

Diptychon

A transcription assistant system for the separation of glyphs in medieval manuscript texts

Björn Gottfried¹, Marius Wegner¹,
Jan-Hendrik Worch¹, Mathias Lawo² *

¹ Centre for Computing Technologies, University of Bremen, Germany

² Akademienvorhaben Monumenta Germaniae Historica,
Berlin-Brandenburgische Akademie der Wissenschaften, Germany

1 Introduction

The analysis and comparison of writing hands is one of the palaeographic applications in the field of document image processing [4, 9, 11]. Due to the many defects found in historical documents, methods are required to enhance those documents [15, 12, 6, 7, 5, 10, 1, 3]. Despite of the large body of work investigated in the enhancement of document images, document enhancement is not always successful and ancient handwritings might be arbitrarily complex. For this reason, automatic methods are frequently approaching their limits. Therefore, the notion *computer assisted transcription* has been coined for methods that enable the interaction between user and machine in order to be able to deal with arbitrary complex handwritings [14, 13]. Whenever algorithms fail, the user gets the opportunity to assist. This is also the philosophy behind the Diptychon system.

Diptychon has been developed in order to support the transcription and comparison of medieval handwritings [8]. The design of Diptychon grants to the user full control of the entire analysis process. In contrast to the conventional palaeographic approach where analysis is restricted to the subjective judgment and decisions of an editor, here a standardised method is implemented: Diptychon leads the editor through the analysis process by automatically suggesting how to extract lines of text and how to separate glyphs. As, however, it is hardly possible to make perfect suggestions automatically when dealing with non-restrictive handwritings, the user can freely adjust the lines of text and glyph separations.

* We gratefully acknowledge the support of the German Research Foundation DFG under grant number GO 2023/4-1 (LA 3066), LA 3007/1-1.

Up to now, Diptychon has been applied to two different writing hands of a wider range: one belongs to the 11th-century chronicler Hugh of Flavigny, the other scribe has been working in the chancery of emperor Charles IV in the third quarter of the 14th century. That is, in terms of Latin palaeography, we deal with a late Carolingian Minuscule used in books ("Buchschrift") by an individual author as well as with a later Gothic charter script ("Urkundenkursive") standardised through long-term chancery usage.

This paper summarises the methodology realised by Diptychon to extract all the glyphs of a document image. While current methodologies are confined to roughly determine single glyphs, for example by determining their bounding box, the presented methods enable the precise extraction of arbitrary complex glyphs. Once all glyphs are available, they can be used for comparison purposes.

2 Processing documents with Diptychon

Diptychon is employed as follows. The image of a document page has to be loaded and enhanced. After that, the system tries to detect all text lines. While the document image is displayed on the left panel, the equally sized right panel provides empty text boxes for each text line found in the document image. The user has the opportunity to correct text lines, to delete them as well as to add new lines. Fig. 1 shows Diptychon with the detected text lines for a document image.

After text lines being detected, the user transcribes the text, line by line. For handwritten texts it is hardly possible to let the system always successfully detect the correspondence between a glyph and the character it is representing. For this reason, the system just proposes more or less accurate correspondences which the user needs to adjust. While these interactions cost some effort, the long-term objective of the system is to learn from the user how she transcribes the manuscript so that the performance of the system grows during the transcription process of a document. Ultimately, the quality of the suggestions which the system prompts to the user will improve.

The core of the system consists in interactive methods which are to correct suggestions made by Diptychon at the level of single glyphs. They are described in the following.

