetted understanding of manuscripts. Even though a wide variety of these methods and tools aim to support scholarly work in computational manuscript research (CMR), only a few of the tools have found widespread and consistent acceptance and use (particularly in the case of digital palaeography). The reasons can be outlined as follows.

Most of the tools that are assumed – or even claimed – to be readily applicable to questions and problems in manuscript research were developed solely from the point of view of information science and often fail to take the specifi requirements of cognizant end-users into account who work in a specific domain of humanities-based manuscript research. The main consequence of involving the humanities to such a small extent in tool design is that users in such fields have difficulty understanding or employing the computer-based solutions (software tools) in their target domain. As a consequence, tools appear as black boxes (Hassner et al. 2014), acutely affecting their usability and usefulness. Frequently, users do not understand how or why the tools behave and perform in a certain way and as a consequence they question the results produced. In addition, they cannot influence or change the way the tools behave, because they cannot understand the way individual tools (or their chained aggregation to a fully-fledge system) work. With only a parochial – or even no – way to understand, test or influence the underlying theories and computational methods, many of the tools end up being unusable in an academic/scientific setting (Stokes, 2012). As an important consequence, an inter-/transdisciplinary approach has to be taken, which first demands a genuine methodology for building bridges between i) humanistic manuscript research primarily grounded in hermeneutics and ii) computational manuscript research grounded in algorithmics.

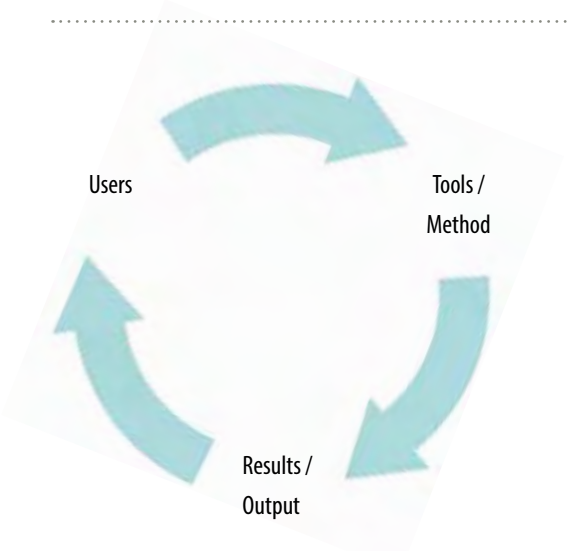


Fig. 1: A simplified user-in-th -loop system.

1.2 Document Image Analysis

According to Kasturi et al. (2002), document image analysis (DIA) consists of 'algorithms and techniques that are applied to images of documents to obtain a computer-readable description from pixel data methods'. In our context, the general term 'documents' refers to either contemporary or historical prints and manuscripts, while 'images' is synonymous with 'digitized manuscripts'. DIA methods such as image pre-processing, keypoint detection, visual feature extraction, text line finding, page layout analysis, writing style analysis and hand identification are employed within computational manuscript research for a wide variety of purposes ranging from simple image quality enhancement to increased legibility to advanced applications such as writer identification for palaeographic dating. Given the stage of research in computational vision, however, it is safe to state that tools for particular tasks and domains cannot be directly derived from published theories and methods for at least two reasons: almost all methods entail parameters that have to be carefully adapted to the particular task and domain at hand and, worth noting, the accuracy, reliability and robustness of their performance greatly depends on the quality of the digitized image (e.g. resolution and contrast, as well as degradations such as noise). In a strict sense, also methods or services from open source repositories (such as GitHub) are hardly applicable to a specific task in a, say, blind fashion – let alone the myth of error-free software. Consequently, for computational vision, the real issue is multifaceted and multi-staged: from theory to algorithms to methods to experiments to tools to services and systems – in total with the aim of supporting the workflow of humanities scholars involved in their hermeneutic research methods. Since computer-assisted manuscript research on the basis of tools also requires thorough validation, evaluation and benchmarking of methods and tools from above, scientifically grounded – but, alas time-consuming – experimentation with synthetic, real and ground truth data is indispensable, if only, e.g. to constrain the space of parameters and understand the propagation of errors through a chain of methods. Last but not least, two more critical aspects have to be mentioned: first, the lack of a clear-cut methodology and thus also any engineering approach to consistently link the realm of theory with the reality of workflow support for scholars and, second, the involvement or even embedding of affected scholars in the whole, notably cyclic, process of requirement analysis, system design, realization, evaluation and re-design, as well as the provision and maintenance of performing, web-based, interoperable and platform-independent tools. Clearly within contemporary informatics, such an integral approach has design thinking including advanced, e.g. agile, software design and development as one of its pillars and lays foundations for a methodology that is not only desirable, but much needed.

Many of the tools, simply speaking, can be considered to be a composite of various DIA methods. However, such a composition is not as simple as it may appear to a layman, and the resulting tools require proper consideration in terms of building blocks, or sub-modules, and their connectedness, parameter regimes, applicability and appropriate presentation of results, among other things (see above). Many of these building blocks are not exposed to end users, and even if exposed, do not provide a meaningful way of interaction to truly keep the user in the loop and in control. This invariably results in a composite tool becoming a monolithic black box that end users have trouble interacting with and trusting.

1.3 User-in-the-loop paradigm

User-in-the-loop as a paradigm (Hassner, 2013) is often suggested as a work-around to resolve many of the critical issues mentioned above (Fig. 1). In general, it attempts to give more control to users by i) opening up the tools to further useful and purposeful interaction, ii) having more control over the range, scale and configuration of methods and iii) even providing feedback for improvement of e.g. overall performance. The main idea here is to i) involve, not to say embed or even immerse, users – who are experts from their respective fields – in the cyclic process sketched in the best possible way and ii) metaphorically speaking, make them part of the algorithm and the methodology. Thus, what is in order is transforming the tools from being computational

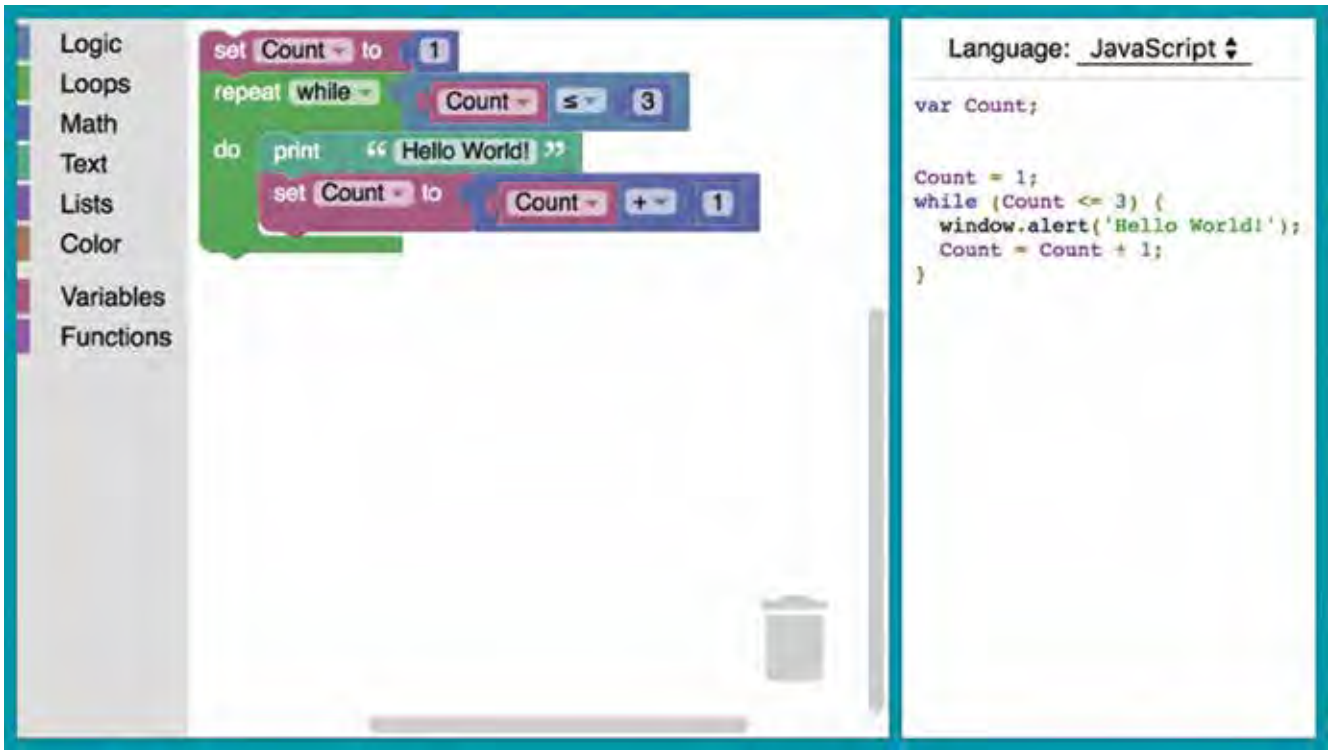


Fig. 2: Google Blockly.

soothsayers that take in data and render the divination of results towards actually performing, reliable and trustworthy systems that work in tandem with users, taking into serious consideration their needs, alterations and feedback. From the point of view of *conditio humana*, a tool (or computer in general) should serve a human and not vice versa. Hence, a scholarly expert in manuscript research should by no means be demeaned or vilified by a technologically allotted role as key puncher or mouse clicker (see below).

While we adopt the user-in-the-loop paradigm as the right and principal approach to open up the seeming black box, the paradigm appears quite late in the software engineering process. We strongly propose that gaining an understanding of the black box must ideally occur quite early in the whole process outlined so far, and in fact, this would mean that, ideally, users must be involved in the initial stages of the process of designing and building tools, as well, and not just in the final stages of using them, be it for experimental or practical purposes. This allows users to i) get an idea of what exactly goes into the tools and ii) garner a better understanding of the final tools or even system, including internal components, performance characteristics, degrees of freedom, parameter sensitivity etc. As will be argued below, the adoption of visual programming (VP) as a tool-building paradigm will enable us to accomplish the goals laid out above in an efficient, effective and user-friendly way.

1.4 Visual programming

Visual Programming (VP) is a novel, pioneering programming paradigm that allows users to develop computer programs by spatially arranging software modules, or computational methods in our case, as graphic symbols, typically in two dimensions, e.g. on a monitor or screen (Myers, 1990). Hence, users are enabled to build programs by putting together computational ‘Lego’-like blocks on a screen, or even a multi-touch table as in our case, in an interactive fashion. The graphic symbols can be either low-level, containing only control structures and variables, or, optionally, high-level, supporting various abstract modules relevant to the application domain. Several types of VP paradigms have been made available so far, the most popular ones being block-based and graph-based or flow-based VP languages. Blockly (Fraser, 2015) and AppInventor (Wolber, 2011) are block-based visual languages that allow users to arrange and fit various blocks into predefined slots and holes to create programs (Fig. 2). Differently, Microsoft Visual Programming Language is a flow-based visual language that allows users to define programs as a graph using nodes and edges, with the nodes usually denoting some sort of processing and the

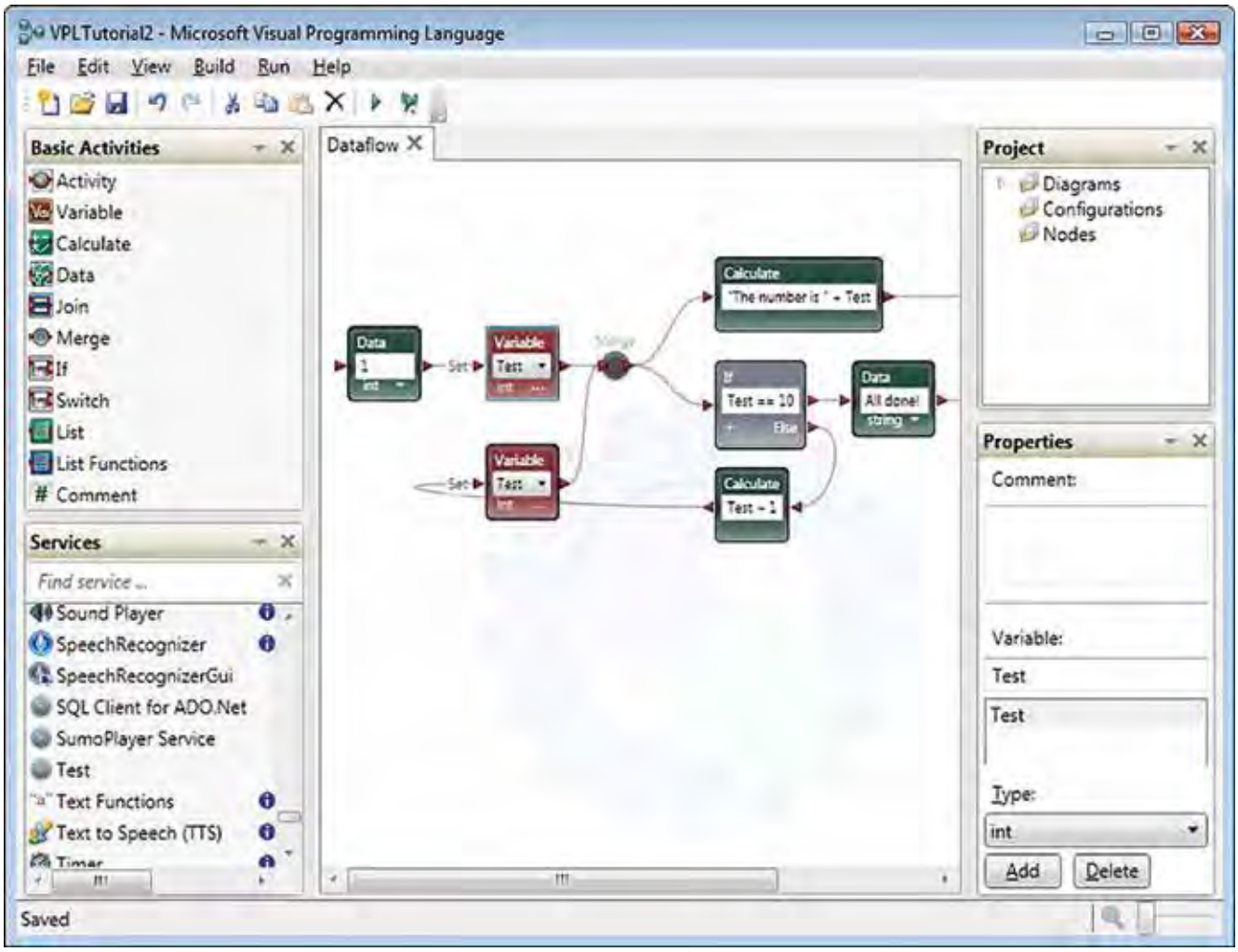


Fig. 3: Microsoft Visual Programming Language.

edges the in- and out-flow of data. Although indeed several other VP paradigms do exist, they are neither very popular nor relevant to our current research span (Fig. 3).

VP is becoming a popular programming paradigm in various domains, particularly in education to teach programming concepts and in Do-It-Yourself (DIY) environments for domain specialists to create both programs and workflows. VP allows specialists to create (or at the minimum, to co-create) solutions to well-defined domain problems in an easy and interactive way without the overhead of learning a mainstream textual programming language. A prominent example of the latter would be AppInventor, which provides a VP language and the necessary environment to develop web apps without the need for any programming experience.

2. Peeking into the black box

To properly eliminate, or at least alleviate, the black box problem, one has to completely understand the mathematical and functional limitations of all the computational methods that make up a tool or system. Often, this is achievable only if the tool developers have a strong mathematics background, particularly when it comes to advanced DIA methods (see above) and even more cutting-edge methods, e.g. stemming from the research field of machine learning (ML) and, particularly, deep learning (DL) in layered neural networks, which are frequently claimed to be black boxes per se due to their lack of accountability and transparency (or, in a strictly algorithmic sense, lack of proof of correctness and uniqueness). While it may not be possible to completely eliminate the black box property itself, it is feasible to deconstruct it into multiple smaller entities, which themselves could nevertheless be black boxes, thereby reducing complexity to a level of better transparency and, thus, understanding. This allows both peeking into the seeming black box and being able to conceive the multitude of interconnected parts contained within it. On a practical level, sometimes it is just enough if one understands

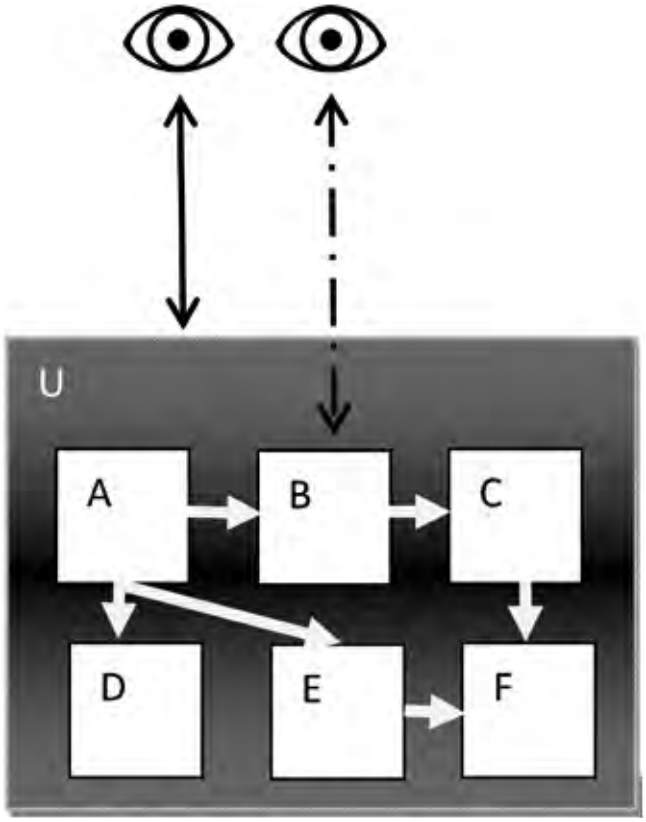


Fig. 4: Users usually only see the overall system *U*. But it is necessary that they also know about the subsystems.

the different components that make up a hierarchical or trivially sequential system and how they interact with each other (e.g. via linking data or cross-effects of parameters, as in the simple case of a typical image processing pipeline for easy tasks). Evidently, in such cases, a great deal of understanding can be easily provided. This is required in order to understand the overall workings of the system and, accordingly, to tailor or streamline it towards compliance with user requirements. The opaque nature of the decomposed subcomponents can further be reduced by exposing their working through appropriate visualization techniques (Fig. 4).

Again, due to the preponderance of non-trivial and non-mainstream scientific problems in interdisciplinary computational manuscript research, e.g. generic layout analysis or hand identification given abundant degradations in digitized manuscripts, sparseness of training data and the lack of ground truth and benchmarks far from a scholar's workflows, the demanded tools and systems are complex in theoretical, methodical, technical and human-factor terms, which is far different from 'toy' problems such as recognizing cats and dogs or preventing a moving robot from ramming a table leg (though these *are* engineering feats).

As already mentioned, with the rise of methods of training data-intensive deep learning (DL), it simply is more difficult or even virtually impossible to open up such systems, as a sufficiently deep neural network is itself a black box from both a theoretical and an algorithmic perspective. However, even DL tools are not always monolithic, since some kinds of pre-processing and post-processing steps almost always go along with the core DL method. Hence, on-going research in the respective communities is about to open these as much as possible again for the sake of a better understanding by both the developers and the users – particularly in the context of applications that require trustworthy computational results, as in signature verification. Although scientific problems with CMR may well be characterized as non-critical (e.g. in comparison with fraud detection), accountability, transparency, trustworthiness etc. are still for good reason key features of tools and systems – otherwise the willingness of scholars to use them in their daily routine will fade or not even be kindled. On top of that, by briefly touching upon issues of accuracy, reliability, replicability etc., an added-value of any CMR tool/system devoted to help solve scholarly research questions must be clearly demonstrated – otherwise it's l'art pour l'art.

As already mentioned, a VP environment can be helpful in deconstructing computational black boxes by taking two different approaches. In the first approach, a VP-based DIY environment can be provided to manuscript scholars in the humanities. It is seldom possible to provide specific computational solutions that are generalizable and applicable to scenarios with problem settings that scholars face on a day-to-day basis. Therefore, we would do better to focus on providing them with genuinely required toolsets that would let them i) explore various methods on their own, ii) deal with digitized manuscripts and iii) create custom-built solutions by themselves in an interactive way using the VP paradigm. They can choose various techniques, explore them and assemble them to produce solutions in accordance with their needs. Needless to say, a certain training of CMR novices and/or IT affinity is presupposed even in this ideal conception. In the second approach, a VP-like environment is used as part of the software engineering methodology to interact and communicate with scholars from the humanities. By using such an environment, experts from informatics are enabled to discuss/develop various solutions effectively by i) providing the scholars with a hands-on interactive experience in solution building

and ii) granting them a co-creating role. Rajan and Stiehl (2018b) coined the term 'interactive Exploration (iX)' for this approach and discussed it in great detail as part of a software/system development methodology (SDM) for tools in CRM. In both approaches, users (in our case manuscript scholars in the humanities) dismantle the monolithic black-box-like solutions either by themselves (in the first playful approach) or by jointly working in tandem or on a team with CMR experts who are sufficientl knowledgeable in computational vision (as in the second, more principled approach by explicitly constructing solutions as an assembly of more basic subcomponents in order to gain an understanding of their interaction and, hence, the overall functionality of the system).

Even in the worst case, in which users are not able to participate in the tool creation process, VP-like features can still be used to i) create an effective visualization of the final tool structure and functionality and ii) communicate the overall working mode and scope of the tool to the end users. Not only will this still help make the black box at least translucent, but it can also be used as a teaching and training environment for students and up-and-coming academics with a strong interest in CMR.

3. Using visual programming to turn the black box white

Below, we elaborate on how various features and properties of VP enable the dismantling process and also facilitate further opening up the black box in various other ways.

3.1 General advantages of visual programming

Visual programming offers several advantages over textual programming. Blackwell (1996) notes that 'typical statements are that VP is more user friendly, helpful, satisfying, intuitive, readable, familiar, appealing, accessible, reliable, pleasant, straightforward, alluring, immediate and obvious than other programming techniques'. As such, VP encourages non-programmers to play around with its visual elements, letting them explore and freely experiment with digital objects to attain their desired programming objectives. A functional computer program can thus be created in a short time by merely placing some graphic objects in an orderly manner. Also, a solution produced in a VP language can (under certain circumstances) be more understandable and communicable than a solution produced in a textual language.

One of the main advantages relevant to our problem setting in CMR is that VPs are better at expressing the problem structure. Their diagrammatic nature coupled with the semantic spatial arrangement enables users to better grasp the structures of a reasonably complex solution. The other major advantage is its resemblance to the real world of a particular application stemming from hermeneutic manuscript research. By mimicking the real world in its visual representation, a VP language can map the manipulation of real-world objects to those of digital objects by choosing an appropriate interaction metaphor. Such factors are advantageous in minimising the black-box problem by enabling users to better understand the intricacies of the tool or system.

3.2 Specific features of visual programming

As a paradigm, visual programming attempts to implement four core features: concreteness, directness, explicitness and liveliness (Burnett, 2002). In the following, we explore how these specific features can help to open up the computational black box with specific references to creating tools in CMR

3.2.1 Concreteness

Concreteness means expressing programmatic aspects using particular instances, e.g. mapping some aspect of semantics to desired behaviour using a specific object or property. A black and white brush realized as a tool could denote a binarization process (i.e. turning colour images into black and white) and the size of the brush could be directly proportional to the threshold of binarization (i.e. pixels below a certain threshold become white). Thus, we are mapping abstract methods such as binarization onto concrete graphic entities in a VP environment. By converting the abstract into concrete, users get a better grasp (also in the physical sense of the word) of the inner workings of a tool.

3.2.2 Directness

Directness can be described as 'the feeling that one is directly manipulating the object', which implies a minimal distance between an objective and the actions required to achieve it. This is usually implemented by choosing an appropriate interaction paradigm that maps the digital objects to appropriate real-world metaphors. To continue the previous example, binarizing a digital image could be implemented by moving a brush over the image. This allows users to intuitively interact with the system directly and make changes.

3.2.3 Explicitness

The internal aspects of a system are made visually explicit, enabling users to infer these aspects intuitively. Particularly in our context, this means making dataflo , e.g. in a chain of computational methods, explicit by visualizing the intermediates and also making parameters associated with various methods explicitly visual for direct control. By exposing the various methods, parameters, dependencies and interconnectedness of the components by the use of graphic objects and visual metaphors, users get a better understanding of the overall tool or system for CMR.

3.2.4 Liveness

The immediacy of feedback that is automatically provided by a program, tool or system is termed liveness. Tanomoto (1990) enumerates four levels of liveness. The first level corresponds to the static visual representation of the system and is by no means interactive. It is meant to be only a diagrammatic representation to help the user understand the structure and flow of a program. CMR tools must strive to provide at least this level of liveness, even if they do not use the VPL paradigm. In the second level of liveness, the system is interactive, and users are able to build the system with graphic elements and run the system to view the results. But users must explicitly execute the setup to view the results. In the third level, the users need not explicitly run the system whenever something is changed, since the system automatically runs in response to changes initiated by users. This encourages scholars to explore and try out different combinations, e.g. sub-modules and parameters and get immediate feedback. In the last level of liveness, the system is always on and provides temporal feedback based on the current state of the system. This can be very relevant for DL systems that typically take a long time to train or any system that handles high-volume data streams. Continuous visual feedback will keep the users engaged and involved by providing a glimpse of the current set and state of running processes, such as computational methods, along with their intermediate results, depending on the parameter settings. Also, e.g. in the case of experimentation with CMR tools in the iXMan_Lab, it is useful to keep full track of progress by compiling a comprehensive lab logbook in order to conform to standards in scientifically grounded experimentation (as known from paragons in experimental physics, psychology and the social sciences).

4. iXMan_Lab

In these CMR, DIA and VP contexts, we now introduce the *iXMan_Lab* (interactive eXploration of Manuscripts Laboratory), whose realization is one of the goals of the Scientific Service Project Z03 of the SFB 950. The underlying motto for the laboratory is to develop concepts, paradigms and prototypes that contribute to the realization of usable and useful CMR tools for manuscript scholars, which they can use in their work activities, as discussed earlier. More specificall , not only methods and tools being developed within Z03, e.g. for word spotting and writing style analysis (see e.g. HAT 2.0), but also open source methods and tools (e.g. from OpenCV; see also OpenX Workshop[1] of June 2018), as well as web-accessible services from various sources (see e.g. DIVAServices of Université de Fribourg), will be integrated in order to enhance the current scope of functionality. Web access to the lab is assured due to interoperability and platform independence – also meaning ubiquity of CMR functionality for scholars outside the Department of Informatics (as current lab host) who are equipped only with standard IT equipment such as desktop computers, laptops or tablets (whereby touch-based devices are preferred).

The lab is driven by an interdisciplinary team using a multi-touch table environment (powered by high-performance computing equipment) as a collaboration, cooperation and communication (C3) medium for a two-fold aim: first, experimentally designing a manageable, feasible and reliable processing chain based on computational vision methods for processing/analysing digitized manuscripts and, second, freezing in a jointly validated (or even evaluated or benchmarked) processing chain by interdisciplinarily reached consensus in order to deliver a useful tool for a broad range of users. In terms of hardware infrastructure, the laboratory is currently equipped with a custom-built 65-inch Multi-Touch Table (MTT) supported by a high-performance multi-core gaming engine. The MTT is additionally adjustable to a wide range of height and angle settings to enable various forms of team collaboration. The laboratory is completely equipped to run both GPU-accelerated image processing/

[1] Centre for the Study of Manuscript Cultures <https://www.manuscript-cultures.uni-hamburg.de/register_openx2018.html>.

Fig. 5: The MTT setup in iXMan_Lab.

analysis algorithms and, if necessary, deep learning methods (Fig. 5).

Even though primarily situated within the Department of Informatics as mentioned above, the lab is uniquely placed within the Centre for the Study of Manuscript Cultures, the host of SFB 950, and is able to web-interact with various scholars from sub-projects of SFB 950. In sum, the laboratory, once fully-fledged, will enable scholars to i) perform meticulous requirement engineering, ii) design and realize experiments and iii) provide workflow-supporting tools due to the close interaction between scholars from Manuscript Studies and Informatics.

5. Advanced Manuscript Analysis Portal (AMAP)

Currently, the main focus of the iXMan_Lab is the further development of the *Advanced Manuscript Analysis Portal* (AMAP), whose conceptual development started in 2015 in the Scientific Service Project Z03 with the beginning of the second funding phase of the SFB. Rajan and Stiehl (2018a) gives a first outline of the design, architecture and functionality of AMAP, which is equipped with an intuitive interaction paradigm in the context of a multi-touch table. In brief, it will allow users to i) intuitively deal with various advanced image processing/analysis methods and other manuscript-related methods and ii) create their problem-specific (and thus customized) chains of computational methods. The general goal is to design and develop AMAP in such a way that even advanced methods can be applied in an easy and intuitive manner by scholars without any programming and only rudimentary technical background. Furthermore, we are designing AMAP to be able to encourage and foster the exploration of various methods, tools, services and workflow

and, at the same time, to enable ease-of-use without any steep learning curve. AMAP ultimately aims to provide a VP-based environment and to support both approaches to deconstructing the black box (as discussed in section 2). It can serve as both a DIY-like environment for scholars working in the humanities and at the same time as an experimental platform for facilitating the interaction and communication of interdisciplinarily constituted teams from the humanities and informatics during the software development process.

We particularly chose to realize AMAP via an MTT, as touch-based technology is gaining huge traction and has the potential of becoming the primary mode of interaction in the near future. Even today, touch-based interfaces are becoming increasingly popular, compared with the traditional Windows/icons/mouse/pointers-based interfaces (WIMP). Also, having a large-scale interaction/interface area available is necessary when interacting with multiple high-resolution images, which is usually the case when analysing digitized manuscripts. The MTT can be further augmented to allow multiple input modalities that could be harnessed to make the system even more natural through its ability to model and mirror physical real-world interaction, e.g. by speech and deixis, with manuscripts as much as possible. An MTT is also an ideal medium to encourage real-time collaboration in a team of scholars from the humanities and informatics through the provision of a sharable, large-scale, monitor-based interaction device.

AMAP currently offers a rich selection of various functionalities such as image filtering, binarization, visual feature detection, word spotting, page layout analysis and writing style analysis (Mohammed et al. 2017). In fact, multiple methods providing the same functionality are also included to enable users to experiment and choose a method that is best suited to their task at hand. Our system also offers the ability to integrate other backend systems that provide DIA techniques as web-based services. This has been realised by integrating methods available at DIVAServices (Würsch et al., 2016) as part of AMAP. Such integrations demonstrate the flexibility of our approach, as well as the ability to assimilate wide-ranging manuscript-related computational methods from the OpenX community (where X stands for data, methods, tools, services etc. in accordance with the Open Science paradigm) into our platform.

5.1 VP and AMAP design

We are currently implementing an innovative hybrid VP language that integrates both a flow-based approach and a block-based approach. The paradigm works on the principle of visualizing the digitized documents and computational methods as virtual objects that can be manipulated spatially in relation to each other in order to aggregate various processing chains, e.g. in the context of experimentation and/or to create task- and problem-specific workflows for scholars from the humanities. The user interface (UI) in particular is designed to reflect real-world metaphors as much as possible in terms of interaction. Users can i) directly attach/detach various methods to/from the digitized manuscripts before and during chaining to processing pipelines and ii) manipulate parameters and subsets of manuscripts to process them even further. Additionally, the UI also supports natural interaction including actions such as piling the pages and turning them to take notes.

Fig. 6 shows a sample screenshot of a rather simple processing chain that has been realized with AMAP for demonstration purposes. It includes a textline detection method that requires a binarization step beforehand as pre-processing. Instead of being a single block, this decomposition explicitly shows that the quality of the textline detection is at least partially dependent on the pre-processing step, and by controlling the pre-processing step the quality of the results can be adjusted to the user's need. Also, by adding several other pre-processing steps, the results can be improved even further through goal-directed experimentation. The user is able to have some control over the system and understand how the process works, as opposed to a single processing block that provides only the output to a given input. Furthermore, a detected single textline from the output can also be seen to be extracted as a subset and a filter can be applied to it, specifically to increase its readability. One can also see a digital logbook recoding all previous operations performed with AMAP along with their timestamp and parameters of the operations.

6. Conclusion

We reported on the current state of computational manuscript research (CMR) and the inherent black box problem that particularly results in low acceptance of computational tools in scientific settings within scholarly manuscript research in the humanities. We proposed to alleviate the black box problem by dismantling the computational black box(es)

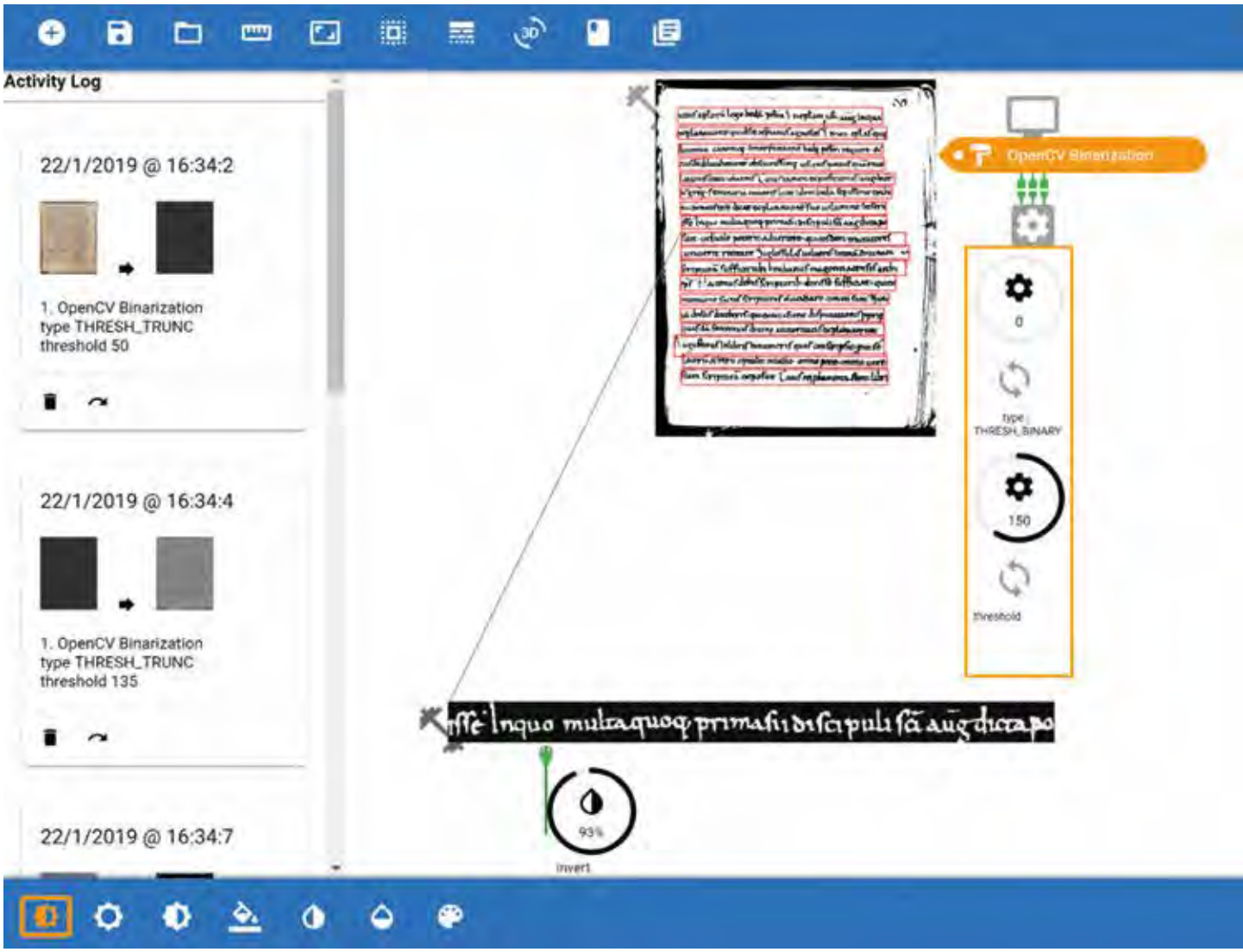


Fig. 6: AMAP Prototype.

into smaller, thus transparent and tractable, elements through visual programming in order to keep the user in the loop and in control – either during experiment-based configuration of processing chains for specific tasks or in the collaboration-driven design of workflows for solving scholarly problems of manuscript research in the humanities. The advantages of VP and its various features that enable the effective dismantling of the black box problem were elaborated in detail. Moreover, we embedded our design and realization approach in the broader context of software development methodology inspired also by design thinking and human computer interaction/communication. We finally introduced our iXMan_Lab concept and AMAP as its tool instance, a web-based prototype tool that attempts to deconstruct the black box by offering potential users from SFB sub-projects various DIA methods in a supportive VP environment. Our adoption of at least some of the principles of the VP paradigm to deconstruct the computational black box is the first step in the long journey ahead to completely eliminate the black box for the sake of truly inter-/transdisciplinary, effective and efficient computational manuscript research.

ACKNOWLEDGMENTS

This work was made possible by support from the Deutsche Forschungsgemeinschaft (DFG). 3M sponsored the necessary hardware for the Multi-Touch Table (MTT). We thank the Z03 team from the Centre for the Study of Manuscript Cultures at Universität Hamburg for providing the necessary support for our research. Andreas Fischer and Marcel Würsch from the University of Fribourg helped integrate DIVAServices into AMAP. The various collaborating SFB sub-projects, too, are to be mentioned here for providing the necessary data and sharing their domain knowledge. Last but not the least, we need to mention our various bachelor and masters students who have contributed to the iXMan_Lab.

REFERENCES

Blackwell, Alan F. (1996), 'Metacognitive Theories of Visual Programming: What Do We Think We are Doing?', *Proceedings IEEE Symposium on Visual Languages*, 240–246. doi: 10.1109/VL.1996.545293.

Burnett, Margaret (2002), 'Software Engineering for Visual Programming Languages', in Shi-Kuo Chang (ed.), *Handbook of Software Engineering and Knowledge Engineering*, vol. 2: *Emerging Technologies* (New Jersey: World Scientific), 77–92

Fraser, Neil (2015), 'Ten Things We've Learned from Blockly', *Proceedings IEEE Blocks and Beyond Workshop*, 49–50. doi: 10.1109/BLOCKS.2015.7369000.

Hassner, Tal, Malte Rehbein, Peter A. Stokes, and Lior Wolf (2015), 'Computation and Palaeography: Potentials and limits', in Oliver Duntze, Torsten Schaßan, and Gregor Vogeler (eds), *Kodikologie und Paläographie im Digitalen Zeitalter 3—Codicology and Palaeography in the Digital Age 3* (Norderstedt: Books on Demand), 1–27.

—— Robert Sablatnig, Dominique Stutzmann, and Ségolène Tarte (eds) (2014), *Digital Palaeography: New Machines and Old Texts (Dagstuhl Seminar 14302)*, (Dagstuhl Reports, 4.7). doi: 10.4230/DagRep.4.7.112.

Kasturi, Rangachar, Lawrence O'Gorman, and Venu Govindaraju (2002), 'Document Image Analysis: A Primer', *Sadhana*, 27.1: 3–22.

Mohammed, Hussein, Volker Märgner, Thomas Konidaris, and H. Siegfried Stiehl (2017), 'Normalised Local Naïve Bayes Nearest-Neighbour Classifier for Offline Writer Identification', *Proceedings 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 9, 1013–1018. doi: 10.1109/ICDAR.2017.168.

Myers, Brad A. (1990), 'Taxonomies of Visual Programming and Program Visualization', *Journal of Visual Languages & Computing*, 1.1: 97–123.

Rajan, Vinodh, and H. Siegfried Stiehl (2018a), 'Bringing Paleography to the Table: Developing an Interactive Manuscript Exploration System for Large Multi-Touch Devices', *13th IAPR International Workshop on Document Analysis Systems (DAS), Short Papers Booklet* (Vienna: Institute of Computer Aided Automation, Computer Vision Lab), [5–6].

—— and H. Siegfried Stiehl (2018b), 'From Eye-to-Eye to Hand-in-Hand: Collaborative Solution Building in Interdisciplinary Manuscript Research', in Manuel Burghardt and Claudia Müller-Birn (eds), *INF-DH-2018*, 1–4. doi: 10.18420/infdh2018-15

Smith, Ray (2007), 'An Overview of the Tesseract OCR Engine', *Proceedings Ninth International Conference on Document Analysis and Recognition (ICDAR)*, vol. 2, 629–633. doi: 10.1109/ICDAR.2007.4376991

Stokes, Peter (2012), 'Modeling Medieval Handwriting: A New Approach to Digital Palaeography', in Jan Christoph Meister et al. (eds), *DH2012 Book of Abstracts* (Hamburg: Alliance of Digital Humanities Organizations), 382–385. <http://www.dh2012.uni-hamburg.de/conference/programme/abstracts/modeling-medieval-handwriting-a-new-approach-to-digital-palaeography.1.html>

Tanimoto, Steven L. (1990), 'VIVA: A Visual Language for Image Processing', *Journal of Visual Languages & Computing*, 1.2: 127–139.

Wolber, David (2011), 'App Inventor and Real-World Motivation', *SIGCSE'11 Proceedings of the 42nd ACM Technical Symposium on Computer Science Education*, 601–606. doi: 10.1145/1953163.1953329

Würsch, Marcel, Rolf Ingold, and Marcus Liwicki (2016), 'DIVAServices — A RESTful Web Service for Document Image Analysis Methods', *Digital Scholarship in the Humanities*, 32.1: i150–i156.

Article

Legally Open: Copyright, Licensing, and Data Privacy Issues

Vanessa Hannesschläger | Vienna

Abstract

In the field of digital science, the methods and approaches of open science are gaining momentum. A vital precondition for applying these methods is knowledge of various aspects of the legal landscape, which this paper aims to address. Specificall , it will discuss the topics of copyright/ intellectual property rights (IPR) in an international context, the possibilities and pitfalls of open licensing and the legal restrictions brought about by the European Union's new data privacy legislation (GDPR).

1. Preliminary remark: disclaimer

Please note that this contribution is not professional legal advice, but a mere collection of thoughts and information compiled from legal sources.

Readers might be wondering why this is explicitly pointed out here, and they would be right to do so: in fact, it is not advisable to include a disclaimer in an online presence (or any sort of publication, really). The reason is that if rights are infringed, the existence of the disclaimer will not affect the judgement of or sentence resulting from the infringement. However, the disclaimer might inspire imposters or predatory companies that specialize in (justified or unjustified) intellectual property right claims to contact the disclaimant and lodge excessive claims. The best way to avoid IPR claims is to avoid using material whose rights holders are undetermined and to make sure that appropriate credit is always given (e.g. when using a licensed image, its license and all other information it requires the user to include should be explicitly stated).

2. Copyright/intellectual property rights (IPR)

The legal frameworks we are embedded in define if, how and how long texts, material, meta-/data and software can be made (and kept) available. The most relevant area of legislation affecting digital research, and especially computational manuscript research, is copyright.

In this context, researchers are always in a Janus-faced position: as creators of content, on the one hand, and as (re-) users of content, on the other. The two sides of the copyright coin are expressed by the *Universal Declaration of Human Rights* (UDHR), which states, 'Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits, but goes on to say that everyone also 'has the right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he [sic] is the author'.[1] It has even been argued that, due to 'the requirement for researchers to make their publicly funded work available to the public',[2] 'copyright is an unsuitable legal structure for scientific works. Scientifi norms guide scientists to reproduce and build on others' research, and default copyright law by its very purpose runs counter to these goals.'[3] Still, copyright law is a reality that contemporary research communities have to face. Especially for researchers working with computational methods, the 'increasingly rapid development of new media continuously leads to new and unanticipated ways of distributing copyrighted works'[4] — which affects researchers both as creators and as users. Knowledge of the basic aspects of national copyright law in the country where the research is carried out is vital in both roles, as it affects the possibilities of working with existing data, on the one hand, and the rights to distribute and be compensated for one's own work, on the other.

[1] UDHR, Art. 27 no. 1 and 2.

[2] Stodden 2009, 40.

[3] Stodden 2009, 35.

[4] Darling 2012, 485.

A second topic of interest in this context is the main principles of national and international copyright legislation. In a European context, employment of the term 'copyright' itself is already problematic, as it refers to a concept from the Anglo-American legal tradition: copyright primarily aims at regulating the right to replicate and reproduce. However, in most European countries, the Germanic legal tradition, which puts a stronger focus on the persona of the creator ('Urheberrecht', 'droit d'auteur'), has shaped 'copyright' legislations. Within (most of) Europe, it is therefore more accurate to speak of 'intellectual property rights' (IPR) rather than 'copyright' law. However: 'As IP law in the European Union is merely harmonized and not unified, the exact scope of copyright and similar rights may differ between Member States (e.g. some Member States recognize an exclusive right for 'scientific and critical editions', while others don't).'[5] It is crucial for researchers to be aware of this fact. However, it is also crucial to know that the EU has taken a first step in the direction of IP law harmonization by accepting the *Copyright in a Digital Single Market Directive* (COD) proposal, by means of which IPR is supposed to be modernized and adapted to the realities of the digital world throughout the EU. From the point of view of the research community, this proposal has its upsides and downsides; while it does aim to implement a general permission of text and data mining in a scientific context,[6] the articles on 'Protection of press publications concerning digital uses'[7] ('link tax article') and 'Use of protected content by information society service providers storing and giving access to large amounts of works and other subject-matter uploaded by their users'[8] ('upload filter article') are being discussed controversially. As the directive proposal was approved by the EU Parliament in September 2018, but has not yet been formalized, the effects on the legal conditions of research in Europe are not yet clear.

3. Open licensing

Due to the territorial limitations of copyright, the digital space that transcends national borders calls for new legal arrangements that are able to protect the researchers' rights, on the one hand, and ensure the reusability of their work, on the other. Open licensing models enable long-term preservation of and international research on data collected in local research projects, thus greatly supporting emerging open approaches in computational manuscript research. However, scholars often lack an overview of the various possibilities to license their findings. The most established model, which has gained great popularity for creative content and is increasingly also applied to research data, is *Creative Commons* licensing (CC). The use of widely known licenses such as the ones provided by CC is advisable (and preferable to writing one's own individual licenses) because it will enable others to understand the conditions under which material is available immediately (without having to read a complicated legal text). Although CC has become a de facto standard for licensing research data, scholars are often unaware of the details of the different CC modules and their consequences; choosing appropriate licenses for software is an even more complex task. Therefore, awareness of the available options of ready-made open licenses and their benefits and potential pitfalls and of license selection (such as license compatibility issues, copyright preconditions and other legal commitments such as work contracts) is crucial.

3.1 The Public License Selector

Creative Commons offers a basic license selection tool that is helpful for researchers who are already sure that a) their content is licensable under CC and b) they have made a conscious decision to use CC. However, in some cases, CC licenses might not be the best choice, for example in the case of code. A very nifty tool that helps select appropriate open licenses for both data and/or code is the *Public License Selector* developed by the European research infrastructure consortium CLARIN ERIC. Users start with a total selection of 22 open, publicly available, ready-made licenses and have to answer a sequence of questions. Each answer narrows down the licenses compatible with the respective preconditions, leaving the user with a final choice of open licenses suited to their specific situation (as well as further information about the individual qualities of all available licenses) at the end of the process.

4. Data privacy and the *General Data Protection Regulation* (GDPR)

As a third main topic, this paper addresses the EU's *General Data Protection Regulation* (GDPR), which came into effect on May 25, 2018. As the GDPR is a regulation, it took legal

[5] Kamocki, Stranák, and Sedlák 2016, 2534.

[6] COD, Art. 3.

[7] COD, Art. 11.

[8] COD, Art. 13.

effect in all EU member states immediately on the day of implementation (in contrast to a mere directive such as the COD, which has to be submitted to and approved by national legislatures before going into effect). The GDPR replaces the EU's *Personal Data Directive* (1995). Although the GDPR does not differ from the *Personal Data Directive* in terms of fundamental concepts, it does establish a few new requirements, as well as tangible punishments (penalties) in case of infringement. While the main aim of the GDPR is to protect citizens and individuals from abuse of their personal information by international corporations, it affects everyone working in digital space (despite several 'research exceptions' such as archiving in the public interest[9]). Hence, the main concepts of the GDPR are outlined below, explaining the most vital points to be considered in the context of computational manuscript research.

4.1 'Data subjects'

Although it might seem obvious to many, not everyone is aware of this basic, but crucial fact: the GDPR applies only to natural persons. This means that the GDPR does not apply to legal bodies (i.e. institutions), but only to real people. It also means that it applies only to living people, not historical ones. As manuscript researchers work with data about dead people more often than with data about living people, this might be a relief for some. However, while the GDPR does not protect data about dead people, other (national) laws might still restrict the collection and publication of data about dead people (this is especially true for relatively recently deceased persons).

The GDPR defines the rights that data subjects have with regard to their personal data. These include the right to information (e.g. about the data themselves, their processing and its purpose, their storage and its duration, their accessibility and their protection), the right to access the data (regardless of whether the data subject was the provider of the data or they were taken from elsewhere) and to rectify them if necessary, to restrict their processing or object to it and, most importantly, the right to erasure. These rights of the data subject can be restricted to a certain extent by the research exceptions defined by the GDPR. However, as this regulation is so relatively new and no court rulings about uncertainties in the GDPR are yet available, it is not advisable to build a research project that relies on these exceptions.

4.2 'Personal data'

According to the GDPR, 'personal data' is 'any information relating to an identified or identifiable natural person'.[10] This means any information about a person that might enable someone to identify that person. Examples include name, date of birth, age, sex, gender, address, pictures of the person, audio recordings of the person's voice etc. Be aware that a combination of data might make a person identifiable even if the person's name is not included in the data collection; e.g. the information that one of the data subjects is male, 80 years old and lives in a village with 300 inhabitants will likely make the data subject identifiable. Also be aware that the form or format of the data is irrelevant. All personal data are protected by the GDPR, even if it is only stored in handwriting on a slip of paper in a desk drawer.

Pseudonymised data (i.e. data that cannot be directly associated with an identifiable individual in their current state, but that could still be transformed into a state that makes it possible to reconstruct whom the data describes) is considered personal data by the GDPR. Anonymised data (i.e. data that cannot be associated with an identifiable individual in their current state, which must be irreversible) is not personal data according to the GDPR.

4.3 'Processing'

According to the GDPR, 'data processing' is 'any operation or set of operations which is performed on personal data'.[11] This means that any action done with data is part of the processing – even collecting, storing and deleting data. This is also true if the data is only in analogue form.

4.4 What does the GDPR require a researcher to do?

Researchers are likely to find themselves in the position of a 'data controller' (a person who 'determines the purposes and means of the processing of personal data'[12]) or a 'processor' (a person who 'processes personal data on behalf of the controller'[13]).

[9] See GDPR, Art. 6, Art. 89, and Rec. 50, Rec. 47, Rec. 113, Rec. 157, Rec. 159.

[10] GDPR, Art. 4 no. 1.

[11] GDPR, Art. 4 no. 2.

[12] GDPR, Art. 4 no. 7.

[13] GDPR, Art. 4 no. 8.

If work with personal data is carried out in a research project, researchers have to make sure to follow the data protection principles the GDPR defines. These include lawfulness (data subjects must give explicit consent to the collection and processing), fairness, transparency (data subjects must be informed about the ways and purposes of data collection and processing), purpose limitation (data can be collected and subsequently used only for purposes specified in advance), data minimization (only the data that are necessary for the defined purpose may be collected), storage limitation (data may be stored only as long as necessary for the defined purpose), accuracy, integrity, confidentiality and accountability (the data controller has to be able to demonstrate that all these principles are met). In addition, the data controller is obliged to protect the data and keep records of all processing activities.

5. Conclusion

Understanding the legal frameworks within which research operates is a vital skill for researchers in the digital age. Not only is it important to understand in what ways one is allowed to work with the findings and material provided by others: researchers also need to understand their own rights to their research in order to open it up for others. A good way to do that is to provide open licenses. In addition to considering intellectual property laws, awareness of data privacy legislation is vital in order to ensure that operations with information about users, data contributors and research subjects are carried out in a legal manner. While the EU has already provided a uniform framework for the latter, the development of the former will have to be observed on a European level over the next years.

REFERENCES

'Common Language Resources and Technology Infrastructure — European Research Infrastructure Consortium' (CLARIN ERIC) <https://www.clarin.eu/> (accessed September 26, 2018).

'Creative Commons' (CC) <https://creativecommons.org/> (accessed September 26, 2018).

Darling, Kate (2012), 'Contracting About the Future: Copyright and New Media', *Northwestern Journal of Technology and Intellectual Property*, 10.7: 485–530 <http://scholarlycommons.law.northwestern.edu/njtip/vol10/iss7/3> (accessed September 26, 2018).

'Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data' (Personal Data Directive) <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A31995L0046> (accessed September 26, 2018).

Kamocki, Paweł, Pavel Stranák and Michal Sedlák (2016), 'The Public License Selector: Making Open Licensing Easier', in Nicoletta Calzolari et al. (eds), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), Portorož, Slovenia*, eds. Calzolari, Nicoletta et al. (Paris: European Language Resources Association), 2533–2538 <http://www.lrec-conf.org/proceedings/lrec2016/pdf/880_Paper.pdf> (accessed September 26, 2018).

'Proposal for a Directive of the European Parliament and of the Council on copyright in the Digital Single Market COM/2016/0593' (Copyright Directive / COD) <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52016PC0593> (accessed September 26, 2018).

'Public License Selector' <https://ufal.github.io/public-license-selector/> (accessed September 26, 2018).

'Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC' (General Data Protection Regulation / GDPR) <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679> (accessed September 26, 2018).

Stodden, Victoria (2009), 'The legal framework for reproducible scientific research: Licensing and copyright', *IEEE Computing in Science and Engineering*, 11.1: 35–40 <https://web.stanford.edu/~vcs/papers/Legal-STODDEN2009.pdf> (accessed September 26, 2018).

'Universal Declaration of Human Rights' (UN General Assembly resolution 217 A / UDHR) <http://www.un.org/en/universal-declaration-human-rights/> (accessed September 26, 2018).

