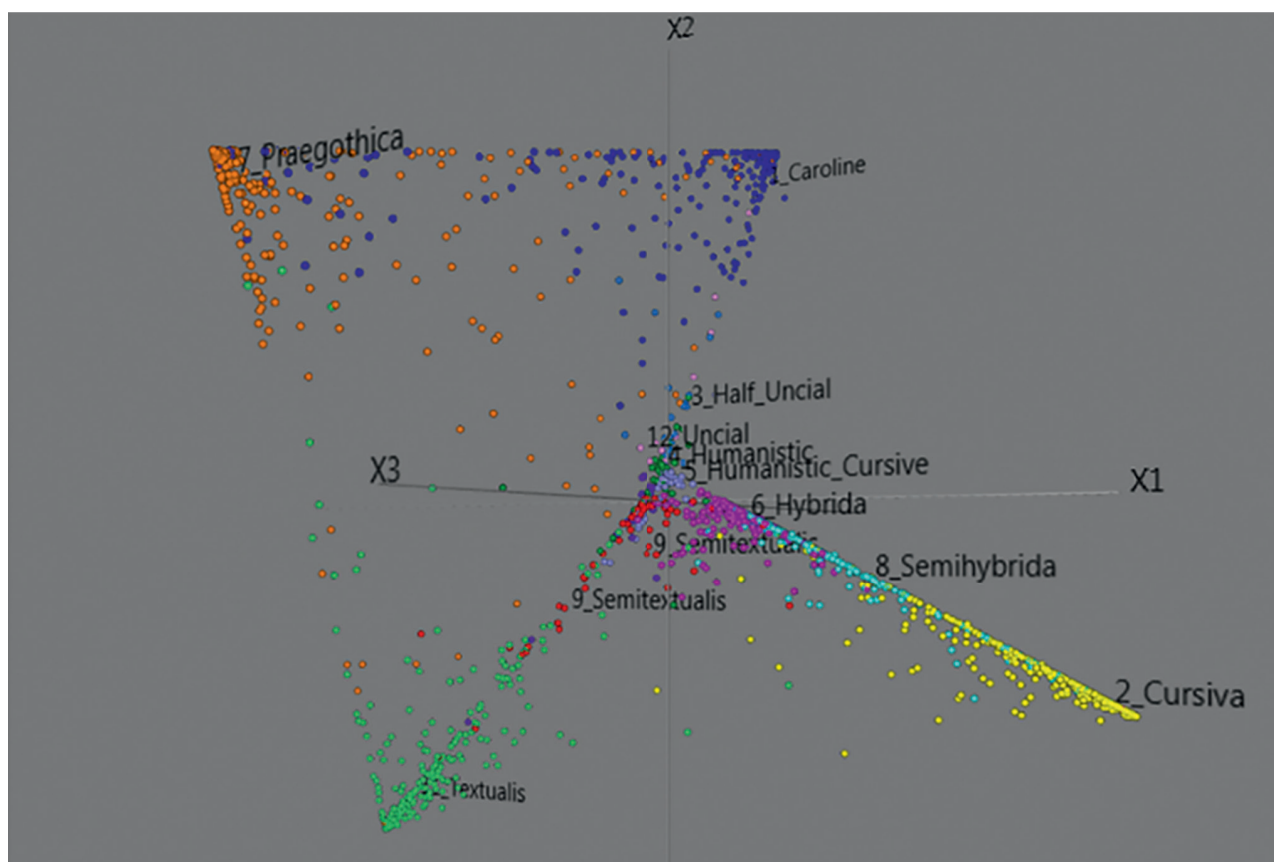


manuscript cultures

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

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

www.csmc.uni-hamburg.de

ISSN 1867–9617

© 2020

Centre for the Study of Manuscript Cultures
Universität Hamburg
Warburgstraße 26
D-20354 Hamburg

CONTENTS

3 | Editorial

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

ARTICLES

5 | On Avoiding Segmentation in Handwritten Keyword Spotting: Overview and Perspectives

Marçal Rusiñol*

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

Dominique Stutzmann, Christopher Tensmeyer, and Vincent Christlein*

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

Daniel Stökl Ben Ezra and Hayim Lapin*

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

Peter A. Stokes*

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

Marcel Gygli (Würsch), Mathias Seuret, Lukas Imstepf, Andreas Fischer, Rolf Ingold*