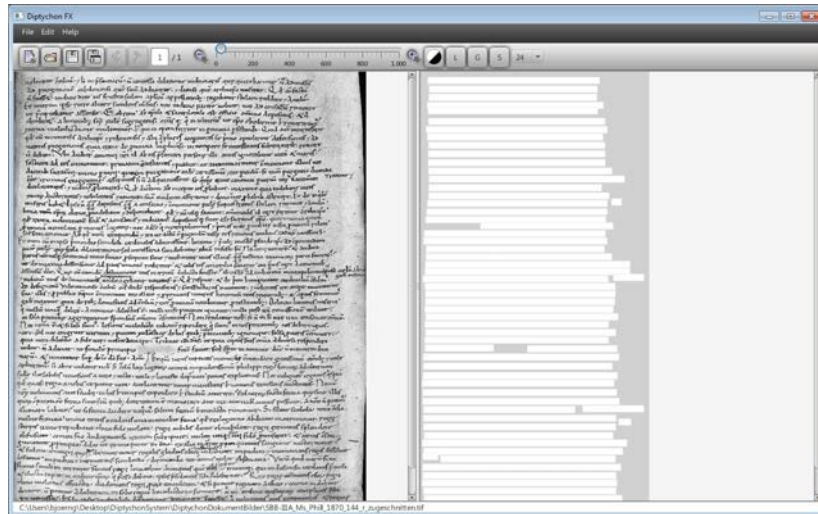


Fig. 1. The Diptychon system.

2.1 Connecting fragments

Provided that two fragments are to be connected to form a single piece (the blue and orange fragments below). For this purpose, one of the two fragments is to be chosen by pressing and holding the left mouse button (the blue one). As a consequence, the colour of that fragment changes to bright green. Keeping the left mouse button pressed, the mouse cursor can be moved over other fragments (the orange one) which will all be added to the former fragment. This is visualised by inking all fragments bright green, which have been touched with the mouse. As soon as all fragments which are to be connected are coloured bright green, the left mouse button can be released.

In the end, every fragment which has been touched by the mouse will be connected to all other touched fragments. To visualise their conjunction as a single fragment, they get the same colour. This also requires to change the colours of all successive fragments, so that again all neighbouring fragments are differently coloured, again in the order orange, red, green, and blue. This example simultaneously shows how even disconnected fragments can be joined to a single fragment, since the letter 'h' in this example consists of two disconnected components.

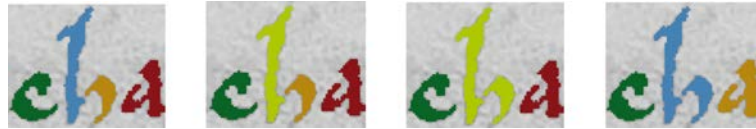


Fig. 2. Joining fragments of a glyph.

2.2 Cutting fragments into pieces

The leftmost image below shows a mistaken fragmentation for the glyph images of 'eu'. Selecting the orange fragment by the left mouse button, which is to be released again immediately after it has been pressed, the fragmentation in the second image is obtained, that correctly separates the 'e' from the 'u'. Though, both the 'e' and the 'u' are now themselves fragmented into two pieces. But as the following images show, these fragments can be easily connected as described in the previous section. The last image eventually shows the correct separation between both characters.



Fig. 3. Separating and joining fragments of a glyph.

2.3 Fragmenting pieces by a rectangular tessellation

In all those cases where the previous methods fail, yet another fragmentation method can be employed that tessellates a region into small squares. The length of the side of a single square can be adjusted by a parameter.

As fragments do hardly form themselves into rectangular areas or even squares, the border areas of fragments will in general be filled with different shapes, which are the remaining parts of those tiles that are too large to fit completely into the border areas. Fig. 4 shows an example and Fig. 5 a blow-up of the critical section.

2.4 Fragmenting pieces by line segmentation

There are a number of cases which are particularly challenging, for example, when glyphs are very close to each other without clear transition,



Fig. 4. Different fragmentation steps: (a) a single fragment consisting of letters ‘e’ and ‘o’, (b) having clicked with the left mouse button into the green area, (c) having clicked into the blue area, (d) having clicked into the orange area; (e) the orange part has been tessellated into small squares by choosing the according function from the context menu; (f) all fragments belonging to the letter ‘o’ have been selected; (g) the remaining fragments have been connected to form the letter ‘e’.