Article

A Comparison of Arabic Handwriting-Style Analysis Using Conventional and Computational Methods

Hussein Mohammed, Volker Märgner, and Tilman Seidensticker | Hamburg, Brunswick, Jena

1. Introduction: the conventional comparison of handwriting in Arabic manuscripts

In general, scholars working with manuscripts often have to decide whether two handwritten pages, quires or even whole codices were written by one and the same person. The identity of the scribe responsible for copying a manuscript or a large section of it may be known in many instances, but this is not always the case. Determining the identity of particular hands allows us to establish when the undated manuscript was written, how it should be contextualised and whether or not it is an autograph, for example – in other words, questions constitutive for the history of text transmission.

To establish or rule out the identity of particular hands, several characteristic features of the writing are compared, in particular its layout (i.e. visual organisation), the general impression of the writing on a page, and the shape of individual graphemes or groups of them. It has to be borne in mind that one and the same scribe may have written texts in various ways, depending on the purpose of the copy, time of day or his/her age and experience. Taking Arabic manuscripts as an example, there is an established distinction between copies made for personal use (*musawwada*, 'draft') and 'clean' copies (*mubayyaḍa*) with a high degree of legibility; there are considerable differences between these two forms of writing, which serve different purposes. Furthermore, a single scribe can write a grapheme in different ways in a text, while different scribes belonging to a common school of writing or following the same writing tradition may write texts in a way that is so similar that the individual hands are hard to distinguish.

The analysis of single graphemes or groups of them yields strong arguments for or against particular hands because this approach allows visual evidence to be produced on the basis of examples of actual writing. Nevertheless, the results are tempered by a certain amount of subjectivity because even a single hand may show a considerable amount of variation over time. If digital image processing can deliver additional plausibility, then this technique is worth using as well. Our article will show the reader how an Arabist's traditional approach to handwriting analysis and the results yielded by a computational method without applying OCR using the Handwriting Analysis Tool (HAT-2 tool) complement one another.

An Arabist's approach will naturally take the general impression of handwriting into consideration, but this has often proved difficult to verbalise. If the scholar wishes to convince others about his theories, it will be easier for him to concentrate on a comparison of single graphemes or ligatures of them. Stating similarities or dissimilarities allows hypotheses to be made about hands' identities.

The handwritten texts on which the dual approach will be tested consist of thirteen 'audience certificates (Arabic *samāʿ*, plural *samāʿāt*) in an Arabic manuscript from the Gotha Research Library (Forschungsbibliothek Gotha) with the shelf mark Ms. orient. A 627 (fols 13b–15b and fol. 37b). An audience certificate, as it is known in English, is a kind of short paratext peculiar to Islamic-Arabic manuscript culture that was particularly used between the twelfth and fifteenth century CE. It contains information on reading sessions during which the text contained in a manuscript was read aloud. In these sessions, short texts or chapters of books on a wide range of topics were read out to small groups of listeners (they could also be read to audiences that were quite large). Five points were especially important in a record of this kind: (i) the name of the attending authority, (ii) the names of the members of the audience, (iii) the name of the reader (if he was not the attending authority), (iv) the name of the 'writer', i.e. the person who was entrusted with writing down the attendees' names[1], and (v) the date of the session.[2]

[1] This office was called 'clerk' in Seidensticker 2015, but here and in Seidensticker 2020 the designation has been changed to 'writer' instead, taking established terminology into consideration. 'Writer' is a literal translation of the Arabic word *kātib* and has been used in quotation marks in this article.

[2] 'The aim of these paratexts is primarily to attest: 1) the participation of either auditors, which entitles them to further transmission, or non-scholars

Why is a comparison of handwriting desirable in the case of our examples? There are three reasons for this:

1) The order of the certificates dates is not a chronological one,[3] although they were certainly written down in chronological order on each page originally, from top to bottom. This mix-up can partly be explained by mistakes that were made when the manuscript was rebound.

2) The (dis)order of certificates #5, 7 and 12, all of which are written at the bottom of the respective pages, which can be explained by the fact that Arabic audience certificates were sometimes copied from other manuscripts containing the same text, which means that they do not necessarily bear witness to a reading from the manuscript to which they were transferred. Although the names of the 'writers' who wrote down the original certificate were copied as well in many instances, the transferred certificates were not written by them, of course, but by the transferring scribes.[4] (Since it is not always explicitly mentioned that a certificate was transferred, the given dates and names generally need to be handled with care when attempting to date a manuscript based on such certificates.)

3) The writer's name is not mentioned in certificates #1 and 13. A comparison of the handwriting may at least show us whether or not the writers were one and the same person. (For a more detailed explanation, see Seidensticker 2015.) This complicated situation can be illustrated by looking at fol. 14a of the Gotha codex (Fig. 1).

The first five lines at the top contain the second half of certificate #1. Certificates #2, 3, 4 and 5 all follow underneath, separated from each other by black lines drawn as borders. The first four certificates are dated in chronological order: Rabīʿ I 486 (= April 1093), Rabīʿ I 487 (= March/April 1094), 6 Ramaḍān 487 (= 19 September 1094) and 27 Shawwāl 490 (= 7 October 1097). Certificat #5 has Dū l-Ḥijja 488 (= December 1095) as its date, which would put it between #3 and #4 chronologically. Neither the writer nor a scribe who might have transferred this certificate is mentioned in #5. As it is plausible that a strict chronological order was originally observed, this irregularity needs to be explained. This was done by comparing the handwriting in each certificate

to record their presence as a pious work; and 2) the correctness of the text transmitted in the manuscript'; Seidensticker 2020, 54.

[3] See table 1 in Seidensticker 2015, 80.

[4] On this habit of transferring, cf. Seidensticker 2015, 80–83 and Seidensticker 2020, section 9.

As a comparison of the attendees has shown, all the certificate but #7 must be considered as pairs because they have a common stock of listeners. This pairwise structure is due to the fact that they refer to readings of two different parts of the book contained in the Gotha manuscript, namely part 6 and part 8. The certificates chronological order is presented in the following table, which puts the certificates added later at the end:

Table 1: The original chronological order of the certifi ates.

Referring to part 6	Referring to part 8
#13	#1
#8	#2
#9	#3
#10	#4
#11	#6
#12 (transferred)	#5 (?)
	#7 (transferred)

Certificate #5 belongs to its counterpart, #12, which was explicitly transferred from another manuscript. Its un-chronological position is a hint that it was transferred, too. To confirm this hypothesis, its writing style was compared to that of #12 and that of #7, which is the second one that was explicitly transferred. One text element was particularly suitable for this purpose: the title and name of the attending authority, which is always mentioned near the beginning of the certific tes, viz. al-Ḥājib Abī l-Ḥasan ʿAlī Ibn Muḥammad Ibn ʿAlī al-ʿAllāf. Figure 2 shows what the names look like in #5, 7 and 12.

In Seidensticker 2015, the writing styles[5] of #7 and #12 were considered to convey an impression of far-reaching agreement in a hand noticeably inclined to the right, and concerning #5, it was stated that 'it seems to be written by the same hand' (p. 82). This analysis appeared to be adequate at the time because there was enough visual evidence provided by the other pairs that showed the same graphemes or ligatures in distinct styles. Some additional comments are made here that strengthen our confidence in the initial assessment:

[5] In this article, '(writing) style' is used to designate the script in a single certificate with all its individual traits

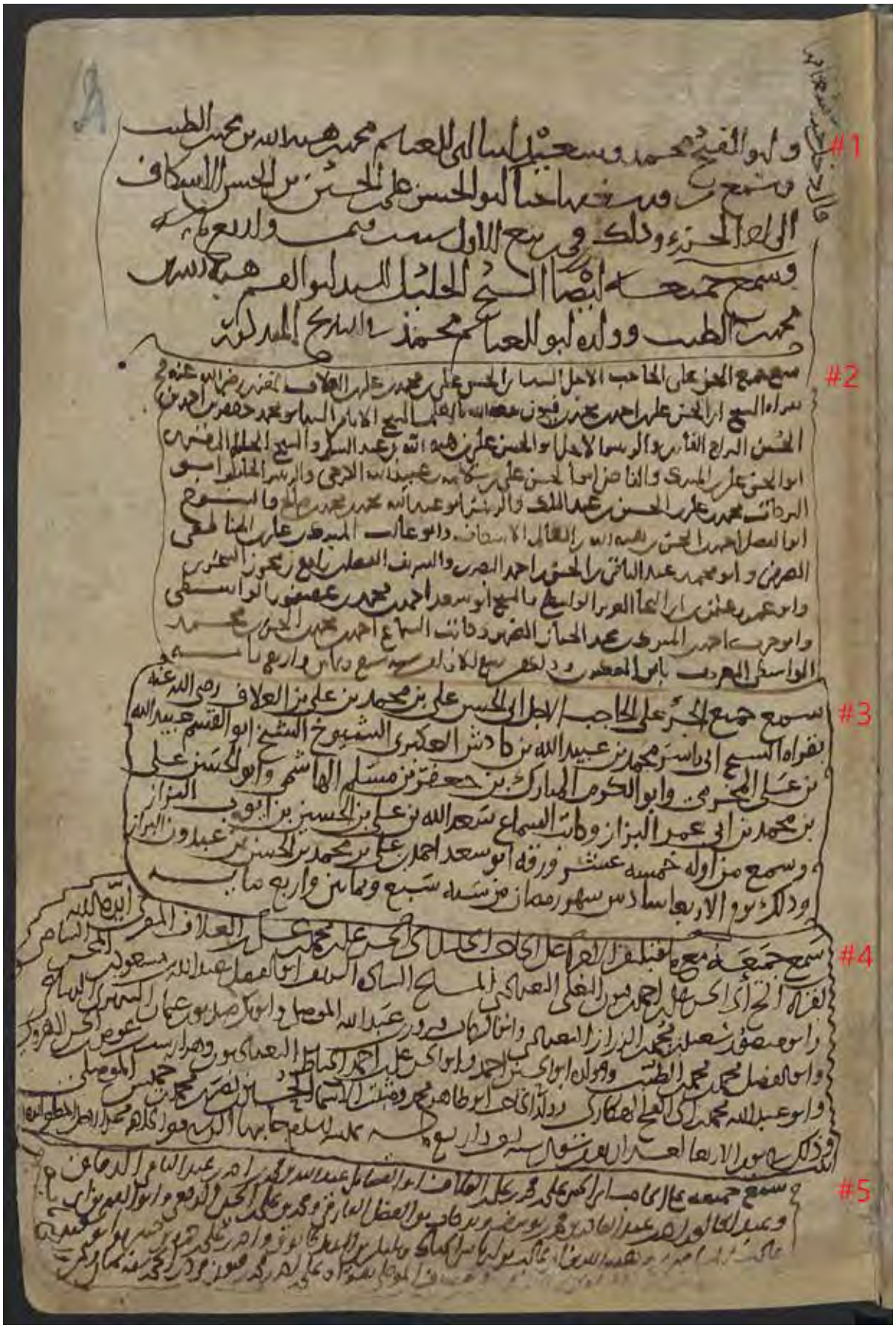
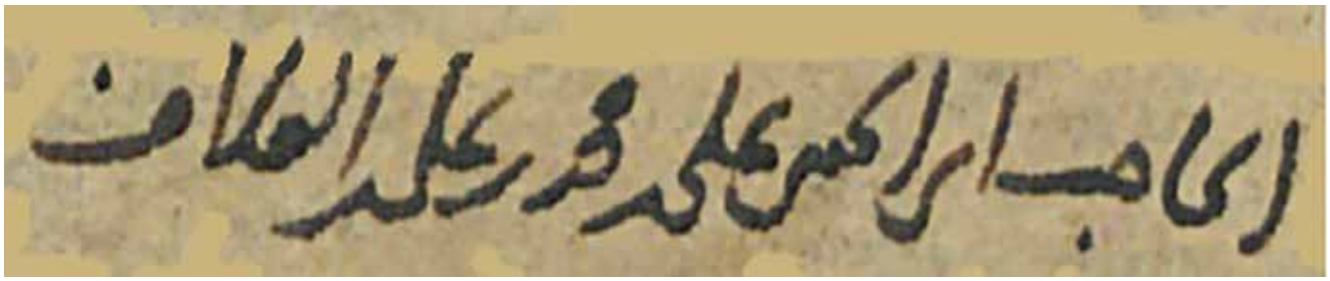
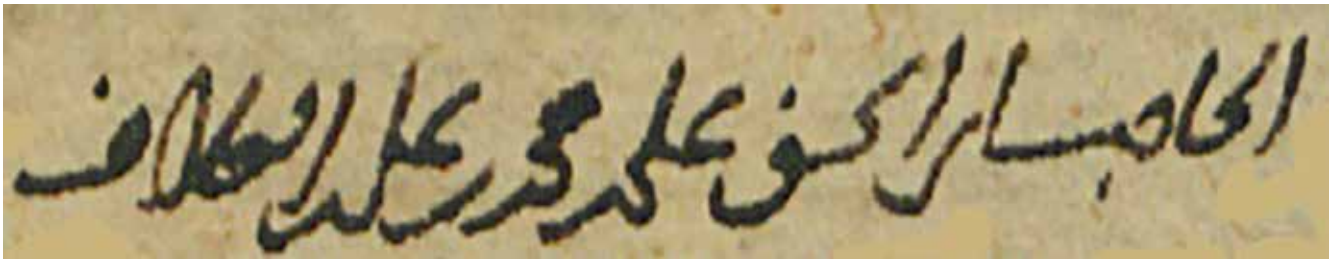


Fig. 1: Fol. 14ª of Gotha, Ms. orient. A 627, showing certifi ates #1 (second part) to #5 from top to bottom.

#5



#7



#12

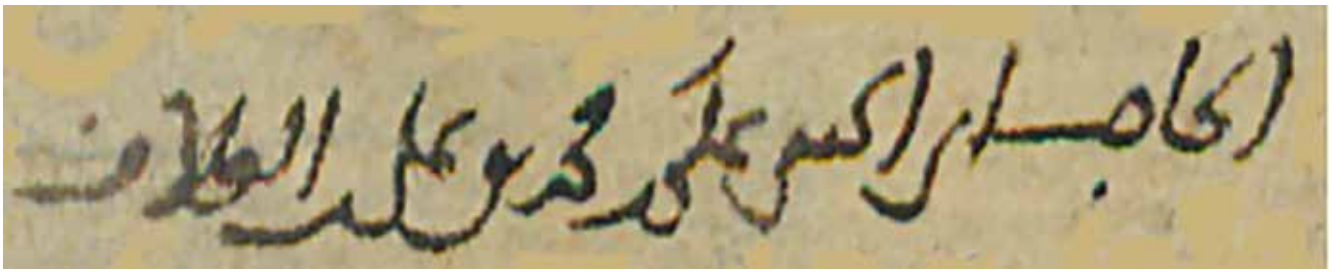


Fig. 2: Enlarged details showing the attending authority's name.

1) The first letter (on the right-hand side), an *alif*, is slightly curved to the left at the bottom in each case;

2) the ligature *Muḥammad* (left of centre) shows a high amount of agreement, especially between #5 and #12;

3) the final letter *fāʾ* (on the left-hand side) is written without an open loop, the dot normally placed above it is shifted towards the left, and its tail is falling and does not rise again at the end.

These points made it seem plausible that certificate #5 was transferred (in other words, that the audience mentioned in it did not hear the book's text being read out *from the Gotha manuscript*) and that it was transferred by the same person as the one mentioned in #7 and #12 as the transferring scribe. Other similar insights on other certificates or pairs of them have been gained with the help of the same kind of comparison.

2. Basic ideas behind the HAT-2 handwriting analysis tool

The basic ideas behind the HAT-2 software tool are presented here without any mathematical equations using sample images to illustrate the method.[6] The main aspects covered by the tool are as follows

- detection of discriminative features in handwriting styles using keypoint detection algorithms
- description of detected features as numerical vectors
- application of probabilistic and statistical approaches in order to measure the similarity between different handwriting styles
- calculation of scores for each handwriting style based on the measured similarities. These score values are presented in an intuitive way in order to facilitate the comparison between handwriting styles.

[6] For more details of the HAT-2 tool and the theory behind the approach implemented in it, see Mohammed 2017 and Mohammed et al. 2017.

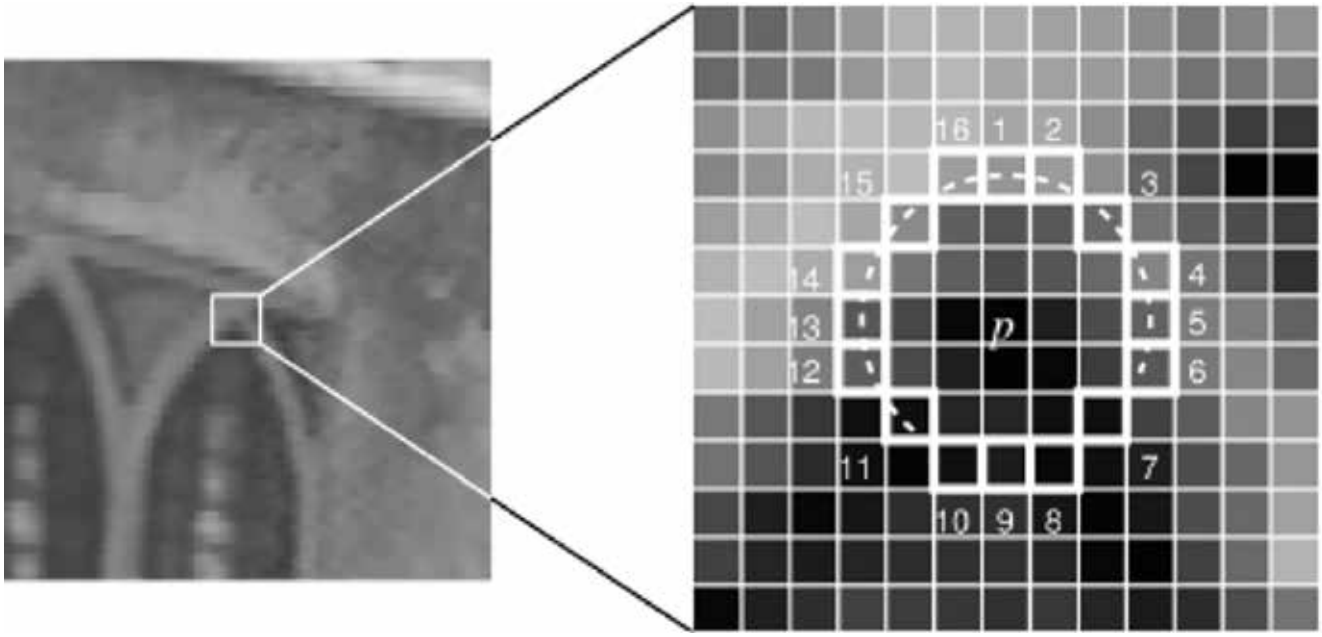


Fig. 3: A small part of an image is shown on the left. On the right, an enlarged section of that is shown with a pixel in the middle (p) and 16 pixels surrounding it, which are used to calculate the strength of the potential keypoint (reproduced from Rosten et al. 2010).

2.1 Keypoint detection

Keypoint detection is a process of detecting and localising salient features in an image. In our particular case, we needed to detect and localise discriminative features in handwriting styles. We followed the recommendation provided in the software manual of HAT-2 (on page 5: Settings; see Mohammed 2017). The FAST keypoint detection algorithm described in Rosten et al. 2010 was used for the tests in this research. A circular neighbourhood of 16 pixels around every pixel p in the image is considered for keypoint detection in the FAST algorithm (see Fig. 3); 'p' is classified as a keypoint if there are nine contiguous pixels in the surrounding 16 pixels (white squares) of the discrete circle with a grey value either larger or smaller than the value of the centre pixel, p. The bigger the difference is between the grey values of all nine pixels and the central pixel, the stronger the keypoint is.

Fig. 4 shows a sample certificate at the top, below which there is an enlarged part of the certificate with keypoints marked by circles in different colours for better distinction between individual keypoints. A smaller part of the middle image (after transformation of the colour to a grey-scale image) with a single keypoint in the central position is shown at the bottom. The numbers represent the grey values of the pixels, which are used to calculate the gradients' magnitude and orientation. It is important to mention that the number of detected keypoints in each certificate is in the order of thousands, and most keypoints are located on the contour of the ink-trace (as can be seen in Fig. 4). A large number of keypoints (ranging from 1,812 to 8,267) were detected in the certificates for just a few lines of text. This high number of detected features enables a reliable statistical estimation to be made about the similarities between different writing styles.

2.2 Discriminative features

The feature vector calculated at each keypoint is the SIFT descriptor. Details about this descriptor are given in Lowe 2004. A brief description of the computation of a descriptor is as follows. The magnitude and orientation of the grey-value gradients are calculated at each pixel in a 16 × 16 square neighbourhood around every detected keypoint. Fig. 5 shows these two values in the left part as the length and orientation of the black arrows at the position of a pixel (a small green square). For each group of 4 × 4 pixels (marked by thick green lines), the magnitudes of each of the eight orientations are summed and grouped in 16 4 × 4 squares, as illustrated in Fig. 5 on the right. Finally, the calculated gradient magnitudes are ordered in a 128-element vector based on their calculated orientation and position. A vector of this kind (called a keypoint descriptor or just descriptor for short) describes the features at every position of a keypoint as a vector of numerical values.

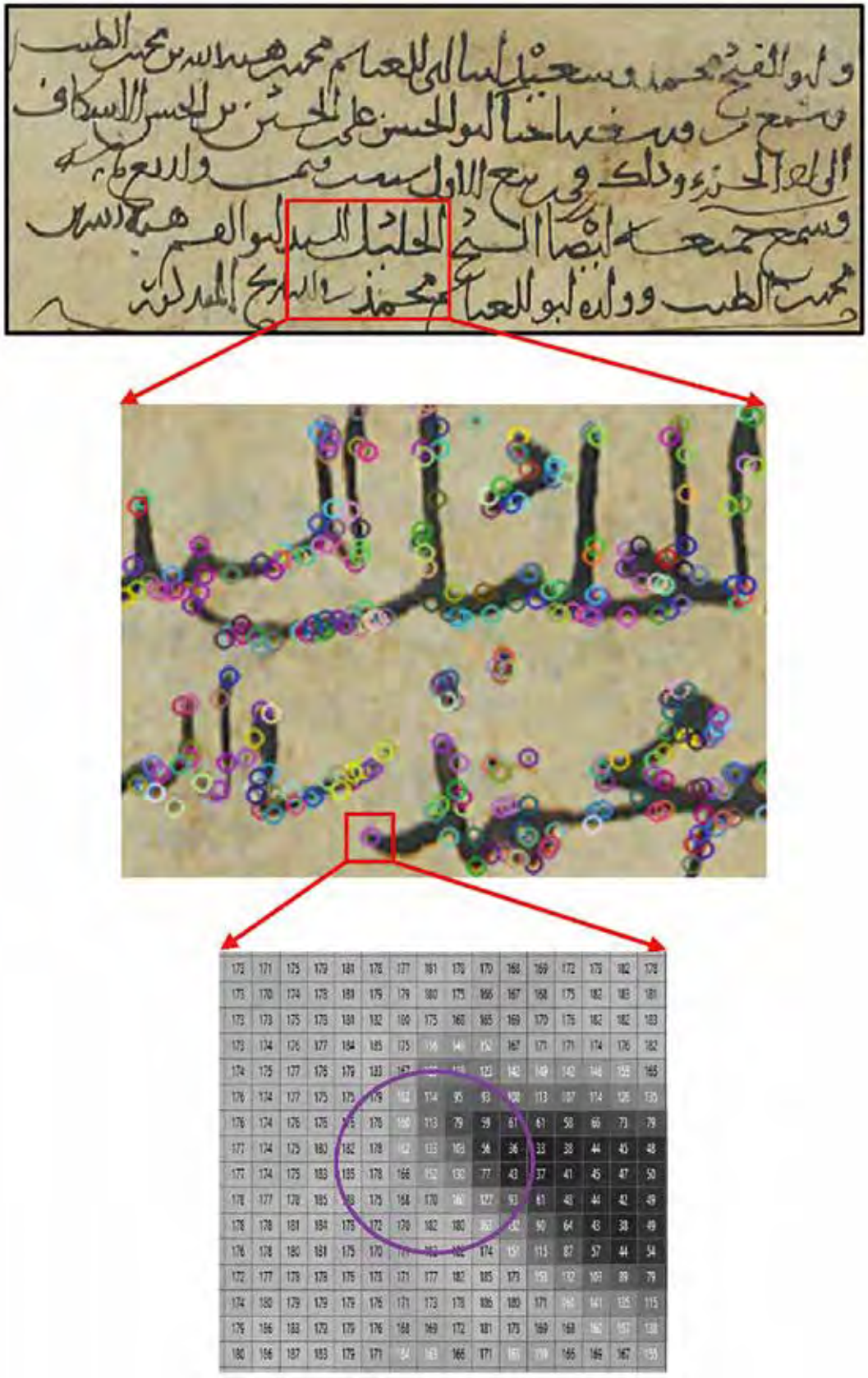


Fig. 4: Visualisation of the keypoints on a sample image of a certifi ate. At the top a certifi ate, in the middle keypoints on the writing shown as circles of different colours, and at the bottom the 16 × 16 pixel surroundings of a keypoint after colour transformation to grey with the grey values shown.

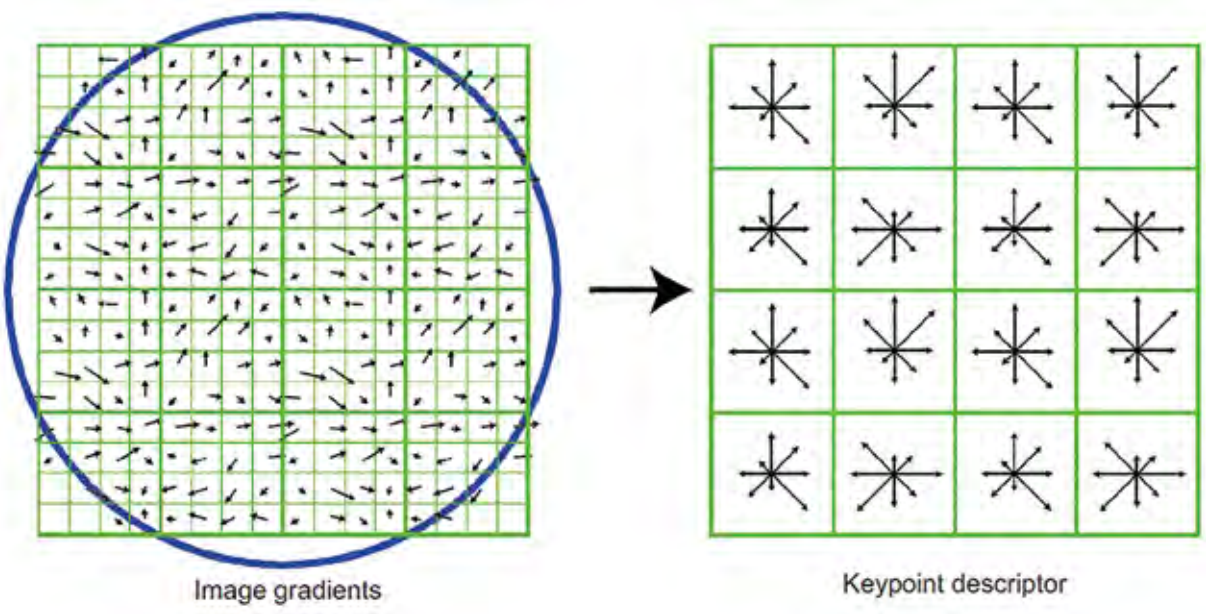


Fig. 5: Graphical representation of the computation of a descriptor at a keypoint position (reproduced from Lowe 2004 with some modifi ations).

2.3 Similarity measurement

Measurements of the similarity between a particular manuscript image and manuscript images of the same style or different styles are performed in the following way:

Keypoints are detected and corresponding descriptors are calculated for all the manuscripts' pages. The similarities between the descriptors of all the manuscript images are calculated using a probabilistic approach called Normalised Local NBNN.[7] The similarities between the handwriting styles are measured by accumulating all the calculated differences between descriptors within each of the manuscript images. Finally, similarity scores are calculated on the basis of the similarities that have been measured. These score values are presented as relative similarities in an intuitive way to facilitate the comparison between handwriting styles. The various styles are then sorted in descending order of their similarity scores. The handwriting style with the highest similarity score is considered the most similar handwriting style to the manuscript in question.

This simplified description of a complex distance measurement only presents the basic idea; more mathematical details can be found in Mohammed et al. 2018.

Table 2: Results of a conventional comparison of handwriting.

Referring to part 6	Referring to part 8	'Writer' or transferring scribe	Result of conventional analysis
#13	#1	Not mentioned	Probably written by the same hand
#8	#2	Identical 'writer'	Identical styles
#9	#3	Different 'writers'	Different styles
#10	#4	Identical 'writers'	Identical styles
#11	#6	Identical 'writers'	Identical styles
#12	#5	Transferring scribe of #12: Ibn Yūsuf	Identical styles
	#7	Transferring scribe: Ibn Yūsuf	Style identical to #12 and #5

[7] See Mohammed et al. 2017 for details.

Table 3: Scores calculated for each certifi ate (style). Each column shows the style on top and all scores calculated for each of the remaining styles.

Style #1			Style #2			Style #3			Style #4			Style #5		
Rank	Style	Score	Rank	Style	Score	Rank	Style	Score	Rank	Style	Score	Rank	Style	Score
1	#13	14.62	1	#8	26.04	1	#4	10.43	1	#10	20.04	1	#12	25.82
2	#9	10.63	2	#3	9.54	2	#13	10.19	2	#3	10.04	2	#7	18.28
3	#3	9.93	3	#4	8.44	3	#1	10.15	3	#13	8.40	3	#10	7.30
4	#11	8.75	4	#1	7.99	4	#10	9.37	4	#8	8.24	4	#8	6.74
5	#8	8.69	5	#10	7.97	5	#2	8.80	5	#1	8.19	5	#3	6.46
6	#4	8.42	6	#13	7.35	6	#9	8.41	6	#6	7.63	6	#4	6.19
7	#6	8.10	7	#5	7.26	7	#8	8.31	7	#2	7.57	7	#13	5.71
8	#10	7.98	8	#9	6.66	8	#12	7.75	8	#5	7.18	8	#9	5.35
9	#2	7.00	9	#12	5.19	9	#5	7.61	9	#7	6.30	9	#6	4.91
10	#7	5.53	10	#7	4.97	10	#7	6.97	10	#12	6.19	10	#2	4.90
11	#5	5.22	11	#6	4.63	11	#6	6.79	11	#11	5.24	11	#1	4.28
12	#12	5.12	12	#11	3.97	12	#11	5.23	12	#9	4.97	12	#11	4.05

Style #6			Style #7			Style #8			Style #9			Style #10		
Rank	Style	Score	Rank	Style	Score	Rank	Style	Score	Rank	Style	Score	Rank	Style	Score
1	#11	17.38	1	#5	27.41	1	#2	23.79	1	#11	12.74	1	#4	20.51
2	#9	9.27	2	#12	22.82	2	#13	11.05	2	#1	11.88	2	#13	8.90
3	#10	9.15	3	#10	6.33	3	#3	9.63	3	#13	9.98	3	#8	8.60
4	#13	9.08	4	#3	5.65	4	#10	9.28	4	#6	9.27	4	#3	8.13
5	#4	8.78	5	#6	5.46	5	#4	8.36	5	#10	8.89	5	#5	7.74
6	#1	8.74	6	#13	5.33	6	#1	7.08	6	#3	8.64	6	#12	7.50
7	#12	7.57	7	#4	5.15	7	#5	6.94	7	#2	7.14	7	#6	7.13
8	#3	7.09	8	#1	4.65	8	#12	5.64	8	#8	6.99	8	#1	6.82
9	#5	7.06	9	#11	4.53	9	#9	5.45	9	#4	6.66	9	#7	6.48
10	#7	6.52	10	#9	4.51	10	#7	4.57	10	#7	6.57	10	#2	6.43
11	#8	5.20	11	#8	4.34	11	#6	4.47	11	#5	6.55	11	#11	6.08
12	#2	4.16	12	#2	3.80	12	#11	3.73	12	#12	4.69	12	#9	5.69

3. Comparison of results

Table 2 shows a simplified list of the results obtained by conventional analysis in Seidensticker 2015. As in Table 1, the sequence is according to the reconstructed chronology of the certificates, not according to their order in the manuscript

To use the HAT-2 tool, each handwritten paragraph is cropped in order to remove any frame lines that are present, which cannot reasonably be analysed as script. The cut-outs are considered to be written by a single writer and are given the same numerical label from style #1 to style #13 as in the conventional comparison. Then the similarity between all the handwriting styles is measured using HAT-2 in order to see if any of the styles are similar, which might indicate that they belong to the same scribe.[8]

Table 3 shows the scores calculated by the HAT-2 tool. The handwriting style of each audience certificate in this particular research case was analysed and compared to the handwriting styles in the remaining certificates. The similarity scores were calculated in each iteration (for each questioned manuscript) with respect to the other certificates, as can be seen in each column. The scores in a column always add up to 100. This means that the score is a relative value, which shows the similarity of a given writing style in a particular manuscript in relation to all the other handwriting styles in that column.

Let us have a look at the scores now in the event of it being likely that the hands are identical, as in #11 and #6. The column for style #6 shows us that style #11 is the most similar one, with a score of 17.38, followed by a big difference to style #9, with a score of just 9.27. The difference between the scores of style #11 and style #9 is quite high (8.11). All the other remaining styles have small differences in their similarity scores. The only big difference in this column is between the first and the second score

Now let us check the other direction. What are the resultant scores if style #11 is the 'base' one? The column of style #11 shows us that style #6 is the most similar style, with a score of 18.42, followed by style #9 with a score of 13.60. In this case, the difference between the scores for style #6 and style #9 in this column is lower (4.82). For style #11, we can see that the similarity to style #6 is higher (18.42) than the similarity between style #6 and style #11 (17.38), but the distance to the following style, #9, is smaller in both cases. This shows us on the one hand that style #11 and style #6 are similar compared to all the other writing styles. On the other hand, we can see that style #9 is more similar to style #11 than to style #6 due to the smaller difference in the scores, in both cases relative to the remaining styles.

Another set of certificates with a striking similarity according to traditional analysis is made up of #5, 7 and 12. Style #5 displays a very high similarity score of 25.82 in relation to style #12, but also a relatively high score of 18.28 for style #7, while the next style, #10, has a score of just 7.30, which is very close to all the remaining scores. Interestingly, the scores for style #7 and style #12 confirm this result. In all three cases, the scores for the first and second style in each column are quite high, whereas the difference in the scores between the second and third line is 10.98 in the case of style #5, 16.49 for style #7 and 11.48 for style #12.

Styles #1 and #13 are a pair with less similarity, which has led Tilman Seidensticker to the tentative conclusion that they were only

[8] Mohammed 2019.

Table 3 continued.

Style #11			Style #12			Style #13		
Rank	Style	Score	Rank	Style	Score	Rank	Style	Score
1	#6	18.42	1	#5	31.86	1	#1	13.99
2	#9	13.60	2	#7	18.42	2	#8	11.60
3	#1	10.87	3	#10	6.94	3	#3	9.87
4	#13	9.81	4	#3	6.51	4	#10	9.28
5	#10	9.05	5	#6	5.69	5	#4	8.26
6	#4	6.37	6	#13	5.55	6	#9	8.23
7	#7	6.04	7	#4	5.40	7	#6	7.76
8	#5	5.86	8	#8	4.83	8	#11	7.34
9	#12	5.60	9	#1	4.19	9	#5	6.70
10	#3	5.45	10	#11	3.98	10	#12	6.06
11	#8	4.83	11	#9	3.53	11	#7	5.52
12	#2	4.11	12	#2	3.11	12	#2	5.40

'probably' written by the same person. The scores confir these doubts: for style #1, the highest similarity score is 14.62 with regard to style #13, which is quite close to the following score of 10.63 for style #9 (a difference of 3.99), and all the remaining scores are very similar. If we compare the scores for style #13, then once again, style #1 has the highest score, but the next one is style #8 with a score difference of 2.39, which is only a small difference, and all the remaining scores are very close to each other again. These results show there is quite a low probability of the hands being the same, which is a good reason to re-evaluate the conventional handwriting comparison again using a bigger basis than just two pairs of details (as was the case in Seidensticker 2015, 81).

Finally, let us also take a look at certificates #9 and #3, which were probably written by two different writers. The visual examination based on three pairs of manuscripts claims there is evidence of two different hands. The resultant scores for style #3 and style #9 using the HAT-2 tool are interesting as there is no clear maximum when comparing the scores, which tells us that none of the remaining styles are similar to style #3 or to style #9. The result of using the HAT-2 tool confirms the visual examination

4. Conclusion

We regard the two approaches as complementary and their interpretation of the data provided by the HAT-2 tool as provisional. One important result is that the information provided in the certificates was largely confirmed and, in the case of certificate #5, it was extended as well.[9] Further research will hopefully show whether the low degree of keypoint similarity between #1 and #13 could be due to the faded writing in #13. Alternatively, the noticeable differences in individual graphemes must be taken into account, which may be the result of higher writing speed or varying performance by the 'writer' over a period of time. In either case, the convergence between the traditional and digital approach is surprisingly high, which is an incentive for conducting future research on the matter.

[9] 'As a result, we can state with a high degree of probability that 1) the handwriting is identical whenever the clerk or copyist has the same name in the text; 2) the handwriting is different whenever different clerks are mentioned; 3) although neither the clerk nor the copyist is mentioned in three cases (nos 1, 5 and 13), the identity of hands 1 and 13 and of hands 5, 7 and 12 is highly plausible. This latter observation makes it also almost certain that no. 5 is a transferred *samā'*, too' (Seidensticker 2015, 83a).

REFERENCES

Lowe, David, G. (2004), 'Distinctive Image Features from Scale-invariant Keypoints', *International Journal of Computer Vision*, 60/2: 91–110.

Mohammed, Hussein A. (2017), 'Handwriting Analysis Tool v2.0 (HAT-2)', <http://doi.org/10.25592/uhhfdm.900>.

—— (2019), *Computational Analysis of Writing Style in Digitised Manuscripts*, PhD thesis, <http://ediss.sub.uni-hamburg.de/volltexte/2019/9730>.

—— et al. (2017), 'Normalised local naive Bayes nearest neighbour classifier for offline writer identification', in *14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 1, 1013–1018. IEEE, Nov. 2017.

—— et al. (2018), 'Writer Identification for Historical Manuscripts: Analysis and Optimisation of a Classifier as an Easy-to-use Tool for Scholars from the Humanities', in *16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, 534–539. IEEE, Dec. 2018.

Rosten, Edward et al. (2010), 'Faster and better: A Machine Learning Approach to Corner Detection', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32/1: 105–119.

Seidensticker, Tilman (2015), 'Audience Certificates in Arabic Manuscripts – the Genre and a Case Study', *manuscript cultures*, 8: 75–91.

—— (2020), 'The Chamberlain's Sessions: Audience Certificates in a Baghdad Manuscript of al-Ḫarā'iṭī's *Iʿtilāl al-qulūb* (Forschungsbibliothek Gotha, Ms. orient. A 627)', *Journal of Islamic Manuscripts*, 11: 53–100.

Article

Illuminating Techniques from the Sinai Desert: Spectral Imaging Inside one of the Oldest Libraries in the World Brings back to Life Re-written and Forgotten Texts. Review of the Imaging Techniques and the Global Advantages on the Field that were Gained from the Sinai Palimpsests Project

Damianos Kasotakis, Michael Phelps, Ken Boydston | Early Manuscripts Electronic Library, California – MegaVision, California

Understanding the location of the very important collection of Saint Catherine's is the first step towards building a framework for the successful digitization of its collection. The Sinai desert is a low humidity environment with granite mountains surrounding the old monastery in this remote area of the Sinai Peninsula. On a first look it appears as an unfriendly place for humans to survive, but the history of the 1500 years old Justinian fortress tells us otherwise. By a closer observation of the monastery's library we fin treasures hidden from the outside world surviving alongside the monks for centuries. Among those treasures we find the palimpsest manuscripts, which are manuscripts that have been erased and re-written, sometimes with more than one layer of script. During the last century there have been more attempts to make the erased text legible with various means. With the Sinai Palimpsests Project, it has been the first time that the palimpsest collection was systematically digitally imaged with non-destructive techniques.

The Sinai Palimpsests Project began identifying technologies for the recovery of those erased texts in 2009 as pilot stage. In December 2011 the project began systematically imaging palimpsests in the Library area, and was concluded in 2016 after five years, with the unreserved support and understanding of the monks of the Sinai Brotherhood, His Eminence the Archbishop, the Librarian Father Justin, and the support of the Egyptian Government.

110 palimpsested manuscripts were know until then, but were not cataloged and studied page by page. After five years, we knew of 160 palimpsest manuscripts, and it is possible that the number may approach 200 if the entire collection of the Monastery is examined page-by-page. During our work and sometimes by mistake due to a simple misread of a manuscript number, we discovered other new palimpsests that were not known until then, with the tools we had developed we were able to catalog those manuscripts as palimpsests in a new digital database.

By August 2016 when the imaging project concluded, 74 palimpsests and the total of 6800 pages had been cataloged and spectrally imaged, resulting in more than 50 TB of data including the processed images.

Metadata describing the manuscript to be imaged are recorded before imaging. A codicological description of the manuscript as a whole was recorded, and then a database entry was created for each page or other manuscript component, including the categorization of each manuscript page by language. Features of the folio are recorded such as: the flesh or hair side, the color of the ink, the attributes of the script, as well as other observations that may be useful during the photography or image processing of the texts. For example, knowing that a page was a flesh side is important for the image processing team to know if good results can be extracted by using transmissive illumination images.

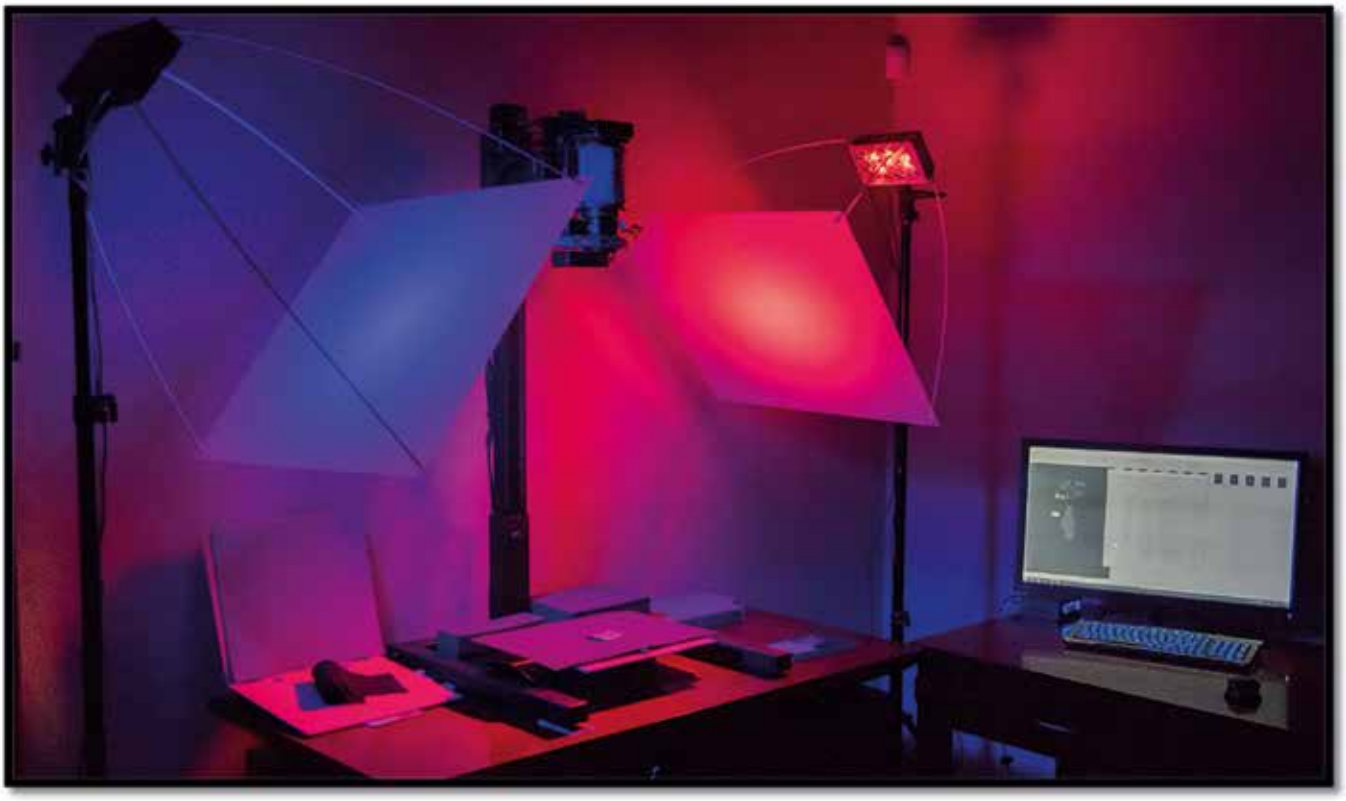


Fig. 1: MegaVision Spectral Imaging System.

Our database called KatIkon was created specifically for the Sinai Palimpsests Project. This database allows us to create a Shotlist that acts as a guidance tool for the photography team. This Shotlist includes not just the order of folios that have to be imaged, but all the necessary metadata for each manuscript folio, so that this information can be embedded during the imaging sequence while the image files are created. It's important to highlight the significance of embedding metadata during the capture process that follows the object with systematic checks of integrity along the data flo .

During the spectral imaging process, we acquire the basic foundation for the image processing of palimpsest texts. It is here that we create the raw-digital-material called RAW photos.

A general setup of the Mega Vision Spectral Imaging System in a copy stand type can be seen in Fig. 1. Visible in Fig. 1 are the LED light sources, the diffusers that help us have an even distribution of light on our scene, the E7 digital back, the filter wheel, the lens, and the capture computer controlling all of the above.

The optical unit of the imaging system is a custom designed 120mm apochromatic lens. This apochromatic lens is corrected in three wavelengths (UV-IR-VIS). Thus we minimize the magnification and focus changes between the UV-IR and VIS spectrum, a change that would introduce the common problem of miss-registration of data between fluorescent and reflectance shots or any need to refocus between IR and UV. The lens is a custom-built lens that was made specifically with multispectral applications in mind and uses a central, Copal 0, Schneider electronic leaf shutter. The lens design utilizes nine elements in seven groups. Six of them are special low dispersion (ED) glass, and the rear element is aspheric.

During our spectral photography in the Sinai, we recorded 33 different black-and-white photos for each manuscript page, using a 50.1 megapixels camera manufactured by MegaVision Inc. that houses an achromatic, silicon, Kodak CCD Imaging Sensor (KAF-50100) with 8176(H) x 6132 (V) pixels. Advantages of using a CCD array vs a CMOS one is seen in the near infrared spectrum where the CCD is slightly more sensitive. More specifically the spectral range of the sensor is between 350-1100nm.

RAW photos are digital files that have not been processed at all, they are uncompressed, unpacked, weight 100 megabytes each, are in -.DNG- format, and hold a dynamic range of 12 bits per channel.

This sensor does not have a Bayer filte . A Bayer filte , encountered in all of modern digital cameras – even our

Fig. 2: Color image produced by combination of visible wavelengths. 7 visible wavelengths are used to produce the color image, in this example only 6 pictures are visible for reasons of spacing.

cellphones – would filter the light only through Red-Green-Blue and would have great losses both in the recording of the remaining colors and in the light intensity (in brightness) which is lost through the filters

With this sensor we record information in the gray-scale, as we record light intensity. As a result, we have a better resolution, and we are truly using the 50,000,000 pixels of our E7 digital back.

This very high resolution gives us the opportunity to study very closely one by one the elements of the text we are trying to retrieve. Many palimpsest manuscripts are very old objects that have been handled again and again, reused and re-written. Many times, the condition of the manuscript is not optimal, with text running into the margins. This is where high special resolution images are important to have solid proof that even a single letter can be identified correctly by a scholar.

For each shot of the 33 of each page, the manuscript is illuminated by a different wavelength of light and in some cases also by different angles. Two raking lights at 15 degrees, two Main lights at 45 degrees, one transmissive light source. More specifically the lights that illuminated the Sinai palimpsest manuscripts had twelve different wavelengths. Modern MegaVision systems use up to 16 or 20 different wavelengths, something that is a result of the Sinai Palimpsests Project demanding more and more data to be recorded for future analysis. This is a highlight for EMEL (Early Manuscript Electronic Library) and its associates, leading developments and contributions on the field of Spectral Imaging. With EMEL's work through the Sinai Palimpsests Project on-site solutions were developed and drove the technology from which the rest of the world is benefiting. Using LED technology, we avoid the production of heat on the surface of the manuscript. The range of light we used, ranges from ultraviolet 365nm to infrared 940nm, covering in between the visible to the human eye spectrum by using seven visible wavelengths ranging from 400nm to 700nm. The necessity of LED illumination in a dark room must be emphasized and our responsibility to perform noninvasive and non-destructive examinations.

There are four different imaging principles in the system: reflectance, fluorescence, raking, transmissiv

Ultraviolet illumination makes the parchment 'fluoresce', since parchment absorbs the short UV wavelengths and re-emits back at longer wavelengths, which are visible to the eye and can be captured by a digital sensor. Areas of the parchment that had previous erased ink on them react differently as a matter of fluorescence from the parchment. This difference is essential in recovering erased writing and is captured in our pictures and can later help us while processing the data. In this way we can differentiate the two-dimensional picture that we have captured into different

groups: parchment with no ink, parchment with erased ink, newer writing ink.

Multiple individual visible wavelengths of light, when combined, produce a very accurate representation of the actual color of the object. Where Bayer-filter digital cameras use only three colors (Red, Green, Blue) to produce a color image, the Spectral Imaging System uses seven colors (Fig. 2) (Royal Blue, Light Blue, Cyan, Green, Amber, Red, and Dark Red) and more recent ones use up to twelve. This is achieved by a combination of seven different grey scale pictures captured under conditions of a dark room with the only illumination coming from those specific wavelengths, configured manually according to L*a*b* readings of the digital pictures, and white balanced according to L*a*b* measurements from color targets (more specifically: the MacBeth color target). Calibration of all of our photos can be traced back using unique serial codes of the color targets which are manufactured under specific standards, thus accurate reproduction of our results can be achieved at any given time. The MegaVision digital back and software can talk to each other in a way that if needed can capture new calibration files, and compare them to older calibrations in tables of numbers for each individual color patch, since what we have as a system is essentially a high-resolution colorimeter.

In cases of manuscripts where the ink has eroded the parchment surface, as we say 'the ink has eaten into the parchment' we use a specially build surface that illuminates the manuscript from its backside. In this way, the area of the letters where the parchment has become thinner the light penetrates it more easily. Our experience has shown that this technique yields significantly better results if the text was written on the flesh side. The transmissive light source used in Sinai was built by William Christens-Barry and featured four visible wavelengths and four IR wavelengths, though only the longest IR wavelength (940 nm) has proven to be consistently useful in processing. For this reason, recent transmissive light sources by MegaVision reduce the number of wavelengths.

The digitization process was made possible using equipment that already existed at St. Catherine's, as well as equipment that was specifically designed for this project. One of the pre-existing equipment was the manuscript support mechanism for the digitization of fragile manuscript codices called the Preservation Book Cradle by Stokes Imaging. The Stokes Cradle is a support mechanism (book-cradle) that is considered to be one of the best options on supporting fragile codices and ideal for the collection of St. Catherine's.

This mechanism mimics the natural movement of the spine of the manuscript codices. In this way sensitive codices are not harmed during digitization or put under any stress, since the book-cradle can support books that open even at a ~90-degree angle (or even a little less).

The processes are lengthy as a matter of time, first with the spectral digital imaging of manuscripts that can take up to 6-10 minutes per page in this specific project (6 minutes of capturing data and 1-4 minutes of material handling), and secondly with the equally demanding task of image processing (less than one minute on batch processing or up to 180 minutes and more for statistical image processing routines). These lengthy steps finally lead us from 33 different black-and-white photographs, to one or two final results where the erased writing becomes as legible as possible with high quality images that can be archived and reexamined objectively by a different number of scholars or scientists around the globe.

PICTURE CREDITS

Fig. 1: © Damianos Kasotakis, MegaVision Spectral.

Fig. 2: © St Catherine's Monastery, Mt Sinai, Egypt. Used with permission.

Article

Image Quality in Cultural Heritage

Tyler R. Peery, Roger L. Easton Jr., Rolando Raqueno, Michael Gartley, and David Messinger | Rochester, NY

Abstract

The application of modern spectral imaging technologies to recover information from cultural heritage objects, such as erased or damaged text in manuscripts, has become quite common in the last two decades. These techniques collect images under different wavelengths and modes of illumination (reflection, fluorescence, and transmission) and then combine them digitally to enhance the readability of low-contrast features. Of course, the imaging technologies (e.g., lighting, lenses, and sensors) continue to improve and manufacturers are always tempting users to purchase their newest advances, often without analyzing the costs and benefits of the system upgrades for the imaging tasks at hand. This observation suggests the need for better analysis of the 'chain' of stages in the imaging system to determine the weak points in the system where improvements would be most beneficial. This paper attempts to begin addressing some aspects of this analysis: the tradeoff between optical diffraction and spatial resolution specified by the pixel size. In the paper, system metrics that have been developed and used in environmental remote sensing are adapted for use in cultural heritage imaging and may help provide insight into the value of the parameters of imaging systems.

1. Introduction

In an age of high-resolution digital cameras and video displays, the acceptance and understanding by users of the idea that 'pixels' represent 'images' and of the meaning of the term 'spatial resolution' are improving. However, several assumptions inherent in the concept of 'resolution' of an image are of significant import in more technical applications. Some of these assumptions are based upon a lack of understanding regarding the human visual system, while others fail to recognize benefits of larger pixel sizes. The key takeaway is the time-old lesson: 'more is not always better'.