53 | Building an Evaluation Framework Researchers Will (Want to) Use

Joseph Chazalon*

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

Vinodh Rajan Sampath and H. Siegfried Stiehl*

73 | Legally Open: Copyright, Licensing, and Data Privacy Issues

Vanessa Hanneschläger*

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

Hussein Mohammed, Volker Märgner, and Tilman Seidensticker

87 | Illuminating Techniques from the Sinai Desert

Damianos Kasotakis, Michael Phelps, and Ken Boydston

91 | Image Quality in Cultural Heritage

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

105 | When Erased Iron Gall Characters Misbehave

Keith T. Knox

115 | 'Dürer's Young Hare' in Weimar – A Pilot Study

Oliver Hahn, Uwe Golle, Carsten Wintermann, and Ira Rabin

123 | Material-Technical Details on Papyrus as Writing Support

Myriam Krutzsch

133 | The Techniques and Materials Used in Making Lao and Tai Paper Manuscripts

Agnieszka Helman-Ważny, Volker Grabowsky, Direk Injan and Khamvone Boulyaphonh

163 | Inks Used to Write the Divine Name in a Thirteenth-Century Ashkenazic Torah Scroll: Erfurt 7 (Staatsbibliothek zu Berlin, Or. fol. 1216)

Nehemia Gordon, Olivier Bonnerot, and Ira Rabin

185 | Contributors

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 (*mubayyada*) 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 *samā*, plural *samā’āt*) 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’ names¹, 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: Rabī' I 486 (= April 1093), Rabī' I 487 (= March/April 1094), 6 Ramaḍān 487 (= 19 September 1094) and 27 Shawwāl 490 (= 7 October 1097). Certificate #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 certificates 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 certificates.

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 unchronological 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 certificates, viz. al-Ḥājjib Abī l-Ḥasan 'Alī Ibn Muḥammad Ibn 'Alī al-'Allāf. Figure 2 shows what the names look like in #5, 7 and 12.

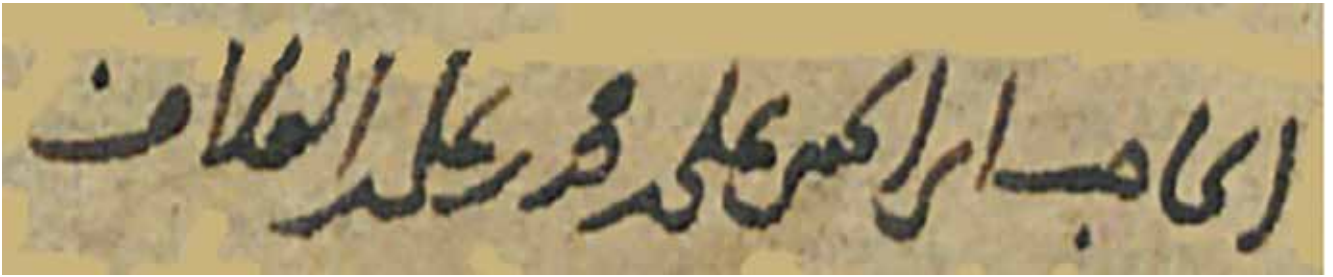
In Seidensticker 2015, the writing styles⁵ of #7 and #12 were considered to convey an impression of far-reaching agreement in a hand noticeably inclined to the right, and concerning #5, it was stated that 'it seems to be written by the same hand' (p. 82). This analysis appeared to be adequate at the 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.

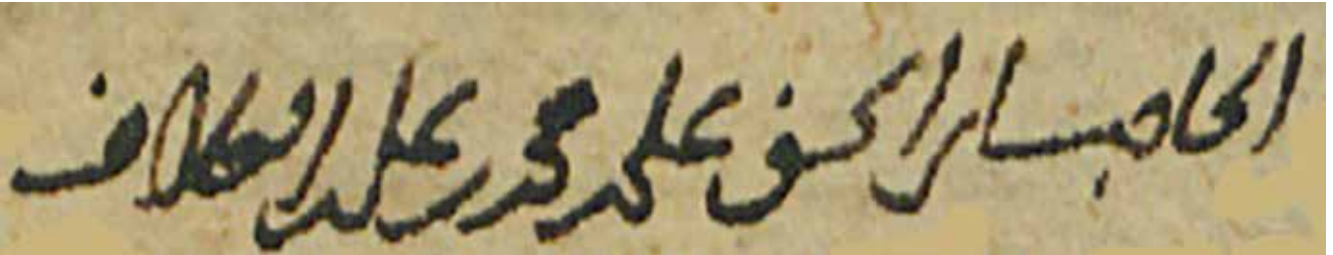


Fig. 1: Fol. 14° of Gotha, Ms. orient. A 627, showing certificates #1 (second part) to #5 from top to bottom.