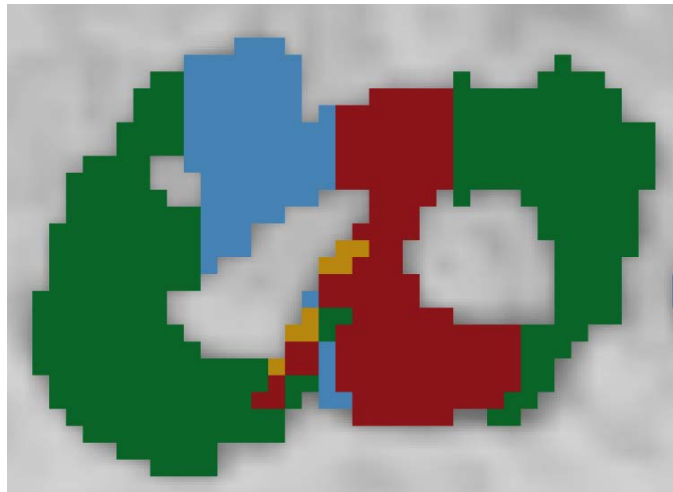


Fig. 5. The enlargement of image (e) from Fig. 4.

when ligatures are to be separated into their constituent parts, or when there are imperfections, such as blobs or holes in the paper, smearing the boundaries of the glyphs. Yet another common problem are nearby text lines. In this case, descenders are overlapping with ascenders of the next text line. The previously described methods are generally not very successful in such cases.

The line segmentation method allows the user to draw an arbitrary line, along which a fragment is supposed to be separated into two pieces, as shown in Fig. 6. In this way, even oblique transitions between glyphs can be dealt with. The line might even cross a couple of adjacent regions all at the same time. As a consequence, they are simultaneously divided into twice as many pieces as there have been regions before.

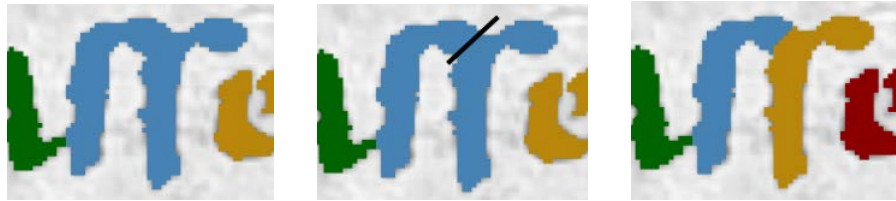


Fig. 6. The line segmentation method.

2.5 Searching for glyphs in the document image

In the end of the transcription process the system represents each single glyph and its corresponding character. Based on this mapping between glyphs and characters, the user can search for single characters or even arbitrary strings in the original document, as shown for a printed document with a Gothic typeface in Fig. 7. Found matches are highlighted by the system, and the user can iterate through all found occurrences of a specific character. This helps the user to get an overview of a document.



Fig. 7. Searching for glyph images. (Excerpt from: 'Die Grenzboten, 28. Jahrgang, 2. Semester 1. Band, Leipzig 1869, Staats- und Universitätsbibliothek Bremen').

Then, it is possible to iterate step by step forward and backward from one instance to the next one. In this way, one gets an idea of the different appearances of a particular character. Additionally, it is possible to distinguish majuscules and minuscules, and Diptychon shows how many instances are found.

Instead of single characters it is also possible to search for strings, to compute how often a particular string appears in the text, and how those occurrences look like. The left hand side of Fig. 8 shows the search for the word ‘solis’.

As it might be of interest to analyse how a particular character looks like at the beginning or the end of a word, it is possible to enter a space character before or after that particular character. As a consequence, only those occurrences will be found which show that character at the beginning and end of a word respectively.



Fig. 8. Left: Searching for the string ‘solis’. Right: Searching for pairs of words ending and starting with specific characters, in this example ‘-e a-’.

Similarly, space characters can also be included into strings. For instance, one might be interested in learning how a scribe wrote an ‘a’ at the beginning of a word when the previous word ended by an ‘e’. The right hand side of Fig. 8 shows this example.

3 Summary and Outlook

An overview of the Diptychon system has been given. Diptychon has been developed in order to extract single glyphs of an arbitrary writing hand from a document image. Up to now, the Diptychon methodology has been applied to Latin characters of the 11th and of the 14th century.