A decision to use a sensor with smaller pixels separated by smaller distances (a smaller 'pitch') to sample the image of the scene or object more finely assumes that there is some value to the consequent improvement in scene spatial resolution. Such a decision may seem to be intuitively obvious, though if the analysis of the subsequent image processing is included in the image chain, the 'quality' of the resulting image determined in part by the 'signal-to-noise ratio' (SNR) of the image data also is important to the final result. Fundamentally, improved spatial 'resolution' results from measuring light over smaller areas, which affects the number of counted light photons and the digital counts assigned to each pixel. This means that the 'integration time' needs to be considered, particularly with regards to the light exposure applied to the object. Similarly, the concept of 'well depth' and its impact on image dynamic range must be considered, as it is directly affected by pixel size. In addition, the impact of the spatial resolution on object size and aliasing concerns should be investigated if fine sampling is of major import. And finall , hard drive storage should also be considered, as single image files can exceed gigabytes of data for high resolution systems.

A point of diminishing returns can be reached with increased resolution, particularly when paired with the human visual system. If an object was created to be viewed by humans, and is being imaged to convey that same information, then imaging beyond the limits of human vision would not likely be beneficial. The benefits of higher resolution systems can include finer detail and sharper images, which may contend with aliasing and human visual system limits. The costs of higher resolution include increased digital storage space and lower signal levels per pixel, which can sometimes be combatted with increased exposure time or illumination. These benefits and drawbacks, along with their caveats, will be the focus of this paper.

2. Spatial resolution and sampling

Spatial 'resolution' is a term that has become familiar to users of smartphone cameras, though the scientific consideration is somewhat more complicated than the naïve concept. The common understanding is that images with more pixels over the same image area will appear 'better' to the viewer. For

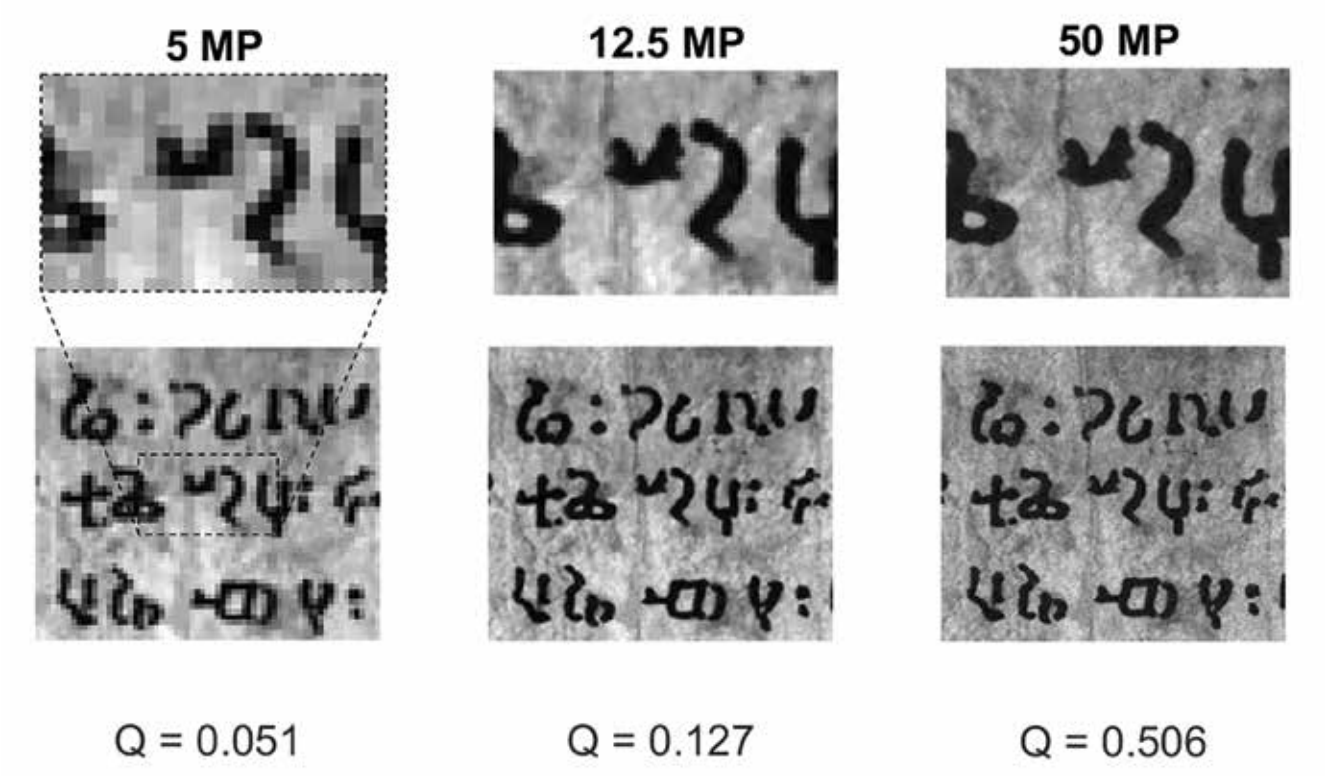


Fig. 1: Effect of decreasing pixel pitch to increase spatial resolution, which produces a 'smoother-looking' image. The gray scale of the images has been scaled by histogram stretching to better illustrate these differences.

example, this will allow the user to 'zoom in' on a feature without seeing the impact of the spatial resolution, often appearing as 'pixilation' or 'blockiness'. The cost of the image having too few pixels is the most intuitive aspect of spatial resolution (Figure 1). The image begins to appear 'blocky' if the user 'zooms in'. However, increasing the resolution can only sharpen an image so much, according to the spatial sampling mentioned in the previous section.

2.1 Target size

In environmental remote sensing, a metric used to determine the required level of image quality based upon target resolution is the 'National Imagery Interpretability Rating Scale' (NIIRS). This scale adjusts linearly with resolution of the intended target of the system, showcasing that increases in resolution and sharpness will be required for successful analysis of finer resolution targets. Just because a system can technically sample at a fine enough resolution, does not mean that the image will be interpretable. For example, if the image is too blurry to distinguish fine detail within texts, then fine sampling will be of little use to the human attempting to interpret it. A further constraint on the system, once the resolution is selected, will be to pair an appropriate diameter of the collection aperture with the pixel separation (known as 'pixel pitch'). The balance of these two will impact sharpness and aliasing of the system, based on a ratio of the cutoff frequency determined by each. Limitations of the human visual system highlight upper limits of sampling (resolution) and impacts of contrast, particularly when the end user is a human attempting to interpret faded texts, and will be reviewed below.

The NIIRS was originally designed for image analysts in the intelligence-gathering community to specify which tasks could be performed using an image with specific collection parameters. As shown in Table 1.[1,2] the NIIRS produces numerical values from 0-9 and specify typical tasks that may be performed on that data, from detecting large buildings to identifying barbs on a barbed-wire fence. For example, a 'low-resolution' image, where the pixel spacing on the ground (the 'ground sample distance' or 'GSD') is large could be useful for identifying large buildings, but not even for counting pedestrians. An image with a smaller GSD might be useful for identifying cars, but not for reading

[1] Irvine 1997.

[2] Fiete 1999.

Table 1: Example targets at given NIIRS levels and their associated approximate sizes.

N	Example	Linear Dimension	Average	Ratio (N/(N+1))
1	Terrain type (urban, forest, water, runways)	7.9–79.0 m	43 m	1.26
2	Large Buildings (hospitals, factories) major high patterns	18.0–35.0 m (Medium) 37.0–45.0 m (Large)	34 m	2.83
3	Houses in residential neighborhood, orchards	10.0–13.0 m	12 m	1.20
4	Sports courts, barns, silos	8.0–11.0 m	10 m	2.00
5	Large tents, large animals (elephants, rhinoceros)	5.0–7.5 m (elephant) 3.5–4.6 m (rhino)	5 m	4.13
6	Sedan or station wagon, utility poles	4.5–5.2 m , requires 1 m to differentiate vehicle type	1.21 m	5.50
7	Steps on stairway, railroad ties	0.20–0.25 m (8–10 in)	.22 m	1.80
8	Baby pigs, windshield wipers	0.15–0.30 m (6–12 in) 12–20 mm (0.5 in – 0.75 in)	.122 m	9.38
9	Barbs on fence, spikes on railroad tie	< 12 mm (< 0.5 in)	.013 m	**Ratio Avg = 3.51**

license plates. Assessment of whether an image can be used for different imaging tasks was the reason for constructing the 'General Image Quality Equation (GIQE)'. The GIQE uses the specifications of a given imaging system to predict the quality of resulting images in the NIIRS.

One image sensor should not be expected to be used for all targets and imaging conditions, just as a microscope camera would be impractical for imaging an entire document. Therefore, the appropriate range of target sizes and required resolutions should be selected for each set of system parameters. It makes more sense to select a system that is most appropriate to the typical application, and later adjust the system as changes require. In addition, there is no value to increasing spatial resolution beyond a certain point, as the human visual system may not be capable of detecting the difference in the image. Granted, computer systems can bring out the finer details to a level that humans could appreciate more, which returns us to a proper selection of NIIRS values.

2.2 Human perception limits

The human visual system has several limiting factors that preclude seeing beyond specific thresholds. Of these, those relevant to "sampling" and "detection" are most pertinent to imaging system design. For sampling, we are limited to the analogue of the Nyquist frequency of the human eye, which is determined by the spacing of photoreceptors upon the retina. This limit specifies an upper limit of the spatial sampling of the system. Any larger spatial frequencies will be 'undersampled' and appear incorrectly due to 'aliasing'. The second factor of 'detectability' is the minimum difference in gray value of the scene required to be sensed by the human visual system. This value is limited by 'spreading' of the light by ocular aberrations and by the subsequent neural processing, which was designed by evolution to concentrate on the most pertinent information required for survival. For an example of the latter, humans readily notice rapid motion of approaching hazards, while ignoring slower

motion or camouflage. The design of the eye-brain system may limit subtle differences in image quality of low and high frequency background noise, which may be very pertinent to computer vision algorithms that do identify and detect those differences.

The Nyquist frequency, ρ_{Nyq}, of the human visual system can be estimated from the spacing of the photoreceptors or by empirical testing. These methods arrive at two similar answers[3], approximately 50-60 cycles per angular degree. To calculate the Nyquist limit for the HVS, the distance d_p between the photoreceptor cones in the region of highest acuity is approximately 3 µm. This angle scales to approximately 5µm per minute of arc or 0.3 mm/degree, on the retina, which translates to about 100 cones/degree. This would allow unaliased sampling of signals of 50 cycles/deg. However, perhaps due to the shape of the cones or neural processing, typical eye tests show that the acuity of an average adult is actually somewhat better: approximately 60 cycles per degree for 20/20 vision[4].

To build upon these factors constrain imaging system design, it is important to specify the purpose of the system. Depending upon the task, the limits of human vision may play a role in image collection. For example, if a multispectral imaging system (MSI) or hyperspectral imaging system (HSI) is to identify handwritten text, resolution beyond that of the human visual system would likely not be necessary. Therefore, given the known distance and focal length of a detector, as well as the average reading distance of a human, a detector's ρ_{Nyq} may be matched to that of the human visual system.

Using an average reading distance of 0.38 m and human acuity limits of 60 cycles per degree, average humans are unable to discern features smaller than 0.111 mm apart.[5] For example, the state-of-the-art MSI system used to image the Enoch palimpsest at the Berlin State Library had a spatial sampling distance determined to be 26 microns/pixel, which is well beyond the extreme limits of human vision and therefore much better than would be expected of handwritten text. If the purpose of the images is to identify and read texts, such high resolution would not be necessary. A lower resolution system with larger pixels and therefore improved dynamic range would exhibit improved SNR or would require less integration time to capture similar SNR values.

[3] Curcio et al. 1990.

[4] Curcio et al. 1990.

[5] Peery and Messinger 2018.

2.3 Optical resolution, diffraction

Light energy emerging from a source physically spreads during propagation, which is known as 'diffraction'. The purpose of the optical system is to collect and 'refocus' this light to create 'images' of the original source points. However, the spreading of the light ensures that the images of two closely spaced point sources in the object will 'bleed' into each other to the point where they may be indistinguishable from the image. This is the ultimate source of the resolution limit of the imaging system. A common approximate rule of thumb for the limiting separation of two point images that may just be resolved is

$$\Delta x \approx \lambda_0 \cdot \frac{f_0}{D_0} = \lambda_0 \cdot f/\# \quad (1)$$

where Δx is the limiting separation of two point images, λ_0 is the dominant wavelength being imaged, f_0 is the focal length of the imaging system, D_0 is the diameter of the lens system, and $f/\#$ is the focal ratio (f/number) of the system:

$$f/\# = \frac{f_0}{D_0} \quad (2)$$

2.4 Q factor

Q is a metric used to compare the spatial frequencies passed by the optics and by the sensor in an imaging system. It can help characterize aliasing as well as general visual sharpness based on that relationship. Q is defined

$$Q = \frac{\lambda_0 \cdot f/\#}{\Delta x} \quad (3)$$

The two values effectively being compared in Q are the cutoff spatial frequencies of the optics and of the sensor of the system. These are frequencies of patterns beyond which higher frequencies cannot pass without being aliased, being recorded as an adjusted, lower frequency. The optical cutoff frequency, ρ_{oco}, is:

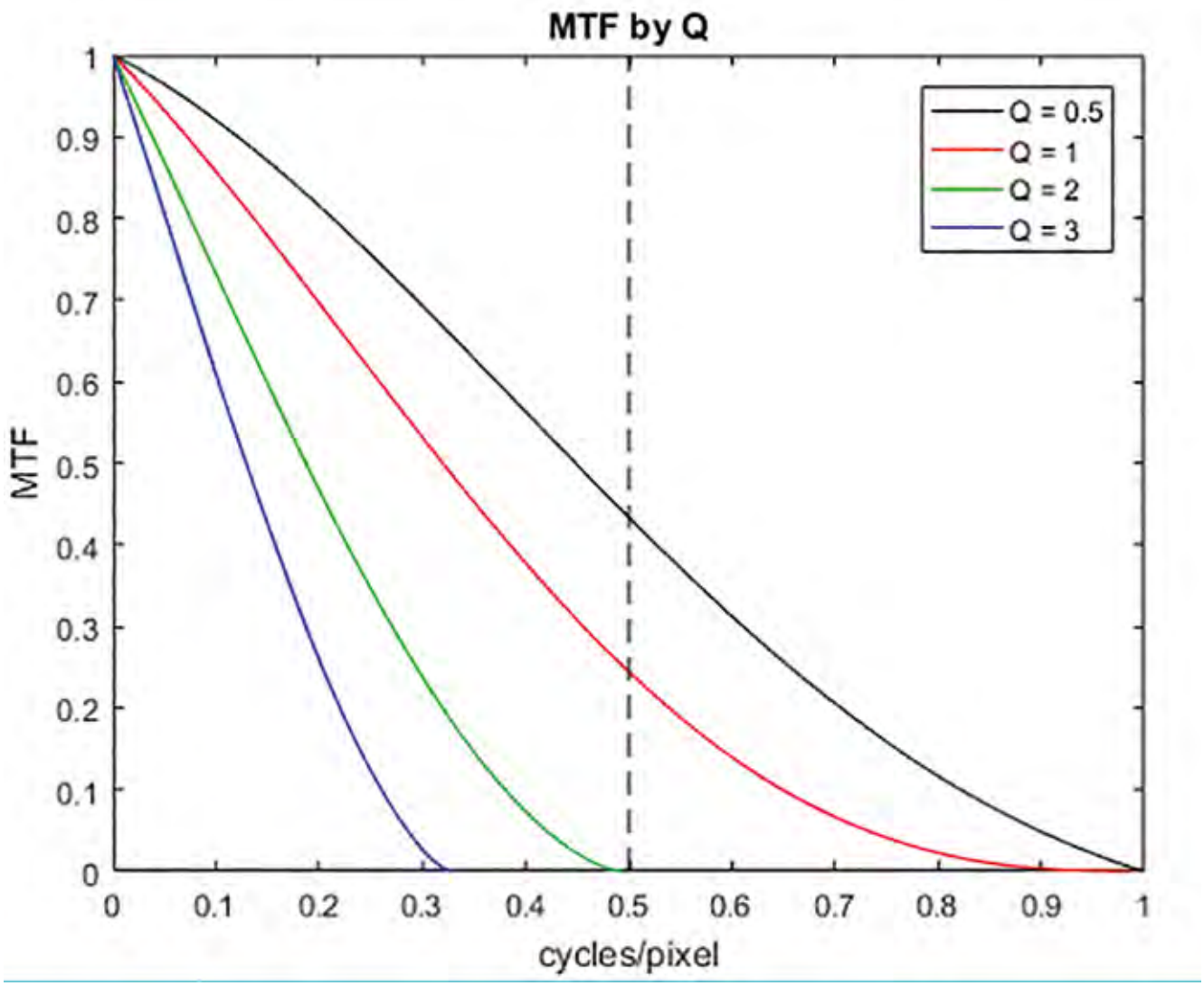


Fig 2: MTF's of various *Q* value systems compared to their Nyquist frequency, normalized to the number of cycles per pixel. Lower *Q* values result in higher frequencies being allowed into the system, resulting in aliasing (right of the dashed line) as well as sharper looking edges.

$$\left(\rho_{\max}\right)_{optics} = \frac{1}{\lambda_0 \cdot f/\#} \quad (4)$$

The sensor samples the spatial signal from the optics at the pixel spacing Δx, and must sample at least twice per period of every spatially oscillating signal in the scene to avoid aliasing, so the maximum spatial frequency passed by the sensor is:

$$\left(\rho_{\max}\right)_{sensor} = \frac{1}{2 \cdot \Delta x} \quad (5)$$

This means that for systems where the two maximum frequencies are matched, the definition of $Q = 2$ represents the balance of ρoco matching the pixel pitch's Nyquist limit of ρnyq.

To better illustrate this relationship of cutoff frequencies with the value *Q*, Figure 2 shows the modulation transfer functions that apply for various *Q* values. For reference, the *Q* value for the MegaVision MSI system as captured in Figure 1 had a *Q* value of 0.5. The Nyquist limit is highlighted by a dashed vertical line.

From Figure 2, it is apparent that the optics in a system with $Q \geq 2$ will not pass spatial frequencies above the Nyquist limit, and thus the image will not be aliased. Systems with smaller values of *Q* allow larger spatial frequencies to enter the system with less modulation (see Figure 2 where the

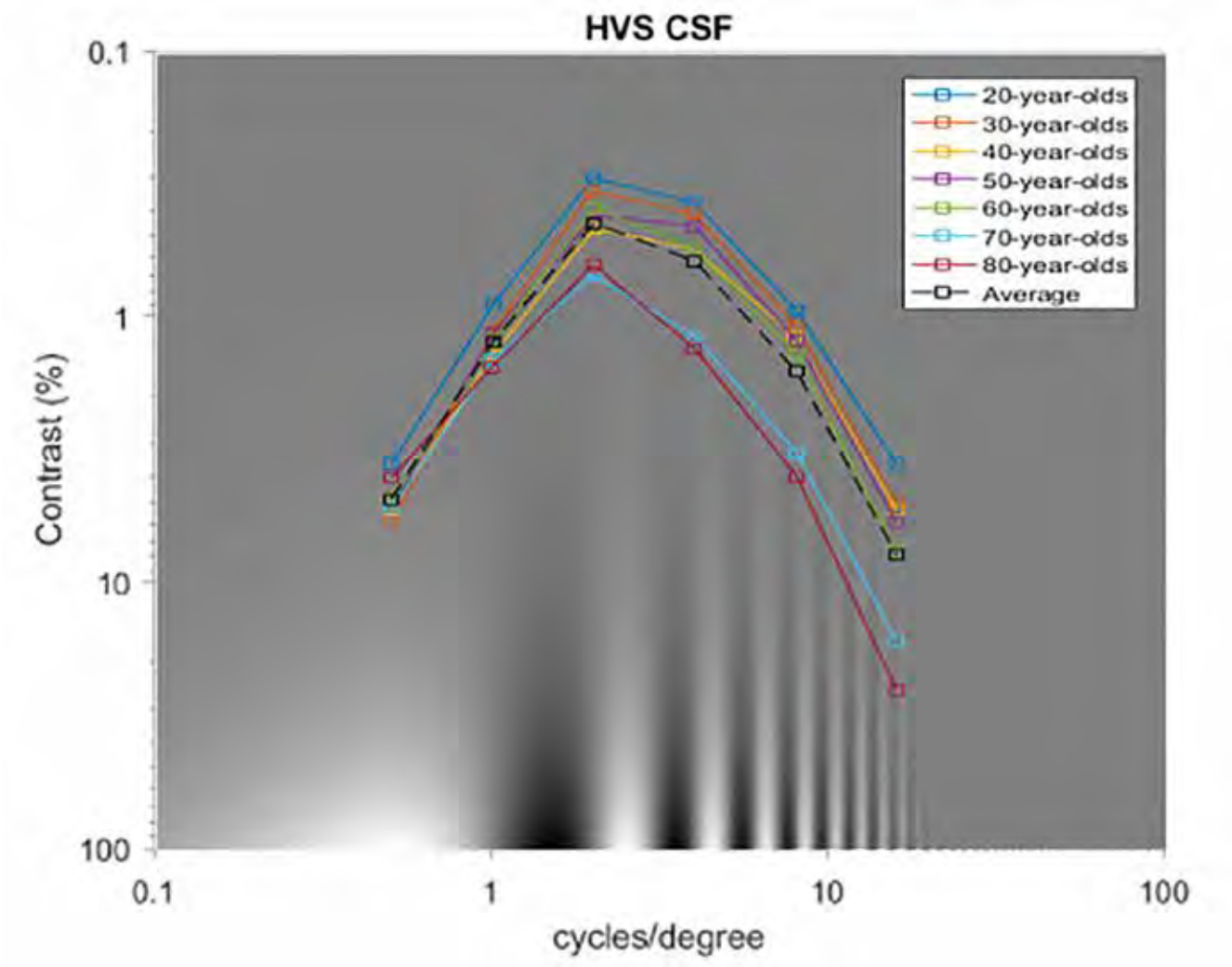


Fig 3: The contrast sensitivity functions of the human visual system by age (after Schieber, 1992) overlaid upon a sine wave of increasing frequency along the x-axis and decreasing contrast along the y-axis.

MTF at the Nyquist frequency is 0 for a $Q = 2$ system and 0.25 for $Q = 1$), resulting in sharper images. Furthermore, because the frequencies above the Nyquist limit are also being modulated, sometimes severely, the aliasing that does occur will also be modulated, diminishing the visibility of those spatial frequencies.

In conclusion, one should make sure they select a range of target sizes for a system to focus on. It may help to make this decision with an understanding of the limitations of the human visual system, especially if the targets are chosen based upon a human interpreting a target. MTF concerns should be considered if extra-fine sampling is determined to be required, as these are the frequencies at which there will be the most aliasing, depending upon the Q design of the system.

3. Resolution and signal level

It is important to consider all implications of higher-resolution imaging systems. It will become apparent that having a sensor with the largest number of pixels is not necessarily most appropriate for a particular situation. A significant portion of this discussion is again based upon the concept of GSD. Smaller pixels in a sensor will measure fewer photons in the same exposure time, thus decreasing the signal-to-noise ratio (SNR). Because doubling the resolution of a detector halves both the x- and y-direction of the pixel pitch, the area of a given pixel to capture photons is reduced to one quarter. Because this also reduces shot noise of the measurement by the square root of the signal, the SNR is reduced in proportion to the increase in linear spatial resolution, and is written as

[6] Schieber 1992.

$$SNR = \frac{N_{sig}}{\sigma_{sys}} = \frac{N_{sig}}{\sqrt{N_{sig}}} = \sqrt{N_{sig}}, \quad (6)$$

where N_{sig} is the signal level (number of electrons, photons, digital counts, etc.), and σ_{sys} is the shot noise of the system.

3.1 Integration time and total power

The decrease in signal due to smaller pixels may be addressed in two ways. The strength of the measured signal could be increased by increasing some combination of the integration time and the amount of light incident upon the target. Increasing integration time is simpler with static detectors, as dynamic systems will have an increase of pixel smear and jitter, which will impact the MTFs. One important concern for cultural heritage documents is that both methods increase the total amount of energy received by the imaging system, but also increase the total energy on the target. These objects usually are sensitive to incident radiation and under various control standards. Documents, inks, and pigments can be damaged by incident light either by chemical reaction or by heating of the target. The British Standards Institution has a standard for recommended lighting of cultural collections, *PAS198:2012 Specification for Managing Environmental Conditions for Cultural Collections*. This standard recommends a document regularly displayed be limited to 50 lux (lumen/m^2). Andrew Beeby (2018) states that typical conservation imaging setups use approximately 1000 lux over a much shorter period.[7] This results in 8000 lux-hr being achievable over a month of exhibition time or a day of conservation imaging.

3.2 Well depth

Because improved resolution decreases the pixel area, and therefore the number of photon-generated electrons that can be measured by a pixel before writing out the data. This maximum number of photoelectrons is known as the 'well depth'. This constraint to the number of possible detected photoelectrons constrains the possible signal-to-noise ratio of the image, which increases as the square root of the number of detected photons, as shown in Eq. (5). Therefore, larger well depth allows for more discrimination in the signal measured as well as allowing more signal to be collected compared to the detector inherent noise level, further increasing the SNR

3.3 Contrast sensitivity function

Recent research has highlighted considerations of the contrast sensitivity function (CSF) with respect to minimum detectable thresholds for both targets (overtext and undertext) and noise in an image.[8] The CSF can be seen in Figure 3, with a CSF curve overlaid on a frequency and contrast varying signal to highlight the region of detection for the HVS. At high frequencies, change detection is reduced by diffraction, chromatic aberration, and spherical aberration within the eye itself. Low frequency detection is reduced due to neural processing optimized for a high pass filter to detect fine details

Therefore, signal should be considered based on target size, especially for large values of spatial frequency. Just because the detector measures a difference in signal does not mean that the human visual system will be able to discern it. The target will only be discernable if, dependent upon the resolution, the contrast between the target and background is of a significant relative magnitude. Methods like principal component analysis and spectral angle mapper may be used to increase the contrast between the target (text) and background (parchment). SNR should be maximized to limit the impact of shot noise (noise based on variability of the signal) on the system, particularly when image processing steps may attempt to histogram stretch an image for better visualization. This is a method that increases contrast of all aspects of the image, including noise.

In conclusion, the optimal resolution will have to be balanced between fine enough to be appealing and usable by human analysts, yet coarse enough to allow quick signal gathering to prevent light damage to targets, which will also increase the SNR for image processing algorithms. The human usability could also be enhanced using pan-chromatic sharpening, which mathematically blends the spatial and spectral resolution of two images into one higher resolution image.

4. Imaging system modeling

Because there are many options when designing an imaging system, and investing in a new one can be expensive, it may be beneficial to attempt to model systems before acquiring them. This methodology is based off radiometry, which follows photons through a well-defined scene, described by a light source, a reflectance target, and an imaging system.

[7] Beeby et al. 2018.

[8] Peery and Messinger 2018.

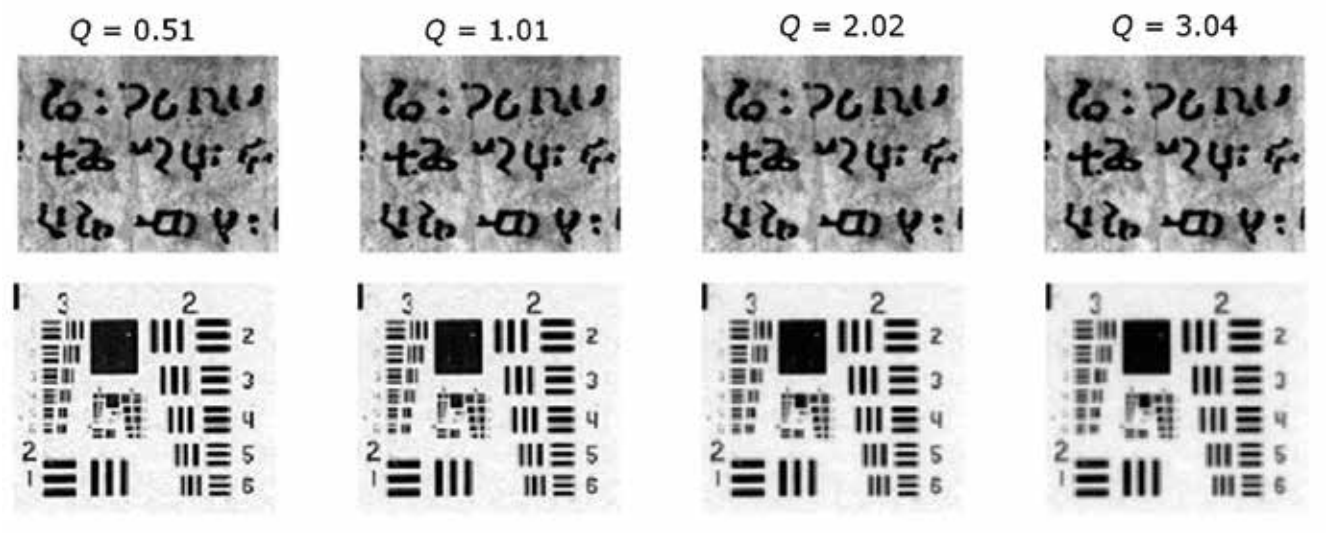


Fig 4: Effects of *Q* on image sharpness. Images with smaller values of *Q* appear sharper as the amplitudes at large spatial frequencies are increased, but this can also exhibit aliasing. Depending on the *Q* value and the frequencies of the image, this may be indiscernible.

The basic idea is to use a high-fidelity image taken with one system, and then modeling an expected output from another system, by tracking pertinent variables.

4.1 Modulation transfer function (MTF)

The modulation transfer function (MTF) of the system was modeled by combining the MTFs of the optics, wavefront error, system jitter, pixel smear, and pixel pitch as

$$MTF_{sys} = MTF_{optics}\ MTF_{sensor} \cdot MTF_{jitter} \cdot MTF_{smear} \cdot MTF_{pitch} \cdot \quad (7)$$

These quantities are specified in accordance to Fiete's system model used for remote sensing[9]

$$MTF_{optics} = \frac{2 \cdot (A + B + C)}{\pi\left(1 - \varepsilon^2\right)} \quad . \quad (8)$$

Though the smear and jitter values may be negligible for a stationary framing camera, added motion from a scanning system would most likely increase these.

The system MTF can now be plotted with respect to the pixel size of a system and thus its Nyquist frequency. By varying the aperture size, and thus adjusting MTFsys, different values of Q can be compared with respect to aliasing, as shown in Figure 2.

4.2 Sampling

Modeling images to view the different impacts of *Q* required modeling the basic Fourier process of image capture. This includes convolving the original input image, $f(x, y)$ with the point spread function of the system (the inverse 2-D Fourier transform of the MTF $H(\xi, \eta)$), and then multiplying by the sampling function $s[x,y]$. This can also be done in the frequency domain by evaluating the Fourier transform of the image and multiplying by the MTF, as

$$g[x,y] = (f[x,y] * h[x,y]) \cdot s[x,y] \quad (9)$$

where '*' denotes the mathematical operation of convolution.[10]

This downsampling process allowed aliasing to be investigated from an existing image. Adjusting MTF alone could not produce new aliasing, as the image had already been sampled by some detector, whether it was a historic document imager, conventional scanner, or any computer generated image (with pixels as the sampling function). As the image was already 'captured' once, all frequencies above the relevant Nyquist frequency have been removed. No MTF adjustment could enhance the frequencies beyond ρ_{Nyq}, as they no longer existed. By downsampling, ρnyq was reduced by increasing the effective pixel spacing according to Equation 6. Various MTFs could then be applied, enhancing or limiting frequencies that reach the sampling function. If the MTF were to cutoff frequencies above the new ρ_{Nyq} then no aliasing would exist.

[9] For full details, including Eq 8 variable values, see Fiete 2002.

[10] Easton 2010.

Table 2: Proposed cultural heritage relevant targets at theoretical NIIRS levels and their associated approximate sizes.

N	Example	Linear Dimension	Average	Ratio (N/(N+1))
7	Large picture (page width)	8.5 in (.215 m)	215 mm	2.11
8	Small Picture (half width+margin)	4 in (.102m)	102 mm	4.86
9	Word of text	5 letters = 5x below ~5/6 = .02 m	21 mm	5.00
10	Character of text	72 pts/in, 12 pt = 1/6 in = .004	4.2 mm	5.60
11	Pen stroke	0.1-1.4 mm fountain pens	0.75 mm	3.95
12	Overlapping strokes	¼ of pen stroke, .025–.35 m	0.19 mm	1.92
13	Hairs breadth	17–181 um (.181–.017 mm)	0.099 mm	2.20
14	Fingerprint	500–590 dpi (.05 -–.04 mm)	0.045 mm	**Ratio Avg = 3.66**

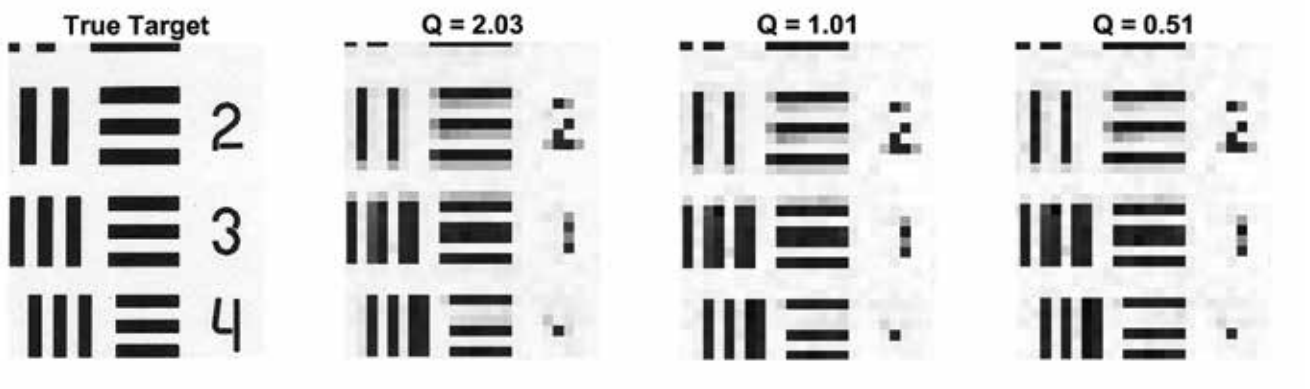


Fig 5: Aliasing can be seen at various *Q* values, increasing as *Q* decreases to the right. The true target was blurred by the MTFs of Figure 2. The sampling size was selected for the horizontal bars around line 3 to be at ρ_{nyq}. Aliasing is most visible on the horizontal bars of target 2, as well as making out the numerals 2–4 in higher *Q*'s and being unable to do so at lower *Q*.

5. Analysis

A list of proposed NIIRS values, their example targets, typical size, and ratio to following levels is shown in Table 2. The scaling ratio of approximately 2x GSD per NIIRS level was followed, per Harrington, et al. (2015). The ratio of GSD's in the final column of Table 1 and 2 can be compared to see relative similarity in scaling, with large uncertainty on some levels due to ambiguity of target definition

The impacts of the modulation transfer function (MTF) on image quality is yet another engineering balancing act. The MTF of a system defines how different wavelengths of

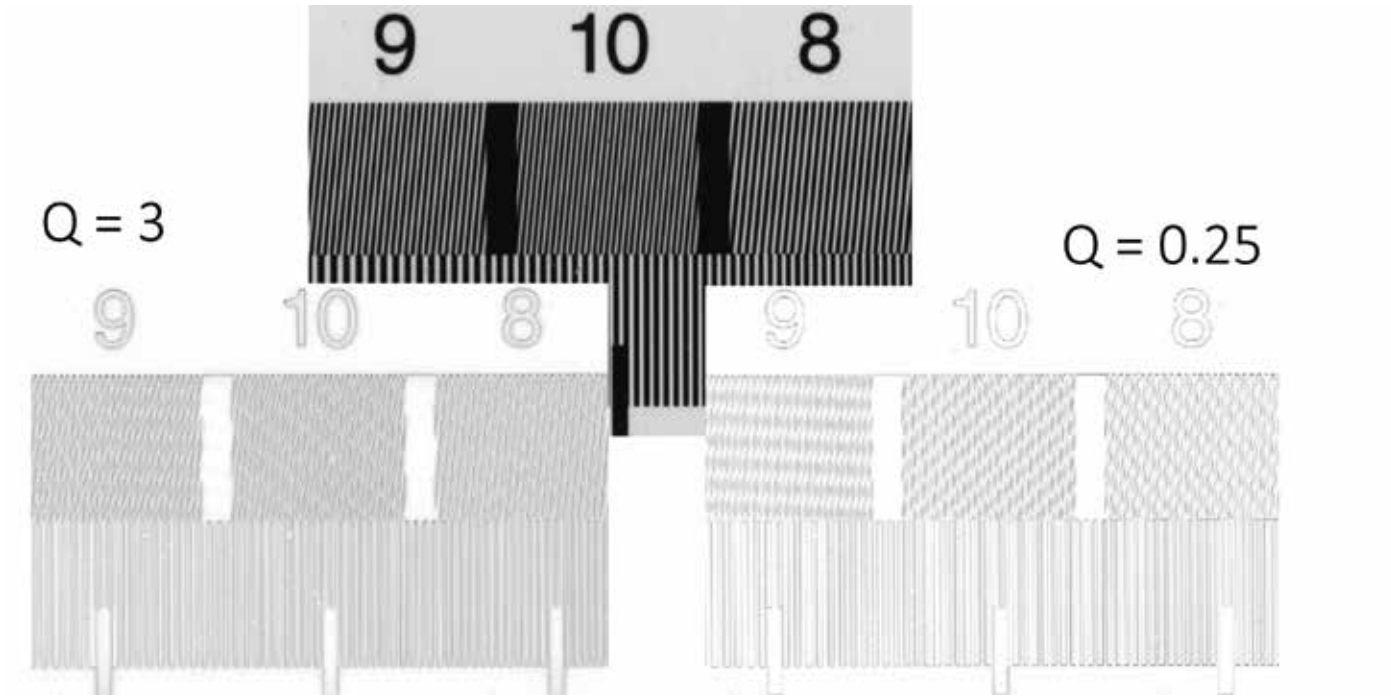


Fig 6: Other options for showing aliasing. Aliasing is highlighted in difference images (bottom) as new patterns. The original image (top) is composed of bars at approximately a 15-degree angle at discrete frequencies, and a lower bar of vertical bars regularly increasing in frequency. Aliasing can be seen in $Q = 0.25$ (right) as horizontal bars in window 9, 45 degree bars in 10, and crossing bars in window 8. Aliasing in the lower bar is highlighted as irregular intervals. $Q = 3$ (left) only shows differences based on blurring and modulation of existing frequencies.

light are modulated through the imaging system, define primarily by the aperture of the lens and the pitch of the pixels, as seen in Equation 8. Because the MTF describes which frequencies, and how much of said frequencies, get through a system, they can be used to determine two important characteristics: how sharp an image is (based on how many high frequencies are passed) and how much aliasing is occurring (based on how many frequencies past the Nyquist frequency enter the system).

The impacts of this aliasing, based on the Q of a system, can be seen in Figure 5. The aliasing effect can most clearly be seen in the horizontal bars of the resolution targets. Sharpness in an image will not only affect human perception, but also computer algorithms. Different algorithms can be affected differently by sharpness,[11,12] therefore pertinent ones should be investigated when designing a system. Most optical systems tend to design to Q = 0.5-1.5, catering to sharper images for the human visual system.[13] The impacts of aliasing on expected uses for the system should also be investigated, to decide whether a lower Q would be acceptable or not. For reference, the Q of the MSI system modeled in this project is approximately 0.5.

Figure 4 was analyzed for CSF calculations based upon overtext, undertext, and shot noise (pixel size). These resolutions were determined based upon the imaging GSD as well as display assumptions, based on a viewing distance of 22.5 in.[14] The results of the target and noise analysis can be reviewed in Figure 9. This shows the overtext and some undertext should indeed be the most discernable, shot noise within the background of the parchment is potentially visible for those with high visual acuity, and noise within the text is not visible.

An analysis of the HVS CSF with regard to the target texts and noise from Figure 7 can be seen in Figure 9. The resolutions were based upon the size of a pixel on a computer display and how many pixels made up a target (pen stroke on a character of text) or noise (a single pixel for Poisson photon arrival statistics). The contrast was measured based on the mean value of the target or signal,

[11] Webster, Anthony, and Scheirer 2018.

[12] Peery and Messinger 2018 (10644).

[13] Fiete 2010.

[14] An in-depth description of the analysis can be seen in Peery and Messinger 2018, where the noise was analyzed based on Poisson arrival statistics of the photons.

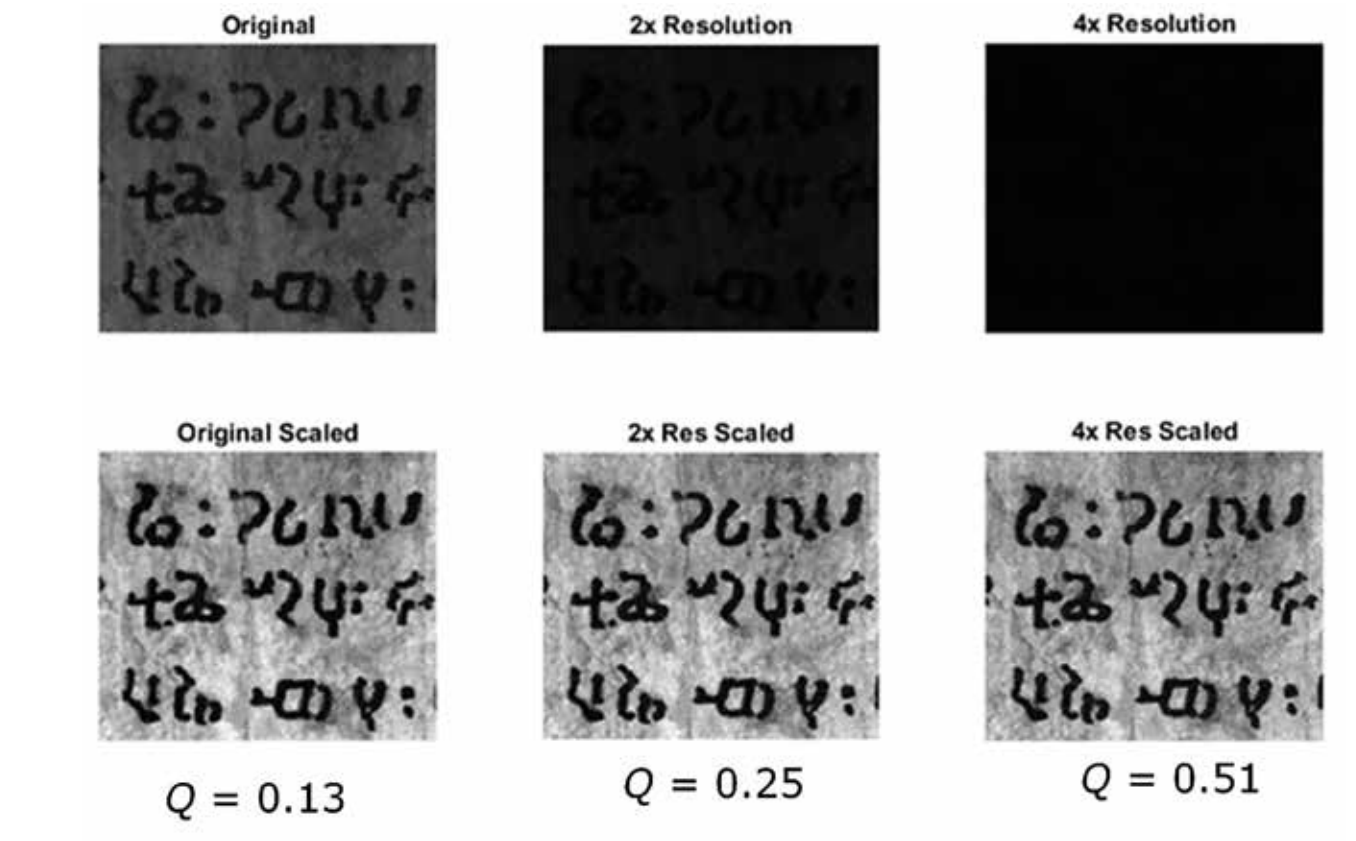


Fig 7: Effect of decreasing pixel size: spatial resolution increases linearly, while sensor area decreases exponentially. Constant scaling is maintained across the top row of images based upon the well depth of the system. The lower images are replicas that have been stretched to use the full dynamic range of the system, but also that highlight the reduced SNR of systems with higher *Q*.

compared with the mean of the background. For the overtext and undertext, this is a relatively straightforward measurement. For the noise levels, the background was either the mean of the parchment or the mean of the text, and the signal noise was based on the variance of the signal based upon shot noise parameters for that mean signal level.

If all other aspects of the system are considered equal, one should also consider the increase in memory required for larger resolution files. Increasing the resolution (decreasing sample size) by a factor of two will increase both the write time and the required amount of storage by the square of the increase, due to having to increase in two dimensions. File size is already a concern of high resolution imaging and combining that with hyperspectral images would only compound the issue further. If storage space for images is a concern, this should be considered by the user.

6. Conclusions

In conclusion, if higher resolution is not required for one's targets of interest, it will cost additional time to maintain the same signal to noise ratio and will also require additional storage space exponentially proportional to the resolution increase. To determine if a target of interest requires higher resolution, NIIRS examples such as the ones listed in Table 2 should be referenced. If the targets were made for human interpretation, human vision system limits should be considered with this. If a higher resolution is required, aliasing could negatively impact higher frequency sampling and should be considered by the *Q* factor of a system. Signal costs can be offset with increases to integration time, but this needs to be weighed against time constraints of the data collect as well as total power limitations on potentially fragile targets. A recap of the costs, benefits, and their respective caveats can be reviewed below:

Fig 8: The same type of modeling as Figure 7, but over a fl t response refle tance target to highlight the 'salt and pepper' noise added to the system due to lower SNR values at higher resolution.

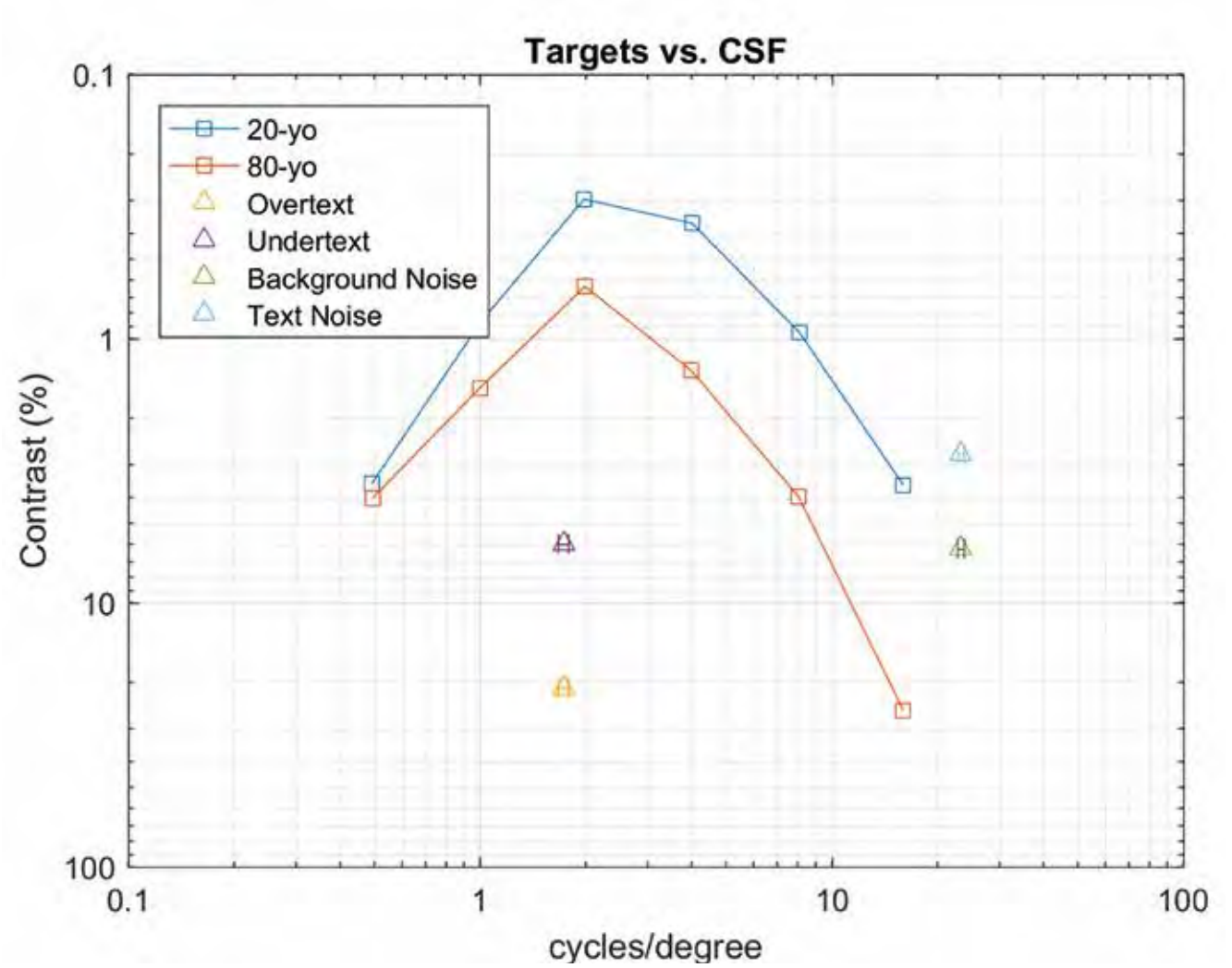


Fig 9: The two limiting CSFs plotted with the resolution and contrast values of overtext, undertext, background noise, and text noise measured at the scale of Figure 7 with a full well depth at 80% refle tance of the R-CHIVE system.[15]

- Benefits of Higher Resolution
  - Finer detail may be seen in the data (assuming it is present and detectable)
    - Consider *Q* aliasing at fi er resolutions
    - Consider HVS limitations for pertinent target levels
  - Potentially sharper images
- Costs of Higher Resolution:
  - Reduction in signal-to-noise ratio (SNR)
    - May require longer integration time or increased signal to counter
    - Consider time/energy constraints on sensitive targets
  - Additional memory requirements for writing/storage

7. Future Work

The General Image Quality Equation (GIQE) has undergone several revisions, leading to the latest GIQE 5. These changes switched the image being evaluated to the unenhanced image, as opposed to one that has undergone image processing techniques to enhance usability. This allows for a master image to be defined on its image quality, and any individual's image processing techniques could be applied to that base. The latest equation relies upon the GSD, signal-to-noise ratio (SNR), and the relative edge response (RER), which defines the sharpness of the image. Given these parameters, a NIIRS value can be output for a given image without requiring specialized analysts to evaluate each image to assign a value they deem appropriate.[16]

A general image quality equation could be developed for cultural heritage targets similar to those proposed in Table 2, such that curators would know what type of system specifications would be required to achieve their imaging goals, based upon target resolution.

[15] Plot taken from Peery and Messinger 2018.

[16] For latest version, see Harrington et al. 2015.

REFERENCES

Beeby, Andrew, L. Garner, D. Howell, and C. E. Nicholson (2018), 'There' s More to Reflectance Spectroscopy Than Lux,' *Journal of the Institute of Conservation*, 41/2: 1–12.

Curcio, C. A., K. R. Sloan, R. E. Kalina, and A. E. Hendrickson (1990), 'Human Photoreceptor Topography', *Journal of Comparative Neurology*, 292/4: 497–523.

Easton Jr, Roger L. (2010), *Fourier Methods in Imaging* (Hoboken: John Wiley & Sons).

Fiete, Robert D. (1999), 'Image Quality and λ fn/p for Remote Sensing Systems,' *Optical Engineering*, 38/7: 1229–1241.

Fiete, Robert D., T. A. Tantalo, J. R. Calus, and J. A. Mooney (2002), 'Image Quality of Sparse Aperture Designs for Remote Sensing,' *Optical Engineering* 41/8: 1957–1970.

Fiete, Robert D. (2010), *Modeling the Imaging Chain of Digital Cameras* (Bellingham, Wash.: SPIE Press; Tutorial texts in optical engineering, 92).

Harrington, Leigh, D. Blanchard, J. Salacain, S. Smith, and P. Amanik (2015), 'General Image Quality Equation; GIQE 5', *Optical Society of America*, 1–12.

Irvine, John M. (1997), 'National Imagery Interpretability Rating Scales (NIIRS): Overview and Methodology', *SPIE*, 3128: 93–103.

Peery, Tyler R., and David W. Messinger (2018), 'MSI vs. HSI in cultural heritage imaging', *Proc. SPIE*, 10768, *Imaging Spectrometry XXII*: *Applications, Sensors, and Processing.* doi.org/10.1117/12.2320671

Peery, Tyler R., and David W Messinger (2018), 'Processes for conducting hsi pan-sharpening with 3d digital flattening, in *Algorithms and Technologies for Multispectral*, *Hyperspectral*, *and Ultraspectral Imagery*, XXIV (International Society for Optics and Photonics), 10644, 106441W.

Webster, Brandon Richard, Samuel E. Anthony, and Walter J. Scheirer (2018), 'Psyphy: a Psychophysics Driven Evaluation Framework for Visual Recognition', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41/9: 2280–2286. doi: 10.1109/TPAMI.2018.2849989

Schieber, Frank (June 2018), 'The Contrast Sensitivity Function (csf)', *USD Internet Psychology Laboratory*, <http://usd-apps.usd.edu/coglab/CSFIntro.html>.

Article

When Erased Iron Gall Characters Misbehave

Keith T. Knox | Kihei, Hawaii

Fig. 1: EurekaVision cultural heritage imaging system, in use at the Staatsbibliothek zu Berlin (Berlin State Library, Prussian Cultural Heritage Collection).

Abstract

Iron gall ink that has been erased from parchment, leaves stains which are the residues of compounds of tannic acid with vitriol, or metallic sulfates. For years, this erased writing has been read by scholars by inspecting the parchment under ultraviolet illumination. This results in increased contrast that enables the erased writing to be read. This is the normal behavior of erased iron gall ink on parchment. When the residue from the erased ink does not behave in this normal manner, other methods need to be developed to reveal the erased text. This paper describes three methods that were developed to reveal erased characters that behave in uniquely different ways.