#5



#7



#12

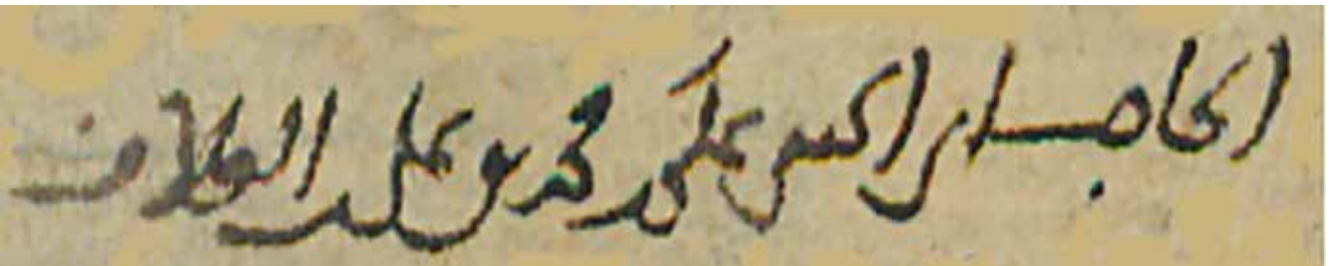


Fig. 2: Enlarged details showing the attending authority's name.

- 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 *fā'* (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.

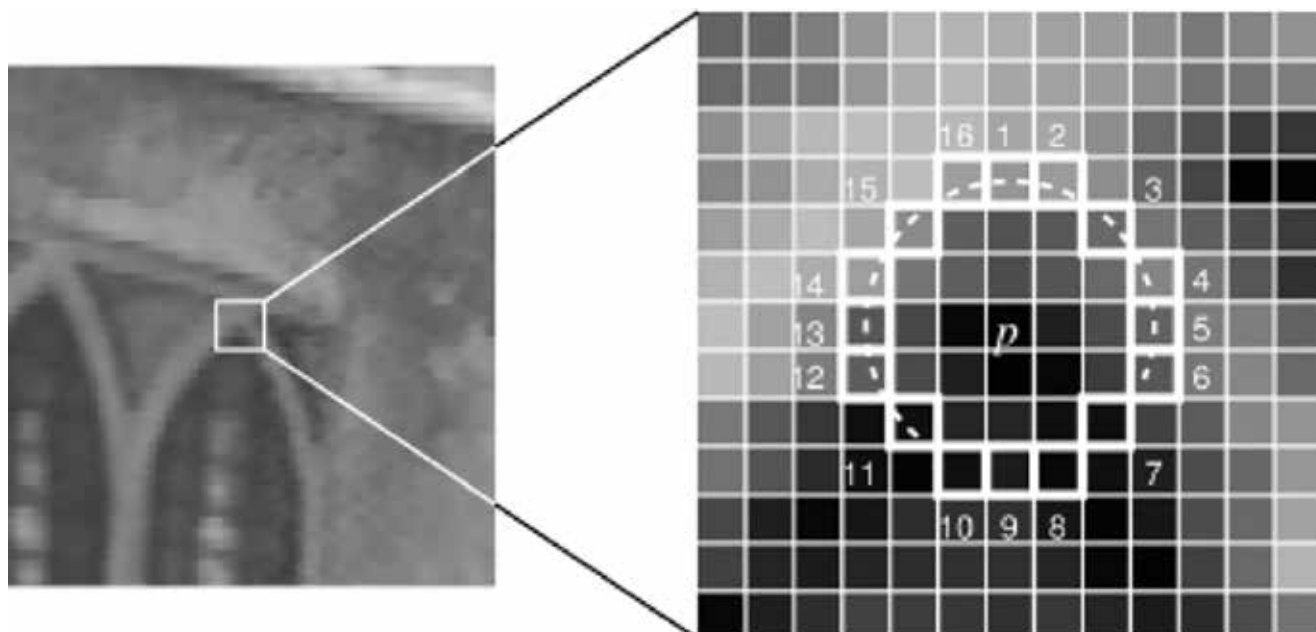


Fig. 3: A small part of an image is shown on the left. On the right, an enlarged section of that is shown with a pixel in the middle (p) and 16 pixels surrounding it, which are used to calculate the strength of the potential keypoint (reproduced from Rosten et al. 2010).

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.

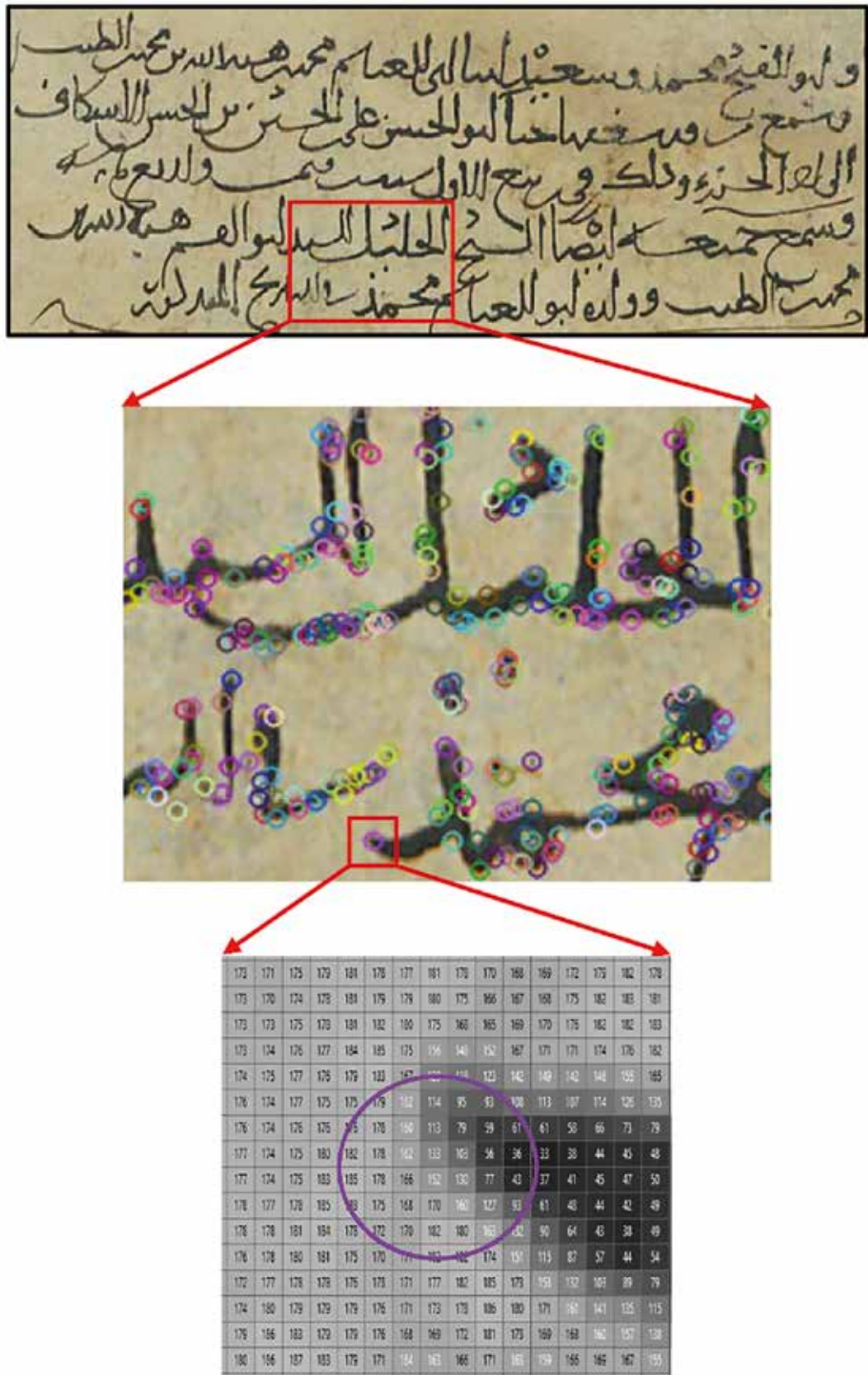


Fig. 4: Visualisation of the keypoints on a sample image of a certificate. At the top a certificate, in the middle keypoints on the writing shown as circles of different colours, and at the bottom the 16×16 pixel surroundings of a keypoint after colour transformation to grey with the grey values shown.