1. Multispectral Imaging of Parchment Manuscripts

The multispectral imaging system used to capture the images for all of the projects, described in this paper, was developed by MegaVision ('Archival and Cultural Heritage Imaging'). One of their imaging systems is shown in Fig. 1. In that image, several LEDs of different wavelengths have turned on at the same time to generate white light for positioning the manuscript. During image capture, LEDs are turned on individually to illuminate the manuscript with only one wavelength of light, at a time. The camera has a 50 mega-pixel, panchromatic sensor that records the response of the manuscript leaf to each wavelength of light. The diffusers ensure a reasonably uniform spread of the light across the manuscript. The LEDs and camera are under computer control, so all of the changes in illumination and camera settings are executed automatically.

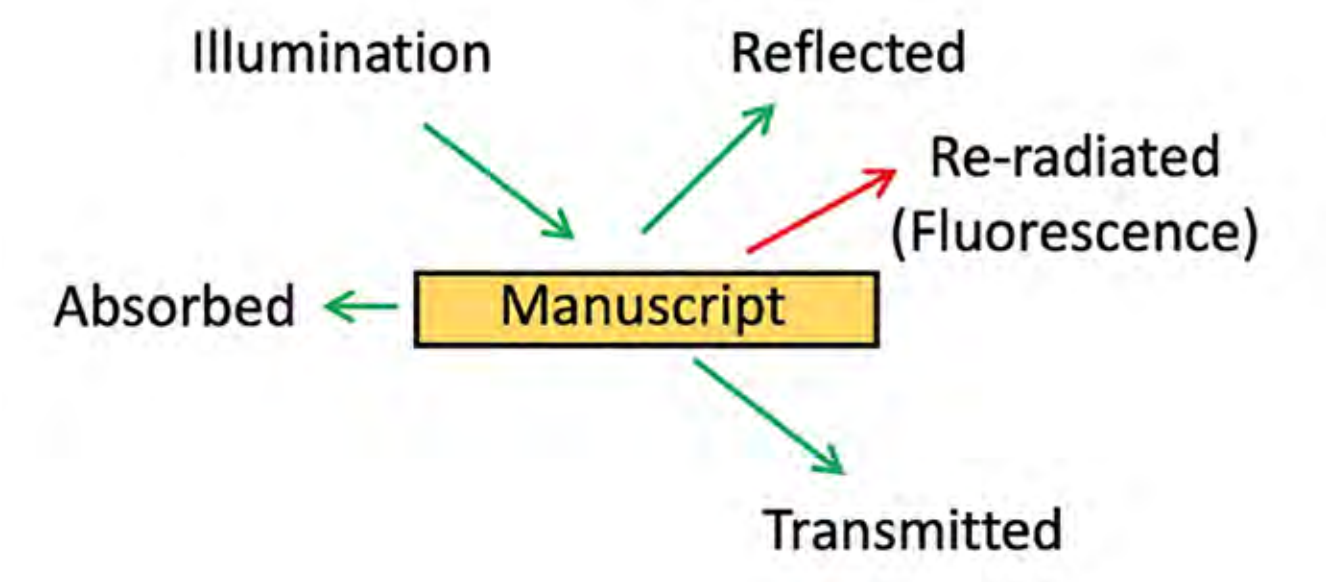


Fig. 2: The light interacts with the parchment in one, or more, of four different ways, refle tion, transmission, absorption, or fluo escence.

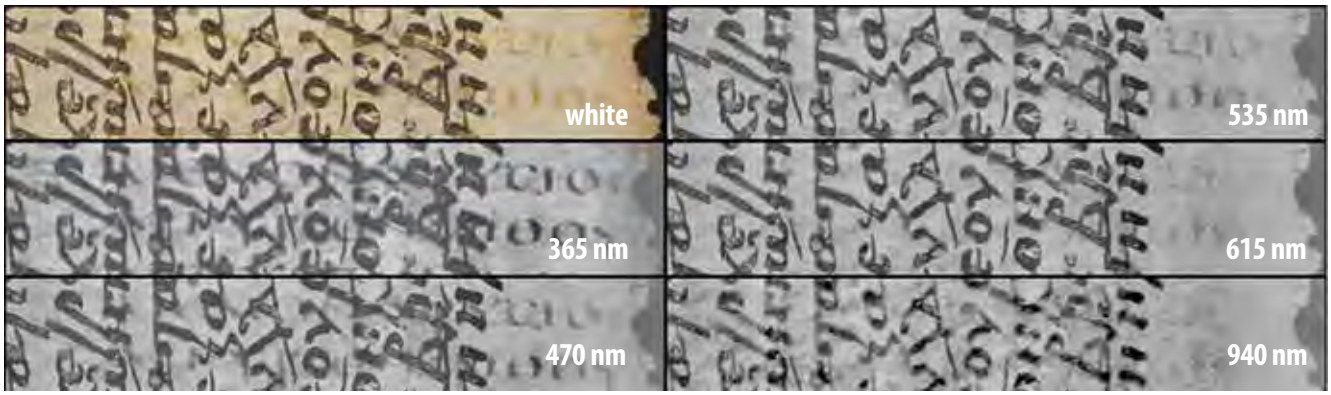


Fig. 3: Erased iron gall ink that has been overwritten. The erased stains are most visible when illuminated with blue or ultraviolet light.

Three types of optical images are captured, as shown in Fig. 2. One is simple reflectance of the light from the parchment surface. The light reflected from the parchment will be at exactly the same wavelength as the illumination. Light of any wavelength, from ultraviolet, through the visible, and into the infrared, can be reflected from the surface

When ultraviolet light is used to illuminate the parchment leaf, the light is absorbed by the parchment, then re-radiated outward, in all directions, at longer wavelengths, typically in the visible region of the spectrum. For example, the ultraviolet light might be absorbed and re-radiated as blue light. This process is called fluorescence. Filters on the camera can be used to determine the colors of the light that are contained in the fluorescence

Lastly, images are taken of light that is transmitted through the parchment leaf. The leaf is illuminated from a light sheet that sits below the leaf. The light travels through the parchment and out the top surface to the camera above. The longer the wavelength of the illumination, the deeper into the parchment the light will go. As a result, the best wavelengths for transmitted-light images are the infrared wavelengths, with visible light also working, depending on the thickness of the parchment. Ultraviolet light does not penetrate deeply enough to transmit through the parchment.

These three different types of images are useful in distinguishing different inks, or in enhancing the regions of erased ink. When the goal is to recover the erased text and make it legible for a scholar to read, the different images can be combined in ways that enhance the desired writings. This study will focus on the response of iron gall ink to multispectral illumination. There are some responses that are well understood, and others that are not. All of them can be used to enhance the erased text (Easton et al. 2010).

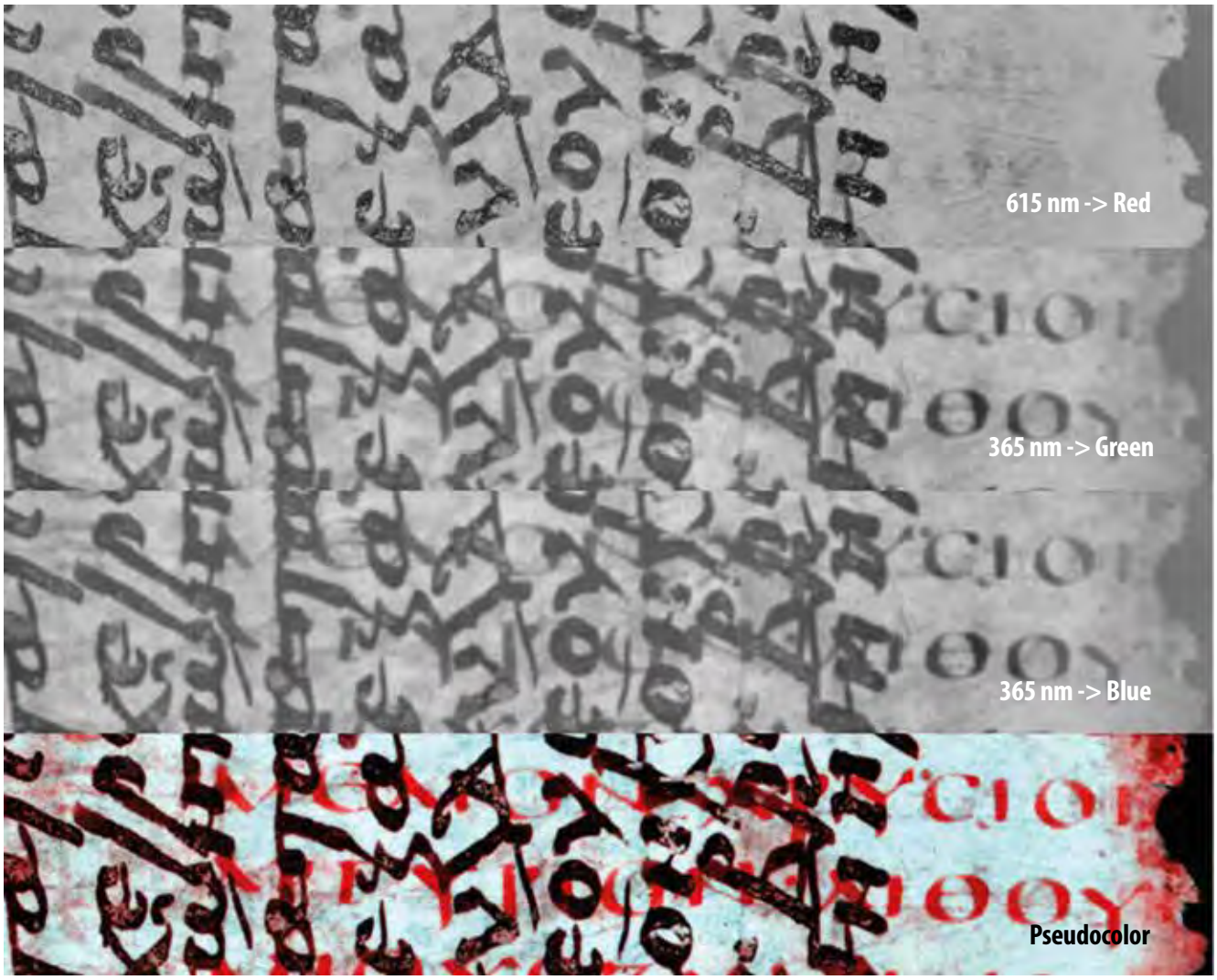


Fig. 4: Greek NF MG 99, fol. 3^{v}–4^{r}. A pseudocolor image is created from a visible image, in which the erased text cannot be seen, and a fluo escence image in which both texts are visible.

Fig. 5: The pseudocolor image (bottom) compared to the natural light image (top). The erased text is much easier to read in the pseudocolor image.

Fig. 6: The Ethiopic overwritten manuscript Petermann II Nachtrag 24 from the Berlin State Library.

2. Normal Optical Behavior

Iron gall ink was made by combining iron salts, such as iron sulfate, with tannic acids extracted from plants, such as galls on trees ('Iron Gall Ink' 2019). When dried, the ink turns a dark color on the parchment. The chemicals from the ink, the tannins, soak into the parchment underneath the characters. To erase the writing, the page can be either washed to remove the ink, or scraped to remove the ink and the top layer of the parchment. In either case, the tannins remain in the parchment in the exact shape of the characters that were removed.

After erasure, with time, the stains left behind darken. Although they become more visible, they can still be difficult to read and are not always clearly shaped (Fig. 3). Their color tends to be a faint reddish-brown. With visible reflectance images, these slight color differences can be detected. Being in the red region of the spectrum, they are more visible under blue light. In the infrared region, they almost completely disappear. In the blue region, they have a high contrast against the parchment.

Under ultraviolet illumination, the erased iron gall characters become very distinct. This is due to the natural fluorescence of the parchment under ultraviolet illumination. The tannins in the parchment, located in the regions of the characters, suppress the fluorescence of the parchment (Rabin 2018), leading to an increased contrast of the characters against the parchment. This increased contrast can even result in sharper edges of the characters, as the faint stains suppress fluorescence more strongly than their color differences are distinguishable. This leads to a simple method to enhance the erased text (Knox 2008). As shown in Fig. 4, the red or infrared image is inserted into the 'red' channel of a pseudocolor image, while the ultraviolet image is inserted into the green and blue channels of the same pseudocolor image. Since the erased text is bright in the red channel and dark in green and blue channels, it shows up as bright red in the pseudocolor image. The overwriting, on the other hand, is dark in all three channels, so it is seen as a neutral dark color. This color difference makes the erased text stand out against the parchment and the overwriting.

When the pseudocolor image is compared to the original natural light image, the results are striking (Fig. 5). In the natural light image, the erased text is faint and barely visible and the characters are not very sharply defined. In the region of the overwriting, the erased characters can be read only with great difficult .

On the other hand, in the pseudocolor image, the erased text is clearly visible, even in the region of the overwriting. The high contrast of the erased text in the ultraviolet, or fluorescence image leads to the high contrast of the edges of the erased text in the pseudocolor image. The color contrast of the erased text, combined with the black & white contrast of the fluorescenc image, gives the erased text in the pseudocolor image a higher contrast than it has in any of the individual single-wavelength images.

3. When Erased Characters Fluoresce

As we have seen already, the normal optical behavior of the stains left behind by erased iron gall ink is to suppress the fluorescence of the parchment, thereby increasing the contrast of the stains against the parchment. When the stains darken

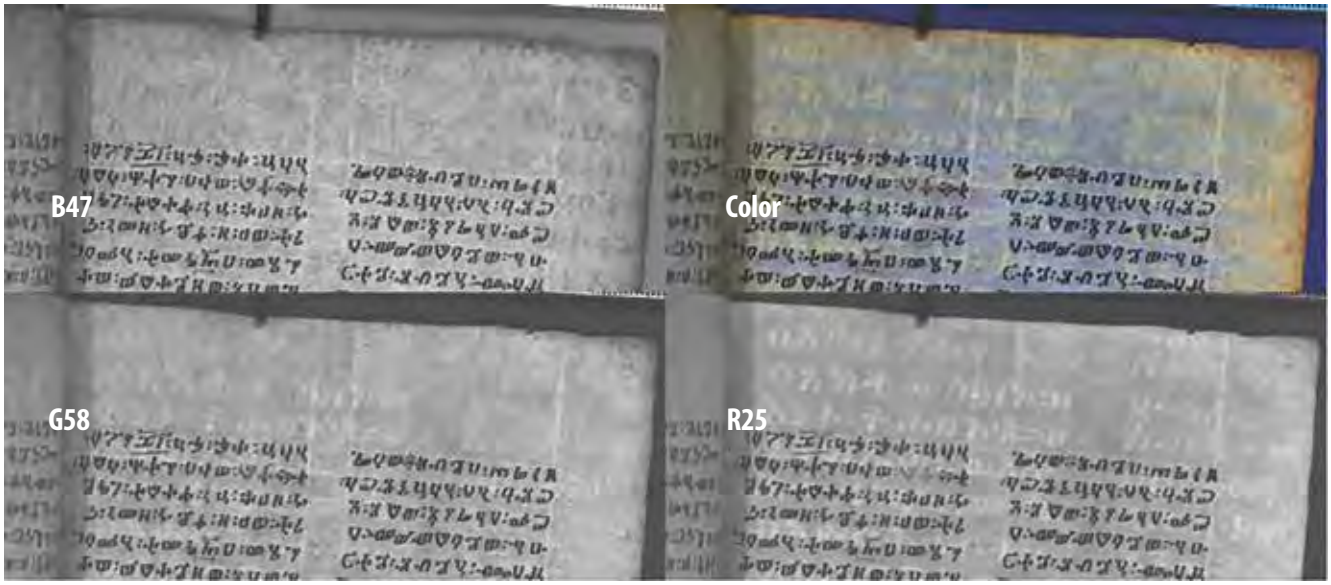


Fig. 7: The fluo escence of the erased text is different through the three color fil ers, B47 (blue), G58 (green) and R25 (red). The color fluo escence image is made from the three fil ered-images.

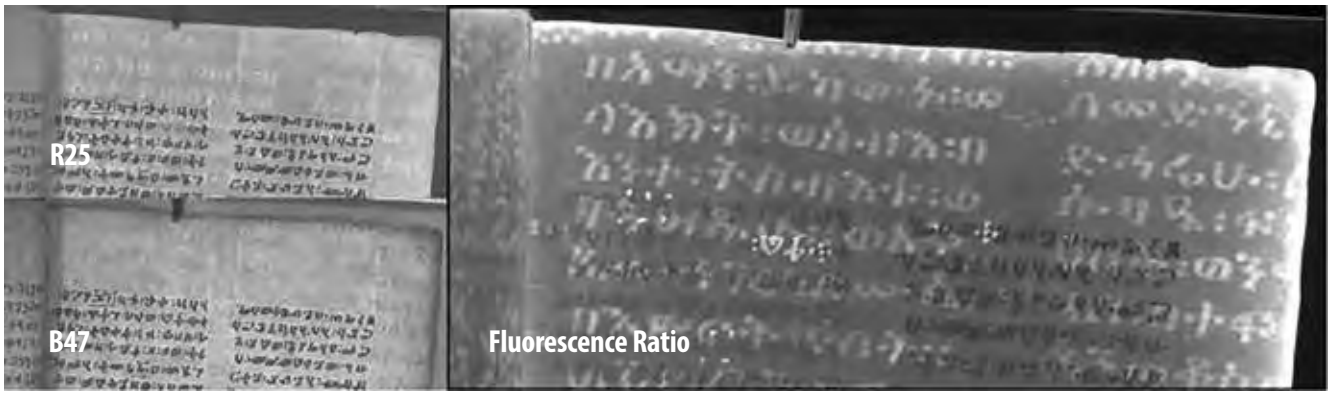


Fig. 8: By dividing the red-fil ered image (R25) by the blue-fil ered image (B47), the erased characters are signifi antly enhanced in the ratio image, shown on the right.

under ultraviolet illumination in this manner, a pseudocolor image can be used to enhance the legibility of the erased characters.

There have been a few times that we have seen cases where the stains do not suppress the parchment fluorescence, but instead fluoresce themselves. It is quite possible that these are not the result of erased iron gall ink, but we do not have sufficient information whether or not this is the case. In any event, the normal processes of combining the infrared and ultraviolet images does not enhance the erased text.

One such case was encountered in the imaging of the Petermann II Nachtrag 24 manuscript that is located in the Staatsbibliothek zu Berlin (Berlin State Library, Prussian Cultural Heritage Collection) (*Orient Digital* 2019). This is an Ethiopic overwritten manuscript written in Ge'ez, see Fig. 6. Because the erased text fluoresces, it is important to look at the colors of the fluorescence. This is accomplished at the camera by filtering the fluorescent light coming from the manuscript. In Fig. 7 are shown the filtered images of the fluorescence The three filters used are red, green, and blue Wratten filters namely, B47, G58 and R25 (*Orient Digital* 2019).

When the images through the three filters are combined in a color image, seen in Fig. 7, one can see that the characters are fluorescing in yellow. From the three individual images, one can see that the characters in the blue-filtered image are dark, while they are bright in the green-filtered image and very bright in the red-filtered image. This is a reasonable expectation from yellow fluorescence. Why these erased characters are fluorescing is not understood, but it is clearly a result of the residues that were left in the parchment after the ink was removed.

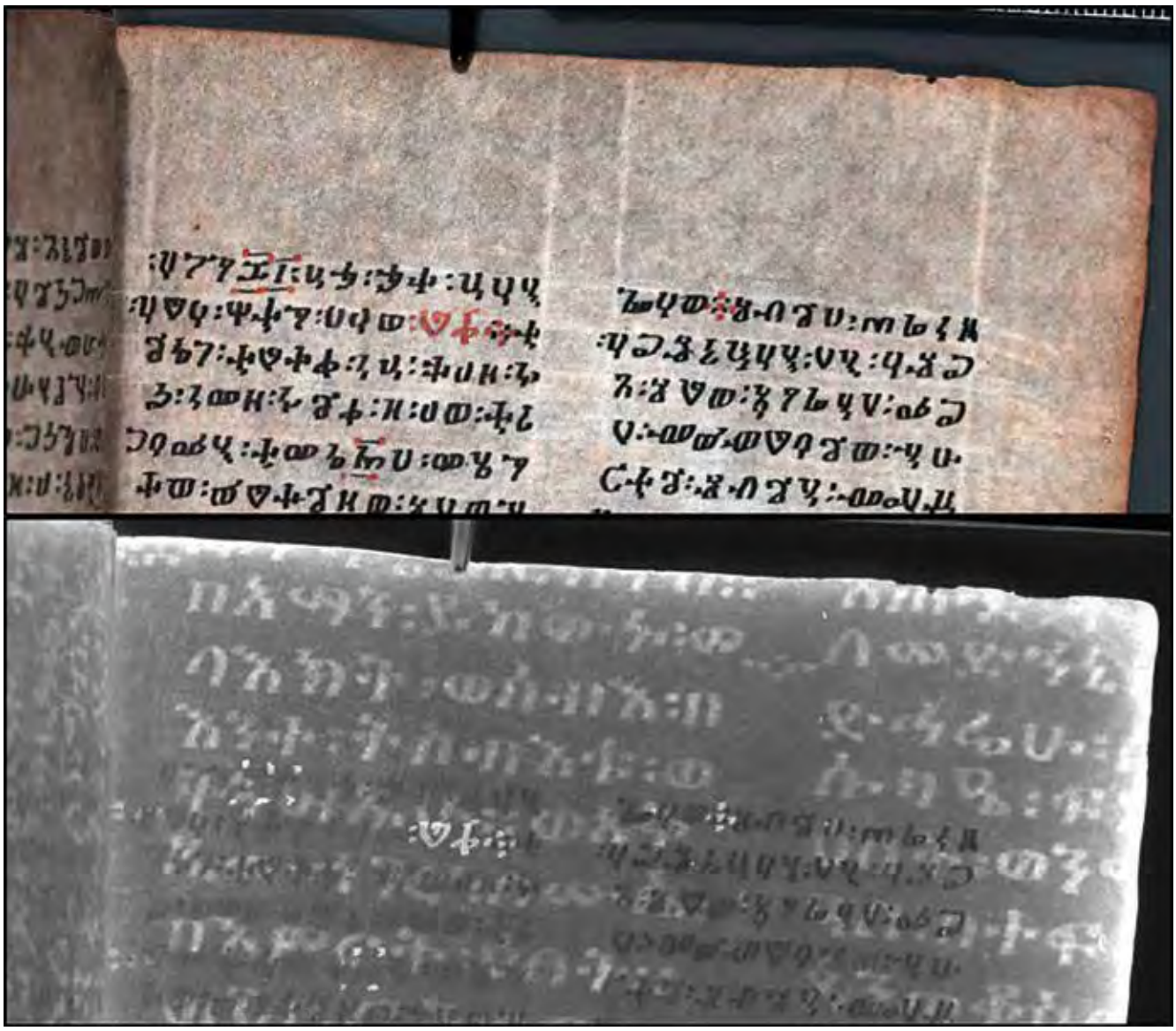


Fig. 9: The fluo escence ratio image (bottom) dramatically reveals the erased characters that fluo esce under ultraviolet illumination, but are barely visible to the eye in natural light (top).

Because the erased characters in the fluorescent images are not all dark, there is an opportunity to further enhance the erased text by properly combining the fluorescence results. Since the digital images are stored in the computer as arrays of numbers, where smaller numbers represent darker values and larger numbers represent brighter values, one can combine the values from different images mathematically. In Fig. 8 is shown the result of dividing the red-filtered image by the blue-filtered image. The result is to dramatically increase the values of the erased text, making them brighter in the resultant fluorescence ratio image

Although the fluorescence of the erased characters in this manuscript is not understood, the color of the fluorescence can be used to significantly enhance the legibility of the erased text. In Fig. 9, the fluorescence ratio image is compared to the same leaf under natural light. Erased characters, that are not visible to the eye, are distinctly revealed in the fluorescenc ratio image.

4. When Characters Erode the Parchment

Iron gall ink, made from tannic acid, can be harsh on parchment. Since parchment is made from skin, there are two sides to a leaf, the hair side and the flesh side. The hair side is fairly resistant to the ink, but the flesh side can be eroded away by the harshness of the ink. This erosion takes the exact shape of the character that was erased. The result is a cavity in the parchment in the shape of the missing erased character. The stains, that would have been left behind by the

Fig. 10: Georgian NF 13, folio 59r. The refle tance response of flesh side to visible and infrared illumination. The image labeled 940 nm TX is an infrared transmission image.

tannins in the parchment, have fallen away leaving behind a small cavity with no tannin stains. This condition happens frequently on the flesh side of the parchment leaves

On the hair side of a piece of parchment, the stains from the erased text respond well to ultraviolet light, yielding enhanced characters. Conversely, on the flesh side, the parchment surface often is eroded by the ink and the stains from the tannins are frequently no longer there. In their place, are small cavities that contain no tannin stains, and therefore do not respond to ultraviolet illumination.

The cavities that result from the eroded characters make the parchment a little thinner, right at the locations of the characters. As a result, when light is shown through the leaf, less light is scattered in the region of the characters, allowing more light to be transmitted through the parchment. This results in bright figures in the shape of the missing characters.

The response of a flesh-side page of a palimpsest from the library of Saint Catherine's Monastery can be seen in Fig. 10. In the color image, dark shapes can be just discerned in horizontal lines across the leaf. Nothing much is seen in reflectance images until the illumination enters the red and infrared region, where the characters turn dark. However, in infrared transmission, the erased text shows up as bright characters, as more light makes its way through the parchment.

Although the erased characters in the 940 nm transmission image look good, they can be improved by dividing the transmission image by the 940 nm reflectance image. Since the erased characters in the reflectance image are darker than the parchment, the division will enhance the bright characters in the transmission image. In addition, the overwriting on leaf is also suppressed by the division. The result of this transmission ratio can be seen in Fig. 11.

When compared to the natural light image, see Fig. 12, the transmission ratio image can reveal erased characters that do not respond to ultraviolet illumination. While we have seen this effect on several manuscripts, we do not see it every time, and we cannot predict when it will occur.

5. When Erased Characters Partially Erode the Parchment

The erosion of the parchment on the flesh side by the remains of the iron gall ink occurs frequently, but not always uniformly. In one project, working on the Jubilees manuscript at the Ambrosiana library in Milan, Italy, we found that the erosion was spotty. On the same leaf, some characters eroded and others nearby remained as tannin stains in the parchment. In other cases, parts of the same character eroded while other parts of the same character did not.

This feature can be seen in the response of one page of the Testament of Moses, folio 109, seen in Fig. 13. In the natural light image, erased characters are visible as stains in some parts of a text line and not visible in other parts of the same line. Even some characters are partially visible and partially not visible. In the blue reflectance image (470 nm), which is similar to an ultraviolet fluorescence image, some parts of the characters are visible. Conversely, in the 940 nm transmission image, the parts of the characters not visible in the blue reflectance image, show up as white characters in the transmission image.

This partial erosion and partial stain led to the development of what we called a Ruby image, which is a ratio of an ultraviolet fluorescence image by a transmission image. To utilize the full variation of the fluorescence and transmission

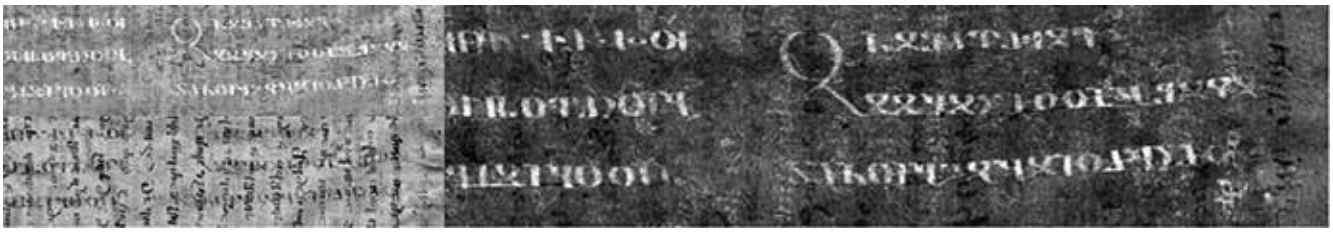


Fig. 11: The transmission ratio image (right) is obtained by dividing the 940 nm transmission image (top) by the 940 nm refle tance image (bottom).

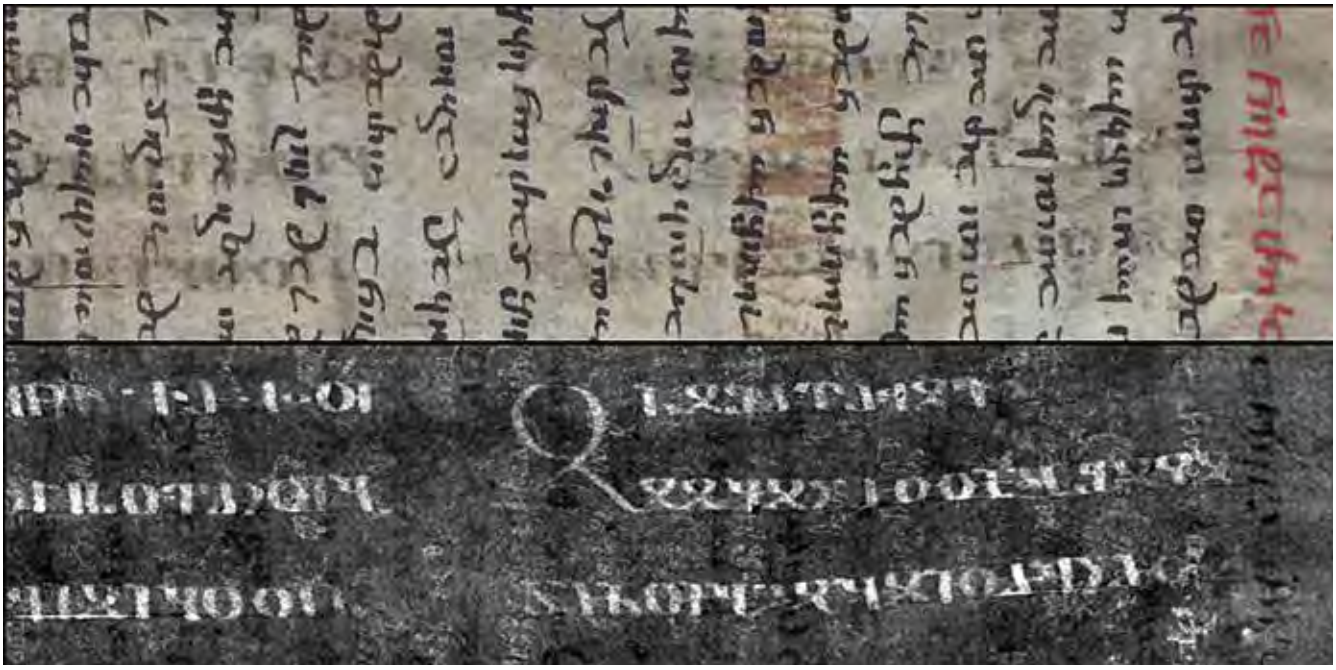


Fig. 12: The transmission ratio image (bottom) reveals many characters that are not visible in the natural light image (top).

responses, color images were made of each before taking the ratio. Fig. 14 shows the ratio of these two images.

By taking a ratio of the fluorescence image and the transmission image, each image fills in the parts of the characters that were missing from the other image. The result is that the ruby image has complete characters and complete text lines. A comparison of the ruby image with the natural light image is shown in Fig. 15.

6. Conclusions

The first tool to use in revealing erased iron gall ink characters is to view the manuscript under ultraviolet illumination. The fluorescing of the parchment, aided by the suppression of the fluorescence by the ink stains, enhances the erased characters. These characters can be further enhanced by combining individual reflectance and fluorescence images together in a pseudocolor image.

Utilizing the suppression of the fluorescence by the tannin stains does not always work. Sometimes, instead, the character residues, themselves, fluoresce. In this case, the color of the fluorescence can be used to enhance the erased characters by dividing two filtered images of the fluorescence. In particular, dividing two images where the erased characters respond differently, but the parchment response is the same. This enhances the contrast of the erased text in comparison to the background parchment around it.

Other times, often on the flesh side of the parchment, the ink may have eroded away some of the parchment, leaving it thinner underneath the writing. This erosion can result in parchment in which there are no stains remaining from the erased text. The lack of tannin stains in the parchment means that the erased characters do not respond to ultraviolet illumination, either in suppression of fluorescence or in fluorescing themselves. This erosion

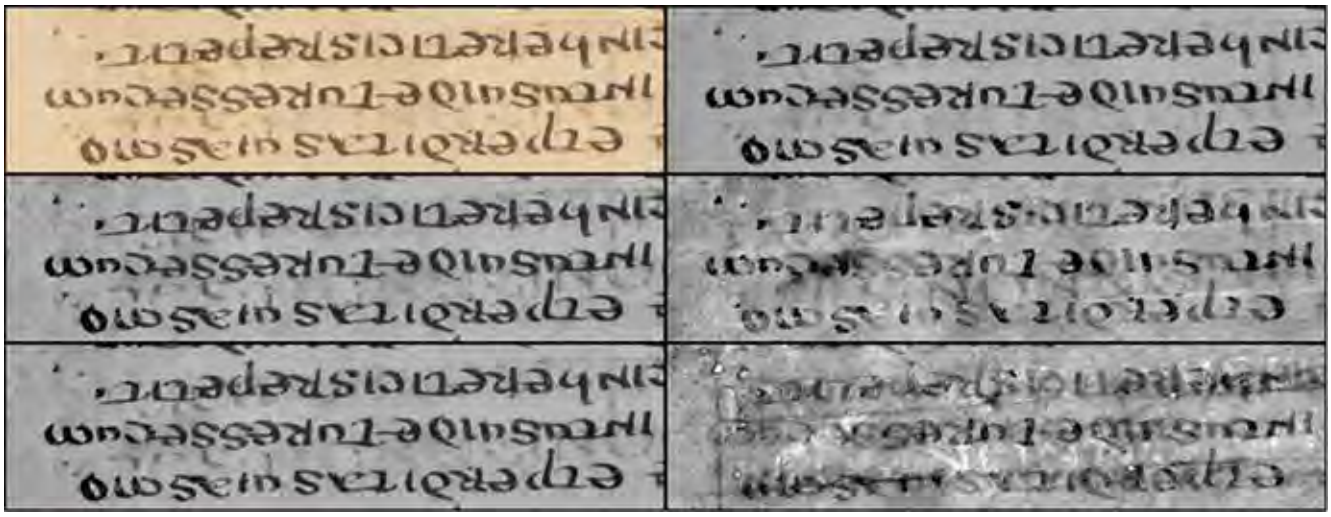


Fig. 13: The natural light image shows traces of the erased text. The refle tance images show some of the characters, but not all. The transmission image shows the parts of the characters not seen in the refle tance images.

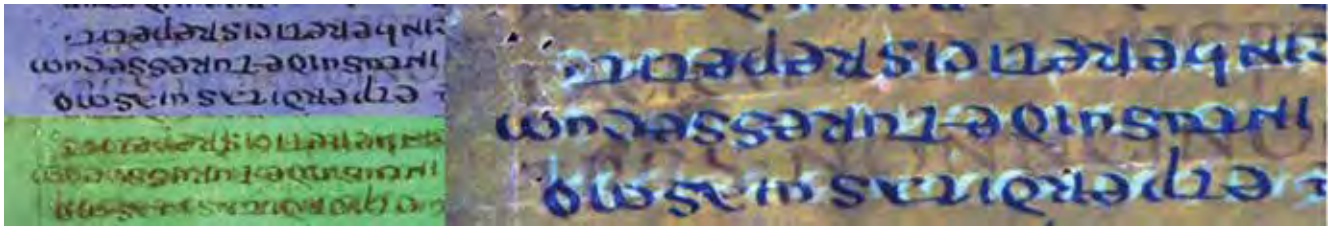


Fig. 14: The ruby image (right) is obtained by dividing a color transmission image (top) by a color fluo escence image (bottom).

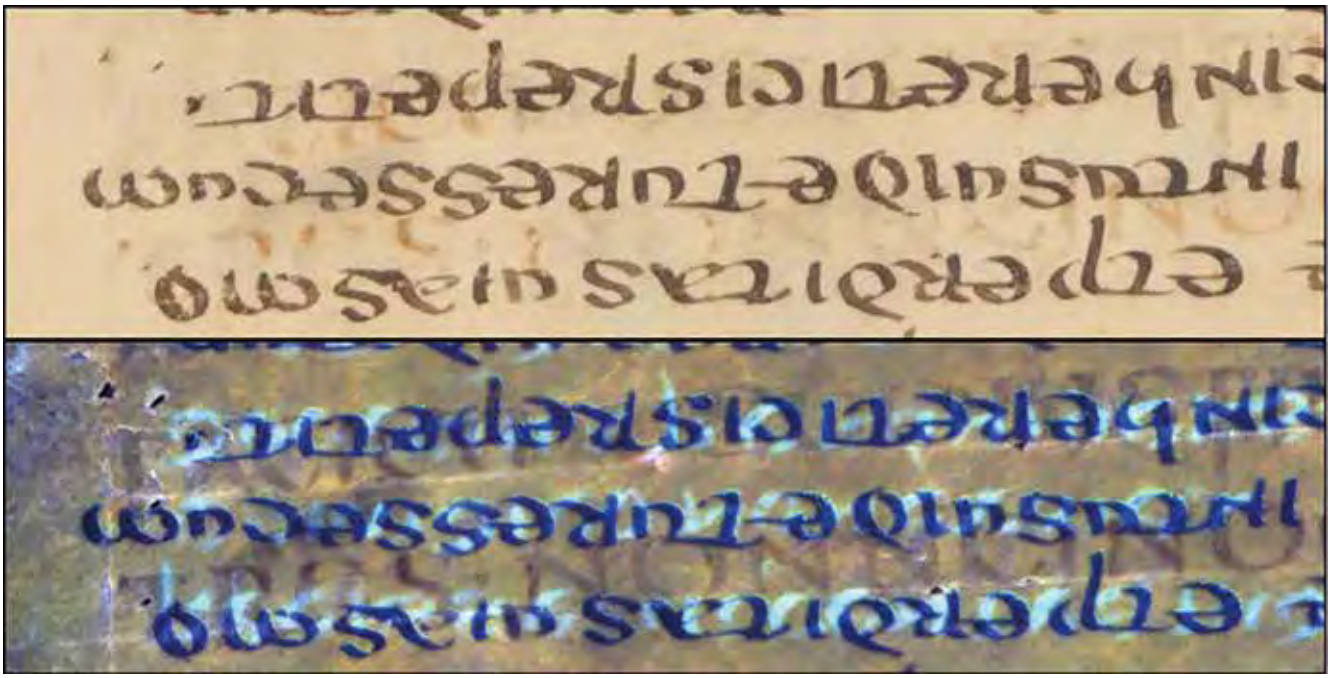


Fig. 15: The natural light image (top) in comparison with the ruby image (bottom).

phenomenon, though, does make the parchment thinner, meaning that more light can be transmitted through the parchment revealing the missing erased characters. At the same time, the infrared reflectance image partially reveals the missing characters as darker than the parchment, rather than lighter as in the transmission image. This may be due to less light being reflected by the parchment, since there is less parchment in the region of the missing characters. As a result, the ratio of infrared transmission image by the infrared reflectance image reveals and enhances the missing text.

In this paper, we have shown four different image processing methods, three of which that were developed over the last decade to handle unique cases where the erased iron gall ink writings do not behave in the normal manner. The three methods are based on combining different types of images captured in the multispectral imaging of the manuscript, depending on which types of images contain the information about the text. We do not believe that these are the only viable methods available, nor the only ways in which erased characters behave. Based on our past experience, we can almost guarantee that additional ways will be found in which erased iron gall ink will misbehave.

7. Acknowledgments

The author would like to acknowledge his colleagues of the past two decades, without whose help these techniques may not have been developed. On the Archimedes project (2000–2010), where the pseudocolor technique was developed, thanks are due to image processing colleagues, Roger Easton, Jr. and William Christens-Barry. The transmission ratio method was developed as part of the Sinai Palimpsests Project (2011-2016), which benefitted from the Director of the Early Manuscripts Electronic Library, Michael Phelps, camera operator Damianos Kasotakis and image processor, David Kelbe. The Petermann palimpsest project (2016), was conducted to help the scholars, Loren Stuckenbruck and Ted Erho. It was greatly aided by Ira Rabin, who loaned us the equipment and made available a member of her staff, Ivan Shevchuk. Lastly, the author's thanks also go to Todd Hanneken who made it possible to work on the Jubilees palimpsest at the Ambrosiana library in Milan, Italy.

REFERENCES

'Archival and Cultural Heritage Imaging' (2019), in *MegaVision* Website <http://www.mega-vision.com/cultural_heritage.html>.

Easton, Roger L. Jr., K. T. Knox, W. A. Christens-Barry, K. Boydston, M. B. Toth, D. Emery, and W. Noel (2010), 'Standardized system for multispectral imaging of palimpsests', *Proceedings of SPIE*, vol. 7531: *Computer Vision and Image Analysis of Art*, (San Jose, CA: SPIE). doi: 10.1117/12.839116

'Iron Gall Ink' (2019), in *Wikipedia*, <https://en.wikipedia.org/wiki/Iron_gall_ink> (last accessed November 2019).

Knox, K. T. (2008), 'Enhancement of overwritten text in the Archimedes Palimpsest', Proceedings of SPIE, vol. 6810: *Computer Image Analysis in the Study of Art*, (San Jose, CA: SPIE). doi:10.1117/12.766679

Orient Digital – Datenbank der orientalischen Handschriften der Staatsbibliothek zu Berlin http://orient-digital.staatsbibliothek-berlin.de/receive/SBBMSBook_islamhs_00025151?lang=en (last accessed November 2019).

Rabin, Ira (2018), Private communication.

'Wratten Number' (2019), in *Wikipedia* <https://en.wikipedia.org/wiki/Wratten_number> (last accessed November 2019).

PICTURE CREDITS

Figs 1–2: © Keith T. Knox.

Figs 3–5: © Holy Monastery of St. Catherine at Mount Sinai, Egypt.

Fig 6: © Staatsbibliothek zu Berlin (Berlin State Library); photography: © Keith T. Knox.

Figs 7–9: © Keith T. Knox.

Figs 10–12: © Holy Monastery of St. Catherine at Mount Sinai, Egypt.

Figs 13–15: © Keith T. Knox.

Article

'Dürer's Young Hare' in Weimar – A Pilot Study

Oliver Hahn, Uwe Golle, Carsten Wintermann, and Ira Rabin | Berlin, Hamburg, Weimar

Abstract

The investigation of physical properties and chemical compositions produces data that is important for answering questions that cannot be solved by historical and philological methods alone. The results presented in this study show the degree to which scientific investigations can supplement and in part amend research on cultural history. Using scanning X-ray fluorescence spectroscopy and additional techniques, the non-invasive investigation of the 'poor-quality' copy of Albrecht Dürer's famous work *Young Hare* (Ger. *Feldhase*) shifts the *Weimar Hare* back into the focus of art historians' interest and could lead to a re-evaluation of this drawing.

Keywords

Non-invasive investigation, imaging techniques, scanning X-ray fluorescence, drawings, Albrecht Dürer

1. Introduction

The Weimar Classics Foundation preserves a coloured drawing on parchment that is generally regarded as a low-quality copy of the *Hare* drawn by Albrecht Dürer (1471–1528) and now on show at Vienna's Albertina Museum (Fig. 1). The drawing displays Albrecht Dürer's monogram. Hans Hoffmann (*c.* 1530–1592), Joris Hoefnagel (1542–1600), Jacob Hofnagel (1573–1632/1633) and possibly Giuseppe Bossi (1777–1813) are discussed as possible draftsmen. The dendrochronological investigation of the beechwood panel on which the parchment is mounted shows that the tree was felled no earlier than 1543. In Figure 2, we overlay two images of the Weimar drawing and the Vienna drawing to illustrate that the Weimar copy has exactly the same dimensions as the Vienna *Hare*.

The *Hare* that Grand Duke Carl August of Saxony-Weimar-Eisenach received from Giuseppe Bossi's heir in 1817 calls for special consideration, mostly because it is one of two alleged copies of the *Hare* from Albertina that were drawn on parchment. The second known copy on parchment is on sale in Cologne (Kunsthaus Lempertz 2008). A similar drawing on paper, but in a much better state

Fig. 1: Copy of *Young Hare* by Albrecht Dürer; Klassik Stiftung Weimar, inv. no. 507379, present state of the drawing.

Fig. 2: Overlaying the *Young Hare* from Albertina Museum (Albertina, Vienna, inv. no. 3073, red colour) with the Weimar copy (blue colour).

Figs. 3a, b: Infrared reflectography of the 'Weimar Young Hare', ~1000 nm, camera: X71, Microbox without a filter. Left (a): IR image; right (b): IR false-colour image.

of conservation than the Weimar one, is in Paris.[1] According to art-historical research, the two copies on parchment are very similar; even the bristly hair along the contours seems to correspond. According to Christof Metzger, the fur structure and the construction of the 'delicate animal' display certain schematisms that make both copies distinct from the original (Metzger 2014 and 2016).

The Weimar sheet is considered to have been proven to be a precise replica (Metzger 2016) that was produced in the following way: the artist transferred the original drawing onto transparent parchment and then painted it with opaque pigments. This required access to Dürer's original, which considerably limits the circle of its possible creators.

We will first consider the provenance of the drawing from the legacy of Giuseppe Bossi. How did it come to be part of the Weimar inventory? After the early death of the artist, art collector and art historian Giuseppe Bossi (1777–1815), whom we will only introduce briefly in this article, strong interest developed in his oeuvre.

Giuseppe Bossi led an extraordinary life. He studied painting at the Accademia di Brera in Milan and became a famous copyist. He performed nature studies (and even dissected bodies in an attempt to grasp how they functioned), he collected art objects along with books and manuscripts for his own library and, finally, he composed poems and wrote treatises. His collection of art objects came to form the core stock of two different museums: The Archaeological Museum in Milan and the Graphic Collection of the Gallerie dell'Accademia in Venice (Zanaboni 2016).

In the summer of 1817, Grand Duke Carl August of Saxe-Weimar-Eisenach (1757–1828) decided to take a trip to Switzerland and northern Italy. One of his most important destinations was Milan. The documents about this journey have been lost, so it is no longer possible to say why precisely Milan aroused Carl August's interest. Perhaps it was due to his acquaintanceship with the Frankfurt and Milanese banker Heinrich Mylius (1769–1854), who then introduced the Grand Duke from Weimar to the Director of Milan's Coin Collection, Gaetano Cattaneo (1771–1841). Cattaneo offered the Grand Duke first choice of some items from Bossi's legacy. Here, too, we have no precise information about the offerings. But it is a fact that the Grand Duke acquired Giuseppe Bossi's carton from 'Parnassus', along with what is called the 'Lucide' – a bundle of three tracings made by Bossi copying Leonardo's *Last Supper*, an acquisition that reflects the Weimar court's fascination with Leonardo da Vinci at that time. The objects were sent from Milan to Weimar. To Carl August's great pleasure, Cattaneo was able to obtain five additional items from Bossi's heirs, including the 'Weimar Dürer Hare', which arrived in Weimar on 13 November 1817 along with the

[1] 'Lièvre couché, tourné vers la droite', Musée du Louvre, inv. no. RF 29072; see Ministère de la culture, *POP*: *la plateforme ouverte du patrimoine* <https://www.pop.culture.gouv.fr/notice/joconde/50350228452>.

Figs. 4a, b: Element distribution images of iron (Fe) and lead (Pb). Left (a): iron; right (b): lead.

other items. Even now, it is still not clear how this drawing came to be in Bossi's possession.

The drawing was already subjected to several individual examinations to determine its colorants. The investigations revealed chalk, white lead, ochre and carbon black as colorants, all of them materials that were already in use in the sixteenth century. In addition, in this first investigation, a conspicuously high zinc content was detected in several spots. Here, additional X-ray imaging scans were performed to clarify whether a zinc-containing pigment was used for the drawing. If this was the case, it might have been zinc white, which would indicate that the drawing was not executed before the first third of the nineteenth century (Hahn 2018)

All in all, the drawing is poorly preserved; the surface is very dark, in part abraded, and it displays mould stains, possibly from water damage. Some sections are no longer discernible (see Fig. 1). Material-science investigations that emphasise imaging techniques should therefore reveal something about the process of the drawing's development (the production of the copy) and form the basis of a re-evaluation of the object. In addition, we wish to illuminate the 'fate' of the drawing to understand the bad state of its preservation.

2. Methods and imaging techniques

Scientific procedures have become an established component of the examination of cultural artefacts and of the development of preservation concepts. Many devices fulfil the conditions of non-destructive analyses, so investigations can be carried out without taking any samples. In particular, imaging techniques like computed tomography, radiography and multi-spectral imaging analysis are increasingly being used to give visibility to contents that are invisible to the naked eye, like preliminary drawings, alterations and inner structures. The resulting 'new pictures' can give us deeper insights into the production process, the structure or the state of preservation. They also document the changes an object has been subjected to.

2.1 Infrared reflectography (IRR)

The wavelengths useful in examining cultural artefacts with infrared light are generally divided into 'near IR', or NIR (700–1000 nm), and 'mid-IR', or MIR (1000–3000 nm). Many materials exhibit different visual appearances under specific wavelengths of IR radiation, absorbing, transmitting or reflecting the radiation. Materials that absorb IR radiation appear dark, those that transmit it are transparent and those that reflect it appear to be white, although the appearance of a material can change with different IR wavelengths. What is most important in these investigations is the performance of certain inks and drawing media. Carbon inks, graphite, charcoal and metal points all absorb IR strongly, while plant inks and iron-gall ink absorb shorter-wavelength IR and become increasingly transparent under longer wavelengths

Fig. 5a: Detail of Dürer's Young Hare from the Albertina.

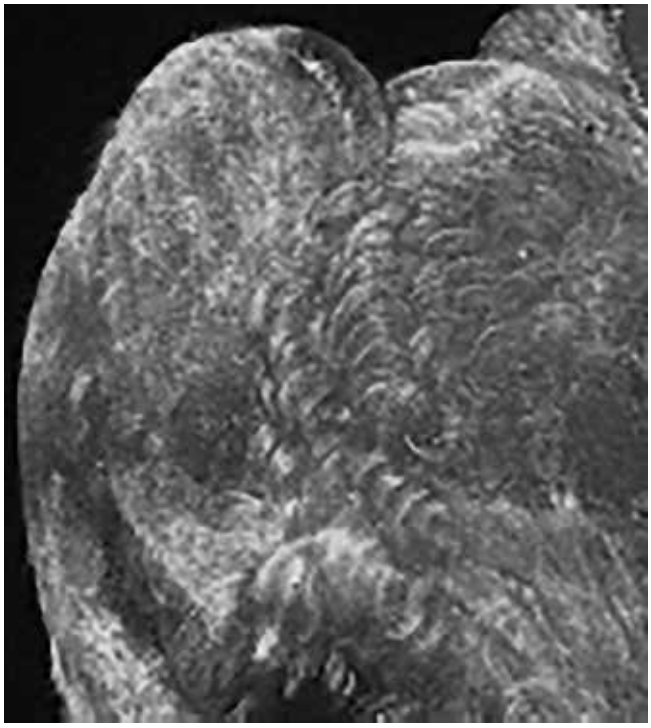


Fig. 5b: Detail of the element distribution image of the lead (Pb) in the Hare from the Klassik Stiftung Weimar.

(Mrusek, Fuchs and Oltrogge 1995). In general, the method was established to reveal underdrawings in paintings in the 1960s (Asperen de Boer 1969), given that many underdrawings were executed with charcoal or metal points.

2.2 X-ray fluorescence analysis (XRF)

X-ray fluorescence analysis (RFA) is a classic method for non-destructively examining the elemental composition of materials, in particular in the field of cultural heritage (Mantler and Schreiner 2000; Mommsen et al. 1996; Klockenkämper 1997; Vandenabeele et al. 2000; Hahn, Reiche and Stege 2006). According to Lahanier, Preusser and Van Zelst 1986, the ideal procedure for analysis should be non-destructive, fast, universal and versatile. It is widely known that the sample is exposed to high-energy X-rays during the measurement. The radiation interacts with the material, whose excited atoms now emit their characteristic radiation. A suitable detector detects this X-ray fluorescence, providing information about the material's elemental composition. The energy of the emitted X-rays is characteristic of a specific chemical element, and the intensity of the signal permits conclusions to be made about the amount present. Using this procedure, an analysis of organic materials is either impossible or only possible to a very limited degree.

In the current investigation, the coloured drawing from Weimar was not subjected to any individual measurements, but analysed with an X-ray fluorescence portal scanner. The RFA scan was conducted with the Jet Stream made by the company Bruker Nano GmbH with rhodium X-ray tubes (Alfeld et al. 2013). During the examination, the probe, which contains the excitation tube (Rh, 50 kV, 0.6mA), two cameras and the detector (Flash™ detector), moves above the object at high speed (diameter of the excitation spot: 50–600 µm, step width: 25–200 µm; dwelling time: 2–900 ms per point). This produces a complete cartography of the various elements, 'element distribution images', within the various sections. The device is conceived in such a way that the measurements are conducted in an ambient atmosphere. The penetration depth of the radiation is between 10_{-6} and 10_{-1} cm, depending on the matrix (Hahn-Weinheimer et al. 1995).

Additional techniques were used for the investigation of the drawing, namely VIS spectroscopy (Fuchs and Oltrogge 1994) and FTIR in diffuse reflection (DRIFTS) (Steger et al. 2019).