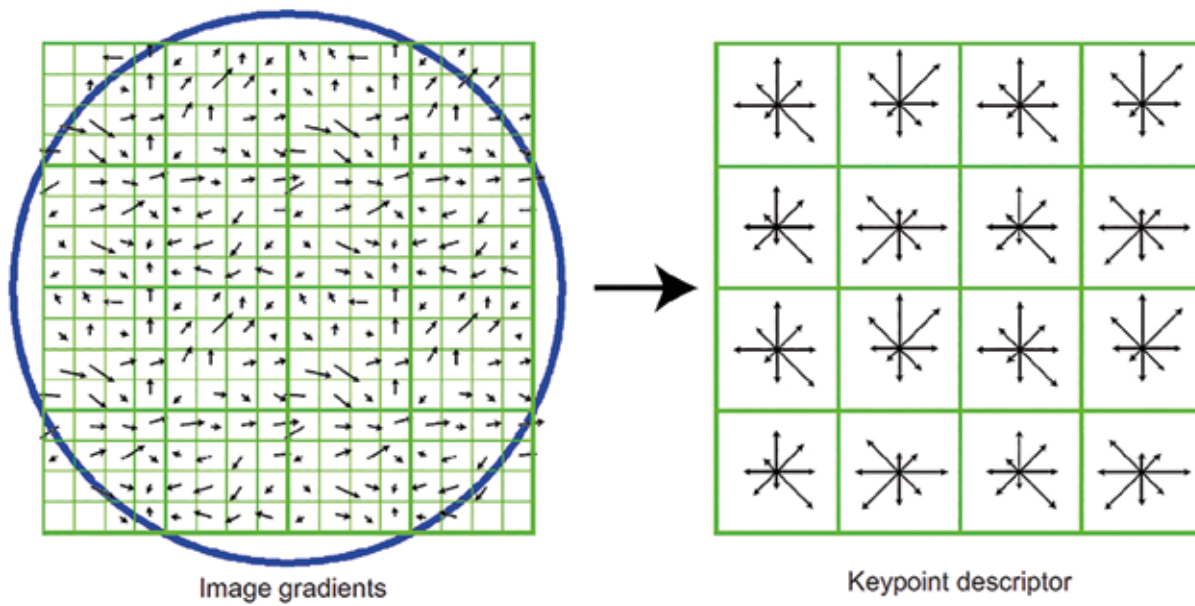


Fig. 5: Graphical representation of the computation of a descriptor at a keypoint position (reproduced from Lowe 2004 with some modifications).

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

⁷See Mohammed et al. 2017 for details.

Table 3: Scores calculated for each certificate (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

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

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

‘probably’ written by the same person. The scores confirm 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-Ḥarā’iṭī’s *I’tilāl al-qulūb* (Forschungsbibliothek Gotha, Ms. orient. A 627)’, *Journal of Islamic Manuscripts*, 11: 53–100.

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 accessible. 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>




Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG



CENTRE FOR THE
STUDY OF
MANUSCRIPT
CULTURES

WRITTEN ARTEFACTS
AS CULTURAL HERITAGE

CENTRE FOR THE STUDY
OF MANUSCRIPT CULTURES
UNIVERSITÄT HAMBURG



**STEP BY STEP GUIDE TO MANUSCRIPT SURFACE
CLEANING AND MAKING E-FLUTE PHASE BOXES
FOR MANUSCRIPTS**

BIDUR BHATTARAI | MICHAELLE BIDDLE

NO 1



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



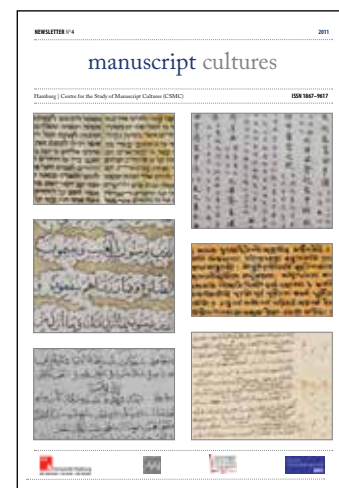
manuscript cultures 6



manuscript cultures 5



manuscript cultures 4



Studies in Manuscript Cultures (SMC)