3. Results and discussion

3.1 Infrared reflectography (IRR)

Figure 3a below displays the IR image that was taken at ~1000nm. It is obvious that the black colorant strongly

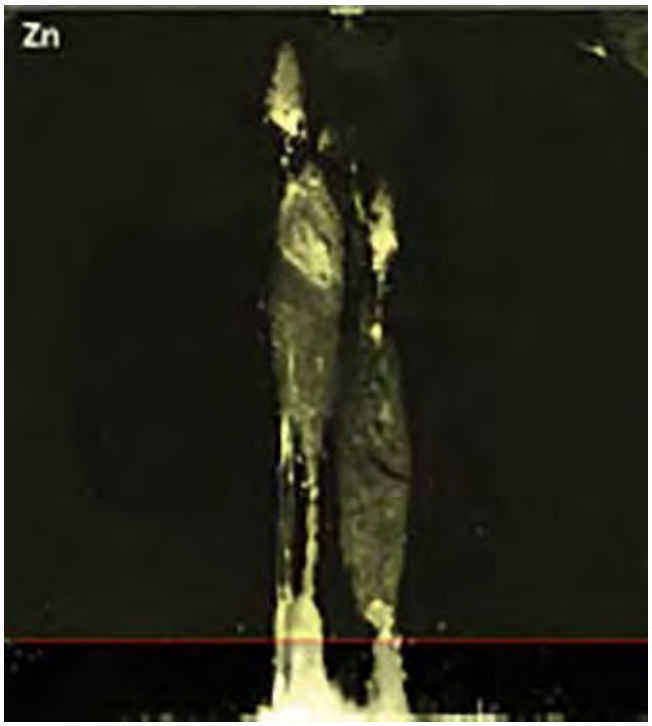


Fig. 6: Element distribution image of zinc (Zn).

absorbs the infrared light and could be identified as carbon black. The other colorants lost their opacity. Taking into account the XRF, the DRIFTS and the VIS results, the colorants have to be considered to be various ochres (iron oxide pigments), lead white and chalk. These pigments were widely used in the fifteent , sixteenth and seventeenth century and thus provide little evidence that the *Weimar Hare* was produced in the sixteenth century. However, the beechwood panel on which the parchment is mounted shows that the tree was felled in the middle of the sixteenth century.

The most important result, however, is that there is no evidence of a preliminary sketch. This is a very important indication about the production of this object. The absence of any preliminary drawing supports the assumption that this drawing is a copy that was prepared by tracing the original drawing onto transparent parchment and then it was painted with opaque pigments afterwards. In contrast to the Albertina *Hare* with the monogram 'AD' that was executed in carbon ink, the monogram 'AD' of the Weimar copy disappears under infrared light – apparently, it was executed with iron-gall ink. The NIR false-colour image in Figure 3b reveals the damage to the drawing. Many areas were abraded – this effect was possibly caused by water damage. Vertical 'structures' become obvious, especially in the lower part of the drawing. Infrared false-colour photography is generally used for the tentative identification of pigments and the discovery of retouching. Congruously, with the help of references, this technique confirms the assumption that, unlike the Vienna drawing, the monogram was written in iron-gall ink.

3.2 X-ray fluorescence analysis (XRF)

The most important results were obtained with X-ray fluorescence analysis. Figures 4a and b show element distribution images of iron (Fe) and lead (Pb). A comparison with the image in Figure 1 makes it clear that the drawing of the fur was much more delicately executed than is visible to the viewer today. This precision is no longer visible in the drawing due to the picture's poor state of preservation.

A closer look at the element distribution images reveals – even to a layperson – that this drawing cannot be a pure copy of Dürer's *Young Hare* despite it having the same dimensions (see Figs 5a, b). The painter's interpretation of the fur is quite different, for one thing. The *Hare* in the Albertina Museum depicts the individual hairs of the animal as straight lines, which probably comes closer to reality, whereas the hairs in the *Weimar Hare* are curled. This indicates that the Weimar drawing differs considerably from the Vienna one (Mildenberger 2018). It may, in fact, have been copied from an earlier drawing by Albrecht Dürer. This is pure hypothesis, though, and awaits verification by comparative art-historical analyses.

Finally, the distribution image of zinc (Zn; Fig. 6) explains the poor state of the drawing's preservation. A substance – probably a paint containing zinc white – ran across the central section of the drawing and damaged it; the streaks and smears can be seen clearly on the element distribution image. Conversely, this picture also shows that no pigment containing zinc was used on large parts of the drawing. The presence of the element is likely to be due to damage, the traces of which were removed inadequately; overall, it had a powerful deleterious effect on the drawing. It is noteworthy that the *Hare* must have experienced this dramatic damage while in the Weimar collection as the appearance of zinc from zinc white paint indicates the earliest possible use was in the 1840s, i.e. around 25 years after the *Hare* had been added to the collection.

4. Conclusion

The results presented here show the degree to which scientifi investigations can supplement and in part amend historical research. Precise imaging techniques can expand our access to

material artifacts because they afford insights into the history of such objects, that are subject to constant change in general. In the meantime, the discovery of preliminary sketches and evidence of corrections and alterations have been established in art history. Processes of corrosion and ageing that gradually alter cultural artifacts and rob them of their original colourfulness are under controversial discussion, however. In this case, material science can provide us with an impetus to correct accustomed ways of seeing things.

The poor state of preservation of the 'Dürer Hare' in Weimar ensured that the drawing had the reputation of being a 'poor-quality' copy for many years. Knowledge derived from material analysis coupled with other results – comparing the dimensions of the drawings in Vienna and in Weimar, for example – shifts the *Hare* back into the focus of art historians' interest and could eventually lead to a re-evaluation of the work. In this context, it could clearly be demonstrated that the drawing was damaged at least 25 years after its arrival in Weimar; in 1817, its state of preservation must have been considerably better.

A number of questions still remain unanswered. Needless to say, the copy cannot possibly be any older than the original. However, why does the *Weimar Hare* display a painting concept apparently older than (or simply different to) the execution of the fur in the *Vienna Hare*? Making such an accurate copy – and transferring its dimensions so well, for instance – will have required access to Dürer's original drawing. How did this drawing become part of Giuseppe Bossi's collection? These and other questions still have to be addressed and resolved by art historians.

ACKNOWLEDGEMENTS

The generous funding by the Deutsche Forschungsgemeinschaft, DFG (SFB 950, 'Manuskriptkulturen in Asien, Afrika und Europa') is gratefully acknowledged.

REFERENCES

Alfeld, M., V. Pedroso, J. van Eikema, M. Hommes, G. van der Snickt, G. Tauber, J. Blaas, M. Haschke, K. Erler, J. Dik, and K. Janssens (2013), 'A mobile instrument for *in situ* scanning macro-XRF investigation of historical paintings', *Journal of Analytical Atomic Spectrometry*, 28, 760–767.

Asperen de Boer, J. R. J. van (1969), 'Reflectography of Paintings Using an Infrared Vidicon Television System', *Studies in Conservation*, 14/3: 96–118.

Fuchs, Robert and Doris Oltrogge (1994), 'Painting materials and painting technique in the Book of Kells', in Felicity O'Mahony (ed.), *The Book of Kells. Proceedings of a Conference at Trinity College Dublin, 6–9 September* 1992 (Aldershot: Scolar Press), 133–171, 147–191, 603.

Hahn, Oliver, Ina Reiche and Heike Stege (2006), 'Applications in Arts and Archaeology', in Norbert Langhoff, Reiner Wedell, Burkhard Beckhoff and Birgit Kanngießer (eds), *Handbook of Practical X-Ray Fluorescence Analysis* (Berlin: Springer-Verlag), 687–700.

—— (2018), *BAM Test Report 2018* (unpublished).

Hahn-Weinheimer, Paula, Alfred Hirner, and Klaus Weber-Diefenbach (1995), *Röntgenfluoreszenzanalytische Methoden: Grundlagen und praktische Anwendungen in den Geo-, Material- und Umweltwissenschaften* (Wiesbaden: Vieweg-Verlag).

Klockenkämper, Reinhold (1997), *Total-Reflection X-Ray Fluorescence Analysis* (New York: Wiley).

Kunsthaus Lempertz (2008), *Auktion 920: Alte Kunst*, 'Lot 1081: Hans Hoffmann – Feldhase', <https://www.lempertz.com/de/kataloge/lot/920-1/1081-hans-hoffmann.html> (last accessed 19 November 2020).

Lahanier, C., F. D. Preusser, L. van Zelst (1986), 'Study and Conservation of Museum Objects: Use of Classical Analytical Techniques', *Nuclear Instruments and Methods Physics Research B*, 14/1: 1–9.

Mantler, Michael, Manfred Schreiner (2000), *'X-Ray Fluorescence Spectrometry in Art and Archaeology', X-Ray Spectrometry*, 29/1: 3–17, <doi: 10.1002/(SICI)1097-4539(200001/02)29:1<3::AID-XRS398>3.0.CO;2-O>.

Metzger, Christof (2014) ,'"Lieben, lächeln und sich erinnern": Albrecht Dürers Hase', in Klaus Albrecht Schröder (ed.), Die Gründung der Albertina: 100 Meisterwerke der Sammlung, (Ostfildern: Hatje-Cantz), 49-56.

—— (2016), 'Kat. 73: Jacob Hoefnagel (Antwerpen 1573-1633/32) nach Albrecht Dürer (1471-1528), Feldhase', in Hermann Mildenberger and Serena Zanaboni (eds), *Von Leonardo fasziniert: Guiseppe Bossi und Goethe* (Berlin: Sandstein Verlag), 194-195.

Mildenberger, Hermann (2018), Personal communication.

Mommsen, H., T. Beier, H. Dittmann, D. Heimermann, A. Hein, A. Rosenberg, M. Boghardt (1996), 'X-Ray Fluorescence Analysis with Synchroton Radiation on the Inks and Papers of Incunabula', *Archaeometry*, 38: 347–357, <doi: 10.1111/j.1475-4754.1996.tb00782.x>.

Mrusek, Ralf, Robert Fuchs and Doris Oltrogge (1995), 'Spektrale Fenster zur Vergangenheit – Ein neues Reflektographieverfahren zur Untersuchung von Buchmalerei und historischem Schriftgut', *Naturwissenschaften*, 82: 68–79, <doi: 10.1007/BF01140144>.

Steger, Simon, Heike Stege, Simone Bretz and Oliver Hahn (2019), 'A complementary spectroscopic approach for the non-invasive in situ identification of synthetic organic pigments in modern reverse paintings on glass (1913–1946)', *Journal of Cultural Heritage*, 38: 20–28.

Vandenabeele, P., A. van Bohlen, L. Moens, R. Klockenkämper, F. Joukes, G. Dewispelaere (2000), 'Spectroscopic Examination of Two Egyptian Masks – a Combined Method Approach', *Analytical Letters*, 33: 3315–3332, <doi: 10.1080/00032719.2000.10399503>

Zanaboni, Serena (2016), '"Seine Schönheit erinnerte an diejenige der römisch-hellenischen Antike", Giuseppe Bossi und die Erwerbungen von Carl August in Mailand 1817–1818: Goethe und die spätklassizistische Rezeption in Weimar', in Hermann Mildenberger and Serena Zanaboni (eds), *Guiseppe Bossi und Goethe: von Leonardo fasziniert* (Berlin: Sandstein Verlag).

PICTURE CREDITS

Figs 1–4: © Carsten Wintermann, Klassik Stiftung Weimar.

Fig. 5a: © Albertina Museum, Vienna.

Fig. 5b: © Uwe Golle, Klassik Stiftung Weimar.

Fig. 6: © The authors.

Article

Papyrus as a Writing Support

Myriam Krutzsch | Berlin

Abstract

Papyrus was used as a writing support for about 4,000 years (*c*.3000 BCE – *c*.1000 CE) and displays varying qualities that can be localised temporally and geographically. It is astonishing that the quality of the oldest papyri from the Fifth Dynasty (the Old Kingdom of Egypt) is characterised by their regular fibre structure, fineness and very thin sheets. The Middle Kingdom also used papyri of high quality; only a practised eye can distinguish the difference in quality between the recto and the verso. Papyri from the time of the pharaohs are translucent, sometimes so much so that the writing on the reverse side can be recognised on the upper side.

This high degree of transparency disappeared in Greco-Roman times as the material became firme , thicker and more compact. The first reason for this is that a different writing implement was used: the *kalamos*, which was much harder and stiffer than the soft, brush-like *bulrush* of pharaonic times. Second, the constantly growing demand for papyrus material led to a decline in its quality. Along with the poor and usually coarse papyrus of Arabian times, though, a conspicuously fine, high-quality material was used. It is unclear how this 'renaissance' of papyrus came about, however.

First of all, the typical structure of papyrus as a writing material will be explained here, which is fundamentally different from paper. The differences lie both in the source material and in the manufacturing method, and therefore result in different consistency, structure and behaviour of the produced writing support.

1. Description of the material

Conservation projects always begin with an assessment of the state of an object's preservation. Over the centuries, the original character of papyrus can change, but still be recognisable. A material description of papyrus consists of the measurement of its dimensions and a number of other criteria that are documented in a conservation protocol along with the current state of the material, namely:

1. its colour
2. the production method used
3. fibre
4. sheets
5. sheet joins
6. roll ends.

1.1 Colour

The original colour of the papyrus sheets that were examined was presumably light ochre, yellow or brown. The colour, and particularly its susceptibility to change, depends on various factors (Table 1).

Since the colour of ancient papyrus can continue to change, depending on what the climatic or storage conditions are like, it cannot be taken into consideration as an indication of the exact location when reconstructing its background.

1.2 Production methods

To date, we only know of one written source mentioning the production of papyrus as a writing support – that of Pliny the Elder.[3] His records are very simple and concise; there are just notes, and yet two different

Table 1: Colour changes in papyrus.

Cause	Eff ct	Example
UV radiation	Bleaches the papyrus and damages the cellulose ↳ consistency becomes soft and rotten	Berlin P 9875
Water[1] and other liquids	Browning and cellulose decay; spots ↳ consistency becomes brittle and often frayed	Berlin P 15995 - 98
Heat and fi e	Browning and destruction of the cellulose ↳ extreme brittleness to the point of disintegration	Elephantine 26774a[2]

[1] Constant, long-lasting influence from water is meant here, for example due to the annual flooding of the Nile

[2] The current location of this papyrus document is unknown.

[3] In his *Naturalis Historia*, Pliny the Elder dedicated several chapters of book 13 to papyrus as a writing material, its production and use (chapters XXI–XXVII); on the production of it, see chapter 13, XXIII.

production methods can be found in them for separating papyrus fibres from the pulp (able 2).

The structure of the two layers of fibres differs, depending on the production method. Unlike the classic method, sheets produced by the peeling method contain what are called expansion gaps, which come from using a needle to inscribe or score the corners of the triangular stalk. This damages the fibres in these areas

1.3 Fibres

Three types of fibres can be distinguished: especially fine ones, coarse ones and ones in the middle. Moreover, individual fibres and bundles of fibres can be recognised. Depending on the number of fibres there are, the result can be a broad or dense structure or one in between. Differences can also be seen in the course of the fibres. Thus, there are not only linear arrangements, but different slanted or wave-like forms. Although all of these phenomena can be found in a single layer of fibre, they can also occur in both layers

Table 2: Papyrus manufacturing techniques.

Classical method	Peeling method
Work with a **knife**	Work with a **needle**
The stalk is cut in lengthwise strips that get narrower from the outside to the inside.	The stalk is unwrapped from the outside to the inside, so to speak, resulting in wider sections.

1.4 Sheet description

According to Pliny the Elder, a new papyrus roll could contain up to 20 individual sheets joined together by overlapping and gluing. The classification of these sheet joins leads to the examination of the forms of sheets used and thus to different types of sheets (Table 3).

1.4.1 Types

There are three different types of sheets, which differ with regard to their lateral edges (Table 4).

1.4.2 Edge forms

It is also striking that there are sheets whose edges have been trimmed, while others display a very irregular course; in sheets of types I and II, there are narrow to broad overlapping recto fibres (approx. 1 cm to 3 cm long), both in the trimmed and the untrimmed sheets.

1.4.3 Proportions

First of all, individual sheets were produced, up to 20 of which were joined in a roll as a rule.[4] How large were these individual sheets, though? Pliny writes that the size of sheets indicates their quality. Moreover, it was also typical of a specific production site, at least in Greco-Roman times [5] Different sheet formats can be found in pharaonic times as well.[6] The degree to which they are connected with the production sites is currently unknown.

1.4.4 Sheet thickness

Measurements of sheet thickness show that there is a developmental series extending from very thin material in the Old and Middle Kingdoms approx. 0.1 mm in size to up to 0.3 mm in Byzantine times (see Table 5).

1.4.5 Consistency

If we ignore the fact that the consistency of papyri has gone through a change over the centuries and millennia along with their colour, we can distinguish three main groups: soft – flexible – brittle. Originally, papyrus material was so flexible that it could be folded easily, not just rolled up or unrolled. Many papyri still have this flexibility today. Others are characterised by their especially soft consistency – these papyri come from Abusir. I count frayed papyri as belonging to this soft group, the consistency of which has changed over long periods of time due to the repeated influence of wate .

At the other end of the scale, we have the large group of brittle or even embrittled papyri. While the embrittled consistency tends to result from the solidity of strong fibre and is mostly found in thicker or coarser papyri, brittleness or fracture susceptibility is the result of the cellulose deteriorating over time, which is also a consequence of aging.

1.4.6 Opacity

Furthermore, three different levels of opacity can be found in papyrus: transparent – translucent – opaque. In pharaonic times, papyrus sheets were usually transparent and sometimes

[4] Pliny the Elder, *Naturalis Historia* 13, XXIII.

[5] Pliny the Elder, *Naturalis Historia* 13, XXIV.

[6] Möller 1917, 6–7.

Table 3: Comparison of historic hand-made paper and papyrus.

Writing support		
Material	**Paper**	**Papyrus**
Period of use	since *c.* the 9th century CE in the Arab world	from *c.*2700 BCE (Old Kingdom), again from *c.* the 10th/11th century CE
Raw materials	- chopped-up fib es from plants or textiles - glue - colourants - filler	- fib es from the stalk of the plant of the same name
	- water	- water from the Nile
Production method	- mixing all the ingredients together	- orientation of the core fib es in a vertical and overlaying horizontal position
	- scooping - couching (draining)	- rapping and pressing
	- drying	- drying
	- surface treatment (of the front side)	---
Result	- one-layer sheet with the fib es and expansion in one direction	- two-layered sheet with the fib es and expansion in two directions
	- material shows the structure of the scooping sieve	- material shows the individual structure and direction of the fib es in both layers
	- sheets have the structure of the scooping sieve	- every sheet has a unique fib e structure
Designation of the sheet side	- front side, higher quality	- recto = horizontal fib e layer, higher quality
	- reverse side, poorer quality	- verso = vertical fib e layer, poorer quality
Depiction of examples	bars and ribs in handmade paper  modern, handmade paper	Berlin P 13583: detail of recto  cross-section:   Berlin P 13583: detail of verso

Table 4: Different sheet edges.

Type I	Recto fib es extend beyond the verso fib es to the right and left
Type II	Recto fib es only extend beyond the verso fib es on one side
Type III	Recto and verso fib es are flush on the rig t- and left-hand side

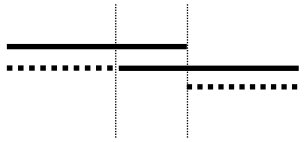
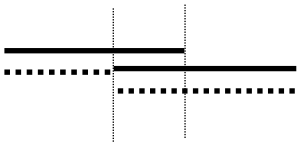
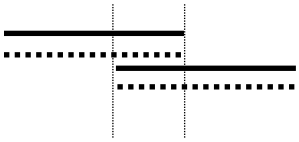
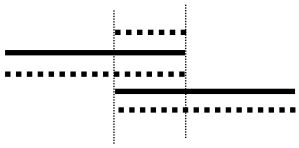
Table 5: Evolution of sheet thickness.

Date	**Thickness of the sheets**	**Examples from the Berlin Papyrus Collection**
Old Kingdom	0.075 – 0.1 mm	0.1 mm: P 15722
Middle Kingdom	0.1 – 0.2 mm	0.1 mm: P 10012
New Kingdom		0.2 mm: P 10487
Late Period		0.25 mm: P 13540
Greek Period	0.25 – 0.3 mm	0.25 mm: P 16985
Byzantine Period	0.3 – 0.4 mm	0.3 mm: P 13275
Arabic Period	0.2 – 0.3 mm	0.2 mm: P 13352

Table 6: The three kinds of join forms.

Manufacturing joins	**Writer's joins**	**File joins**
Carried out in manufacturing as the final s ep in producing the writing surface	Carried out by the writer before or during writing	Carried out in the office to put individual documents together
The arrangement and quality are even	The order of the sheets and diligence in execution varies	Made with little diligence and often even glued over the writing
 Berlin P 3003 H-K, detail	 Berlin P 3005, detail	 Berlin P 11652 B, detail

Table 7: Classifi ation of the join forms.

Type I	Type II	Type III	Type IV
			
Two-layered	Three-layered	Four-layered	Five-layered

even translucent, i.e. the writing on the reverse side was visible on the front. This translucency increasingly disappeared from Greco-Roman times onwards and the sheets became markedly denser. Later, in Arab times, sheets of papyrus were made in a transparent form again and a high level of quality was achieved in their manufacture, hence we can speak of a 'renaissance' of the material as a writing support.

1.4.7 Surface texture

The surface of the two sides of a papyrus sheet is determined by its fibres, their type, density and arrangement, and not least the skill of the papyrus-maker. Consequently, the surface can vary from being rough to flat and smooth. In addition to that, a matt surface and a silky, glossy surface achieved by polishing can be distinguished.

1.5 Sheet joins

Individual sheets of papyrus were joined together, ultimately creating a roll or scroll. The joins can be classifie [7] as one of three kinds (Table 6) and four different forms (Table 7), the latter being subdivided into basic, special and mixed.

Sheet join II is the most common type, followed by type III. In contrast, types I and IV are rarely encountered. We can distinguish between basic, special and even dual mixed forms, depending on how accurately the overlapping was carried out and what the course of the recto fibres was (in particular) on the lateral margins (cf. the different sheet types). Sometimes five-layered sheet joins can be encountered, as in type IV. The latter result when a sheet from the end of a roll is attached to a sheet of type III. This form of roll end is characterised by a narrow verso fibre (approx. 1 cm wide) glued to the end of the roll as a reinforcing safeguard. This way, the papyrus has three layers at such spots. Sheet joins can also be distinguished in terms of the order of the sheets, the care taken in production, and the concentration of the glue in execution (Fig. 1).

The width and thickness of the sheet joins show a development similar to that of sheet thicknesses. Thus, the narrowest joins (of under 1 cm) are found on papyri from the Old and Middle Kingdoms. In the New Kingdom, the width of the joins is about 1.5 cm, in Greco-Roman times about 2.5 cm and in Byzantine times it finally reaches a gluing width of 3 to 3.5 cm, and in some cases even up to 4 cm.[8] Hence the width of the sheet joins can be an important criterion for dating papyrus as a writing support.

1.6 Roll ends

Along with the roll ends mentioned in section 1.5 in the form of a glued-on verso fibre strip, there is another completely different kind of roll end as well. In this case, a whole sheet was attached with the verso side facing up, i.e. it was attached to the recto side (cf. Berlin P 3147; Fig. 2).

1.7 Special forms of rolls

Occasionally, special forms are encountered as well. These appear very rarely, however, and include what I call a single-sheet roll. These sheets, which are usually excessively wide (80 cm or more), have narrow verso strips on both lateral edges as a conclusion and were thus conceived as a roll. Berlin P 10482 is an especially striking example. This papyrus is 82 cm wide and was rolled, which is clearly recognisable by its regularly recurring flaws

2. Origin

When it comes to the origin of the writing support, I take this to mean the place where the material was first produced, which is presumably identical to or close to the place where the plant itself was cultivated, as opposed to the site where

[7] Krutzsch 2017, 217–218.

[8] Krutzsch 2017, 218.

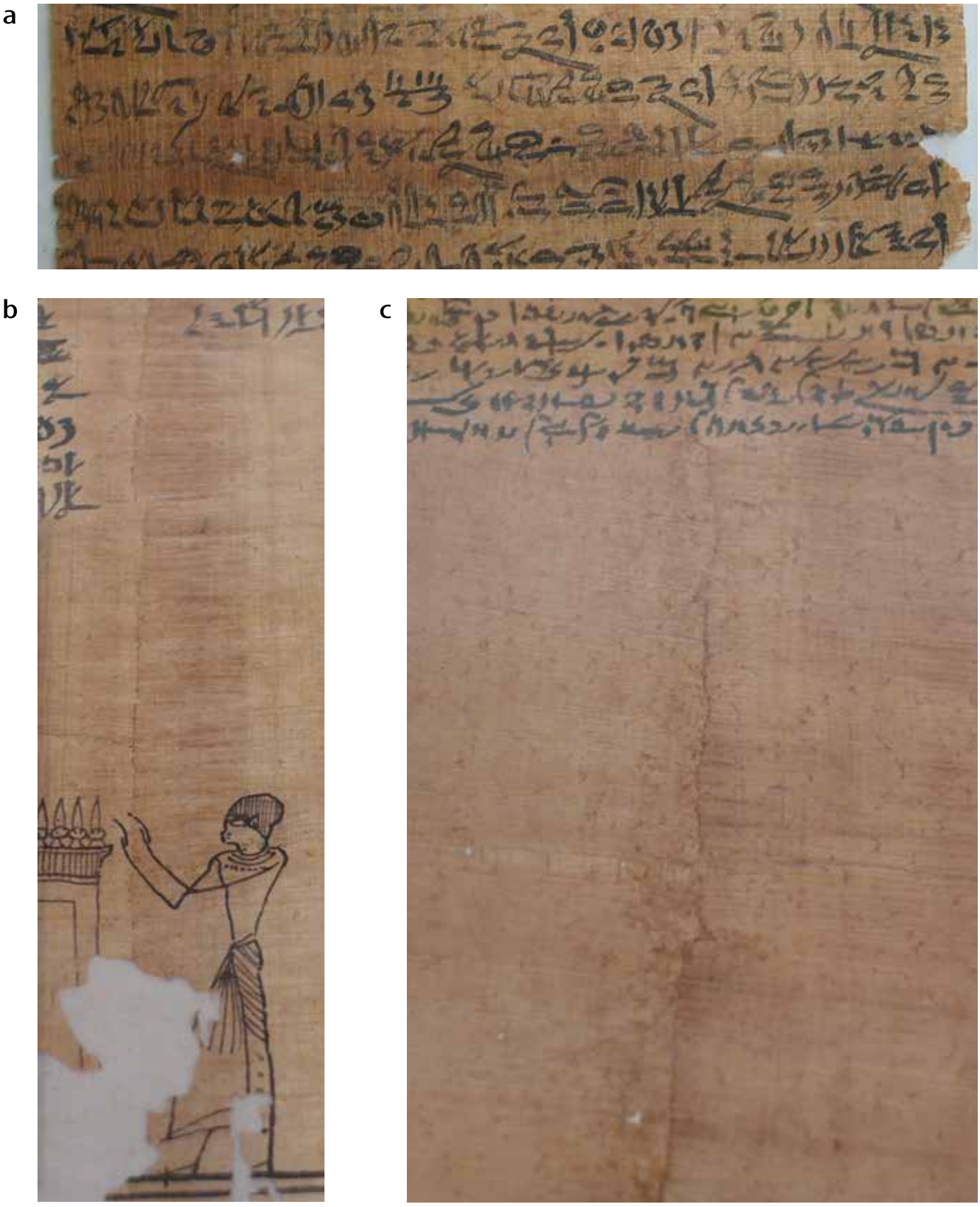


Fig. 1: Examples of the execution of different sheet joins: a) Papyrus Berlin P 10463, detail showing a practically invisible horizontal sheet join in the middle; b) Papyrus Berlin P 10478, detail showing the glue that penetrated to the front; c) Papyrus Berlin P 3100, detail showing glue that seeped to the edge.

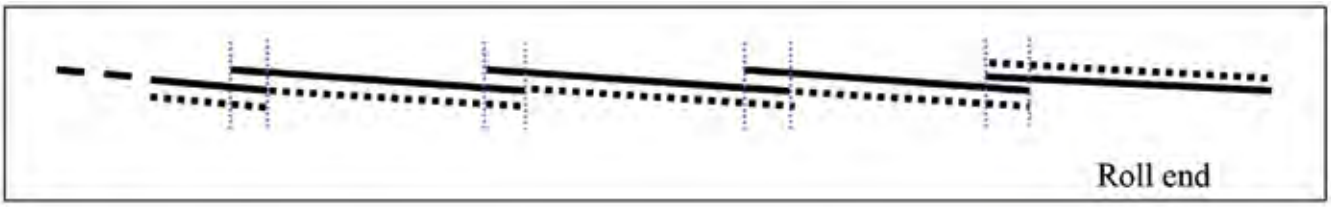


Fig. 2: Cross-sectional drawing of a roll end (right) with a complete verso sheet.

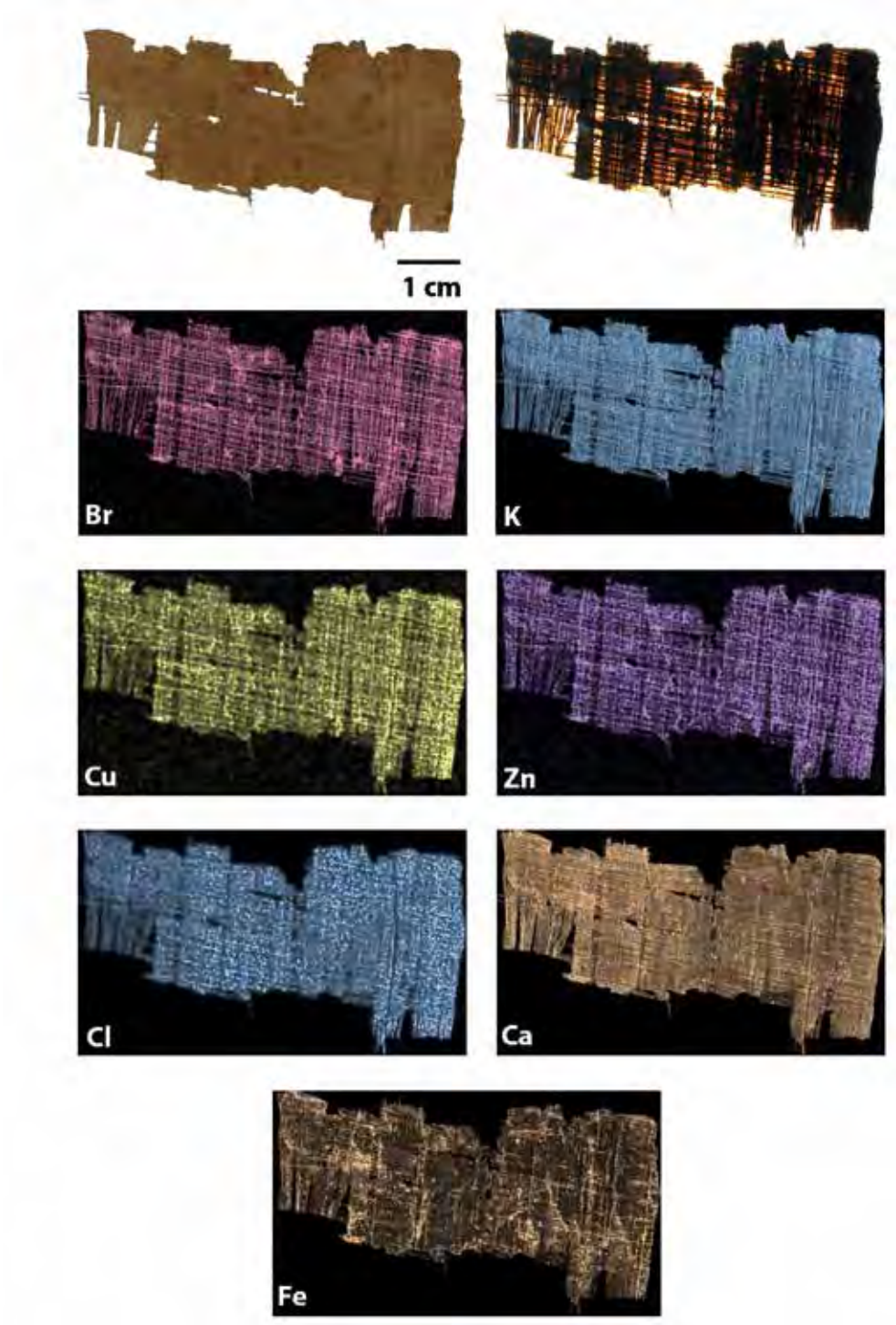


Fig. 3: Elemental distributions from an XRF scan of a fragment from Dime (Arabic times).

Table 8: Three classes of ink under different illuminations.

Illumination	Carbon ink	Mixed ink	Iron-gall ink
Normal daylight			
NIR			
UV			
Example	Berlin P 15771, detail	Berlin P 8500, detail	Berlin P 8283, detail

the written objects were found or acquired. In addition to that, there is the place where the text was written, which may be mentioned in the text. This site may not have anything in common with the site where the papyrus material originated,[9] though, so we may be dealing with four different places that need to be distinguished.

It is therefore evident that the text and the material on which it is written (the 'support') are of equal importance in codicology, so both need to be investigated. A material analysis can be an aid not only to dating objects, but to localising them as well. The first scientific investigations using the XRF Jet Stream device at the Federal Institute for Materials Research and Testing (BAM Bundesanstalt für Materialprüfung und Untersuchung, Berlin)[10] entailed an elemental analysis of papyrus fragments from different archaeological sites, which revealed some local differences. In future, we aim to compare these results with the composition of the soil samples from the corresponding archaeological sites in order to filter out the traces from the excavations.

In a fragment from Dime from Arabic times (see Fig. 3), four elements (bromine, potassium, copper and zinc) can be found in the papyrus sheet we examined that display its structure, which indicates that these elements are in the material. In contrast, the distributions of chlorine, calcium and iron are less ordered and the elements seem to be attached to particular fibres, i.e. they are likely to be contaminants on the surface of the material.

3. Inks and pigments

The three-colour Dino Lite microscope used in my analysis makes it easy to see what kind of inks are present in a sample. Soot inks never change their opacity and colour, iron-gall inks are less visible under NIR light, but are well-pronounced under UV illumination, while the mixed ink, in contrast, can be seen under both UV and IR light.

[9] Rabin and Krutzsch, 2019.

[10] My cordial thanks go to Prof Ira Rabin (BAM, Berlin and CSMC, Hamburg University) and Mr Greg Nehring for their collaboration in analysing and investigating the materials.

Table 9: Coloured drawings under different illuminations.

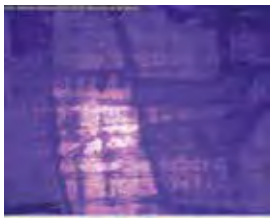
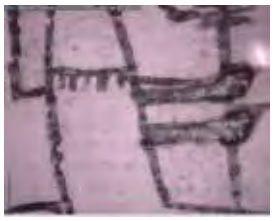
Pigments	Normal daylight	UV	NIR
(1) Red and green pigments; carbon ink (Papyrus Berlin P 3166 A)			
(2) white, green and cinnabar pigments; carbon ink (Papyrus Berlin P 3166 A)			
(3) Egyptian blue, cinnabar and yellow pigments; carbon ink (Papyrus Berlin P 3166 A)			
(4) yellow pigment with orpiment, green pigment; carbon ink (Papyrus Berlin P 3158 II)			

Table 9 shows that in the coloured drawings, soot-based pigment can clearly be seen under the NIR light producing the clear outlines, whereas no difference between the individual pigments can be discerned under UV light.

While white, red and yellow are not visible under IR light, green and blue-green look 'gravy-like'. Under UV light, white is clearly visible and blue and yellow pigments reflect. The latter yellow pigment is occasionally found with an admixture of orpiment, which is not visible under UV light, but appears white under NIR light.

This means that the investigation and identification of inks and colour pigments can be just as helpful for dating and clarifying the origins of objects as precise material analysis.

REFERENCES

Krutzsch, Myriam (2107), 'Einzelblatt und Rolle', in Frank Feder, Gunnar Sperveslage, and Florian Steinborn (eds), *Ägypten begreifen: Erika Endesfelder in memoriam* (Berlin/London: Golden House Publications; IBAES 19) <http://www2.rz.hu-berlin.de/nilus/net-publications/ibaes19>.

Möller, Georg (1927), *Hieratische Paläographie: die aegyptische Buchschrift in ihrer Entwicklung von der fünften Dynastie bis zur römischen Kaiserzeit*, vol. 1: *Bis zum Beginn der achtzehnten Dynastie*, 2nd edn (Leipzig: Hinrichs).

Pliny the Elder (1960), *Natural History*, edited and translated by Harris Rackham, vol. 4: *Libri XII–XVI* (Cambridge, Mass.: Harvard University Press; Loeb Classical Library 370).

Rabin, Ira, and Myriam Krutzsch (2019), 'The Writing Surface Papyrus and Its Materials: 1. Can the writing material papyrus tell us where it was produced? 2. Material study of the inks', in Alberto Nodar and Sofía Torallas Tova (eds), *Proceedings of the 28th International Congress of Papyrology, Barcelona, 6 August 2016* (Barcelona: Publicacions de l'Abadia de Monserrat, Universitat Pompeu Fabra; Scripta Orientalia 3) 773–781.

PICTURE CREDITS

Article

The Techniques and Materials Used in Making Lao and Tai Paper Manuscripts

Agnieszka Helman-Ważny, Volker Grabowsky, Direk Injan, and Khamvone Boulyaphonh | Hamburg, Chiang Mai, Luang Prabang

Abstract

This article discusses the materials and techniques used in the stab-stitched binding of paper manuscripts from Laos and northern Thailand. Eighteen bound manuscripts selected from the collections of the Buddhist Archives in Luang Prabang, Laos, and the Chiang Mai Rajabhat University Library in Thailand were studied in terms of their bookbinding style, form and the materials used to make them. Historical and scientific methods were combined to understand this particular binding style better along with the functional and historical aspects of these manuscripts.

1. Introduction

The history of books in South-east Asia is yet to be studied systematically and currently seems to include more theory and opinion than evidence. There are some important cultural studies that relate to manuscripts from the region in various ways, however. One of the first modern scholars to focus on Thai manuscripts was David K. Wyatt, an American historian who published an immense body of literature on Thai history and culture.[1] He also pioneered the study of manuscripts, which culminated in his last book, *Manuscripts, Books, and Secrets*.[2] Various Shan manuscripts kept in German libraries have been edited by Barend Jan Terwiel and Chaichuen Khamdaengytodtai, who wrote a comprehensive introduction on the production, use and transmission of Shan paper manuscripts.[3] Prior to that, Henry Ginsburg, who was then the curator of the British Library's South-east Asian collection, published two excellent books on South-east Asian manuscripts: *Thai Manuscript Painting*[4] and *Thai Art and Culture: Historic Manuscripts from Western Collections*.[5]

Scholarly interest in the rich manuscript tradition of South-east Asia was revived in the 1980s and gained further momentum in the early 2000s, indicating that Tai manuscripts provide an extensive, fascinating and rewarding field of research to scholars from a wide range of disciplines, such as philology, history, religious studies, social anthropology, material culture and preservation. One of the key people whose research shaped the mainstream at that time is Volker Grabowsky, who has developed a number of projects devoted to cataloguing, documenting and editing Tai manuscripts at the University of Hamburg in recent years. He has written on the revival of the Tai Lü manuscript culture in South-west China in collaboration with Apiradee Techaririwan, whose dissertation on the function and historical development of paratexts in Tai Lü manuscripts from Yunnan and northern Laos also discusses material aspects.[6] Together with Khamvone Boulyaphonh, Grabowsky has edited several catalogues of collections of Lao manuscripts in Luang Prabang.[7] Bounleuth Sengsoulin's PhD thesis from 2016 is a detailed study of Lao manuscript culture. A comprehensive analysis of these manuscripts was recently published in the *Journal of the Siam Society*.[8] Modern studies of northern Thai (Lan Na) manuscript collections are also considered in it, including the material aspect of these manuscripts.[9]

Over the last two decades, Jana Igunma, Ginsburg Curator at the British Library in London, has also published a number of excellent articles on Thai, Lao and Cambodian manuscript cultures.[10] She has contributed to various exhibition projects

[1] Wyatt 1984, 2002, 2003.

[2] Wyatt 2006.

[3] Terwiel and Chaichuen Khamdaengytodtai 2003.

[4] Ginsburg 1989.

[5] Ginsburg 2000.

[6] Grabowsky and Apiradee 2013, Grabowsky 2019a; also see Apiradee 2019.

[7] See Khamvone and Grabowsky 2017 and 2018.

[8] Grabowsky 2019b.

[9] Direk 2015 and 2016.

[10] Igunma 2007, 198–211, Igunma 2010, Igunma 2013a, 25–32, Igunma 2013b, 629–634, Igunma 2014, 35–50, Igunma 2015, 65–81, Igunma 2017, 22–81.

Figs 1a and b: The traditional technique of writing on palm leaf with a sharp tool (a), then rubbing ink into the scratches (b).

promoting the art and material cultural approach to South-east Asian manuscripts. Her *Buddhism Illuminated: Manuscript Art from Southeast Asia* (co-authored with San May) focuses on a variety of types and forms of manuscripts, writing tools, bindings and storage cabinets, but also sheds light on other aspects of South-east Asian manuscript culture.[11] In 2017, a special issue of *Manuscript Studies* was published that was devoted entirely to the history and collections of Thai manuscripts kept in museums and libraries.[12] This excellent survey of collections clearly shows the increase in interest in Thai manuscript culture. That same year, David Wharton published his PhD thesis *Language, Orthography and Buddhist Manuscript Culture of the Tai Nuea*, which focuses on local practices that are part of the endangered scribal tradition of the Tai Nüa and a number of closely related Tai groups which have generally been overlooked in the field of Buddhist Studies.[13] Wharton undertook a detailed study of a single manuscript in a Tai Nüa village near Müang Sing in north-west Laos and used it as an entry point for a broader investigation of Lik (Lik Tho Ngòk) manuscript culture, as found in Müang Sing today, including the distinct roles of the Lik and Tham orthographies, scribal vocation, manuscript production, uses and functions and the contents of the texts. A vast number of Thai manuscripts have been digitised and made available to the public online in recent years.[14]

There are thousands of palm-leaf and paper manuscripts on shelves and in trunks waiting to be discovered and studied. They are excellent sources of information on the historical reconstruction of craftsmanship, such as manuscript production. The last few years have seen a growing appreciation of old manuscripts as material objects, and preliminary studies of materials have also been carried out. There are no detailed studies on the process and technology involved in making manuscripts yet, however, other than sparse promotional information for tourists. From a scholarly point of view, this subject is largely unexplored, even in China and other countries in South-east Asia. Research on palm-leaf production is even rarer. One of the few Thai academics who have devoted their energy to this subject is Kongkaeo Wiraprachak, a senior scholar who works at the Manuscript Department of the National Library of Thailand. [15]

Originally, Lao and Tai manuscripts were written on palm leaves; paper is only recorded as being used as a writing support from the seventeenth century onwards.[16] The usage of a new material, among other things, stimulated a change in the form of books. Palm leaves were primarily used for religious texts. This does not mean that secular texts were only written on paper, though. Paper – a much more flexible material – simply made a greater variety of formats possible and was consequently employed for various types of texts, including religious, astrological and medical subjects. The history of manuscripts from this region of Asia is complex

[11] May and Igunma 2018.

[12] *Manuscript Studies*, vol. 2 (2017), no. 1, edited by Justin McDaniel.

[13] Wharton 2017.

[14] Examples of such databases can be found on websites like *Thai Manuscripts Online: Digital Collections of Thai & Tai Manuscripts and Rare Books*, ed. by Jana Igunma <http://www.thaimanuscripts.de/>.

[15] Kongkaeo 1990.

[16] The oldest extant paper manuscripts from Thailand and Laos are from the seventeenth century, while the oldest dated extant palm-leaf manuscript (from northern Thailand) is from 1471.

Fig. 2: An example of a leporello book from the Chiang Mai Rajabhat University Library in Thailand. The manuscript, which contains part of the *Suttantapitaka*, is written in Pali in Mon script.

Fig. 3: An example of a leporello book collection in-between wooden covers, from the Buddhist Archives of Luang Prabang, Laos.

Figs 4a–c: Manuscript DS 0062 00257 showing typical illustrations in Lao and Tai manuscripts concerned with astrological and divinatory rituals.

Fig. 4a: Fol. 28^{v} showing a drawing of various animals: a tiger, serpent (*naga*), rat, elephant, horse, *garuda* and cat, which the Tai-Lao regard as *tua phoeng* (animals one can rely on). A man counts the number of years he has lived, starting with the tiger as the animal he relies on the most and continuing until he reaches his present age and the animal that corresponds to it. A woman would do the same, starting with the cat. These 'dependable animals' are used for an exorcising ceremony for someone whose destiny is deteriorating.

Fig. 4b: Fol. 1^{v} showing a diagram used when building a new house to determine whether the year will be appropriate for the endeavour. It contains eight drawings, including a bamboo tree, a crystal altar, a coffin a crystal staircase, a crystal castle, a house, a house upside down, and a *kalakini* (representing 'misery'). The divinatory procedure works in the same way as the one mentioned above, counting the person's age starting at the bamboo tree and moving in the direction of the coffin If the counting ends at the coffin upside-down house or *kalakini*, then the year will not be a good one for building. If the age ends at other drawings, it will be a good year.

Fig. 4c: Fol. 5^{v} shows a drawing called *Roek saphao* ('The Occasion of the Junk') or *Roek fa* ('The Occasion of the Sky') and shows a junk in the shape of a swan. It contains twenty-seven numbers and is used to determine the fate or destiny of a newborn child.

Fig. 5: Examples of old stab-stitched books found in Luang Prabang in Laos.

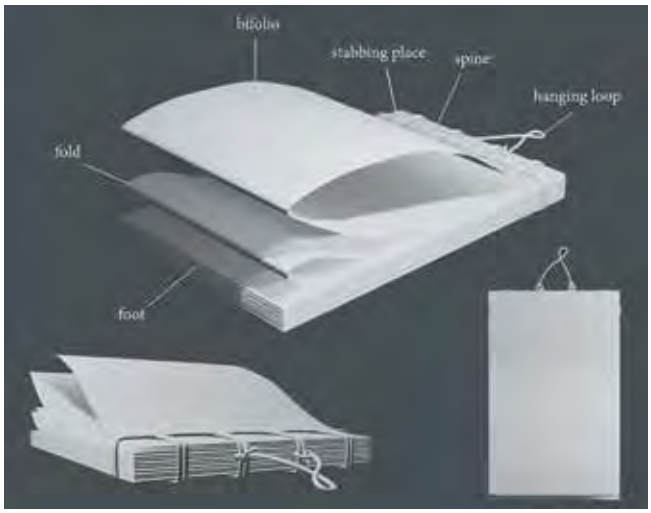


Fig. 6: Image of a 3D model of a stab-stitched book based on books found in Laos and northern Thailand.

and conditioned by both local and global factors, such as the availability of materials, social and economic change, cultural habits and patronage, and, of course, the purpose they were meant to serve.

Palm leaves were readily available in the tropical climate zone and were durable and resistant to insects (more so than paper). However, they were not as convenient as paper when it came to making notes or writing down other less formal texts; it is easier to make a note in ink on paper than scratch the text onto a palm leaf and then rub ink into the scratch (see Figs 1a and b). A piece of paper was also easier to fold, put in one's pocket or carry than thick, heavy and brittle palm folios. The next disadvantage of palm leaves is that they can easily fall out if the thread joining them together happens to break, which may lead to the book leaves getting put back in the wrong order, obviously making it hard to read the text from 'page' to 'page'.

As soon as paper was adopted in the area in the seventeenth century, palm-leaf books (Thai/Lao: *khamphi bailan*) started to be replaced by leporello books (also known as concertina or accordion-style books; see Figs 2 and 3). The thick paper – a new writing support used in a leporello format – was comparatively durable and rigid enough to serve the same writing purposes as palm leaves. At the same time, its greater flexibility and size, which was not limited by the shape of the palm-leaf, made it the most convenient material for book production. Moreover, although some impressive examples of illustrated palm-leaf manuscripts do exist,[17] especially from northern India, paper was much more convenient for drawing the kinds of illustrations that are common in Lao and Thai manuscripts (Figs 4a–c).

Since paper was used as a new writing material, there is no doubt that a major change in book technology occurred. Other book forms were also developed, such as the leporello mentioned above. Loose leaves were also bound using a stab-stitched technique particularly employed in Laos and northern Thailand (Fig. 5), although it was also used in adjacent Tai-speaking areas in eastern Burma (Shan State) and south-western China (Sipsòng Panna, Dehong, Menglian/Moeng Laem, etc. in Yunnan). In Asian bookbinding history, we do not have any exact dates to determine when a particular style was created or when it fell into disuse. Most styles actually co-existed for long periods of time, and we can only assume when the peaks of their popularity were in particular periods, specific regions or when they were used for specific purposes. The techniques and materials employed for binding are rarely mentioned in a historical context, being considered too common to be noted, as opposed to those related to the mounting of paintings or used in art genres regarded as more prestigious than manuscripts (which often are not even perceived as art because of their literary nature).

Stab-stitched books are the least well known of all the books made in Tai manuscript culture. It is hard to say how many of those works have been preserved, which areas they were most often produced in and how this format evolved historically. The general technique, however unique it may have been, varies considerably in terms of its technological details and the materials that were used. This variety resulted

[17] See Guy 1982, 18.

in a diversity of views about these books, which can only be understood through detailed documentation of their forms, material components and production techniques.

It should be noted here that English terminology on the binding of Asian books is not uniform and can be vague at times. Stab-stitched books, for example, can also be described as four-hole stitched, side-stitched,[18] threaded, double-leaved,[19] stabbed, stitched, pouched[20] or sewn. These terms are simply related to a variety of typical binding elements, but some of them are more specific than others. When such books are described in English, the terms used generally follow Western bookbinding terminology, which can lead to misunderstandings sometimes. Despite similar objectives, the specific binding traditions evolved from different origins, concepts and ideas. As Minah Song explains in her article 'The History and Characteristics of Traditional Korean Books and Bookbinding', the complexity of Asian book terminology and original terms is often lost in translation.[21] She states that most vocabulary describing the parts of a book used by Chinese, Korean and Japanese people originates from ancient Chinese terminology, which adopted anthropomorphic terms related to the human head, for example. Specific names were created from the function, location, shape or even symbolic importance of particular elements.[22]

A stab-stitched binding (Fig. 6) is characterised by using a thread-stitch. There are many variations of it, depending on the book's size, the materials used and the local binding tradition. The common denominator, however, is that a thread is pulled through the stabbed holes in the book and wrapped around the 'back' of it to produce a spine. Traditional Asian books bound this way are usually significantly lighter in weight than leporello and loose-leaf books due to the specific

Fig. 7: The leaves and spine of a stab-bound manuscript in the Buddhist Archives of Luang Prabang (Nan Savaeng Ms. 1).

properties of the paper used to make their leaves. The best paper for the production of folded books was made of bark from paper-mulberry trees, referred to locally as *sa* trees (*Broussonetia papyrifera*). This paper is softer and more flexible than paper made from the *khòi* tree (*Streblus asper*), which is used further south in central and southern Thailand as well as in Cambodia.[23] The folios of a stab-bound book never open all the way, which is why a book of this kind will not lie flat when open. One of the most typical features of stab-stitched books is that the head and tail of the spine are sewn around, giving it a very distinct appearance (Fig. 7).

The precise origin of the stab-stitched binding used in Lao and Tai manuscript culture is unknown.[24] However, in view of how book- and papermaking technologies spread in Asia, this book format may have been adapted indirectly from China via communities living on or near the Laos-Yunnan-Myanmar border.

In Laos, northern Thailand and adjacent Tai-speaking regions in Shan State (Myanmar) and south-western Yunnan, books made of paper-mulberry paper (called *chia sa* in Lao and *kradat sa* in Thai) are known by the generic term *phap sa*, sometimes pronounced *pap sa*, or by its variant, *pòp sa* (in the Tai Lü tradition). *Sa* is the indigenous name of the paper-mulberry tree and *phap* means 'to fold'. This reflects the dominant use of the leporello as the standard

[18] Song 2009.

[19] Martinique 1973.

[20] Ikegami 1979 [1986].

[21] Song 2009, 63.

[22] For details of the Chinese stab-stitched binding style, see Li and Wood 1989, 114–117. The authors refer to material and written sources. In terms of material evidence, they describe the selection of Buddhist devotional works from the British Library collection written on both sides of folded and stitched leaves of coarse hemp paper. As for written evidence, Li and Wood refer to an account by Wang Zhu (who lived in Henan province during the Northern Song in the first half of the eleventh century) where the stitching method of bookbinding was mentioned as being used to restore old books (making them more durable) even before Wang Zhu was born. Additionally, the colophons in some of the booklets tell us that they were copied and bound during the Tang (618–907 CE) and Northern Song (907–960 CE) Dynasties.

[23] Regarding the production of *sa* and *khòi* paper manuscripts, see Kongkaeo Wiraprachak 1990. Also see Bounleuth Sengsoulin 2016, 45–46. *Khòi* paper is made of *Streblus asper* and has been used in Thailand for centuries for Buddhist folding books and official records. The paper is less refined than *sa* paper, but is more durable and resistant to yellowing and insect damage.

[24] Regarding the history of the stitched binding in China, Korea, Japan and Tibet, see Helman-Ważny et al., forthcoming.

Fig. 8: Manuscripts hung up on a wall in the house of Ai Saeng Kham, an elderly scribe from Ban Mòng Mangrai near Chiang Rung in Sipsòng Panna.

book format of paper-mulberry paper manuscripts in that region. Although *phap sa* covers all kinds of mulberry-paper manuscripts regardless of their binding, special terms exist which differentiate the more widespread leporellos from the stab-stitched bound books. The Lao call the latter *phap nyip*, which literally means 'stitched/sewn folding books', while the leporellos are known as *phap lan*. The Shan call the leporellos – which resemble the Burmese *parabaik* in terms of the size and decoration of their covers – *pap tup*, whereas the bound books are named *pap kiñ*.[25] Apart from the Tai Nüa, the Tai Lü – in their Yunnanese homeland and in neighbouring areas in Myanmar, Laos and northern Thailand – are the only major Tai ethnic group where stab-stitched binding predominates. These bound books are called *pap hua* in the vernacular, a term that is also used in northern Thailand and refers to sewing the folios at their top (Tai: *hua*, 'head'). Terwiel and Chaichuen have observed that '[t]here is a special kind of *pak kiñ*, made of a single sheet of paper,

[25] For details of the typology of Shan books, see Terwiel and Chaichuen 2003, 20–28.

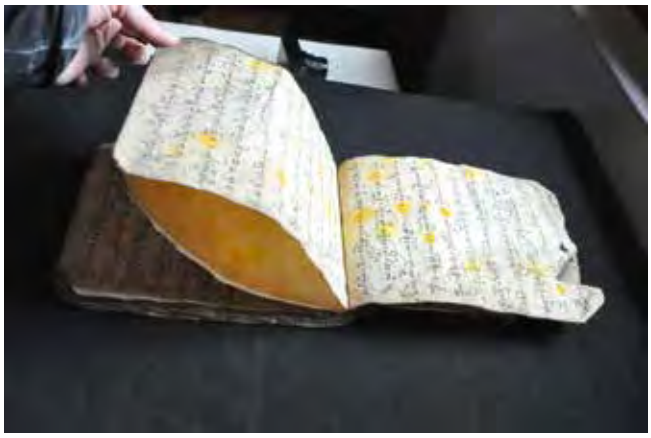


Fig. 9: Stab-stitched binding of a manuscript from the Buddhist Archives in Luang Prabang, which is archived as Nan Savaeng Ms. 1. The fold of the bi-folio is at the bottom of the book.

which may consist of several pieces that have been glued together to form [a] single sheet. This paper is fastened at the upper side to a wooden rod that has been exactly cut to the width of the paper'.[26]

More than three dozen stab-stitched manuscripts were recently found by the staff of the Buddhist Archives in Luang Prabang during a survey of home collections belonging to villagers living in a rural area in Luang Prabang province (in northern Laos) (the study was supervised by Khamvone Boulyaphonh). The villages in Pak U and Nambak are all inhabited by Tai Lü whose ancestors migrated from the Laos-Yunnan-Myanmar borderlands generations ago.[27] It is the manuscript culture of the Tai Lü rather than that of the mainstream Lao which is best known for the production of such sewn paper-mulberry paper manuscripts. This underscores the fact that the tradition of making such stab-stitched bound manuscripts is still alive in northern Laos, at least in Tai Lü communities, as is the case in the Tai Lü heartland of Sipsòng Panna (Yunnan) and Müang Sing (in the Lao province of Luang Namtha). Usually, two to five families in the villages comprising local healers or astrologers owned manuscripts. These books were all in a specific rectangular format, usually bound at the top with folded folios stitched through the stabbed holes that had been made by a sharp tool, hence the name 'stab-stitched binding'.

[26] Terwiel and Chaichuen 2003, 26.

[27] These manuscripts are going to be inventoried and digitised in a sub-project carried out by Dr Khamvone Boulyaphonh as part of the *Digital Repository of Endangered and Affected Manuscripts* (DREAMSEA) project run by Prof Omar Fathurahman (Jakarta) and Prof Jan van der Putten (Hamburg) with financial support from the Arcadia Fund, London.

Figs 10a and b: Stab-stitched binding of two manuscripts from the Chiang Mai Rajabhat University Library in Thailand. Manuscript (a) is archived as ARP.Tai. Nüa.001 and has bi-folios folded at the right side-edge, and manuscript (b) is ARP.TAI.001, which is folded at the left side-edge.

The manuscripts studied contain incantations, astrological treatises, including horoscopes and divination texts, along with medical treatises. The manuscripts tend to be informal, having been used to keep notes on topics like the calendar or records of weather patterns.

In February 2020, the authors of this article conducted a survey of manuscript collections in the Buddhist Archives of Luang Prabang, Laos, and at Chiang Mai Rajabhat University Library in Thailand. We were able to select eighteen sewn manuscripts (see Table 1 in the Appendix) and study their bindings, form and materials in an effort to understand the nature of the developments in bookbinding and the functional and historical aspects of these manuscripts.

2. Techniques of binding

The stab-stitched manuscripts we surveyed were generally produced locally for home use, which is the main reason why they have this particular form: to serve the specific purpose of being carried easily or hung up on a wall or on the pillars of the traditional wooden-stilt houses the Tai build (Fig. 8). The rectangular format of the manuscripts varied from 23.5 to 59 cm in height and 16 to 49 cm in width. We did not observe any repetition of particular sizes, which suggests that the format of books of this kind has not been standardised. Their leaves (folios) are folded in half, thus their outer edges were all sewn together at the end opposite the fold, which also prevented the ink from bleeding through the paper (like the Chinese stab-stitched manuscripts mentioned earlier; paper in Lao/Tai manuscripts of that type is usually thicker, however). In our sample, the number of bi-folios in the books ranged from 8 to 187. Usually, bi-folios were bound at the top opposite the fold (see Fig. 9), but two manuscripts from the Chiang Mai Rajabhat University Library consisted of bi-folios folded on side edges still sewn at the top (CMRU ARP.Tai.Nüa.001 folded at the right side-edge and CMRU ARP.TAI.001 folded at the left side-edge) (Figs 10a, b). Most folded bi-folios were usually stabbed at intervals of 2.5 to 4 cm using a sharp metal tool approximately 1 to 3 cm from the edge of the binding (opposite the fold) (Figs 11a, b). The distance depends on the size of the manuscript. The sewing process usually starts at the central hole (or one of the central holes) on the inside cover. Whenever the binder wanted to pull the cord around the spine of the book-block, he would switch to the other side of the block first and make sure the stitches were already done and mirrored on the other side. The half-point sewing process is demonstrated in Figures 12 and 13.

In the manuscripts we studied, there were between 4 and 13 stabbing points, depending on the manuscript's size, where the string had been pulled through to bind the book (Fig. 14). However, the manuscripts that were intended to be hung up typically have 5 or 7 stabbing points, and those that were wrapped in a roll and tightly bound usually have even more. The binding was done using a strong cotton string and was protected by a layer of cloth (Fig. 15). Such manuscript

Figs 11a and b: The holes are 'stabbed' by using a sharp metal punch with an eye in it through which a cotton cord is then threaded. The technique is being demonstrated here by Direk Injan from the Chiang Mai Rajabhat University Library in the city of Chiang Mai, Thailand.

rolls were usually stored below the ceiling of the owner's house (Fig. 16).

In general, the folios would only be stab-stitched once the whole text had been written down and finalised. An experienced scribe would leave enough space before the first line of the text at the top of a page to ensure that this line would still be fully visible after binding. In some manuscripts, the first line of at least some of the folios is hardly visible at all, though, providing evidence that backs up the hypothesis that the writing proceeded the binding. However, there are also cases where the binding proceeded the writing, as several of the last few folios have been left blank.

3. Materials

3.1 Thread

In our sample, the string was usually made of thickly wound thread, sometimes doubled. Thin thread was occasionally employed as well, however. The thread was usually cotton twine, possibly a type also used for tying up palm-leaf manuscripts. The thread of a manuscript made of twisted paper was also used sometimes, as exemplified by manuscripts archived as Nan Savaeng Ms. 3 (cotton paper) and DS 0062 00295 (paper/mulberry paper). Occasionally, the stitching ended in a loop so the manuscript could be hung up on the wall, as can be observed in all of Nan Savaeng's manuscripts as well as DS 0062 00293, DS 0062 00294, CMRU ARP 060, CMRU ARP 067, CMRU ARP 094 and CMRU Ms. 1 (Fig. 17). The manuscripts which were supposed to be hung up on the wall usually had an odd number of stabbing points, so the hanging loop was exactly in the middle.

3.2 Covers

Some precious manuscripts were bound together with a cloth cover (which could be coloured or patterned), as exemplified by manuscripts CMRU ARP.Tai.Nüa.001, CMRU Ms. 2 and CMRU Ms. 3. The covers used in our sample were usually larger than the manuscript folios, as Wharton also observed in his own study of Tai Nüa manuscripts.[28] Terwiel and Chaichuen also confirmed this for Shan stab-stitched bound books, which are called *pap kiñ* in the local vernacular, in

[28] Wharton 2017, 51.

Fig. 12: The sewing process observed in stab-stitched books in Laos and northern Thailand.

contrast to concertina-style folding books (*pap tup*).[29] If cloth covers are used, they are attached to the front and back of the manuscript as part of the binding process or only at the back in some cases. The cloth covers are folded in over the spine on both sides to protect the edges of the folios when the manuscript is rolled up (see the images in Table 1). Covers are more common for larger manuscripts, which are sometimes rolled up for storage, and often have floral patterns on them, which is particularly common in Tai Nüa manuscript culture. An additional length of cotton cord is attached to these manuscripts to hold the rolled-up manuscript together instead of a loop. This can be seen in CMRU RP.Tai.Nüa.001, CMRU Ms. 2 and CMRU Ms. 3.

3.3 Paper

One of the most important features distinguishing these manuscripts is the use of local *sa* paper made of the bast (*phloem*) of various types of mulberry tree. While only the paper-mulberry tree (*Broussonetia papyrifera*) was grown in northern Thailand and Laos until the early 1970s, Japanese paper-mulberry trees – notably *Broussonetia kazinoki* and *Broussonetia kurzii* – were introduced in the late 1970s as well. Their share of the total production of *sa* paper is still

[29] Terwiel and Chaichuen 2003, 21, 24.

Fig. 13: Stab-stitched method with double threading, demonstrated by Direk Injan, Chiang Mai Rajabhat University Library.

Fig. 14: The leaves and spine of a stab-bound manuscript archived as CMRU ARP. TAI.006 at the Chiang Mai Rajabhat University Library in Thailand.

Fig. 15: Manuscript ARP.Tai.Nüa.001, Chiang Mai Rajabhat University Library, Thailand, has a stab-stitched binding and a fl wery cotton cover. It is bound up in a roll here.

Fig. 16: The house of Nan Chai Saeng, Ban Nam Kaeo Luang (a Tai Nüa village) in the district of Müang Sing, Luang Namtha province, Laos, where manuscripts are kept in protective bags and stored below the ceiling of the main living-room.

limited, however.[30] Despite the common assumption that the bound manuscripts were always made using paper-mulberry (*sa*) paper, the fibre analyses[31] we performed revealed that other kinds of paper such as industrial paper made of bamboo and other grass-type fibres were used instead; our microscopic analyses showed that one was made of bamboo and other types of grass as opposed to seventeen made of paper mulberry (*sa*) (see Table 1).

The long paper-mulberry fibres were well preserved. This was observed in all the samples and is shown in Figure 18. The fibres are characterised by having thick walls, narrow, slightly irregular lumens, irregular cross-markings, a transparent membrane enveloping the fibres, and blunt and rounded fibre ends (Figs 19 and 20). Marja-Sisko Ilvesallo-Pfäffli presented micrographs of the bast fibres separated

[30] See Siriporn 2001 and Aubertin 2004; Chaiyapol 2013.

[31] The fibre analysis was conducted by Agnieszka Helman-Ważny, Centre for the Study of Manuscript Cultures, University of Hamburg. An Olympus BX51 BF/DF Transmitted/Reflected Light Microscope with polarised light was used for fibre identification. An Olympus UC30 camera and Olympus Stream software were used for separate photographic documentation and image analysis. A range of magnification was used from 40× to 400× with both plain and polarised light. The paper samples were immersed in distilled water in a small beaker and boiled for ten to twelve minutes. The water was then decanted and the samples were drained off. About 0.2 g of paper pulp was placed on a microscopic slide and separated into a fine suspension of individual fibres. The fibres were then observed in a water solution using polarised light. The selected samples were stained with two drops of Herzberg staining reagents (zinc chloride-iodine) and observed through an optical microscope. The colour of the resulting stain depends on the lignin content of the fibre and helps to distinguish the species and morphological characteristics of the fibres as well as other cells and elements in the paper pulp. The results were compared to reference samples collected earlier by Agnieszka Helman-Ważny and to available fibre atlases

Fig. 17: Manuscript archived as CMRU ARP.094 with a loop so it could be hung up on a wall (from Shan State, Burma); Office of Arts and Culture, Chiang Mai Rajabhat University Library.

from the inner bark of this fast-growing tree, highlighting the transparent membrane enveloping many of the observed fibres [32] In early stages of the process of pulp beating, the membrane, which indicates a primary wall enveloping the fibre, is disrupted, becoming clearly visible [33]

As stated above, despite the majority of fibres being identified as paper mulberry, we also found minor amounts of other fibres that could not be identified in the pulp. These fibres can be seen in the paper mulberry pulp in CMRU Ms. 1, shown in Figure 20. Furthermore, the paper of the manuscript archived as Ms. 3 from the Chiang Mai Rajabhat University Library does not contain any paper-mulberry paper at all, but is composed of bamboo and other grass-type fibres (Figs 22–23)

The paper mulberry fibres are almost unlignified. Detailed data on the inner bark of paper mulberry, also regarding the

Fig. 18: Well-preserved paper-mulberry fib es observed under the microscope at 40× magnifi ation in manuscript DS 0062 00294 from the Buddhist Archives of Luang Prabang, Laos.

parameters of its chemical pulping, is available in a research report published by Wikhan and Buapun in 2001.[34] Both the low content of lignin (3.32%) and the high content of holocellulose (71.03%) in the inner-bark are worthy of note. Thanks to the low lignin content, the cooking process of the paper-mulberry inner bark can be run under mild conditions with regard to the admixture of alkali and the cooking time. Besides identification of the raw material used for paper production, the next important material feature which allows paper to be characterised comes from the technology and tools used during its production. Generally speaking, there are two types of mould used in papermaking: a floating mould and a dipping mould. Both kinds are additionally characterised by the type of sieve they employ. A floating mould is placed on the surface of the water and paper pulp is poured onto the sieve in the frame of the mould (Figs 24a, b). With the dipping mould, however, the pulp is mixed with water before the mould is dipped into it (Figs 25a, b). As a result, paper made with a floating mould on which the pulp is poured is usually thicker and the fibres are more unevenly distributed in the sheet of paper compared to paper made with a dipping mould, where the pulp is mixed in the water tank. This is an important distinguishing feature between the two methods of papermaking.

In addition, a fixed sieve made of woven cotton, hemp or flax is often attached to the floating mould, while a movable sieve made of bamboo, reed or another kind of grass is attached to the dipping mould. The different sieves each

[32] Ilvesallo-Pfäffli 1995, 348–349

[33] Helman-Ważny 2006, 3–8.

[34] Wikhan and Buapun 2001.

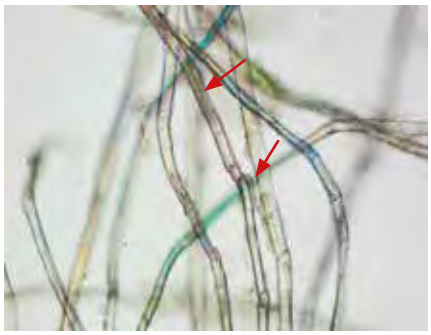


Fig. 19a: The arrows show some thick-walled fib es of paper mulberry enveloped by a transparent membrane. Observed at 200× magnifi ation in the micro-sample from manuscript CMRU ARP.TAI 001 from the Chiang Mai Rajabhat University Library, Chiang Mai, Thailand.

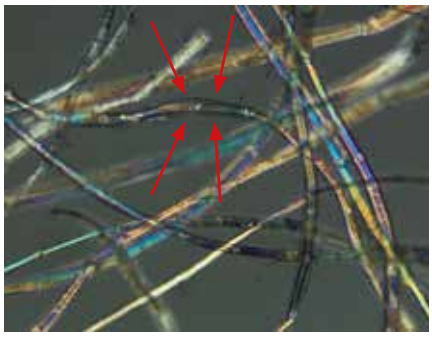


Fig. 19b: The arrows show some long and slightly rigid bent fib es of paper mulberry enveloped by a transparent membrane. Observed at 200× magnifi ation in manuscript CMRU ARP.TAI 006 from the Chiang Mai Rajabhat University Library, Chiang Mai, Thailand.

Fig. 19c: A fib e of paper mulberry enveloped by a transparent membrane after being treated with a Herzberg stain. Observed at 400× magnifi ation in manuscript CMRU ARP.TAI 001, the Chiang Mai Rajabhat University Library, Chiang Mai, Thailand.

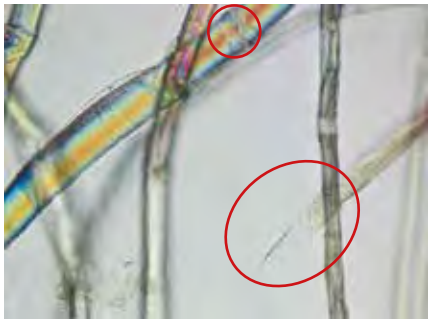


Fig. 20a: paper-mulberry fib es, cross-markings and a naturally rounded fib e end enveloped by a transparent membrane (red oval). Observed under the microscope at 400× magnifi ation in manuscript DS 0062 00294 from the Buddhist Archives of Luang Prabang, Laos.

Fig. 20b: The natural blunt end of a paper-mulberry fib e after treatment with a Herzberg stain. Observed under the microscope at 400× magnifi ation in manuscript BAD 13-2-021 from the Buddhist Archives of Luang Prabang, Laos.

Fig. 20c: The natural fib e end enveloped by a transparent membrane, also with a natural round end, observed at 200× magnifi ation in manuscript CMRU ARP.060 from the Chiang Mai Rajabhat University Library, Thailand.

leave their own distinct imprint on the paper. A textile sieve leaves faint woven patterns (sometimes barely perceptible) and a movable sieve leaves a laid-line pattern (bamboo sieves leave regular laid lines, while reed or other grass sieves leave irregular ones). With them being the impressions of stitches that tie the strips of bamboo, reed or other grass together, the chain lines are sometimes visibly perpendicular to the laid lines. At some point when plastic started to be used in the region, possibly in the middle of the twentieth century, the textile sieve used in the fixed mould was replaced by a plastic one, which was more durable and resistant to microorganisms (which significantly extends the time it can be used for in a humid environment). The specific imprint it leaves on paper after it has dried on the screen (shown in Fig. 26) can be seen in manuscripts (Fig. 27). If we know the date when it was first made, we can also date the manuscript accordingly.

The paper used in some of the manuscripts in our sample was treated with various insect-repellent substances such as the oily resin collected from the wood of the *Yang Na* tree *(Dipterocarpus alatus* Roxb. ex G. Don). It resulted in the brownish colouration observed in these manuscripts, which is shown in Fig. 28.

4. Conclusions

In general terms, the stab-stitched method of bookbinding documented in Lao and Tai manuscripts is similar to the method commonly used in East Asian books. However, the technique of sewing with a double thread, making a loop (i.e. a stitch started in the middle) and the rolled form are unique

Fig. 21: A 'knot' of other fib es stained violet in paper-mulberry-based pulp. Observed at 40× magnifi ation in the manuscript archived as CMRU Ms. 1 from the Chiang Mai Rajabhat University Library in Thailand.

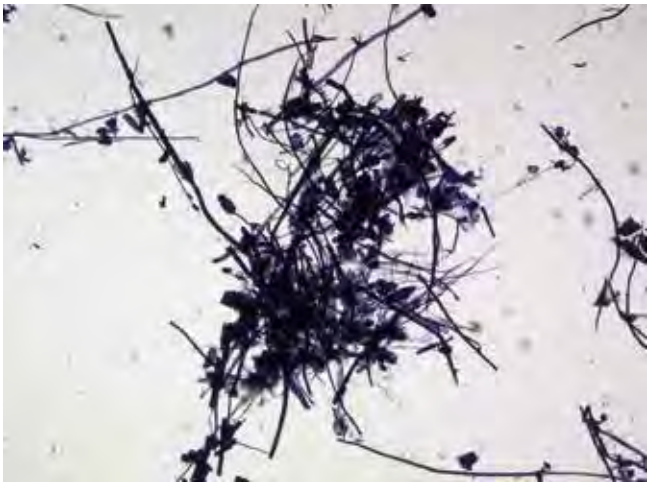


Fig. 22: The bamboo and grass-type pulp observed at 40× magnifi ation in manuscript CMRU Ms. 3 from the Chiang Mai Rajabhat University Library in Thailand.

Fig. 23: The wide bamboo vessel observed at 100× magnifi ation in manuscript CMRU Ms. 3 from the Chiang Mai Rajabhat University Library in Thailand.

elements of Lao and Tai book culture. Interestingly, this local feature of bookmaking is commonly associated with printed book culture in East Asia, and the method was used to bind manuscripts in our own sample. It should be noted here that we also observed some variation within Lao and Tai bookbinding in our sample: there were two main types, depending on the size of the manuscripts. One type was usually smaller and contained a loop in the middle, which allowed such manuscripts to be hung up on the wall. Relatively large manuscripts were usually stored in a rolled-up form. Different dimensions and materials as well as the style of this binding created a unique sense of aesthetics. The purpose is also related to the type of text and determines how a specific manuscript is supposed to be used

The relationship between the function of a manuscript, the type of text it contains, the style of binding used and the format was also clearly apparent in our sample. The stab-stitched manuscripts contained incantations, horoscopes and divination and were either astrological or medical in nature. They all tended to be informal, being used to keep notes on the calendar, records of the weather or as ritual guidelines, for example. These manuscripts were produced in a geographical region inhabited by communities with shared cultural traditions that went (and still go) beyond national borders. In their Yunnanese homeland and neighbouring areas of Burma, Laos and northern Thailand where they also reside, the Tai Lü are one of the few ethnic groups where stab-stitched bindings still predominate. The Tai Nüa ('Chinese Shan') and the Shan (Tai Yai) are two others. Their bound books are called *pap hua* in the vernacular, a term also used in northern Thailand to refer to sewing the folios at their 'heads' (Tai: *hua/ho*). The villages in Pak U and Nambak in a rural area of Luang Prabang province in northern Laos where our stab-stitched manuscripts were found are all inhabited by Tai Lü whose ancestors migrated from the Laos-Yunnan-Myanmar borderlands generations ago. It is the manuscript culture of the Tai Lü rather than that of the mainstream Lao which is best known for producing such sewn paper-mulberry manuscripts. This underlines the fact that the tradition of making stab-stitched bound manuscripts is still alive in northern Laos, at least in the Tai Lü community, as is the case in the Tai Lü heartland of Sipsòng Panna (Yunnan) and Müang Sing (the Lao province of Luang Namtha).

Figs 24a, b: The paper-sheet formation method using a fi ed mould with a plastic screen attached to a wooden frame. Observed at Say Namkhan Company in Luang Prabang, Laos. The flo ting method is characterised by pouring paper pulp directly onto the sieve in carefully measured scoops (similar to a technique widely used in the Himalayas).

Figs 25a, b: The paper-sheet formation method using a fi ed mould with a plastic screen attached to the wooden frame at Sa Paper & Umbrella Handicraft Factory, Chiang Mai. The dipping method is characterised by the paper pulp being scooped from the water tank (as opposed to the technique above, observed in Laos, where the pulp is poured directly onto the sieve in carefully measured scoops).

The manuscripts in our sample also contain materials that are typically used in this kind of format. The paper in the various books we examined is of similar thickness; thinner than in leporello books. The manuscript leaves are made of folded sheets of paper that was made by hand using a papermaking mould consisting of a wooden frame with a finely meshed textile or plastic net spread over it. Some manuscripts were treated with insect-repellent substances, which left brownish traces on the paper. Our microscopic analysis shows that the majority of the manuscripts are composed of paper mulberry

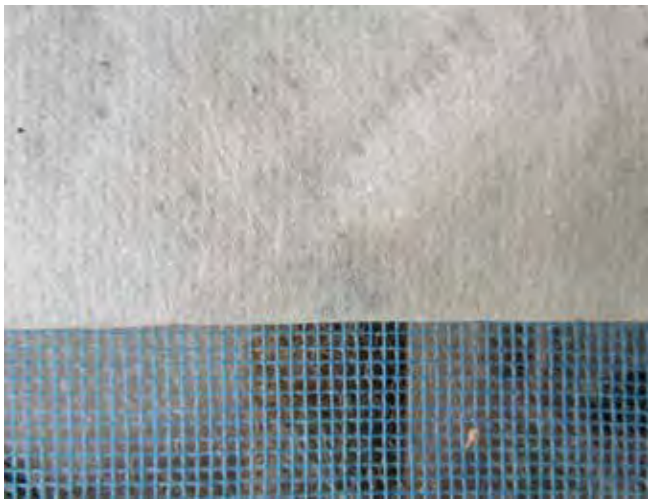


Fig. 26: The sieve imprint left on paper made using a plastic sieve attached to a wooden frame. The photograph was taken in a papermaking factory near Chiang Mai in 2020.

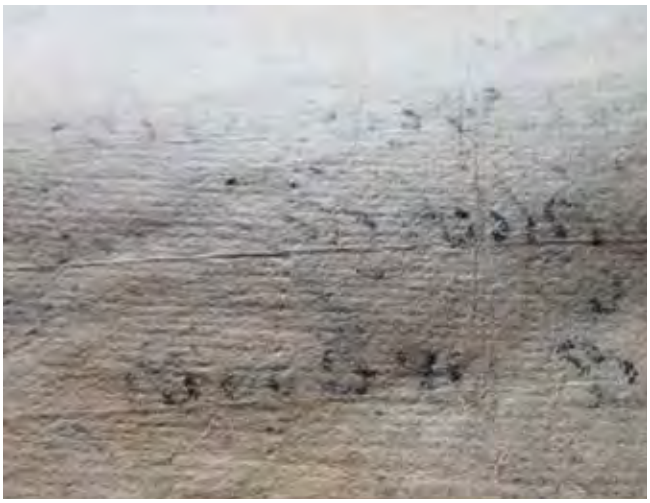


Fig. 27: The texture of paper showing the imprint of a papermaking sieve, seen in a manuscript from the Buddhist Archives in Luang Prabang archived as DS 0062 00292.

Fig. 28: Paper treated with insect repellent, leaving a brownish discolouration. Observed in the manuscript archived as CMRU_Ms. 1 from the Chiang Mai Rajabhat University Library, Thailand.

fibres. Local people commonly associate the stab-stitched manuscript format with *sa* paper by tradition. However, our study shows that other types of paper were also used for producing the stab-stitched manuscripts. Both detected materials and technologies have the potential to help in dating and finding the place where manuscripts originated, especially when these pieces of information are combined with information gleaned from examining the text. Further research on the history of technologies employed during book production could help researchers to achieve even more precision in provenance studies of these manuscripts. It would be helpful to discover the dates when certain materials started to be used by craftsmen (such as plastic screens on papermaking moulds) or when materials other than mulberry fibres were used

ACKNOWLEDGEMENTS

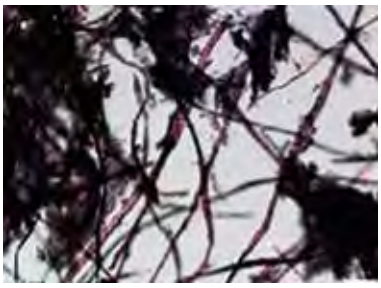
The research for this article was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2176 'Understanding Written Artefacts: Material, Interaction and Transmission in Manuscript Cultures', project no. 390893796. The research was conducted within the scope of the Centre for the Study of Manuscript Cultures (CSMC) at Universität Hamburg.

Table 1: The selection of eighteen manuscripts with stab-stitched bindings preserved in the Buddhist Archives of Luang Prabang, Laos and the Rajabhat University Library in Chiang Mai, Thailand.

No.	Institution's/owner's accession number; image of artefact showing the front cover or the front page	Provenance	Category of text and function	Title
1	Nan Savaeng Ms. 1 	Ban Pak Chaek village, U district, Luang Prabang province	Traditional medical treatise	ຢາ ໜານພັດຂຽນໄວ້ ຜອດຄົນທັ້ງຫຼາຍ *Ya Nan Phat Khian Vai Pot Khon Than Lai* (medicine; *Nan Phat* composed for all people)
2	Nan Savaeng Ms. 2 	Ban Pak Chaek village, U district, Luang Prabang province	Secular ritual and ceremony; for reciting on various occasions of this kind	ຄຳສັບພະຄຳ *Kham Sappha Kham* ('All Kinds of Words')
3	Nan Savaeng Ms. 3 	Ban Pak Chaek village, U district, Luang Prabang province	White magic; the texts are recited by a ceremonial master in a ritual to heal a sick person.	ທໍລະນີສານຫຼວງ ນາງ ດຳ ແລະສານຈີດ *Thòlani San Luang, Nang Dam* and [Thòlani] *San Chüt* ('Great Message of the Earth', 'Black Lady' and 'Tasteless Message of the Earth')

Size (h × w × t) cm	Number of bi-folios	Form of binding style	Number of lines of text on page	Materials
23.5 × 16.5 × 2.0	20	Stab-stitched	11	Woven type of paper composed of paper-mulberry fib es Cover 
40.5 × 19.5 ×1.5	25	Stab-stitched	22–23	Woven type of paper composed of paper mulberry Fol. 1  Cover 
37.5×16.0×3.0	71	Stab-stitched; 5 stabs; double thread; fabric cover measuring 86 × 24 cm; additional paper cover	23	Woven type of paper (plastic sieve print) made of paper mulberry; various modern inks, blue ball pen Paper layer from the cover  Fol. 2  Paper string made of cotton 

No.	Institution's/owner's accession number; image of artefact showing the front cover or the front page	Provenance	Category of text and function	Title
4	DS 0062 00292 	Vat Manorom Sattharam, Luang Prabang	Astrology; traditional Lao calendar and tables for determining the compatibility of lovers before their wedding.	Untitled [ໂຫຣາສາດ *Horasat*] (astrology)
5	DS 0062 00293 	Vat Manorom Sattharam, Luang Prabang	Astrology; traditional Lao calendar and calculations of auspicious or inauspicious days for daily activities.	Untitled [ໂຫຣາສາດ Horasat] (astrology)
6	DS 0062 00294 	Vat Manorom Sattharam, Luang Prabang	Astrology; traditional Lao calendar and calculations of auspicious or inauspicious days for daily activities.	Untitled [ໂຫຣາສາດ Horasat] (astrology)

Size (h × w × t) cm	Number of bi-folios	Form of binding style	Number of lines of text on page	Materials
34.0×24.5 ×1.5	29	Stab-stitched; 7 stabs; fabric cover measuring 33.5–34 × 28–29 cm	20–21	Raw paper made of paper mulberry 
33.0×26.0×3.0	37 + fragment	Stab-stitched; 5 stabs	19–21	Raw paper made of paper mulberry 
39.0 ×28.0×3.0	46	Stab-stitched; 5 stabs	20–23	Raw paper made of paper mulberry 

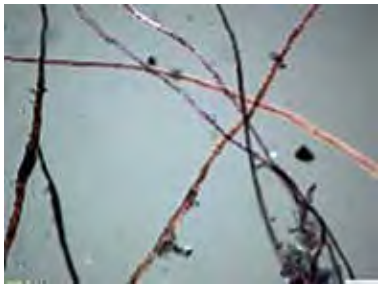
No.	Institution's/owner's accession number; image of artefact showing the front cover or the front page	Provenance	Category of text and function	Title
7	DS 0062 00295 	Vat Manorom Sattharam, Luang Prabang	Astrology; traditional Lao calendar and calculations of auspicious or inauspicious days for daily activities.	Untitled [ໂຫຣາສາດ *Horasat*] (astrology)
8	BAD-13-2-021 	Vat Saen Sukharam, Luang Prabang	The texts are used at Buddhist and secular rituals and ceremonies.	Untitled [ພິທີກຳ *Phithikam*] (ceremony)
9	DS-0056-00640-017v 	Ban Mano village, Luang Prabang, Laos	Astrology; traditional Lao calendar and calculations of auspicious or inauspicious days for daily activities.	Untitled [ໂຫຣາສາດ *Horasat*] (astrology)

Size (h × w × t) cm	Number of bi-folios	Form of binding style	Number of lines of text on page	Materials
34.0–35.0 ×37.5×2.0	28	Stab-stitched; paper string; 16 stabs	17–20	Raw paper made of paper mulberry Fol. 1  Paper thread made of paper mulberry 
38.5 ×36.5–39.0 ×2.3	38 (front part 10 + bottom part 28)	Stab-stitched; 13 stabs; single thick string; rebound (previous stabs visible)	17–18	Woven type of paper made of paper mulberry 
29.0 ×49.0 ×1.0	19	Stab-stitched; 13 stabs; thin yellow thread; whole book folded in half	16 Leftside margin clearly marked in pencil, irregular right margin Drawings	Woven type of paper (plastic sieve print) made of paper-mulberry fib es Fol. 3 

No.	Institution's/owner's accession number; image of artefact showing the front cover or the front page	Provenance	Category of text and function	Title
10	CMRU_ARP.060 	Shan State, Burma Office of Arts and Culture, Chiang Mai Rajabhat University (donated by Prof. Anatole-Roger Peltier, 2018)	*Yanta* (magic spells) written on a piece of cloth or on paper to be carried by the person who wants to use it to ward off dange . Sometimes it is also kept at home for apotropaic purposes.	*Yanta* ("ยันต์" magic letters)
11	CMRU_ARP.067 	Shan State, Burma Office of Arts and Culture, Chiang Mai Rajabhat University (donated by Prof. Anatole-Roger Peltier, 2018)	These blessings are recited when various donations are made to the Buddhist Sangha as part of rituals and ceremonies. The second text is a collection of local Tai Lü songs and verses recited at Buddhist ceremonies.	Untitled (คำเวนทาน และคำขับ) (Blessings for donations / songs)
12	CMRU_ARP.094 	Shan State, Burma Office of Arts and Culture, Chiang Mai Rajabhat University (donated by Prof. Anatole-Roger Peltier, 2018)	These blessings are recited when various donations are made to the Buddhist Sangha as part of rituals and ceremonies.	Untitled (คำเวนทาน) (Blessings for donations)

Size (h × w × t) cm	Number of bi-folios	Form of binding style	Number of lines of text on page	Materials
27.5 ×31.0 ×1.5	12	Stab-stitched at the top; 7 stabs	9–13 Drawings	Woven type of paper made of paper mulberry treated with Yang oil (Dipterocarpus alatus Roxb. ex G. Don) to protect it from insects 
29.0 ×26.5 ×0.5	8	Stab-stitched at the top; 5 stabs	18	Woven type of paper made of paper mulberry 
35.5 ×28.0 ×1.5	25	Stab-stitched at the top; 7 stabs	17–18	Woven type of paper made of paper mulberry 

No.	Institution's/owner's accession number; image of artefact showing the front cover or the front page	Provenance	Category of text and function	Title
13	CMRU_ARP.Tai.Nüa.001 	Northern Shan (borderland between Myanmar and China) Office of Arts and Culture, Chiang Mai Rajabhat University (donated by Prof. Anatole-Roger Peltier, 2018)	Non-canonical Jataka tale (used in preaching during various Buddhist ceremonies).	*Paet Laeng Òk Yòt* (แปดแล้งออกยอด) (*Bunding Kalae*, the joining of two ridges on a double-sloped roof, which point to the top by extending beyond their crossing point)
14	CMRU_ARP.TAI.001 	Office of Arts and Culture, Chiang Mai Rajabhat University (donated by Prof. Anatole-Roger Peltier, 2018)	On the principles of Buddhism (explanations of matters related to the human body and other phenomena of the body). Used on various occasions.	*Kham Sòn Rüang Khan Ha* Didactic poetry on the fifth *khandha* (คำสอนเรืองขันธ์ 5) (Body and emotion)
15	CMRU_ARP.TAI.006 	Shan State, Burma Office of Arts and Culture, Chiang Mai Rajabhat University (donated by Prof. Anatole-Roger Peltier, 2018)	On the lives of the previous Buddhas, the Buddha of our era and the future Buddha. Used in preaching.	Untitled (*Phutthawong* พุทธวงศ์) (On the lives of the Buddhas)

Size (h × w × t) cm	Number of bi-folios	Form of binding style	Number of lines of text on page	Materials
59.0 ×35.0×5.0	84 bifolios folded at the right side-edge	Stab-stitched at the top; 5 stabs; fabric cover measuring 73.8 × 51.5 cm	20–21	Woven type of paper made of paper mulberry 
53.5 ×35.0 ×3.5	29 bifolios folded at the left side-edge	Stab-stitched at the top; 7 stabs; frag-ment of fabric cover measuring 20.3 x 40.8 cm	22–23	Woven type of paper made of paper mulberry 
53.0 ×31.6 ×7.5	187	Stab-stitched bound at the top; 7 stabs	21	Woven type of paper made of paper mulberry 

No.	Institution's/owner's accession number; image of artefact showing the front cover or the front page	Provenance	Category of text and function	Title
16	CMRU_Ms. 1 	Shan State, Burma Office of Arts and Culture, Chiang Mai Rajabhat University (donated by Prof. Anatole-Roger Peltier, 2018)	Principles of the Teachings of the Buddha. To be read or preached at various ceremonies.	*Sumaṇadevo* (สุมณเทโว) (Buddhist doctrine)
17	CMRU_Ms. 2 	Disangpani, Assam, India Office of Arts and Culture, Chiang Mai Rajabhat University	Buddhist text pertaining to the rules of monastic discipline.	*Vinaya Pitaka - Mahāvagga* *(*วินัยปิฎก มหาวรรค*)*
18	CMRU_Ms.3 	Disangpani, Assam, India Office of Arts and Culture, Chiang Mai Rajabhat University	The second of the three *Baskets* of the Theravada Buddhist Canon, also called the *Basket of Discourse*. It is used by monks as a manual for explaining the Dhamma.	*Suttanta Pitaka* (สุตตันตปิฎก)

Size (h × w × t) cm	Number of bi-folios	Form of binding style	Number of lines of text on page	Materials
34.5 ×18.0 ×2.2 Cover 41.0 × 19.0 cm	62	Stab-stitched bound at the top; 7 stabs; fabric cover measuring 41 × 19 cm		Woven type of paper made with paper mulberry (with minor addition of other fib es) 
51.0×32.0×2.0	34	Stab-stitched, bound at the top; 4 stabs; kept in a roll; sewn with thin thread; fabric cover	36	Woven type of paper made of paper mulberry 
45.5×28.5×0.5	15	Stab-stitched, bound at the top; 7 stabs; kept in a roll; fabric cover measuring 68.7 × 66.7 cm	26–27	Woven type of paper made of bamboo and some type of other grass (plus one unknown pink fib e) 

REFERENCES

Apiradee Techasiriwan (2019), *Tai Lü Manuscripts from Southern Yunnan and Northern Laos: The Function and Development of Paratexts in a Recently Revived Manuscript Culture*, PhD thesis, Faculty of Humanities, University of Hamburg.

Aubertin, Catherine (2004), 'Paper Mulberry (*Broussonetia papyrifera*) in Lao PDR: a successful example of forest product domestication', in Koen Kusters and Brian Belcher (eds), *Forest Products, Livelihoods and Conservation: Case Studies of Non-Timber Forest Product Systems*, vol. 1: Asia, Jakarta: Center for International Forestry Research, 227–245.

Bounleuth Sengsoulin (2016), Buddhist Monks and Their Search for Knowledge: *An Examination of the Personal Collection of Manuscripts of Phra Khamchan Virachitto (1920–2007), Abbot of Vat Saen Sukharam, Luang Prabang*, PhD thesis, Faculty of Humanities, University of Hamburg.

Chaiyapol Sakunluangaram (2013), *Art & Craft Wisdom of Paper Mulberry: A Case Study of Ban Ton Pao Paper Mulberry Cluster, San Kamphaeng District, Chiang Mai Province*, Master's thesis, Srinakharinwirot University, Bangkok (in Thai).

Direk Injan (ed.) (2015), *Wat Nam Cam: bai lan ngan sin lae prawattisat* (Wat Nam Cam: Palm-leaf [manuscripts], art work and history), Chiang Mai: Chiang Mai Rajabhat University.

—— (ed.) (2016), *Rai chü khamphi bai lan lae phap sa: aksòn tham lan na nai phak nüa, lem 1*, BE 2554–2558 (Catalogue of palm-leaf manuscripts and mulberry-paper folded books: the Dhamma script of Lan Na in the northern region, vol. 1, CE 2011–2015), Chiang Mai: Chiang Mai Rajabhat University.

Ginsburg, Henry David (1989), *Thai Manuscript Painting*, London: British Library.

—— (2000), *Thai Art and Culture: Historic Manuscripts from Western Collections*, London: British Library.

Grabowsky, Volker (2019a), 'The Ethno-Religious Identity of the Tai People in Sipsong Panna and Its Resurgence in Recent Manuscripts', in Desley Goldston (ed.), *Engaging Asia: Essays on Laos and Beyond in Honour of Martin Stuart-Fox*, Copenhagen: NIAS Press, 290–322.

—— (2019b), 'Thai and Lao Manuscript Cultures Revisited: Insights from Newly Discovered Manuscript Collections in Luang Prabang', *Journal of the Siam Society*, 107/1: 79–118.

—— and Apiradee Techasiriwan (2013), 'Tai Lue Identities in the Upper Mekong Valley: Glimpses from Mulberry Paper Manuscripts', *Aséanie*, 31: 11–54.

Guy, John (1982), *Palm-leaf and Paper: Illustrated Manuscripts of India and Southeast Asia*, Melbourne: National Gallery of Victoria.

Helman-Ważny, Agnieszka (2006), 'Asian paper in works of art: a comparative fiber analysis', *Hand Papermaking*, 21/2: 3–9.

—— (2016), 'More than meets the eye: fibre and paper analysis of the Chinese manuscripts from the Silk Roads', *STAR: Science & Technology of Archaeological Research*, 2/2: 127–140.

—— Direk Injan, Khamvone Boulyaphonh and Volker Grabowsky (2021), 'The Stab-stitched Binding of Tai Manuscripts: A Survey of History, Technique, and Function,' *Journal of the Siam Society*, 109/1: 129–168.

Hunter, Dard (1978), *Papermaking. The History and Technique of an Ancient Craft*, New York: Dover Publications.

Igunma, Jana (2007), 'Thai local knowledge: the hidden wisdom of palmleaf manuscripts and folding books', in *Lokales Wissen und Entwicklung*, ed. Sabine Miehlau and Frank Wickl, Bad Honnef: Horlemann, 198–211.

—— (2010), 'The history of the book in Southeast Asia (2)', in *The Oxford Companion to the Book*, ed. Michael F. Suarez, S. J. Woudhuysen and H. R. Woudhuysen, New York/Oxford: Oxford University Press.

Igunma, Jana (2013a), 'Aksoon Khoom: Khmer heritage in Thai and Lao manuscript cultures', *Tai Culture. Interdisciplinary Tai Studies Series*, 23: 25–32.

—— (2013b), 'The history of the book in Southeast Asia: the mainland', in *The Book – A Global History*, ed. M. F. Suarez and H. R. Woodhuysen, New York/Oxford: Oxford University Press, 629–634.

—— (2014), 'The beautiful art of Tai palm-leaf manuscripts', *Southeast Asia Library Group Newsletter*, 46: 35–50.

—— (2015), 'Meditations on the Foul in Thai Manuscript Art', *Journal of the International Association of Buddhist Universities*, vol. V, Bangkok: IABU, 65–81.

—— (2017), 'Henry D. Ginsburg and the Thai Manuscripts Collection at the British Library and Beyond', *Manuscript Studies,* 2/1: 22–81.

Ikegami, Kōsanjin (1979 [1986]), *Japanese Bookbinding: Instructions from a Master Craftsman / Kōjirō Ikegami*, adapted by Barbara B. Stephan, New York: Weatherhill.

Ilvessalo-Pfäffli, Marja-Sisko (1995), Fibre Atlas: *Identification of Papermaking Fibres*, Berlin: Springer.

Khamvone Boulyaphonh and Volker Grabowsky (2017), *A Lost Monastery Library Rediscovered: A Catalogue of Lao Manuscripts Kept in the Abbot's Abode (kuti) of Vat Si Bun Hüang, Luang Prabang, Hamburger Südostasienstudien*, vol. 15, New York/Luang Prabang: Anantha Publishing.

—— and —— (2018), *Written Treasures from Northern Laos: A Catalogue of Manuscripts from Vat Xiang Thòng, Luang Prabang, Hamburger Südostasienstudien*, vol. 16, New York/Luang Prabang: Anantha Publishing.

Kongkaeo Wiraprachak (1990), *Kan tham samut thai lae kan triam bai lan* (The Making of Thai Folding Books and the Preparing of Palm-leaf Manuscripts), Bangkok: National Library of Thailand, Fine Arts Department.

Li, Zhizhong and Frances Wood (1989), 'Problems in the history of Chinese bindings', *The British Library Journal*, 15/1: 104–119.

Manuscript Studies, vol. 2 (2017), no. 1, Special Issue: Thai and Siamese Manuscript Studies, University of Pennsylvania Press <https://repository.upenn.edu/mss_sims/vol2/iss1/> (accessed 15 March 2021).

Martinique, Edward (1973), 'The Binding and Preservation of Chinese Double-Leaved Books', *The Library Quarterly: Information, Community, Policy*, 43/3: 227–236.

Pan, Jixing (1998), *The History of Chinese Science and Technology: Papermaking and Print*, Beijing: Science Press.

San San May and Jana Igunma (2018), *Buddhism Illuminated: Manuscript Art from Southeast Asia*, London: The British Library.

Siriporn Jangsutthivorawat (2001), 'Paper mulberry identification using molecular genetic techniques', Master's thesis, Kasetsart University, Department of Genetics, Bangkok (in Thai).

Song, Minah (2009), 'The history and characteristics of traditional Korean books and bookbinding', *Journal of the Institute of Conservation*, 32/1: 53–78.

Terwiel, Barend Jan (with the assistance of Chaichuen Khamdaengyodtai) (2003), *Shan Manuscripts, Part 1*, Stuttgart: Franz Steiner.

Thai Manuscripts Online: Digital Collections of Thai and Tai Manuscripts and Rare Books <http://www.thaimanuscripts.de/> (accessed 28 January 2021).

Tsien, Tsuen-Hsuin (1985), *Chemistry and Chemical Technology*, Part 1: *Paper and Printing*, ed. by Joseph Needham = vol. 5 of *Science and Civilisation in China*, Cambridge: Cambridge University Press.

Wharton, David (2017), *Language, Orthography and Buddhist Manuscript Culture of the Tai Nuea: an apocryphical jātaka text in Mueang Sing, Laos*, PhD thesis, Faculty of Arts and Humanities, University of Passau.

Wikhan Anapanurak and Puangsin Buapun (2001), 'Paper Mulberry Pulp Properties by Various Alkaline Processes', <http://posaa.kapi.ku.ac.th/Document/PDF/FinalRep2001_V2/Full_2A-2-1(3).pdf> (accessed 3 October 2020).

Wyatt, David K. (1984, 2003), *Thailand: A Short History*, New Haven/London: Yale University Press.

Wyatt, David K. (2002), *Siam in Mind*, Chiang Mai: Silkworm Books.

Wyatt, David K. (2003), *Reading Thai Murals*, Chiang Mai: Silkworm Books.

Wyatt, David K. (2006), *Manuscripts, Books, and Secrets* (unpublished manuscript).

INTERVIEWS

Nan Savaeng, son of Mai Saengkham, a local healer from Pak Chaek village, interviewed on the telephone on 21 August 2020.

Nan Pan, a ceremonial master from Pak Chaek village, interviewed on the telephone on 21 August 2020.

PICTURE CREDITS

Figs 1–3: © Agnieszka Helman-Ważny.

Fig. 4: © Khamvone Boulyaphonh.

Fig. 5: © Agnieszka Helman-Ważny.

Fig. 6: © Olga Ważny.

Figs 7: © Agnieszka Helman-Ważny.

Fig. 8: © Volker Grabowsky, 2013.

Figs 9–11: © Agnieszka Helman-Ważny.

Fig. 12: © Olga Ważny.

Figs 13–15: © Agnieszka Helman-Ważny.

Fig. 16: © Volker Grabowsky.

Figs 17–28: © Agnieszka Helman-Ważny.

Article

Inks Used to Write the Divine Name in a Thirteenth-Century Ashkenazic Torah Scroll: Erfurt 7 (Staatsbibliothek zu Berlin, Ms. or. fol. 1216)[1]

Nehemia Gordon, Olivier Bonnerot, and Ira Rabin | Ramat Gan, Berlin, Hamburg

Abstract

The scribe of 'Erfurt 7', a thirteenth-century Torah scroll now kept in Berlin, initially left blank spaces for the divine appellations *Elohim* (אלהים) and *YHWH Elohim* (יהוה אלהים), which were filled in during a second stage of writing. The appearance of the ink employed to write the appellations was significantly darker than that of the surrounding ink. X-ray fluorescence analysis (XRF) has shown that the light and dark brown inks had similar elemental compositions, but contained different ratios of iron to potassium, which could be explained by the use of different batches of ink. According to some medieval sources, the divine appellations were sometimes filled in during a second stage of writing in the presence of ten men from the Jewish community. In Erfurt 7, the two-stage procedure was only performed in the first 1.5 columns of the original sheets, suggesting it may have been part of a public ceremony inaugurating the writing of the divine names in the scroll. Erfurt 7 emerges from this study not only as a ritual object used for liturgy, but as a rallying point for the Jews of Erfurt to come together as a community to express their reverence for the written form of God's name. The divine name *YHWH* (יהוה) was written in a smaller script than the surrounding text on three replacement sheets using the same two-stage procedure. The ink used on the replacement sheets contained zinc, which is characteristic of other Erfurt manuscripts as well. This suggests that Erfurt was the place where the scroll was used, cherished and eventually repaired.

1. Background

Erfurt 7 is one of a cache of fifteen Hebrew manuscripts seized during a massacre of the Jewish community of Erfurt in March 1349. The manuscripts have been in non-Jewish custody ever since, so any Jewish scribal interventions must have pre-dated that event. Fourteen of these manuscripts are currently housed at the Staatsbibliothek zu Berlin (Berlin State Library), including four Torah scrolls designated Erfurt 6 (Ms. or. fol. 1215), Erfurt 7 (Ms. or. fol. 1216), Erfurt 8 (Ms. or. fol. 1217) and Erfurt 9 (Ms. or. fol. 1218). Jordan Penkower has dated Erfurt 7 to the thirteenth century and Erfurt 6, Erfurt 8 and Erfurt 9 to the fourteenth century (albeit with some reservations in the case of Erfurt 8).[2]

2. Scribal features of Erfurt 7 (Ms. or. fol. 1216)

Erfurt 7 consists of fifty sheets of parchment with three columns per sheet. Efraim Caspi discovered that it matches the manuscript Munich, Bayerische Staatsbibliothek, Cod. hebr. 212, which may have served as its *tiqqun soferim*, the codex from which the scroll was copied.[3] The scroll contains sixty lines per column, while the codex contains thirty lines per page, so each column in the scroll corresponds to two pages in the codex. All but six columns in the scroll begin with a word that starts with the letter *vav* at the beginning of a verse (see below). Out of the one hundred and forty-four columns in Erfurt 7 (including nine columns from replaced sheets, which are mentioned below), a hundred and thirty-seven begin with the same word as the corresponding verso

[1] This article is based on a chapter of Nehemia Gordon's doctoral thesis, the research for which was conducted at the Bible Department at Bar-Ilan University, Israel and carried out under the supervision of Prof. Yosef Ofer. The X-ray fluorescence (XRF) tests, data treatment and statistical analysis described here were carried out by Dr Olivier Bonnerot under the supervision of Prof. Ira Rabin. We would like to express our warmest thanks to the staff of the Staatsbibliothek zu Berlin, in particular Petra Figeac, Christoph Rauch and Melitta Multani, for their assistance throughout the analysis of the scroll. This research was partly funded by the German Research Foundation (DFG) in conjunction with the Federal Excellence Strategy and the Cluster of Excellence EXC 2176, 'Understanding Written Artefacts: Material, Interaction and Transmission in Manuscript Cultures', project no. 390893796. It was partly carried out at the Centre for the Study of Manuscript Cultures (CSMC) at the University of Hamburg.

[2] Penkower 2014, 118–119.

[3] Caspi 2014, 234–236.

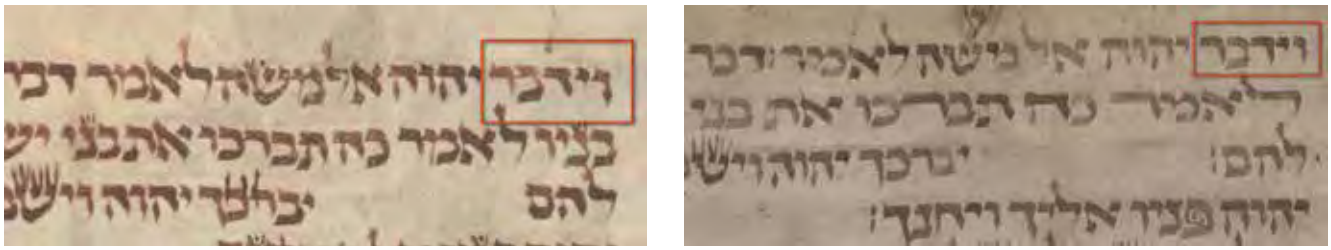


Fig. 1: Comparison of Erfurt 7 (sheet 33, col. 2 [col.98]), left image, with the corresponding verso page of Cod. hebr. 212 (fol. 98ᵛ), right image. Both begin with the same word, וידבר, the first word in Num. 6:32. The correspondence of the layout between the scroll and the codex applies to the beginning of columns/ verso pages, but not to individual lines.

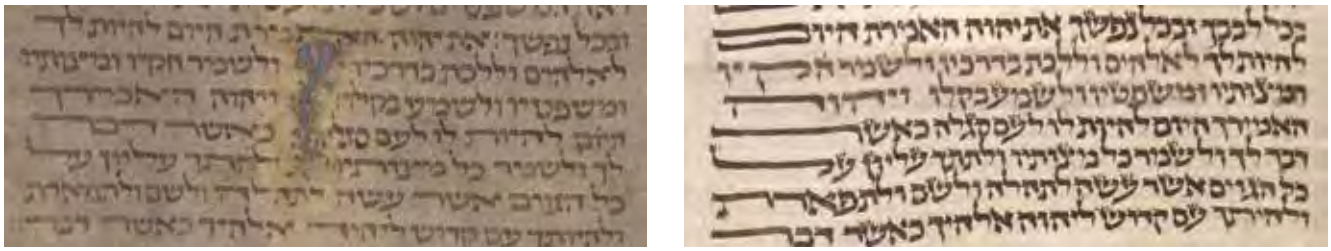


Fig. 2: The scribe of Cod. hebr. 212 (left) made moderate use of dilating (stretching) letters at the end of fol. 143ʳ to begin the next page with the word ויצו, the first word in Deut. 27:1. This word begins with a *vav* in accordance with the practice of *vave ha-ʿamudim*. The scribe of the scroll Erfurt 7 (right), made extensive use of dilating letters at the end of the corresponding column (sheet 48, col. 1 [col. 142]) in order to begin the next column with the same word. Cod. hebr. 212 had to be written around a tear in the parchment, which was sewn up before the codex was written (the thread or sinews have been removed or fallen out since then). As there was no tear in Erfurt 7, the scribe had to dilate more letters than the scribe of Cod. hebr. 212 to begin the next column with the same word.

page of Cod. hebr. 212, six columns differ from Cod. hebr. 212 by one verse, and one column differs by four verses. For example, Erfurt 7, sheet 34, col. 2 [col. 98] and Cod. hebr. 212, fol. 98ᵛ both begin with the word וידבר ('And he spoke'), the first word in Num. 6:32 (see Fig. 1).[4] Erfurt 7 may have been copied directly from Cod. hebr. 212, from another scroll that was copied from Cod. hebr. 212 or the two manuscripts may actually derive from a common source.