Ed. by Michael Friedrich, Harunaga Isaacson, and Jörg B. Quenzer

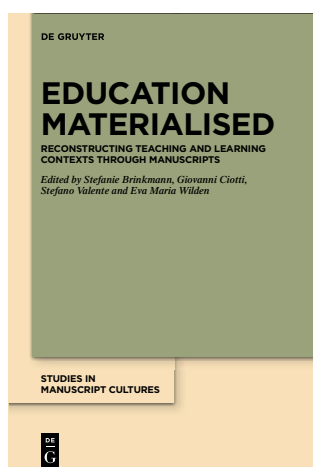
From volume 4 onwards all volumes are available as open access books on the De Gruyter website:

<https://www.degruyter.com/view/serial/43546>

<https://www.csmc.uni-hamburg.de/>



Forthcoming



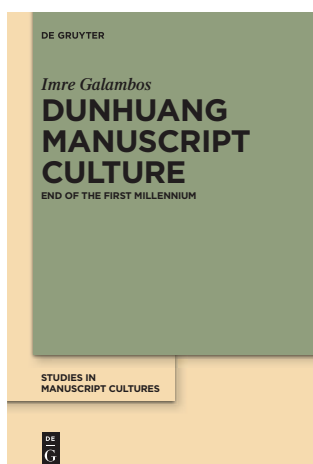
23 – *Education Materialised: Reconstructing Teaching and Learning Contexts through Manuscripts*, edited by Stefanie Brinkmann, Giovanni Ciotti, Stefano Valente and Eva Maria Wilden

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 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

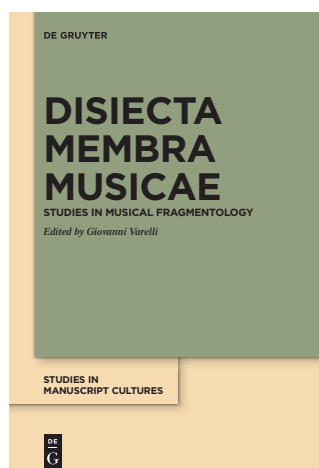
From volume 4 onwards all volumes are available as open access books on the De Gruyter website:

<https://www.degruyter.com/view/serial/43546>

<https://www.csmc.uni-hamburg.de/>



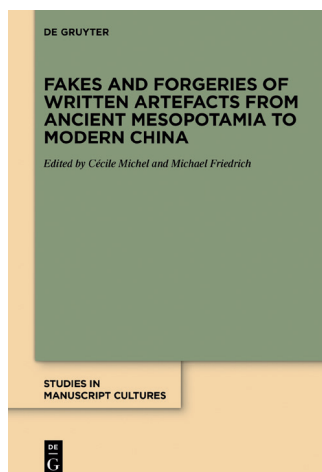
New release



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 series of thirteen articles devoted to fakes and forgeries of written artefacts from the beginnings of writing in Mesopotamia to modern China. The studies emphasise the subtle distinctions conveyed by an established vocabulary relating to the reproduction of ancient artefacts and production of artefacts claiming to be ancient: from copies, replicas and imitations to fakes and forgeries. Fakes are often a response to a demand from the public or scholarly milieu, or even both. The motives behind their production may be economic, political, religious or personal – aspiring to fame or simply playing a joke. Fakes may be revealed by combining the study of their contents, codicological, epigraphic and palaeographic analyses, and scientific investigations. However, certain famous unsolved cases still continue to defy technology today, no matter how advanced it is. Nowadays, one can find fakes in museums and private collections alike; they abound on the antique market, mixed with real artefacts that have often been looted. The scientific community's attitude to such objects calls for ethical reflection.

ISSN 1867–9617

© 2020

Centre for the Study of Manuscript Cultures

Universität Hamburg

Warburgstraße 26

D-20354 Hamburg

www.csmc.uni-hamburg.de