The custom of beginning each column of a Torah scroll with the letter *vav* was commented on by the German legal scholar Rabbi Meʾir bar Jekuthiel ha-Kohen in the late thirteenth century (*c*.1260–1298):

> That which some of the ignoramus scribes have a custom to do – to start each column with a *vav*, which they call *vave ha-ʿamudim* – appears to be absolutely forbidden.[5]

The phrase *vave ha-ʿamudim* is a pun that can mean both '[the letters] *vav* in the columns [in a scroll]' and 'hooks of the pillars'. The latter meaning is used to describe an architectural feature of the Tabernacle in the Book of Exodus (27:10, 11; 38:10, 11, 12, 17). The architectural meaning in relation to the Tabernacle was cited in the Babylonian Talmud as an explanation of the name of the letter *vav* and the reason for its shape. The resemblance of the letter *vav* to a hook atop a pillar was deemed proof that the Torah was originally written in the so-called Assyrian script (the script used in Talmudic times, the Middle Ages and today) rather than Palaeo-Hebrew script.[6] This Talmudic reference may have led to the custom of writing Torah scrolls with *vavs* at the head of each column as an allusion to the script in which the Torah was supposedly written.

According to a *responsum* written by Meʾir bar Jekuthiel's teacher, Rabbi Meʾir of Rothenburg (*c*.1215–1293), known as Maharam, the scribal practice of *vave ha-ʿamudim* 'is not from the Torah or a rabbinical enactment, but rather there was a specific scribe, Rabbi Leontin of Mühlhausen, who

[4] The numbers in square brackets throughout this study refer to the overall column number from the beginning of the scroll [col. 1] to the end [col. 150].

[5] Meʾir bar Jekuthiel ha-Kohen, *Haggahot Maimuniyyot*, 1524, 56ʳ (*Hilkhot Tefillin u-Mezuzah ve-Sefer Torah* § 7:7). On *vave ha-ʿamudim*, also see Penkower 2019, 136–138; Penkower 2020, 44.

[6] Babylonian Talmud, *Sanhedrin* 22a.

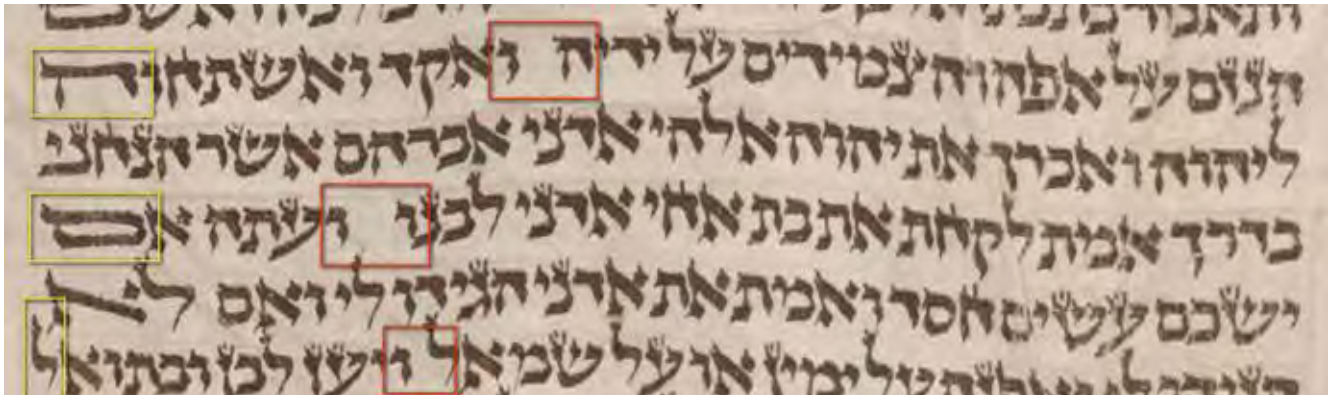


Fig. 3: Dilation and contraction of letters to justify the left margin (yellow). Spaces between verses (red). Erfurt 7, sheet 6, col. 1 [col. 16].

was showing off his skill'.[7] Both Meʾir bar Jekuthiel and Maharam forbade the use of *vave ha-ʿamudim* because it often required scribes to contract or dilate letters in order to begin a column with a specific word. These rabbis viewed dilation and contraction of letters as distorting the legally mandated shapes of the letter. Scribes would generally dilate or contract letters towards the end of a column when they realised they were going to reach the designated word for the beginning of the next column too early or too late, respectively (Fig. 2).

The objections of these thirteenth-century German rabbis notwithstanding, *vave ha-ʿamudim* are found in Erfurt 7 and two other thirteenth-century Ashkenazic Torah scrolls: Washington D.C., Museum of the Bible, SCR.4820 and Jerusalem, National Library of Israel, 34° 8421.[8] Contrary to Maharam's claim that the technique was invented by a twelfth- or thirteenth-century scribe in Germany, the scribal custom of *vave ha-ʿamudim* can already be found in an Oriental Torah scroll from *c*.1000 designated St Petersburg, Russian National Library, Parchment Scroll 3. The left-hand margin of each column in Erfurt 7 was justified by dilating and contracting letters near the end of the line, irrespective of any *vave ha-ʿamudim* (see Fig. 3).

Medieval Torah scrolls were generally written without any punctuation. However, some scribes used spaces to separate verses. According to Rabbi Isaac of Corbeil (d. 1280), '[a scribe must leave] the size of a small word between one verse and the next'.[9] Erfurt 7 usually followed this practice, separating verses with spaces that tended to be larger than those that separated words, as Fig. 3 shows.

Another scribal custom involved beginning six columns with specific words, the first letters of which form the acronym ביה שמו (*beyah shemo*), 'in Yah, His name'.[10] This custom goes back to at least the eleventh century, with some variation regarding which words were chosen to represent the letters *yod, shin and mem*. The six letters that form the acronym in Erfurt 7 are the *bet* of בראשית in Gen. 1:1 (sheet 1, col. 1 [col. 1]), the *yod* of יהודה in Gen. 49:8 (sheet 13, col. 1 [col. 37], the *he* of הבאים in Exod. 14:28 (sheet 17, col. 1 [col. 49]), the *shin* of שפטים in Deut. 16:18 (sheet 46, col. 1 [col. 136]), the *mem* of מוצא in Deut. 23:24 (sheet 47, col. 3 [col. 141]) and the *vav* of ואעידה in Deut. 31:28 (sheet 50, col. 1 [col. 148]). The beginning of these six columns in Erfurt 7 matches the corresponding verso pages in Cod. hebr. 212 (fols 1ᵛ, 37ᵛ, 49ᵛ, 136ᵛ, 141ᵛ and 148ᵛ). According to the tradition followed by Cod. hebr. 212 and Erfurt 7, the words chosen to represent *yod*, *he* and *vav* are not at the beginning of verses. Virtually all Torah scrolls written since the thirteenth century incorporate the scribal custom of *beyah shemo*, while *vave ha-ʿamudim* was widespread, but not universal.

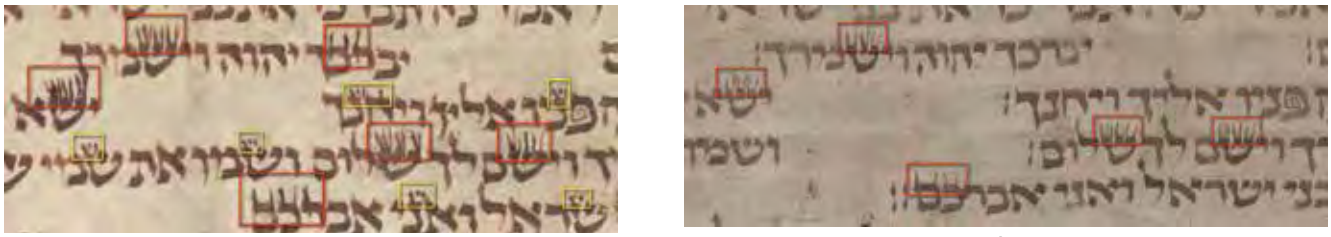
Erfurt 7 contains tagin, that is, decorative or mystical 'crowns' on many of the letters. There are two types of crowns in this manuscript. The first type consists of three small strokes added to the tops of the seven letters *shin*, *ʿayin*, *ṭet*, *nun*, *zayin*, *gimel* and *ṣadi*, which form the acronym *shaʿaṭnez geṣ*. This first type of crown also adorns the letter *ḥet*. The second type consists of larger lines added

[7] Maharam's *responsum* was published in Meʾir ben Barukh of Rothenburg, *Teshuvot Pesaqim u-Minhagim*, ed. Kahana 1959/1960, vol. 2, 150 (§ 158). According to Israel Ta-Shema, the *responsum* was actually written by Maharam's teacher rather than Maharam himself, see Ta-Shema 1977, 24.

[8] Penkower 2019, 137.

[9] Isaac of Corbeil, *ʿAmude Ha-Golah*, 1556, 58a (§ 155); Strauss 2012.

[10] On the scribal practice of *beyah shemo*, see *ʿAdat Devorim*, ed. Baer and Strack 1879, § 66; Menaḥem ben Solomon ha-Meiʾri, *Qiryat Sefer*, ed. Hershler 1956, 53 (§ 2.3); Yedidyah Norzi, *Minḥat Shai*, ed. Betzer and Ofer 2005, 59 (Gen. 1:1), 151 (Gen. 49:8), 180 (Exod. 14:28), 251 (Lev. 16:8), 312–313 (Num. 24:5), 352 (Deut. 12:28), 364 (Deut. 23:24), 375 (Deut. 31:28); Caspi 2014, 235; Penkower 2019, 135–138.



a b

Fig. 4a: Two types of crowns in Erfurt 7 (sheet 33, col. 2 [col. 98]). Small crowns (yellow) adorn the letters *shin*, *ʿayin*, *ṭet*, *nun*, *zayin*, *gimel* and *ṣadi*, known by the acronym *shaʿaṭnez geṣ*, as well as *ḥet*. Large crowns (red) adorn specific instances of specific letters, largely based on Cod. hebr. 212 or a common source. Fig. 4b: Large crowns (red) on specific instan es of specific let ers in Cod. hebr. 212 (fol. 98^{v}). The small crowns on the letters *shin*, *ʿayin*, *ṭet*, *nun*, *zayin*, *gimel*, *ṣadi* and *ḥet* are absent.

Fig. 5: Crowns added to Erfurt 7 by a later scribe stand out against the background of the original ink, which has been worn away.

lamed final *nun* *pe*

Fig. 6: Examples of regular letters (top) and 'strange letters' (bottom) in Erfurt 7.

to the tops of specific letters of specific words, largely based on Cod. hebr. 212 (Fig. 4).[11] The small crowns are withheld when the large crowns are placed on the *shaʿaṭnez geṣ* letters or *ḥet*, suggesting that the large crowns may have been added before the small crowns (or at the same time).

Some crowns were added by a later scribe. This is particularly obvious when the ink of the original scribe has flaked off, but the ink of the later scribe who added the crowns is well preserved. A clear example of this can be found in Lev. 19:19 (sheet 28, col. 3 [col. 84]), which contains the word שעטנז (an obscure term that refers to a mixture of wool and linen). This word was given special attention because it is the first part of the phrase *shaʿaṭnez geṣ*, the acronym formed from the seven letters normally marked with the first type of crown mentioned above. The later scribe marked this word with the second type of crown (Fig. 5), which is consistent with the corresponding passage in Cod. hebr. 212 (fol. 85^{r}).

Some of the letters were decorated with special paratextual features known as *otiyot meshunot* (literally, 'strange letters'), which may have had a mystical significance (Fig. 6) [12]

Various corrections were made in Erfurt 7 to resolve some common scribal errors, with most erasures carried out by abrasion, that is, scratching the ink off the parchment with a sharp implement. These sorts of corrections were done by both the original scribe and later scribes. For example, in Lev. 23:16 (sheet 29, col. 3 [87]) a second hand abraded over half a line and wrote the correction in a different type of ink than that used by the original scribe (see Fig. 7). Most of the correction is written over the abrasion, with the exception of the last three letters, which were written over an unused section of parchment. Corrections in different hands and different inks are apparent throughout Erfurt 7 (see Fig. 8).

One of the most common types of scribal errors is the confusion of graphically similar words, especially when a rare word is confused with a common one. Such an error occurred in Erfurt 7 in Num. 30:4 (sheet 39, col. 3 [col. 117]), where the scribe initially wrote the common word בנערה ('with the girl') instead of the relatively rare word

[11] Cf. Michaels, 2020, 6–7.

[12] Caspi and Veintrob 2011; Olszowy-Schlanger, 2019, 124–134; Michaels 2020, 1–20.

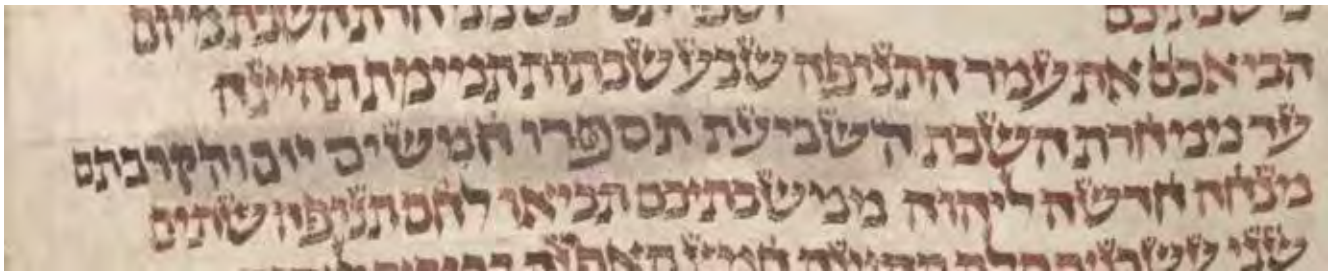
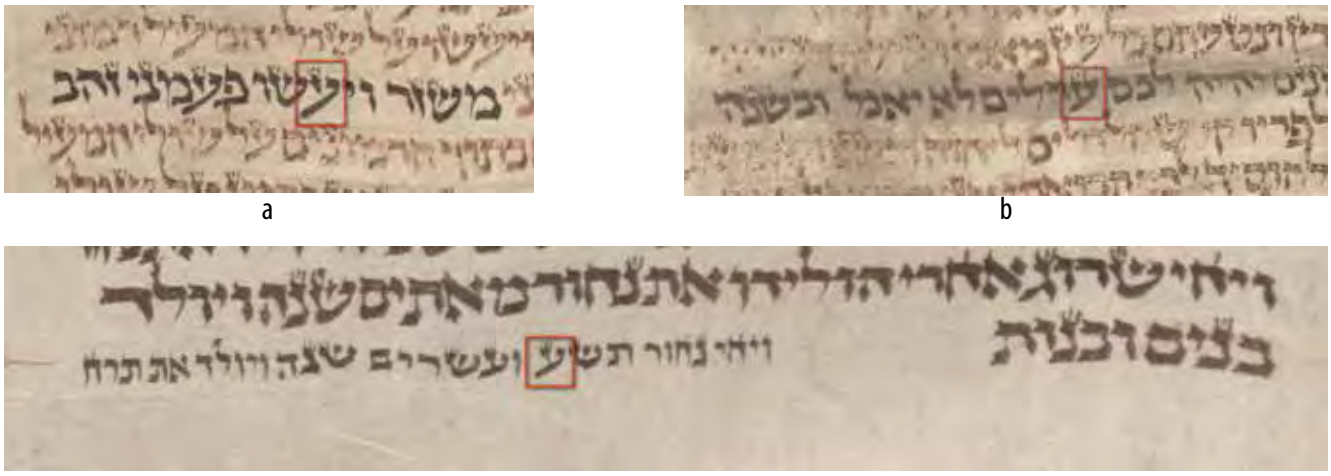


Fig. 7: Signs of abrasion using a sharp instrument to erase text in Lev. 23:16 (sheet 29, col. 3 [87]).



a b c

Fig. 8: Corrections made in Erfurt 7 by three different scribes, as evident from the appearance of the inks and the forms of the letters. The letter *ʿayin* has a different form for each of the three scribes (see the red box). Correction (a) was probably done by the original scribe of Erfurt 7 using a different batch of ink than for the main text on the sheet (sheet 23, col. 2 [col. 68]). Correction (b) was written by a second scribe over an erased section of text (sheet 28, col. 3 [col. 84]). Correction (c) was done by a third scribe who supplied a missing verse that was accidentally omitted due to *homoioarkton*, that is, two passages with similar beginnings (sheet 3, col. 1 [col. 7]).

בנעריה (‘in her youth’) (Fig. 9). While the two words have quite different pronunciations (*ban-naʿarah* vs. *bi-neʿureha*), vowel points are not written in a Torah scroll, making these graphically similar words easy to confuse. The original scribe apparently caught the error as he was writing, and stopped before completing the left leg of the *he*. It seems he then left a blank space and continued writing the next word while he waited for the ink to dry. This is consistent with the instructions of Rabbi Menaḥem ben Solomon ha-Meʾiri (1249–1315), who said that ‘[a scribe] should not wipe away wet ink because the blackness of the ink is very pronounced and remains [on the sheet], which is not [very] elegant…’.[13] Once the ink dried, the scribe of Erfurt 7 resolved the error in Num. 30:4 by first abrading the stunted left leg of the *he* with a sharp instrument. Next, he added a horizontal stroke to the remnants of the *he*, turning it into the required *resh* (technically it looks more like a *dalet*, but the scribe was satisfied). Finally, the scribe added the missing *yod* and another *he* in the blank space he had left earlier.

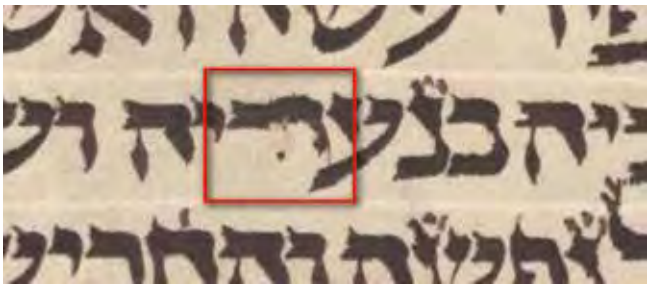


Fig. 9: A scribal error fi ed by the original scribe using abrasion. The scribe originally wrote בנערה (‘with the girl’) instead of the graphically similar בנעריה (‘in her youth’).

Sometimes the abraded text is still clearly legible. For example, in Gen. 18:19 (sheet 4, col. 2 [col. 11]) the scribe originally wrote ושמרו את דרך (‘and they will keep the way’) with an extraneous instance of the direct-object marker את. A

[13] Menaḥem ben Solomon ha-Meʾiri, *Qiryat Sefer*, ed. Hershler 1956, 53 (§ 2.3).

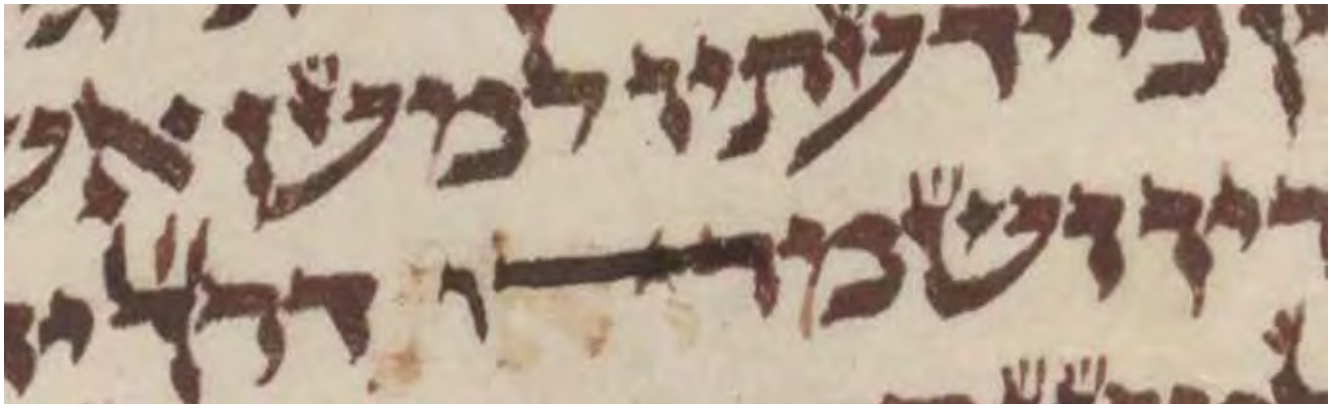
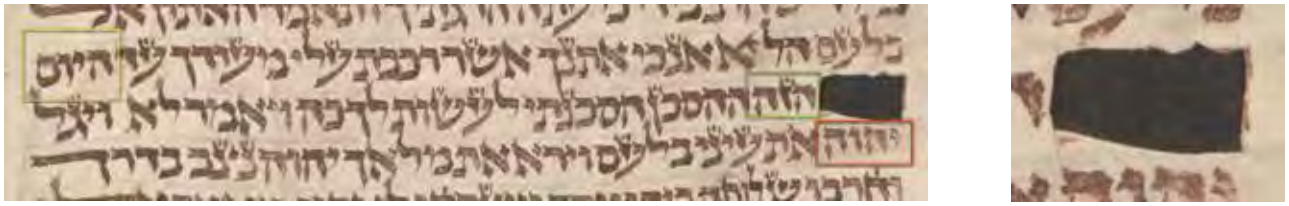


Fig. 10: The word את (a directobject marker) erased by abrasion is still clearly visible, as is the bottom portion of an erased *vav*.



a b

Fig. 11: Erfurt 7, sheet 37, col. 3 [col. 111]. a: the scribal error of *parablepsis* occurred when the scribe finished a line of text (see the yellow box at the far left) and looked back at his source. Instead of reading and copying the word הזה ('this'; marked by a green box), his eye jumped to the next line where he read the graphically similar *Tetragrammaton* יהוה (marked by a red box here). This caused him to initially skip an entire line of text. The scribe immediately realised his mistake, however, and copied the correct word הזה following the wrongly placed instance of the *Tetragrammaton*. Due to rabbinical strictures on erasing the *Tetragrammaton*, the scribe surrounded the erroneous word with an ink rectangle and later it was excised from the parchment. b: a close-up of the hole left from excising the *Tetragrammaton*.

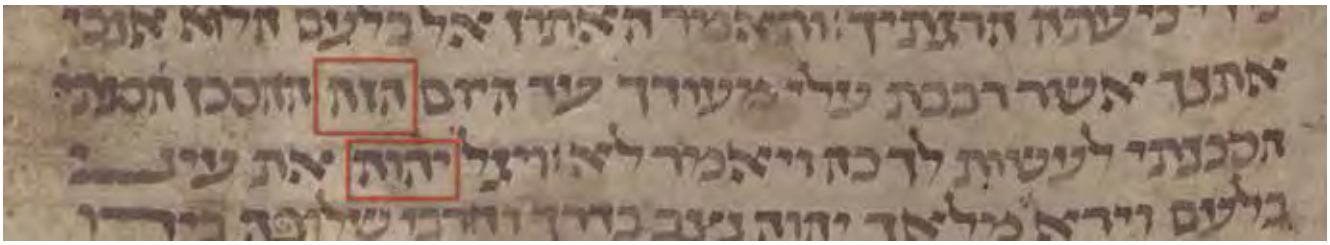


Fig. 12: Cod. hebr. 212, fol. 111ᵛ, detail showing the graphically similar words הזה ('this') in Num. 22:30 and יהוה (God's name) in Num. 22:31 on the next line.

later corrector erased the word את, although remnants of the letters are still clearly legible (Fig. 10). To remove the empty space left by the erased word, the corrector first erased the descending line of the *vav* (damage done to the parchment by the abrasion is still visible). Next, he added a horizontal line through the roof of the *resh* and the remaining top of the *vav*. Finally, he rewrote the *vav* following his dilated *resh*.

Another common type of scribal error is *parablepsis*, that is, when a scribe's eye jumps from one section of text to another as he is looking back and forth between his source and the text he is writing. The result is the scribe skipping a section of text. In Erfurt 7, this happened in Num. 22:30 (sheet 37, col. 3 [col. 111]; see Fig. 11) when the scribe's eye jumped from the word הזה ('this') on one line to the *Tetragrammaton* – a Greek term for the four-letter name of God – יהוה on the next line (Num. 22:31). In Cod. hebr. 212 (fol. 111ᵛ) the words הזה and יהוה are graphically similar (Fig. 12). The scribe of Erfurt 7 must have realised his mistake immediately because he copied the correct word הזה on an unused section of parchment, immediately afterwards. Rabbinic strictures deemed it sacrilegious to deface the *Tetragrammaton* by abrading it.[14] To solve this problem, the scribe initially surrounded the divine name with an ink rectangle, traces of which are still visible. In

[14] Gordon 2020.

a b c d e

Fig. 13: The scribe of Erfurt 7 worked around pre-existing holes in the parchment, usually leaving ample space before and after the hole. a: A pre-existing hole in the parchment that was sewn up (sheet 3, col. 1 [col. 7]). The word בנו (Gen. 11:5) is squeezed in on the third line, to the right of the stitches. b: A tear that was sewn up with an adjacent hole (sheet 12, col. 3 [col. 36]). c: The scribe contracted the letters *tav* (ת) and *he* (ה) of the word ועתה (Gen. 50:5) on the second line to fit them in to the right of the tear (sheet 13, col. 1 [col. 37]). d: Sheet 30, col. 3 [col. 90]. e: Sheet 36, col. 1 [col. 106].

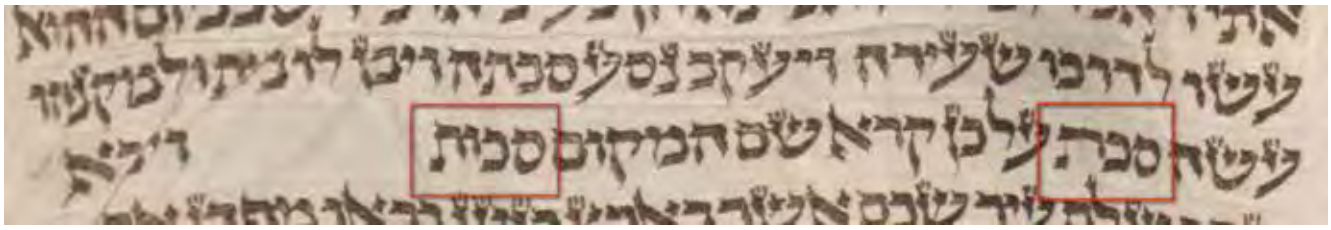


Fig. 14: The word *sukkot* spelt *plene* with a *vav* – סכות (left) – and in a 'defective' form without a *vav* – סכת (right) – in the same verse (Gen. 33:17; sheet 8, col. 3 [col. 24]).

a later phase, the original scribe or a different one excised the divine name, that is, he cut it out of the parchment using a razor or sharp knife, leaving behind a rectangular hole. In Erfurt 7, excision was used to remove twenty-two erroneous instances of the *Tetragrammaton* or another divine appellation, leaving a rectangular hole in the parchment in each case.[15] In some cases, the original scribe worked around pre-existing holes in the parchment as well as tears in it that were sewn up (Fig. 13).[16]

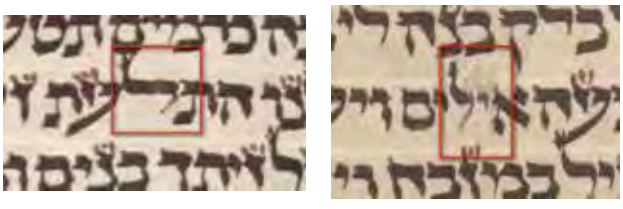
In terms of its typology, Erfurt 7 exhibits Ashkenazic characteristics that distinguish it from the standard Tiberian version of the Masoretic Text in numerous ways.[17] The Tiberian text was transmitted and standardised in Tiberias over a period of several centuries, culminating in the Aleppo Codex around the year 925. The influence of the Aleppo Codex gradually spread throughout the Jewish world over the course of several centuries, resulting in the 'correction', or rather standardisation, of Bible codices and Torah scrolls. One area of this standardisation was matters of *plene* and defective orthography. The four Hebrew letters *alef* (א), *he* (ה), *vav* (ו) and *yod* (י) can be used to indicate the presence of a vowel, in which case they are referred to as *ʾimot qeriʾah*, translated by the Latin term *matres lectionis* ('mothers of pronunciation'). When a word was written with a *mater lectionis* (sing.), it was called *male*, translated by the Latin term *plene* ('full'); when it lacked a mater lectionis, it was called *ḥaser* ('defective'). Biblical Hebrew did not have uniform orthography when it came to *matres lectionis*, so the same word could be written in a *plene* way or a 'defective' one without changing the meaning. Take the word *sukkot* ('booths' or 'name of a town'), for example, which is written as סכות (with a *vav*) in a *plene* way and as סכת (without a *vav*) in a defective way in Gen. 33:17 (Fig. 14).

The Tiberian scribes developed an intricate system to fix the precise orthography of every word in the Bible, including the use of *matres lectionis*, for every instance of every word. Erfurt 7 contains numerous corrections of *plene* and defective orthography that bring it in line with the standard Tiberian text. Changing a word from *plene* to defective spelling usually involved erasing an extraneous *vav* or *yod* as

[15] Sheet 14, col. 2 [col. 41], sheet 17, col. 2 [col. 50], sheet 27, col. 3 [col. 81] *bis*, sheet 29, col. 1 [col. 85], sheet 29, col. 3 [col. 87], sheet 32, col. 3 [col. 96], sheet 33, col. 1 [col. 97] *bis*, sheet 35, col. 1 [col. 103], sheet 37, col. 1 [col. 111], sheet 39, col. 2 [col. 116], sheet 41, col. 3 [col. 123], sheet 42, col. 1 [col. 125] *bis*, sheet 43, col. 1 [col. 127], sheet 43, col. 2 [col. 128], sheet 43, col. 3 [col. 129], sheet 44, col. 1 [col. 130], sheet 44, col. 3 [col. 132], sheet 45, col. 3 [col. 135] and sheet 50, col. 1 [col. 148].

[16] See cols 36, 65, 90 and 106, for example.

[17] Penkower 2015, 124–128.



a b

Fig. 15: Corrections of orthography in Erfurt 7. a: The *plene* התולעת was changed to the defective התלעת in Deut. 29:39 (sheet 48, col. 3 [col. 144]). b: The defective אלים was changed to the *plene* אילים in Num. 23:1 (sheet 38, col. 1 [col. 112]).

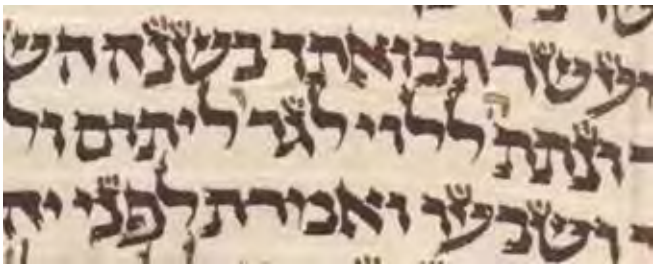
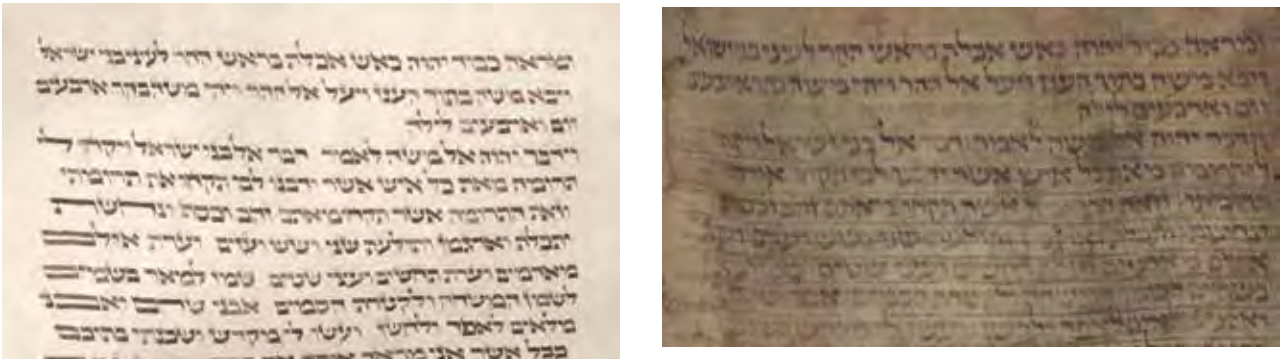


Fig. 16: The letters *he* (right) and *vav* (left) added by a later scribe as supralinear corrections, that is, above the line, in Deut. 26:12 using a different ink than the original scribe's (sheet 48, col. 1 [col. 142]).

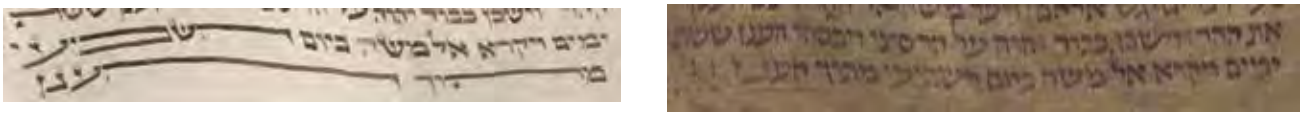


a b

Fig. 17: A replacement sheet in Erfurt 7 (sheet 19, col. 2 [col. 56]), a, compared with the corresponding verso page of Cod. hebr. 212 (fol. 56v), b. The correspondance between the two manuscripts' layout applies to the beginning of columns and usually to paragraph spacing, but not to individual lines. The replacement sheets may have been copied from the original sheets of Erfurt 7, which were presumably damaged or worn out at the time, or directly from the codex.

well as part of an adjacent letter. The partially erased adjacent letter was then dilated to cover up the erased *vav* or *yod*. For example, in Deut. 29:39 the scribe originally wrote the *plene* התולעת ('the worm') contrary to the standard Tiberian text, which has the defective התלעת. The corrector erased the extra *vav* (traces of which are still visible) as well as part of the adjacent *lamed* (ל). He then dilated the *lamed* to cover up the missing *vav* (Fig. 15a). Changing a word from defective to *plene* involved the opposite process, that is, erasing part of a letter to make room for a missing *vav* or *yod*. The partially erased letter would then be rewritten in a contracted form (as in Fig. 15a). In Num. 23:1, for example, the scribe wrote the defective word אלים ('rams') contrary to the standard Tiberian text, which has the *plene* form אילים. A later corrector, using a different ink than the original scribe's, erased part of the *lamed* and then rewrote it in a contracted way. He then added a small *yod* between the *alef* and contracted *lamed*.

Occasionally, a later scribe would add missing letters supralinearly, that is, above the line. For example, in Deut. 26:12 a later scribe using a different ink than the original scribe's added the letter *he* to the word ונתתה ('and you will give') and the letter *vav* to וליתום ('and to the orphan') (Fig. 16). The *he* was a *mater lectionis*, so its addition did not change the meaning. However, the *vav* in וליתום functioned as a conjunction with the meaning 'and'. The first of these corrections brings Erfurt 7 in line with the standard Tiberian text, whereas the second is actually contrary to it. This means there was a scribe in thirteenth- or fourteenth-century Germany with a manuscript that he believed to be authoritative and that he used to correct Erfurt 7, but this manuscript did not conform with the standard Tiberian text in every detail. The conjunction *vav* added by the later scribe is found as a textual variant in several other Ashkenazic manuscripts, which confirms this



a b

Fig. 18: The scribe of the replacement sheets dilated the final lines of sheet 19, col. 1 [col. 55], shown in snippet (a), in order to begin the following column with the word ומראה in Exod. 24:7 (Fig. 17). The corresponding section of Cod. hebr. 212 (fol. 56r) – snippet (b) – has no signifi antly dilated letters in it even though it also begins the next column with the same word as the replacement sheet in Erfurt 7. This is due to the horizontal density of the handwriting added by the scribe who wrote the replacement sheets.

Fig. 19: Erfurt 7, sheet 1, col. 2 [col. 2], detail with divine appellations (יהוה אלהים) in dark brown ink on lines two, four and fi e, surrounded by non-sacred words in lighter ink. The 0.5 cm blank space to the left of יהוה אלהים (Gen. 2:7) on the second line is due to the scribe's overestimation of how much space he would need to fill in the divine appell tions in the second stage of writing.

was a version known in medieval Germany, even though it is contrary to the standard Tiberian text.[18]

Three of the fifty sheets (19, 20 and 26) of Erfurt 7 were replaced in the late thirteenth or early fourteenth century.[19] They are roughly the same size and have the same layout as the original sheets, with three columns and sixty lines per column, and each column beginning with the letter *vav* at the beginning of a verse. Every column starts with exactly the same words as the corresponding verso page in Cod. hebr. 212, suggesting that they have the same column layout as the sheets they replaced. For example, Erfurt 7, sheet 19, column 2 (col. 56) and Cod. hebr. 212, fol. 56v both begin with the word ומראה ('and the appearance'), the first word in Exod. 24:17 (Fig. 17). The handwriting of the scribe who produced the replacement sheets was horizontally denser than the handwriting of the scribe who wrote the original sheets. This obliged him to dilate letters extensively in order to begin each column with the same word as in the original sheets; two letters alone (*tav* [ת] in מתוך and *he* [ה] in הענן) take up about seventy-five per cent of the last line of sheet 19, col. 1 [col. 55], for example (see Fig. 18a). The script and overall appearance of the replacement sheets closely resemble those of Erfurt 6 and Erfurt 8, which indicates they were produced in the same region and period and perhaps even by the same school of scribes.

3. Procedure for writing divine appellations

A Jewish custom among medieval scribes involved leaving blank spaces for divine appellations, which were filled in during a second stage of writing. This procedure has parallels in the Second Temple period and is still followed today by some scribes. According to certain medieval rabbis such as David ben Solomon ibn Abi Zimra (*c*.1479–1573), it only applied to the *Tetragrammaton YHWH*, which was considered the unique name of God. Other rabbis, such as

[18] See Kennicott, vol. 1, 417.

[19] Caspi 2014, 234–236.

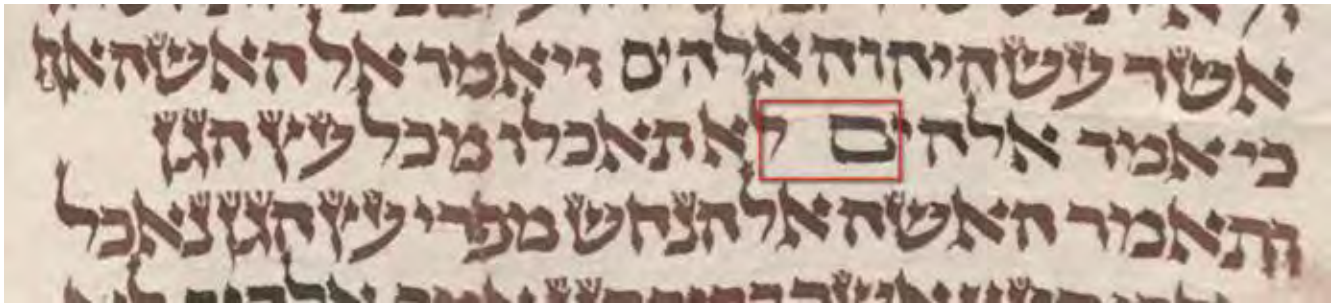


Fig. 20: The scribe misjudged how much space he would need to fill in the divine appellation *Elohim* (sheet 1, col. 2 [col. 2]). To fix this, he dilated the final *mem*, but an extra 0.5 cm of space still remained anyway. This sort of space in the middle of a line normally indicates a separation between verses in Erfurt 7. However, in this case, it is in the middle of a verse.

a b

Fig. 21: The dark brown ink of divine appellations superimposed on lighter brown ink from the line below (a: sheet 1, col. 1 [col. 1]; b: sheet 1, col. 2 [col. 2]). This could only have happened if the divine appellations were written after the text on the line below. Photographed at about 20x in visible light. Contrast enhanced.

Joseph Caro (1488–1575), applied it to all divine appellations that are not allowed to be erased according to rabbinical law, such as *YHWH*, *Adonai* ('Lord') and *Elohim* ('God').[20]

In the Middle Ages, three reasons were given for following this practice. First, divine appellations had to be written with specific intent, hence a distracted scribe might delay writing them until he could concentrate on them properly. Second, divine appellations could only be written by a scribe following his immersion in a *mikveh* (ritual bath), and so the scribe might delay writing them until he had done this. Third, there was a custom to write divine appellations in the presence of ten adult Jewish men (or according to one version, ten saintly Jewish men), hence the scribe might delay writing them until he could muster the required audience.[21]

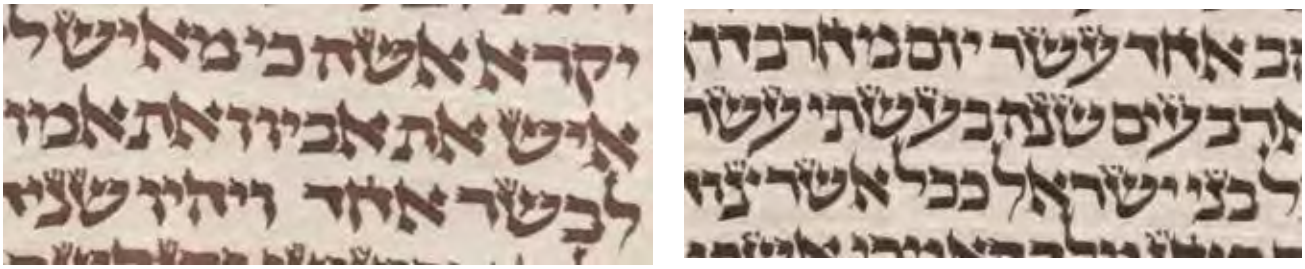
In Erfurt 7, the scribe left blank spaces for the divine appellations *Elohim* (אלהים) and *YHWH Elohim* (יהוה אלהים) throughout Gen. 1:1–3:5, corresponding to the first 1.5 columns of the first sheet.[22] The same scribe then filled them in himself (judging from the script) during a second stage of writing. The inseparable preposition *kaf* (כ) in כאלהים (Gen. 3:5) was also written during this second stage. The ink used to write these divine appellations is a darker shade of brown than that of the surrounding text (see Fig. 19).

[20] See Gordon 2020, 14, n. 23; Gordon forthcoming.

[21] Gordon forthcoming.

[22] *Elohim* and *YHWH Elohim* are the only divine appellations in this section of the biblical text.

Fig. 22: An erroneous instance of the *Tetragrammaton* יהוה (*YHWH*) was surrounded by an ink rectangle and later excised. It was replaced with אני יהוה ('I am *YHWH*') on an unused section of parchment, proving the *Tetragrammaton* was written at the same time as the rest of the text in Exod. 6:2.



a b

Fig. 23: Two shades of brown ink used by the original scribe of Erfurt 7; a: light brown (sheet 1, col. 2 [col. 2]), b: dark brown (sheet 41, col. 2 [col. 122]. The two inks were probably black originally.

In some instances, the spacing around the divine appellation is awkward, indicating that the scribe wrongly estimated how much space would be required to fill in the name in the second stage of writing; יהוה אלהים in Gen. 2:7 is followed by about 0.5 cm of blank space, for example (Fig. 19, line 2). Similarly, אלהים in Gen. 3:1 is followed by 0.5 cm of blank space and the final *mem* (ם) has been dilated to about 0.75 cm (Fig. 20). To put this into perspective, the blank space between words in this scroll is usually about 0.1 cm wide, the space between verses is about 0.5 cm wide and the final *mem* is about 0.5 cm wide.

Further evidence of the two-stage procedure can be found in places where some of the ink used for a divine appellation on one line is superimposed on the ink of a word on the line below it. In sheet 1, col. 1 [col. 1], line 20 (Gen. 1:14), the dark brown ink of the base of the final *mem* of אלהים ('God') is superimposed on the lighter ink of the mast of the second *lamed* in הלילה ('the night') on line 21 (Fig. 21a). This could only have happened if אלהים on line 20 was written after הלילה on line 21. Similarly, in sheet 1, col. 2 [col. 2], line 20 (Gen. 2:19), some of the dark brown ink used for the left leg of the first *he* of יהוה (*YHWH*) is superimposed on the lighter ink used for the mast of the *lamed* of כל ('all') on line 21 (Fig. 21b). Again, this could only have happened if יהוה on line 20 was written after כל on line 21.

The divine appellations, which begin in Gen. 3:8 (sheet 1, col. 2 [col. 2], line 45), were written at the same time as the rest of the text, as is evident from the correction of scribal errors. In Exod. 6:2, for example (sheet 14, col. 2 [col. 41]), the scribe initially made a mistake that needed erasing and a portion of the verse had to be rewritten. It was not permissible to abrade the *Tetragrammaton*, so the scribe invalidated it by drawing a rectangle around it in ink. In a later phase, the original scribe or a later one excised the *Tetragrammaton*, leaving a hole in the parchment, although traces of the ink rectangle are still visible (Fig. 22).[23] The correction אני יהוה ('I am *YHWH*') was written by the original scribe directly after the hole on a clean, unabraded section of parchment, followed by a space between the verses. Had the scribe skipped writing the divine appellations and only filled them in later, it would not have been possible to write the correction אני יהוה on a clean section of parchment immediately after the erroneous instance of the *Tetragrammaton*. This space would have been filled with the following words (וארא אל), which the scribe would have needed to erase by abrasion. Such an

[23] Gordon 2020, 124, n. 139.

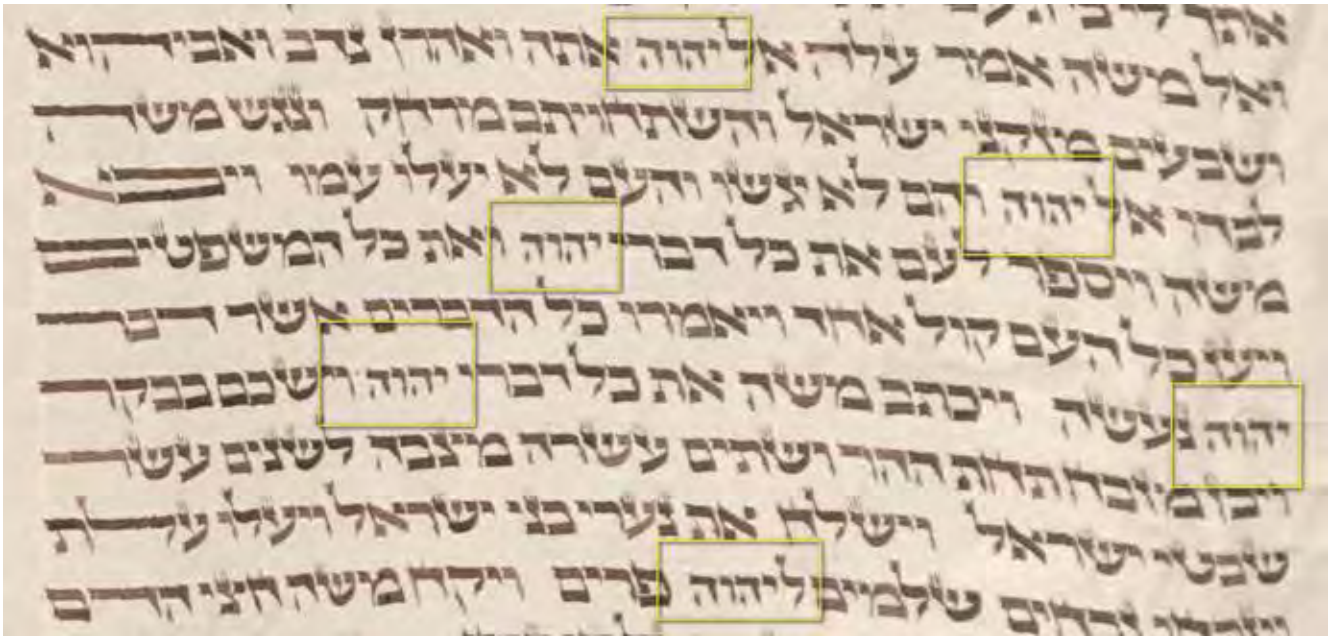


Fig. 24: Divine appellations filled in during a second phase of writing in a replacement sheet in Erfurt 7 (sheet 19, col. 1 [55]). The appellations here are slightly smaller than the surrounding text.

erasure would have left obvious marks, though, which is actually the case earlier on the same line of sheet 14, col. 2 [col. 41] where the words אל משה ויאמר אליו are written over an abrasion.[24]

The reason for only performing the two-stage procedure in the first 1.5 columns (of the original sheets) may have been hinted at by Rabbi Judah ben Samuel he-Ḥasid of Regensburg (*c*.1150–1217):

> Some say [the scribe] was required to write [the divine name] in the presence of many people to warn [him] to write it with proper intent.[25]

Thus it is possible that the scribe of Erfurt 7 performed the second stage of this procedure in the presence of members of the Jewish community of Erfurt as a way of reminding both himself and the community of the profound sanctity of the written form of the divine name. It would have been impractical to maintain a large gathering of community members for the time it took to add all the divine appellations required in a scroll, of course. Hence, the first 1.5 columns may have been chosen as a way of 'inaugurating' the writing of the divine name in a public ceremony.

The dark brown shade of ink used to write the divine appellations in the first 1.5 columns of the first sheet (Gen. 1:1–3:5) has the same colour and appearance as the ink used to write entire sheets elsewhere in the scroll (such as sheets 2, 3 and 38–44; see Fig. 23b). Similarly, the light brown shade of ink used to write the 'non-sacred' words (that is, anything other than the divine appellations) in the first 1.5 columns also has the same colour and appearance as the ink used to write whole sheets (like sheets 5, 14, 17 and 21–25; see Fig. 23a). All the inks were probably black when fresh and turned different shades of brown over time; the degradation of the Fe complex into the soluble constituents Fe^{2+} and gallic acid is known to cause brown discolouration in iron-gall inks.[26]

The divine name יהוה (*YHWH*) was inserted into blank spaces in the three replacement sheets (19, 20 and 26), sometimes in a smaller script than the surrounding text. It does not seem that the scribe miscalculated the space needed for the *Tetragrammaton* as it was sometimes written in a smaller script than the surrounding text even when there was

[24] Similarly, in Num. 12:14, the scribe initially neglected to insert an open *parashah* (a space between sections). To remedy this, he erased the verse by abrading it, with the exception of the *Tetragrammaton*, which he excised, leaving a rectangular hole. The corrected verse was then written on the following line on a clean, unabraded section of the sheet. In Deut. 4:2, the scribe initially wrote the *Tetragrammaton*, excised it and then wrote the correction immediately after the rectangular hole on a clean, unabraded section of the parchment. The fact that these corrections were written on clean, unused parts of the respective sheets indicates that the scribe caught the errors as he was writing. This means he must have been writing the *Tetragrammaton* at the same time as the rest of the words in these sections.

[25] *Sefer Hasidim*, ed. Wistinetzki 1924, 420 (§ 1762).

[26] Krekel 1999; Rabin et al. 2012.

ample room (Fig. 24). The appearance of the script suggests the divine name was added by a second scribe. It is possible that the main text was produced by an apprentice whose master filled in the *Tetragrammaton*. The inseparable preposition *lamed* (ל) was written in the second stage, often being dilated to fill in the blank space. The two-stage procedure was not followed for other divine appellations such as אלהים ('God') that appear in the three replacement sheets. The scribe(s) of the replacement sheets treated the *Tetragrammaton* with more sanctity than other divine appellations, whereas the original scribe treated all the divine appellations with equal sanctity.

The primary aim of this material analysis was to determine the relationship between the divine appellations (DN) in the first 1.5 columns (Gen. 1:1–3:5) of the original sheets (OS) and the surrounding non-sacred words (NS). Initially, it was hypothesised that the difference in shade between the dark brown (DK) divine appellations (OS.DK.DN) and the surrounding light brown (LT) non-sacred words (OS.LT.NS) occurred because of different elemental compositions and possibly even different types of ink, viz. carbon ink and iron-gall ink. A secondary aim was to compare the divine appellations written in light brown ink (OS.LT.DN) after Gen. 3:8 with the surrounding non-sacred words written in what appears to be the same light brown ink used for non-sacred words in the first 1.5 columns (OS.LT.NS). It was assumed that these two subcategories of ink (OS.LT.DN, OS.LT.NS) would have identical characteristics.

Some sheets used a dark brown ink (OS.DK) of the same colour and appearance as that used for the divine appellations in the first 1.5 columns. Similarly, other whole sheets used a light brown ink (OS.LT) of the same colour and appearance as that used for non-sacred words in the first 1.5 columns. To complete the picture of the original sheets, we proposed exploring the relationships between these two main shades of ink used throughout the scroll (OS.DK, OS.LT), the corrections performed by various hands (OS.CR) and the crowns added to the letters (OS.CW).

What the relationship was between the divine appellations (RS.DN) and the surrounding non-sacred words (RS.NS) was a key question regarding the replacement sheets (RS). Another question concerned the relationship between the inks used for the replacement sheets and the different inks used to write the original sheets (OS.LT, OS.DK, OS.CR, OS CW), along with their connection to other manuscripts in the Erfurt collection, especially scrolls that appear to be similar from a palaeographical perspective, such as Erfurt 6 and Erfurt 8.

4. Experiment

Testing was conducted using ultraviolet-visible-near-infrared (UV-vis-NIR) reflectography followed by X-ray fluorescence analysis (XRF) to determine the relationship between the inks used for the main text written in the first stage and the divine appellations written in the second stage in the original sheets and the three replacements. This two-step procedure, which has now been used successfully by BAM for more than a decade, allows a reliable, non-destructive and non-invasive investigation of inks.[27]

4.1 UV-vis-NIR reflectography

Carbon-based, plant and iron-gall inks belong to different typological classes of black ink. Soot ink is a fine dispersion of carbon pigments in a water-soluble binding agent, whereas plant-based ink consists of a solution of tannin extracts and a binding agent. Iron-gall ink combines water-soluble components (iron sulphate and tannin extract from gall nuts) with insoluble black material that evolves when the components undergo a chemical reaction. Each ink class has distinct optical properties: the colour of soot ink/carbon ink is independent of the wavelength between 300 and 1700 nm; iron-gall ink gradually loses its opacity towards long wavelengths (that is, 750–1400 nm) and becomes transparent at 1400 nm, whereas plant ink is already transparent at ~750 nm.[28] We used a portable microscope (a Dinolite AD4113T-I2V USB) with illumination from the ultraviolet (UV, 395 nm), visible (VIS) and near-infrared (NIR, 940 nm) regions of the electromagnetic spectrum and magnifications of x50 to x200 to determine the ink type and tannin distribution.

4.2 X-ray fluorescence (XRF)

Elemental analysis by X-ray emission techniques relies on the study of characteristic patterns of X-ray emissions from atoms irradiated with high-energy X-rays or electrons: X-ray fluorescence (XRF) and energy-dispersive X-ray spectroscopy (EDX) respectively. When the external excitation beam interacts with an atom within the sample, an electron is ejected from the atom's inner shell, creating a vacancy. In the next step, another electron from an outer

[27] Hahn et al. 2004, Rabin et al. 2012, Cohen et al. 2017, and Rabin 2017.

[28] Mrusek, Fuchs and Oltrogge 1995.

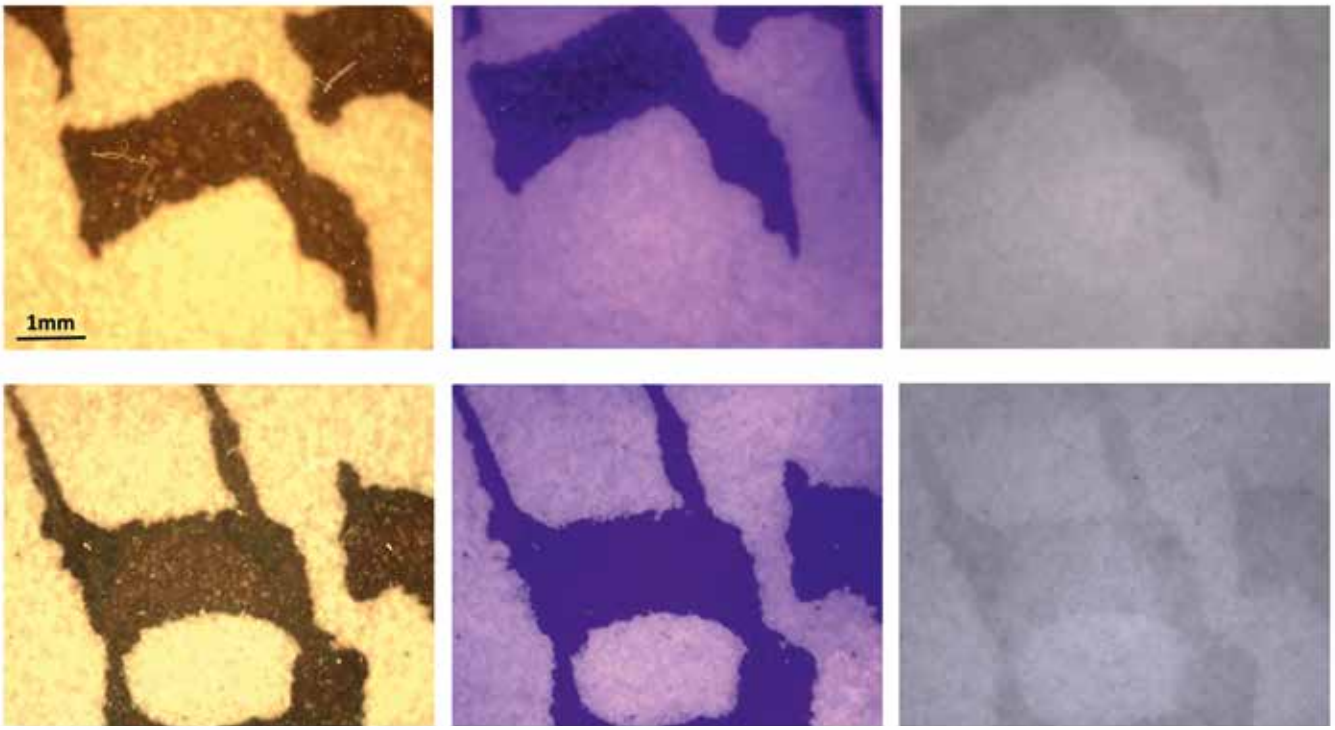


Fig. 25: Details of spots from sheet 1, col. 1, line 20 corresponding to the roof of the *resh* (ר) of ויאמר (top row) belonging to the OS.LT.NS category and the roof of the fina *mem* (ם) of אלהים (bottom row) belonging to the OS.DK.DN category under visible light (left), UV light (395 nm, middle) and near-infrared light (940 nm, right) at x50 magnifi ation. The top ink corresponds to the lighter shade of brown observed and the bottom one to the darker shade of brown. Both inks appear to have faded, but are still visible at 940 nm, which indicates that they are of the iron-gall type. The slightly yellow hue under visible light is an artifact of the LED lamps used.

shell fills the vacancy, emitting X-rays in the process. The energy of the emitted X-ray fluorescence is characteristic of a certain element, whereas the signal intensity (and a number of other factors) is related to the amount of the element in the volume sampled. It is worth noting that each technique has its limits in terms of applicability and different penetration depths. Excitation by electrons (in the case of X-ray spectroscopy, EDX), which is conventionally used in electron microscopy, is limited to the study of surfaces (but capable of detecting lightweight elements), whereas excitation by X-rays (XRF) has greater penetration power. However, conventional portable instruments are limited to detecting elements where $Z > 11$, that is, elements heavier than sodium.

The X-ray fluorescence (XRF) technique is commonly used for analysing the elemental composition of various objects concerned with cultural heritage. More specificall , earlier studies of carbon and iron-gall inks conducted by BAM and the Centre for the Study of Manuscript Cultures resulted in a database of possible metallic contaminants being created. In the case of iron-gall ink, the most common source of iron necessary to form the iron-gallate complex responsible for the black colour of the ink is vitriol, which contains different metallic contaminants (such as manganese, copper and zinc) in different quantities. An analysis of these contaminants allows researchers to compare and discriminate between different inks.

Due to time and place limitations, the X-ray analysis we performed was done *in situ* at the Staatsbibliothek using ELIO, a portable micro-XRF spectrometer from Bruker-XGLab with a 4W low-power rhodium tube, a 25 mm^2 large-area silicon drift detector (SDD) and an interaction spot of 1 mm. All of our measurements were conducted under the following experimental conditions: a spot analysis of 120 seconds and excitation parameters of 40 kV and 80 µA.

4.3 X-ray fluorescence (XRF) data treatment

After measurement, the spectra were processed with Spectra (ARTAX) software from Bruker to identify the elements and determine their net peak intensities. The contribution of the support was then subtracted. The thickness of ink can vary considerably, depending on the spot analysed, so making a direct comparison of net peak intensities can lead to incorrect interpretations. To compare iron-gall inks

successfully, we used iron as a standardisation parameter for all the other inorganic components to be quantified, that is, the intensities due to the contaminants were normalised to that of iron, following the semi-quantitative adaptation of the method described by Hahn, Malzer, Kanngießer and Beckhoff (2004; see note 27 above).

Finally, to determine whether the differences observed between several groupings of data were statistically significant, Welch's t-test was performed (an adaptation of Student's t-test to compare the means of two independent groups without requiring the variances to be the same).[29] In Welch's test, which is similar to Student's t-test, but not used as often, the null hypothesis that two groups of data have equal means is tested. Unlike Student's t-test, the two tested groups can have different variance values. A t-value is computed, which depends on the difference between the two means and the standard deviations, and then compared with a table of values depending on the degree of freedom of the dataset.[30] If the computed t-value is bigger than the tabulated critical t-value for a given significance level α, then the null hypothesis can be rejected and the probability that the means are statistically different is (1-α/2)*100%.

The degree of freedom is approximated by the following formula,[31] where σ_A and σ_B are the standard variations of samples in groups A and B respectively and n_A and n_B are the number of samples in groups A and B respectively (the result is then rounded up or down to the closest integer):

$$df = \frac{\left(\frac{\sigma_A^2}{n_A}+\frac{\sigma_B^2}{n_B}\right)^2}{\frac{\sigma_A^4}{n_A^2(n_B-1)}+\frac{\sigma_B^4}{n_B^2(n_A-1)}}.$$

The t-value is calculated as follows, where μA and μB are the mean values of sample groups A and B respectively, σ_A and σ_B are the standard variations of samples in groups A and B respectively and n_A and n_B are the number of samples in groups A and B respectively:

$$t = \frac{|\mu_A-\mu_B|}{\sqrt{\frac{\sigma_A^2}{n_A}+\frac{\sigma_B^2}{n_B}}}.$$

5. Results

All of the inks investigated via reflectograph , both the light brown (OS.LT) and dark brown (OS.DK) ones from the original sheets and the inks on the replacement sheets (RS), were found to be of the iron-gall type. When illuminated by near-infrared light, the iron-gall ink faded, remaining slightly visible, as seen in Figure 25 (right). In contrast, carbon ink would have remained unfaded under near-infrared light. No carbon ink was detected anywhere in the scroll.

Following the reflectographic survey, we conducted XRF measurements to determine the elemental composition and relative intensities of the inks. To that end, 35 spots in the ink were examined and 11 spots were analysed on the parchment support so that the contribution of the parchment could be subtracted.

The original aim of the investigation was to compare eight subcategories of inks based on their appearance and function within the text. These subcategories distinguished between the two primary inks used throughout the original sheets (OS), specifically light brown ink (LT) and dark brown ink (DK). They also differentiated between divine appellations (DN), non-sacred words (NS), corrections (CR) and crowns (CW). Additionally, two subcategories were delineated in the replacement sheets (RS), namely divine appellations (DN) and non-sacred words (NS). The result was six subcategories in the original sheets (OS.LT.DN, OS.LT.NS, OS.DK.DN, OS.DK.NS, OS.CR and OS.CW) and two subcategories in the replacement sheets (RS.DN and RS.NS). However, limitations in terms of time, place and equipment prevented us from collecting enough XRF data points to be able to draw any reliable conclusions about all the subcategories. As a result, the statistical analysis had to employ broader categories based on the visual appearance of the ink (OS.LT, OS.DK, RS) as representative of some of the most important subcategories. The difference between the inks used in the original sheets and replacements was obvious. Although inks from the original sheets did not contain any copper or zinc (or only had traces of them), the inks from the replacement sheets all exhibited traces of copper and high counts for zinc, as Figure 26 shows.

As for the two shades of ink observed in the original sheets, OS.LT and OS.DK, the distinction is less obvious since both shades contain the same elements and the spectra look similar at first glance. Counts for iron, potassium and calcium vary from one spot to another, but cannot be directly correlated with differences in the appearance of the inks. However, the two groups are clearly

[29] Welch 1951. The variance is the average of the squared differences from the mean, and the standard deviation is the square root of the variance.

[30] Otto 2017, 32.

[31] Otto 2017, 35.

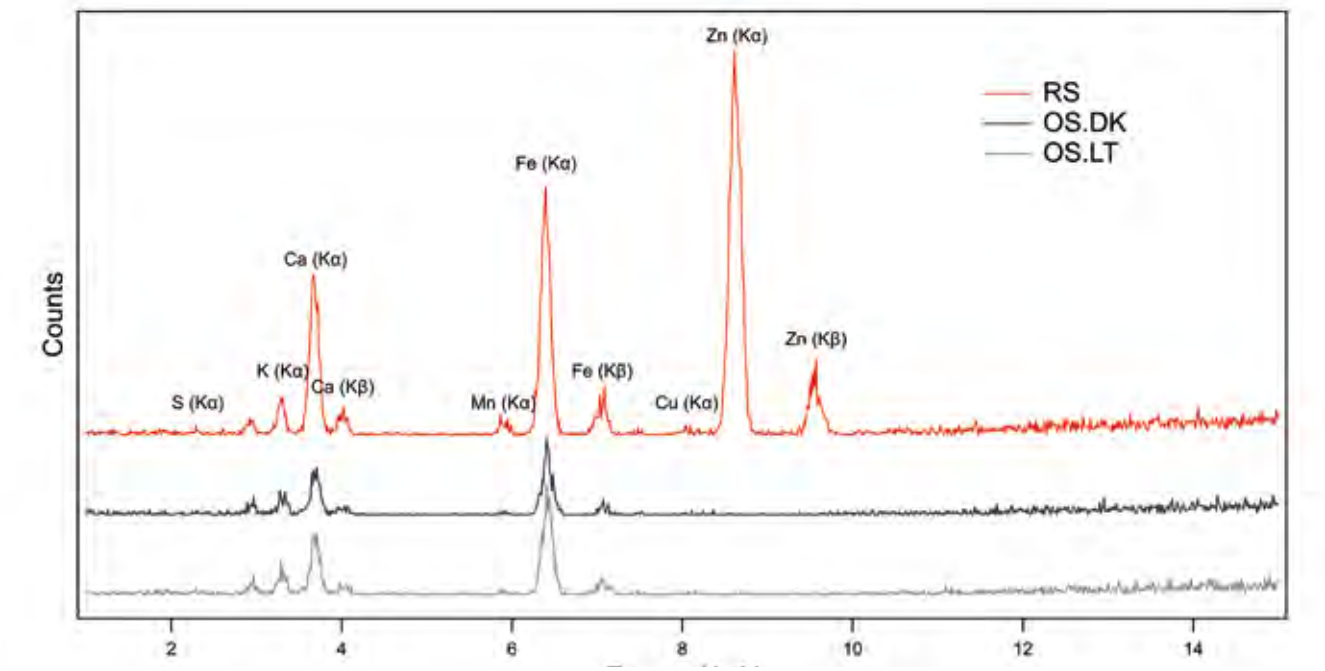


Fig. 26: XRF spectra of ink from OS.LT (sheet 1, column 1, line 20, roof of *resh* from ויאמר, bottom), OS.DK (sheet 1, column 1, line 20, roof of final *mem* from אלהים, middle) and RS (sheet 19, column 1, line 41, *yod* from דברי, top). In addition to the elements indicated, a peak at 2.96 keV corresponds to the Kα of argon (Ar), which was detected because the analysis was performed in open air. This element is not relevant to our study.

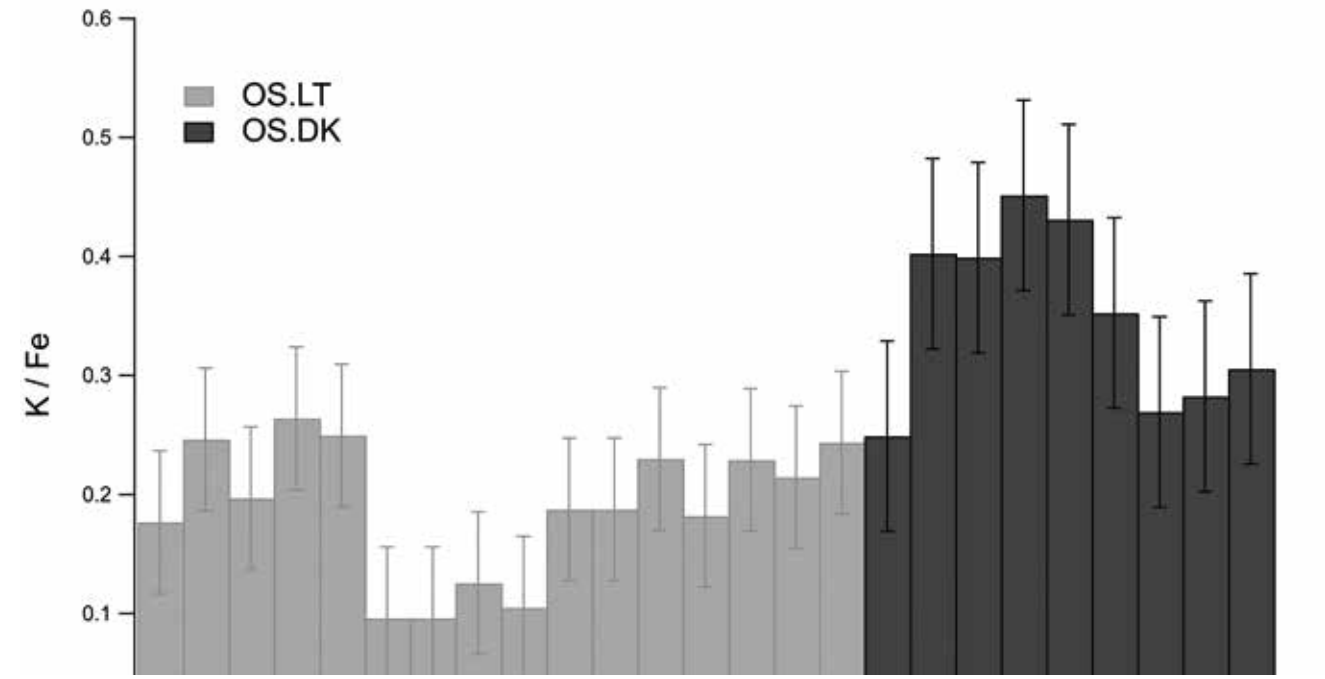


Fig. 27: Comparison of K/Fe for both shades of ink observed on the original sheets.

distinguished from the potassium normalised to iron (Fig. 27). Inks with the lighter shade have a lower potassium-to-iron ratio, with an average value of 0.19 (a standard deviation of 0.06), whereas inks with the darker shade have an average K/Fe value of 0.35 (a standard deviation of 0.08). Table 1 below shows the list of ink spots analysed and summarises the XRF results:

To determine whether the difference between the K/Fe mean value of OS.LT and OS.DK data was statistically significant, Welch's t-test was performed. The t-value is well above the critical threshold of the 0.5 per-cent significance

level (Table 2). Therefore, the probability that the means are statistically different is 99%. We tentatively attribute the two shades of ink from the original sheets to different writing sessions for which the scribe prepared different batches of ink with different ratios of potassium to iron.

6. Discussion

Iron-gall ink was used throughout the scroll, both on the original sheets and the replacement ones, whereas there was no sign of any carbon ink. This finding is a significant one because there are discussions in rabbinic literature about the permissibility and desirability of adding vitriol (*calcanthum*) to ink (for example, in the *c*. third-century minor tractate *Sefer Torah* 1:5, the *c*. eighth-century minor tractate *Soferim* 1:6 and the twelfth-century *Mishneh Torah* of Maimonides (*Tefillin, Mezuzah and the Torah Scroll* 1:4).[32]

The primary question in this study was what the relationship was between the ink used for the divine appellations in Gen. 1:1–3:5 (the first 1.5 columns) and what was used for the surrounding text. It was clear from visual observations that the two inks were consistently of different shades, indicating that they were probably employed during two separate writing sessions. Unfortunately, limitations in time, place and equipment prevented us from collecting enough XRF data points in Gen. 1:1–3:5 to be able to arrive at unequivocal conclusions about the inks in this section. However, some tentative conclusions can be reached by regarding the two shades of ink used throughout the original sheets (OS.DK and OS.LT) as representative of the visually similar darker and lighter inks used to write the divine appellations and non-sacred words in Gen. 1:1–3:5 respectively (OS.DK.DN and OS.LT.NS). Based on this approach, it can be cautiously concluded that the two shades of ink have the same basic elemental composition and correspond to different batches of ink with different ratios of potassium to iron. The two inks having the same elemental composition is consistent with the divine appellations having been added by the same scribe who wrote the non-sacred words, which appears to be the case on palaeographical grounds.

Another question we answered concerned the relationship between the three replacement sheets (19, 20 and 26) and the original ones. Although the ink on the original sheets did not contain any copper or zinc (or just traces of the elements), the ink from the replacement sheets is characterised by traces of copper and high zinc counts. Such a high signal for zinc may be characteristic of Erfurt as high zinc counts were already found in previous investigations of Erfurt 1 (Staatsbibliothek zu Berlin, Or. fol. 1210-1211)[33] and more recently in a number of manuscripts produced or annotated in Erfurt.[34] Finally, it is highly significant that the three replacement sheets in Erfurt 7 not only closely resemble Erfurt 6 from a palaeographical perspective, but also have the same zinc-rich ink. The latter finding strongly supports the attribution of the replacement sheets to the Erfurt area.

7. Conclusions

This survey was conducted with limitations in terms of time, location and equipment. More specificall , the tests were limited to a two-day period at the facility of the Staatsbibliothek zu Berlin. The examinations required a small portable XRF spectrometer with an interaction spot of 1 mm. In view of these confines, there was only enough time to analyse forty-six 1 mm spots (thirty-five in ink and eleven in the background parchment). Despite our data set being so limited, some important results were obtained:

1. The two shades of ink used throughout the scroll contain the same chemical composition of iron-gall ink.
2. It was tentatively confirmed that the first 1.5 columns of the scroll were written in two stages.
3. A relationship was shown to exist between the zinc-rich ink used for the replacement sheets and other Hebrew manuscripts and Latin annotations from the Erfurt collection.

These results give us a glimpse into the life of the Jews of Erfurt who gathered as a community to inaugurate the writing of the divine names in the Torah scroll, which would serve them in public liturgy for about a century. It seems that the scroll was also repaired using three replacement sheets in this very same community.

These conclusions notwithstanding, some of the textual and paratextual elements could not be reliably analysed. These included ‘crowns’ and other paratextual features for which the

[32] Higger 1930, 22; *Masekhet Soferim*, ed. Higger 1937, 99–100; Maimonides, *Mishneh Torah*, ed. Qafih 2007, vol. 2, 309. For the English translation, see Higger 1930, 9–10.

[33] Hahn et al. 2008.

[34] BAM performed an experimental analysis of the inks used in these manuscripts in 2014, followed by a more detailed investigation of inks from Erfurt 2 in 2019. The results of both experiments are discussed in Martini et al. forthcoming a, and in Martini et al. forthcoming b.

Table 1: Ink spots analysed, with the corresponding category/subcategory, sheet (Sh), column (C), line (L), overall column number (#), word, letter and counts for the elements of interest (after subtracting the parchment's contribution, arbitrary units): sulphur (S), potassium (K), calcium (Ca), manganese (Mn), iron (Fe), copper (Cu) and zinc (Zn). Counts for potassium normalised to iron are also indicated.

Category	Sh	C	#	L	Word	Letter	S	K	Ca	Mn	Fe	Cu	Zn	K/Fe
OS.LT.NS[1]	1	1	1	20	ויאמר	roof of *resh*	11	123	177	17	656	0	0	0.19
OS.DK.DN[2]	1	1	1	20	אלהים	roof of final *mem*	9	101	73	11	375	0	0	0.27
OS.LT.NS	1	1	1	21	הלילה	roof of second *lamed*	10	98	100	17	523	0	0	0.19
OS.LT.NS	1	1	1	54	ויברך	roof of *bet*	20	131	160	22	570	0	0	0.23
OS.DK.DN	1	1	1	54	אלהים	roof of *he*	14	191	105	21	676	0	0	0.28
OS.LT.NS	1	1	1	54	את	foot of *tav*	23	113	143	18	621	0	0	0.18
OS.LT.NS	1	1	1	55	מכל	roof of *lamed*	24	119	79	19	555	0	0	0.21
OS.DK.DN	1	2	2	40	כאלהים	roof of *kaf*	0	334	181	23	1093	0	0	0.31
OS.LT.NS	1	2	2	41	למאכל	roof of first *lamed*	0	189	188	12	776	0	0	0.24
OS.LT[3]	1	2	2	52	ויאמר	roof of resh	5	70	0	0	396	0	0	0.18
OS.LT	1	2	2	52	אלהים	roof of *lamed*	21	160	36	7	650	0	0	0.25
OS.LT	1	2	2	52	אל	oblique line of *alef*	10	111	36	5	564	0	0	0.20
OS.LT	1	2	2	54	ויאמר	roof of *resh*	13	137	22	26	519	0	0	0.26
OS.LT	1	2	2	54	יהוה	roof of first *he*	14	130	0	13	521	0	0	0.25
OS.LT	1	2	41	40	אלהים	roof of *he*	29	143	409	20	1492	0	0	0.10
OS.CR[4]	14	2	41	40	ויאמר	oblique line of *alef*	16	111	425	39	1315	0	0	0.08
OS.DK[5]	14	2	41	40	יהוה	roof of first *he*	19	282	143	18	1133	0	0	0.25
OS.LT	14	2	41	40	וארא	top of *vav*	29	143	409	20	1492	0	0	0.10
OS.LT	14	2	53	36	וגנב	base of *bet*	20	200	166	29	1593	0	0	0.13
RS[6]	18	1	55	41	דברי	*yod*	0	167	362	83	1658	44	3152	0.10
RS	19	1	55	41	יהוה	*yod*	7	122	161	79	1111	82	2863	0.11

[1] OS.LT.NS: light ink from original sheets corresponding to non-sacred words.

[2] OS.DK.DN: dark ink from original sheets corresponding to divine appellations.

[3] OS.LT: other light ink from original sheets.

[4] OS.CR: ink of corrections from original sheets.

[5] OS.DK: other dark ink from original sheets.

[6] RS: ink from replacement sheets (19, 20, 26).

Category	Sh	C	#	L	Word	Letter	S	K	Ca	Mn	Fe	Cu	Zn	K/Fe
RS	19	2	77	53	אני	*yod*	24	150	45	87	1349	96	3349	0.11
RS	26	2	77	53	יהוה	*yod*	27	167	14	116	1624	132	3848	0.10
OS.CW[7]	26	3	83	36	שעטנז	crowns	0	129	81	18	544	2	0	0.24
OS.LT	28	3	83	36	כלאים	oblique line of *alef*	9	136	151	3	1296	0	0	0.10
OS.CR	28	3	83	43	לכם	roof of *kaf*	23	128	1505	55	1470	19	31	0.09
OS.CR	28	3	83	42	להטיבך	roof of *he*	25	89	1624	30	532	4	5	0.17
OS.DK	43	3	129	55	דבר	roof of *resh*	8	214	189	14	532	0	0	0.40
OS.DK	43	3	129	55	יהוה	roof of first *he*	1	158	146	11	396	0	0	0.40
OS.CR	43	3	129	55	אלהיך	erasure line	2	19	0	0	76	0	0	0.25
OS.DK	43	3	129	55	אלהיך	leg of final *kaf*	0	102	0	6	226	0	0	0.45
OS.DK	43	3	129	55	לך	roof of final *kaf*	3	187	145	3	434	0	0	0.43
OS.CW	43	3	129	56	לפניך	right crown of *nun*	4	25	49	6	57	0	0	0.44
OS.DK	43	3	129	56	לפניך	base of *nun*	1	111	65	7	315	0	0	0.35

[7] OS.CW: ink of crowns from original sheets.

inked area is smaller than the interaction-spot size of the micro-XRF spectrometer (1 mm). Furthermore, additional tests are warranted with analysis of more samples from each subcategory of ink in the scroll. This could be accomplished by using an XRF imaging spectrometer that has an adjustable interaction spot ranging from 50 to 650 μm and would allow whole columns of the scroll to be imaged. This would provide us with better statistics as well as readily available information from spatial maps of the elements' distribution. Future tests using this approach would complement the preliminary results presented in this article. The following studies could be performed in this manner:

1. A definitive comparison of the dark brown and light brown inks used to write the divine appellations and the main text respectively in Gen. 1:1–3:5 (the first 1.5 columns) without recourse to a comparison with other sections of the scroll.

2. Investigating the relationship between the divine appellations in the replacement sheets and the surrounding text, which our visual observations indicate were written in two stages.

3. Examining what relationship corrections in the original sheets (written in different scribal hands) have to each other, to the main text of the original sheets and to the text on the replacement sheets.

4. Examining the relationship between the main text and the 'crowns'; the latter were clearly added to the letters, but it is unclear whether the original scribe or a later one made these additions.

5. Looking at how crowns that were added later are related to the original crowns and other corrections.

Analysis with an XRF imaging spectrometer of the aforementioned type would give us an unprecedented glimpse of the life of a medieval Ashkenazic Torah scroll and the Jewish community that once produced it, cherished it, maintained it and revised it.

Table 2: Comparison of K/Fe for the two shades of ink observed on the original sheets.

	OS.LT	OS.DK
Number of spots analysed (n)	16	9
Average K/Fe (μ)	0.19	0.35
K/Fe standard deviation (σ)	0.06	0.08
Degree of freedom from Welch's t-test (df)	$df = \frac{\left(\frac{\sigma^2_{OS.LT}}{n_{OS.LT}} + \frac{\sigma^2_{OS.DK}}{n_{OS.DK}}\right)^2}{\frac{\sigma^4_{OS.LT}}{n^2_{OS.LT}(n_{OS.DK}-1)} + \frac{\sigma^4_{OS.DK}}{n^2_{OS.DK}(n_{OS.LT}-1)}} \approx 22$	
t-value from Welch's t-test	$t = \frac{\lvert\mu_{OS.LT} - \mu_{OS.DK}\rvert}{\sqrt{\frac{\sigma^2_{OS.LT}}{n_{OS.LT}} + \frac{\sigma^2_{OS.DK}}{n_{OS.DK}}}} = 17.6$	
Critical t-value for a signifi ance level α of 0.02	2.508	

REFERENCES

ʿ*Adat Devorim* (1879), in Seligman Baer and Hermann Leberecht Strack (eds.), *Die Dikduke ha-Teamim des Ahron ben Moscheh ben Ascher und andere alte grammatisch-masorethische Lehrstücke zur Feststellung eines richtigen Textes der hebräischen Bibel*, Leipzig: Fernau.

Caspi, Efraim (2014), 'Burnt *Gevil* on the Sabbath of New Moon: Concerning the Torah Scrolls in the Erfurt Collection', *Yerushatenu*, 7: 231–249 (in Hebrew).

—— and Mordechai Veintrob (2011), 'Concerning the Date of Sefer Tagin', *Yerushatenu*, 5: 300–310 (in Hebrew).

Cohen, Zina, Judith Olszowy-Schlanger, Oliver Hahn and Ira Rabin (2017), 'Composition Analysis of Writing Materials in Geniza Fragments', in Irina Wandrey (ed.), *Jewish Manuscript Cultures*, Berlin: De Gruyter, 323–337.

Gordon, Nehemia (2020), 'Blotting Out the Name: Scribal Methods of Erasing the *Tetragrammaton* in Medieval Hebrew Bible Manuscripts', *Textus*, 29: 8–43, 111–155.

—— (forthcoming), 'Scribal Procedures for Writing the *Tetragrammaton* in Medieval Hebrew Bible Manuscripts', *Journal of Jewish Studies*.

Hahn, Oliver, Wolfgang Malzer, Birgitt Kanngießer and Burkhard Beckhoff (2004), 'Characterization of Iron-Gall Inks in Historical Manuscripts and Music Compositions Using X-ray Fluorescence Spectrometry', *X-ray Spectrometry*, 33: 234–239.

——, Timo Wolff, Hartmut-Ortwin Feistel, Ira Rabin and Malachi Beit-Arié (2008), 'The Erfurt Hebrew Giant Bible and the Experimental XRF Analysis of Ink and Plummet Composition', *Gazette du livre médiéval*, 51: 16–29.

Higger, Michael (ed. and trans.) (1930), *Seven Minor Treatises*, New York: Bloch.

—— (ed.) (1937), *Masekhet Soferim*, New York: Debe Rabanan.

Isaac of Corbeil (1556), ʿ*Amude ha-Golah*, Cremona: Vincenzo Conti.

Kennicott, Benjamin (1776–1780), *Vetus Testamentum Hebraicum cum Variis Lectionibus*, 2 vols, Oxford: Clarendon.

Krekel, Christoph (1999), 'The Chemistry of Historical Iron Gall Inks', *International Journal of Forensic Document Examiners*, 5: 54–58.

Maimonides, Moses, *Mishneh Torah*, ed. by Yosef Qafih (2001–2007), 25 vols, 4th edn, Jerusalem: Machon Mishnat ha-Rambam.

Martini, Annett, Zina Cohen, Olivier Bonnerot, Oliver Hahn, and Ira Rabin (forthcoming a), 'New Perspectives on the Hebrew Manuscripts from the Erfurt Collection on Grounds of Experimental XRF Ink Analysis'.

——, Olivier Bonnerot, Grzegorz Nehring, Zina Cohen, and Ira Rabin (forthcoming b), 'The Genesis and Reception of the Bible Codex Ms. or. fol. 1212 in the Light of Ink Analysis'.

Meʾir ben Barukh of Rothenburg, *Teshuvot Pesaqim u-Minhagim*, ed. by Isaac Zeʾev Kahana (1957–1977), 4 vols, Jerusalem: Mosad ha-Rav Kook.

Meʾir ha-Kohen of Rothenburg (1524), *Haggahot Maimuniyyot*, Venice: Bragadini.

Menaḥem ben Solomon ha-Meʾiri, *Qiryat Sefer: ʿal Hilkhot Sefer Torah*, *Tefillin u-Mezuzah*, ed. by Moshe Hershler (1956), Jerusalem: Hamasorah.

Michaels, Marc (2020), *Sefer Tagin Fragments from the Cairo Genizah*, Leiden: Brill.

Mrusek, Ralf, Robert Fuchs and Doris Oltrogge (1995), 'Spektrale Fenster zur Vergangenheit. Ein neues Reflektographieverfahren zur Untersuchung von Buchmalerei und historischem Schriftgut', *The Science of Nature*, 82/2: 68–79.

Olszowy-Schlanger, Judith (2019), 'The Making of the Bologna Scroll: Palaeography and Scribal Traditions', in Mauro Perani (ed.), *The Ancient Sefer Torah of Bologna: Features and History*, Leiden: Brill, 107–134.

Otto, Matthias (2017), *Statistics and Computer Application in Analytical Chemistry*, 3rd edn, Weinheim: Wiley-VCH Verlag GmbH & Co.

Penkower, Jordan (2014), 'The Ashkenazi Pentateuch Tradition as Reflected in the Erfurt Hebrew Bible Codices and Torah Scrolls' in Frank Bussert, Sarah Laubenstein and Maria Stürzebecher (eds), *Zu Bild und Text im jüdisch-christlichen Kontext im Mittelalter* (Erfurter Schriften zur jüdischen Geschichte, vol. 3), Jena: Bussert & Stadeler, 118–141.

—— (2019), 'The 12th–13th Century Torah Scroll in Bologna: How It Differs from Contemporary Scrolls' in Mauro Perani (ed.), *The Ancient Sefer Torah of Bologna: Features and History*, Leiden: Brill, 135–166.

—— (2020), 'Graphic Fillers and Marginal Corrections in the Bologna Torah Scroll' in Yitzhak Berger and Chaim Milikowsky (eds), '*In the Dwelling of a Sage Lie Precious Treasures*': *Essays in Jewish Studies in Honor of Shnayer Z. Leiman*, New York: Yeshiva University Press, 33–49.

Rabin, Ira, Roman Schütz, Anka Kohl, Timo Wolff, Roald Tagle, Simone Pentzien, Oliver Hahn, and Stephen Emmel (2012), 'Identification and Classification of Historical Writing Inks in Spectroscopy: A Methodological Overview', *Comparative Oriental Manuscript Studies Newsletter*, 3: 26–30.

—— (2017), 'Building a Bridge from the Dead Sea Scrolls to Mediaeval Hebrew Manuscripts', in Irina Wandrey (ed.), *Jewish Manuscript Cultures*, Berlin: De Gruyter, 309–322.

Strauss, Abraham (2012), 'Space between Verses in Writing Torah Scrolls, Tefillin and Mezuzahs', *Yerushatenu*, 6: 207–222 (in Hebrew).

Ta-Shema, Israel (1977), 'Towards a Characterisation of 14th-century German Rabbinic Literature', *Alei Sefer*, 4: 20–41 (in Hebrew).

Welch, Bernard Lewis (1951), 'On the Comparison of Several Mean Values: An Alternative Approach', *Biometrika*, 38: 330–336.

Wistinetzki, Jehuda (ed.) (1924), *Yehuda he-Ḥasid, Sefer hasidim = Das Buch der Frommen nach der Rezension in Cod. de Rossi No. 1133*, mit einer Einleitung von Jakob Freimann, Frankfurt: Wahrmann.

Yedidyah Norzi, *Minḥat Shai*, ed. by Zvi Betzer and Yosef Ofer (2005), Jerusalem: World Union of Jewish Studies.

PICTURE CREDITS

Contributors

Olivier Bonnerot

BAM Federal Institute for Material Research and Testing

Division 4.5

Unter den Eichen 44-46

12203 Berlin, Germany

Universität Hamburg

Centre for the Study of Manuscript Cultures

Warburgstr. 26

20354 Hamburg, Germany

eMail: olivier.bonnerot@bam.de

Khamvone Boulyaphonh

Buddhist Archives of Luang Prabang

Sala Thammavihan, Vat Suvannakhili

06000 Luang Prabang, Lao PDR

eMail: khamvone.boulyaphonh@gmail.com

Ken Boydston

MegaVision Inc.

P.O. Box 60158

Santa Barbara, CA 93160, USA

eMail: ken@mega-vision.com

Joseph Chazalon

EPITA Research and Development Laboratory (LRDE)

14-16 rue Voltaire

94270 Le Kremlin-Bicêtre, France

eMail: joseph.chazalon@lrde.epita.fr

Vincent Christlein

Friedrich-Alexander-Universität Erlangen-Nürnberg

Department of Computer Science

Martensstraße 3

91058 Erlangen, Germany

eMail: vincent.christlein@fau.de

Vincent Christlein

Friedrich-Alexander-Universität Erlangen-Nürnberg

Department of Computer Science

Martensstraße 3

91058 Erlangen, Germany

eMail: vincent.christlein@fau.de

Roger L. Easton Jr.

Chester F. Carlson Center for Imaging Science

Rochester Institute of Technology

54 Lomb Memorial Drive

Rochester, NY 14623-5604, USA

eMail: rlepci@cis.rit.edu

Andreas Fischer

University of Fribourg

Department of Informatics

PER 21 bu. B429

Boulevard de Pérolles 90

1700 Fribourg, Switzerland

eMail: andreas.fischer@unifr.ch

Michael Gartley

Chester F. Carlson Center for Imaging Science

Rochester Institute of Technology

54 Lomb Memorial Drive

Rochester, NY 14623-5604, USA

eMail: mxgpci@rit.edu

Uwe Golle

Klassik Stiftung Weimar (KSW)

3.6 Abteilung Restaurierung, Direktion Museen

Burgplatz 4

99423 Weimar, Germany

eMail: uwe.golle@klassik-stiftung.de

Marcel Gygli (Würsch)
Fachhochschule Nordwestschweiz –
Hochschule für Technik
Institut für Interaktive Technologien
Bahnhofstrasse 6
5210 Windisch, Switzerland

eMail: marcel.gygli@fhnw.ch

Nehemia Gordon
Bar-Ilan University
Bible Department
Ramat Gan, 5290002, Israel

eMail: ngordon4@gmail.com

Volker Grabowsky
Universität Hamburg
Fakultät für Geisteswissenschaften
Asien-Afrika-Institut
Edmund-Siemers-Allee 1, Flügel Ost
20146 Hamburg, Germany

Universität Hamburg
Centre for the Study of Manuscript Cultures
Warburgstr. 26
20354 Hamburg, Germany

eMail: volker.grabowsky@uni-hamburg.de

Oliver Hahn
Universität Hamburg
Centre for the Study of Manuscript Cultures
Warburgstr. 26
20354 Hamburg, Germany

eMail: oliver.hahn@uni-hamburg.de

BAM Federal Institute for Material Research and Testing
Division 4.5
Unter den Eichen 44-46
12203 Berlin, Germany

eMail: oliver.hahn@bam.de

Vanessa Hannesschläger
Österreichische Akademie der Wissenschaften
ACDH-CH – Austrian Centre for Digital Humanities and Cultural Heritage
Sonnenfelsgasse 19
1010 Wien, Austria

eMail: vanessa.hannesschlaeger@oeaw.ac.at

Agnieszka Helman-Ważny
Universität Hamburg
Centre for the Study of Manuscript Cultures
Warburgstr. 26
20354 Hamburg, Germany

eMail: agnieszka.helman-wazny@uni-hamburg.de

Lukas Imstepf
University of Fribourg
Department of Informatics
PER 21 bu. B429
Boulevard de Pérolles 90
1700 Fribourg, Switzerlan

eMail: Lukas.imstepf@unifr.ch

Rolf Ingold
University of Fribourg
Department of Informatics
PER 21 bu. B429
Boulevard de Pérolles 90
1700 Fribourg, Switzerland

eMail: rolf.ingold@unifr.ch

Direk Injan
Chiang Mai Rajabhat University
Office of Arts and Culture
202 Chang Phueak Road, Tambon Chang Phueak
Mueang Chiang Mai District
Chiang Mai, 50300, Thailand

eMail: direkcmru@gmail.com

Damianos Kasotakis
Early Manuscripts Electronic Library (EMEL)
904 Silver Spur Road, #254
Rolling Hills Estates, California 90274, USA
<http://emel-library.org/>

eMail: damian07@gmail.com

Keith T. Knox
Early Manuscripts Electronic Library
<www.emel-library.org>
2739 Puu Hoolai Street
Kihei, Hawaii 96753, USA

eMail: knox@cis.rit.edu

Myriam Krutzsch
Egyptian Museum and Papyrus Collection SMB
Geschwister-Scholl-Str. 6
10117 Berlin, Germany

eMail: aemp@smb.spk-berlin.de

Volker Märgner
Technische Universität Braunschweig
Institut für Nachrichtentechnik
Schleinitzstraße 22
38106 Braunschweig, Germany

Universität Hamburg
Centre for the Study of Manuscript Cultures
Warburgstr. 26
20354 Hamburg, Germany

eMail: maergner@ifn.ing.tu-bs.de

David Messinger
Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology
54 Lomb Memorial Drive
Rochester, NY 14623-5604, USA

eMail: dwmpci@rit.edu

Hussein Mohammed
Universität Hamburg
Centre for the Study of Manuscript Cultures
Warburgstr. 26
20354 Hamburg, Germany

eMail: hussein.adnan.mohammed@uni-hamburg.de

Tyler R. Peery
Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology
54 Lomb Memorial Drive
Rochester, NY 14623-5604, USA

eMail: tp5281@rit.edu

Michael Phelps
Early Manuscripts Electronic Library (EMEL)
904 Silver Spur Road, #254
Rolling Hills Estates, CA 90274, USA
<http://emel-library.org/>

eMail: mphelps@emelibrary.org

Ira Rabin
Universität Hamburg
Centre for the Study of Manuscript Cultures
Warburgstr. 26
20354 Hamburg, Germany

eMail: ira.rabin@uni-hamburg.de

BAM Federal Institute for Material Research and Testing
Division 4.5
Unter den Eichen 44-46
12203 Berlin, Germany

eMail: ira.rabin@bam.de

Rolando Raqueno
RDSC – Rochester Data Science Consortium
260 East Main Street, Suite 6108
Rochester, NY 14604, USA

eMail: rraqueno@cs.rochester.edu rdsc@rochester.edu

Vinodh Rajan Sampath
Universität Hamburg
Department of Informatics
Vogt-Kölln-Str. 30
22527 Hamburg, Germany

eMail: vinodhrajan.sampath@uni-hamburg.de

Marçal Rusiñol
Univ. Autònoma de Barcelona
Computer Vision Center,
Dept. Ciències de la Computació, Edifici O
08193 Bellaterra (Barcelona), Spain

eMail: marcal@cvc.uab.es

Tilman Seidensticker
Friedrich-Schiller-Universität Jena
Institut für Orientalistik, Indogermanistik, Ur- und Frühgeschichtliche Archäologie
Seminar für Orientalistik
Zwätzengasse 4
07743 Jena, Germany

Universität Hamburg
Centre for the Study of Manuscript Cultures
Warburgstr. 26
20354 Hamburg, Germany

eMail: tilman.seidensticker@uni-jena.de

Mathias Seuret
University of Fribourg
Departement of Informatics
DIVA Research Group
Boulevard de Pérolles 90
1700 Fribourg, Switzerland

eMail: mathias@seuret.com

H. Siegfried Stiehl
Universität Hamburg
Department of Informatics
Vogt-Kölln-Str. 30
22527 Hamburg, Germany

eMail: stiehl@informatik.uni-hamburg.de

Peter A. Stokes
École Pratique des Hautes Études – Université PSL
Section des Sciences Historiques et Philologiques
4-14 rue Ferrus
75014 Paris, France

eMail: peter.stokes@ephe.psl.eu

Daniel Stökl Ben Ezra
École Pratique des Hautes Études (EPHE)
Section des Sciences Historiques et Philologiques
4-14 rue Ferrus
75014 Paris, France

eMail: Daniel.Stoekl@ephe.psl.eu

Dominique Stutzmann
Institut de Recherche et d'Histoire des Textes, CNRS
14, cours des Humanités
93322 Aubervilliers Paris, France

eMail: Dominique.stutzmann@irht.cnrs.fr

Christopher Tensmeyer
Brigham Young University
Dept. of Computer Science
Provo, UT 84602, USA

eMail: tensmeyer@byu.edu

Carsten Wintermann
Klassik Stiftung Weimar (KSW)
3.6 Abteilung Restaurierung, Direktion Museen
Burgplatz 4
99423 Weimar, Germany

eMail: carsten.wintermann@klassik-stiftung.de

Written Artefacts as Cultural Heritage

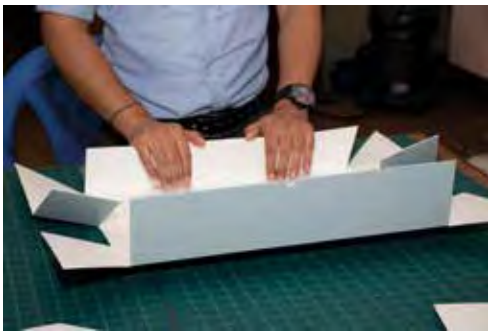
Ed. by Michael Friedrich and Doreen Schröter

Written Artefacts as Cultural Heritage was established in 2020. The series is dedicated to the double role of written artefacts as representations and generators of humankind's cultural heritage. Its thematic scope embraces aspects of preservation, the identity-defining role of artefacts as well as ethical questions.

The mix of practical guides, colloquium papers and project reports is specifically intended for staff at libraries and archives, curators at museums and art galleries, and scholars working in the fields of manuscript cultures and heritage studies.

Every volume of *Written Artefacts as Cultural Heritage* has been peer-reviewed and is openly accesible. There is an online and a printed version..

If you wish to receive a copy or to present your research, please contact the editorial office
https://www.csmc.uni-hamburg.de/publications/cultural-heritage.html



manuscript cultures (mc)

Editors: Michael Friedrich and Jörg B. Quenzer
Editorial office: Irina Wandrey

CSMC's academic journal was established as newsletter of the research unit 'Manuscript Cultures in Asia and Africa' in 2008 and transformed into a scholarly journal with the appearance of volume 4 in 2011. *manuscript cultures* publishes exhibition catalogues and articles contributing to the study of written artefacts. This field of study embraces disciplines such as art history, codicology, epigraphy, history, material analysis, palaeography and philology, informatics and multispectral imaging.

manuscript cultures encourages comparative approaches, without regional, linguistic, temporal or other limitations on the objects studied; it contributes to a larger historical and systematic survey of the role of written artefacts in ancient and modern cultures, and in so doing provides a new foundation for ongoing discussions in cultural studies.

Every volume of *manuscript cultures* has been peer-reviewed and is openly accessible:
https://www.csmc.uni-hamburg.de/publications/mc.html

If you wish to receive a copy or to present your research in our journal, please contact the editorial office:
irina.wandrey@uni-hamburg.de

manuscript cultures 14

manuscript cultures 13

manuscript cultures 11

manuscript cultures 10

manuscript cultures 9

manuscript cultures 8

manuscript cultures 7

manuscript cultures 6

manuscript cultures 5

manuscript cultures 4

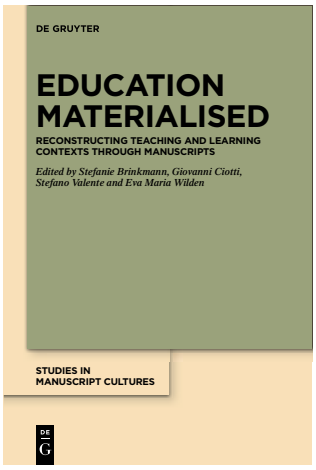
Studies in Manuscript Cultures (SMC)

Ed. by Michael Friedrich, Harunaga Isaacson, and Jörg B. Quenzer

From volume 4 onwards all volumes are available as open access books on the De Gruyter website:
https://www.degruyter.com/view/serial/43546
https://www.csmc.uni-hamburg.de/

DE GRUYTER

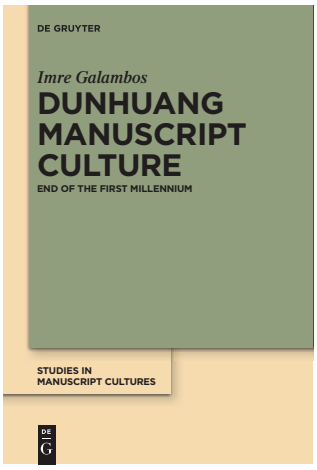
Forthcoming



23 – *Education Materialised: Reconstructing Teaching and Learning Contexts through Manuscripts,* edited by Stefanie Brinkmann, Giovanni Ciotti, Stefano Valente and Eva Maria Wilden

Manuscripts have played a crucial role in the educational practices of virtually all cultures that have a history of using them. As learning and teaching tools, manuscripts become primary witnesses for reconstructing and studying didactic and research activities and methodologies from elementary levels to the most advanced.
The present volume investigates the relation between manuscripts and educational practices focusing on four particular research topics: educational settings: teachers, students and their manuscripts; organising knowledge: syllabi; exegetical practices: annotations; modifying tradition: adaptations.
The volume offers a number of case studies stretching across geophysical boundaries from Western Europe to South-East Asia, with a time span ranging from the second millennium BCE to the twentieth century CE.

New release



22 – *Dunhuang Manuscript Culture: End of the First Millennium*, by Imre Galambos

Dunhuang Manuscript Culture explores the world of Chinese manuscripts from ninth–tenth century Dunhuang, an oasis city along the network of pre-modern routes known today collectively as the Silk Roads. The manuscripts have been discovered in 1900 in a sealed-off side-chamber of a Buddhist cave temple, where they had lain undisturbed for for almost nine hundred years. The discovery comprised tens of thousands of texts, written in over twenty different languages and scripts, including Chinese, Tibetan, Old Uighur, Khotanese, Sogdian and Sanskrit. This study centres around four groups of manuscripts from the mid-ninth to the late tenth centuries, a period when the region was an independent kingdom ruled by local families. The central argument is that the manuscripts attest to the unique cultural diversity of the region during this period, exhibiting – alongside obvious Chinese elements – the heavy influence of Central Asian cultures. As a result, it was much less 'Chinese' than commonly portrayed in modern scholarship. The book makes a contribution to the study of cultural and linguistic interaction along the Silk Roads.

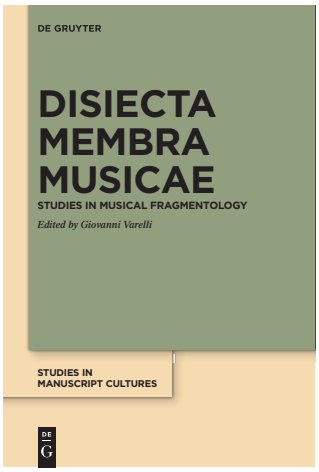
Studies in Manuscript Cultures (SMC)

Ed. by Michael Friedrich, Harunaga Isaacson, and Jörg B. Quenzer

From volume 4 onwards all volumes are available as open access books on the De Gruyter website:
https://www.degruyter.com/view/serial/43546
https://www.csmc.uni-hamburg.de/

DE GRUYTER

New release



21 – *Disiecta Membra Musicae: Studies in Musical Fragmentology*, edited by Giovanni Varelli

Although fragments from music manuscripts have occupied a place of considerable importance since the very early days of modern musicology, a collective, up-to-date, and comprehensive discussion of the various techniques and approaches for their study was lacking. On-line resources have also become increasingly crucial for the identification, study, and textual/musical reconstruction of fragmentary sources. *Disiecta Membra Musicae. Studies in Musical Fragmentology* aims at reviewing the state of the art in the study of medieval music fragments in Europe, the variety of methodologies for studying the repertory and its transmission, musical palaeography, codicology, liturgy, historical and cultural contexts, etc. This collection of essays provides an opportunity to reflect also on broader issues, such as the role of fragments in last century's musicology, how fragmentary material shaped our conception of the written transmission of early European music, and how new fragments are being discovered in the digital age. Known fragments and new technology, new discoveries and traditional methodology alternate in this collection of essays, whose topics range from plainchant to *ars nova* and fifteenth- to sixteenth-century polyphony.

New release

20 – *Fakes and Forgeries of Written Artefacts from Ancient Mesopotamia to Modern China,* edited by Cécile Michel and Michael Friedrich

Fakes and forgeries are objects of fascination. This volume contains a serie of thirteen articles devoted to fakes and forgeries of written artefacts from th beginnings of writing in Mesopotamia to modern China. The studies empha sise the subtle distinctions conveyed by an established vocabulary relating t the reproduction of ancient artefacts and production of artefacts claiming t be ancient: from copies, replicas and imitations to fakes and forgeries. Fake are often a response to a demand from the public or scholarly milieu, or ever both. The motives behind their production may be economic, political, reli gious or personal – aspiring to fame or simply playing a joke. Fakes may be revealed by combining the study of their contents, codicological, epigraphi and palaeographic analyses, and scientific investigations. However, certain fa mous unsolved cases still continue to defy technology today, no matter hov advanced it is. Nowadays, one can find fakes in museums and private collec tions alike; they abound on the antique market, mixed with real artefacts tha have often been looted. The scientific community's attitude to such objects calls for ethical reflection

ISSN 1867–9617

Centre for the Study of Manuscript Cultures
Universität Hamburg
Warburgstraße 26
D-20354 Hamburg

www.csmc.uni-hamburg.de

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

CENTRE FOR THE STUDY OF MANUSCRIPT CULTURES

DFG