Although the quality of the documents is low in the sense of many defects, such as pages which are bleached out, the glyphs can be extracted

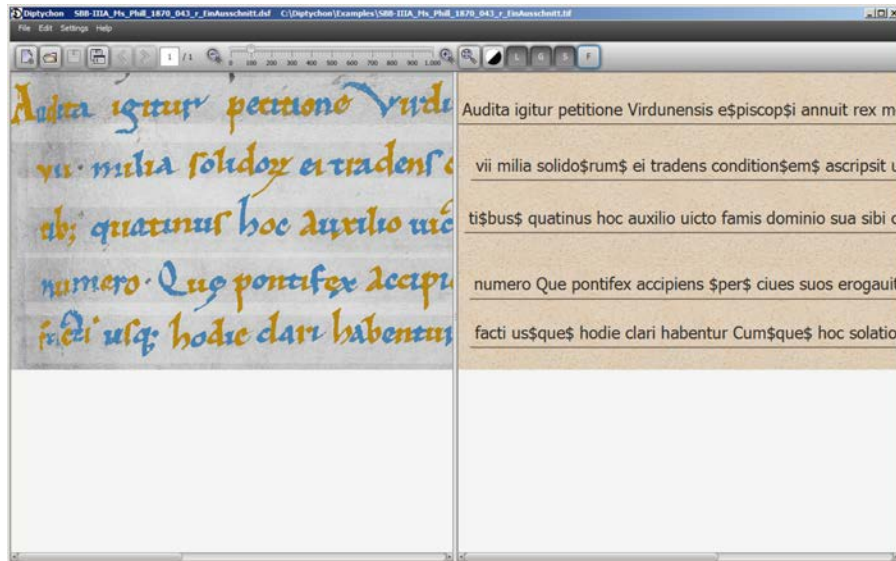


Fig. 9. The Diptychon system with a transcribed document image. The original document image is shown on the left hand side, while the transcription is found on the right hand side. Each text line of the latter is approximately at the same vertical level as the text line in the original document.

quite well with the system due to available interaction methods. These enable the user to determine precisely the border between glyphs as well as between the glyphs and the background. Once a document has been transcribed with Diptychon, search methods enable the user to get an overview over the given handwriting. Fig. 9 shows part of a transcribed document image.

Future extensions are underway. They include the improvement of automatic transcription suggestions based on methods such as those described in [16], the export of transcriptions as TEI-files [2], and the export of further document details, like statistics about the document, in particular regarding the occurrences of glyphs and their visual characteristics.

References

1. H. K. Aghajan, B. H. Khalaj, and T. Kailath. Estimation of skew angle in text-image analysis byslide: Subspace-based line detection. *Mach. Vis. Appl.*, 7(4):267–276, 1994.
2. J.R. Ahronheim. Descriptive metadata: Emerging standards. *Journal of Academic Librarianship*, 24(5):395, 1998.

3. P. A. Chochia. Image enhancement using sliding histograms. *Computer Vision, Graphics, and Image Processing*, 44(2):211–229, 1988.
4. C. De Stefano, F. Fontanella, M. Maniaci, and A. S. di Freca. A method for scribe distinction in medieval manuscripts using page layout features. In G. Maino and G. Foresti, editors, *ICIAP*, volume 6978 of *LNCS*, pages 393–402. Springer, 2011.
5. K.-C. Fan, Y.-K. Wang, and T.-R. Lay. Marginal noise removal of document images. *Pattern Recognition*, 35(11):2593–2611, 2002.
6. U. Garain, T. Paquet, and L. Heutte. On foreground–background separation in low quality document images. *IJDAR*, 8(1):47–63, 2006.
7. B. Gatos, I. Pratikakis, and S. J. Perantonis. An adaptive binarization technique for low quality historical documents. In Simone Marinai and AndreasR. Dengel, editors, *DAS VI*, volume 3163 of *LNCS*, pages 102–113. Springer, 2004.
8. B. Gottfried, M. Wegner, and M. Lawo. Towards the interactive transcription of handwritings: anytime anywhere document analysis. *International Journal on Document Analysis and Recognition (IJDAR)*, doi 10.1007/s10032-014-0234-7, pages 1–15, 2014.
9. W. Hofmeister and A. Hofmeister-Winter. Schriftzüge unter der High-Tech-Lupe. Theoretische Grundlagen und erste praktische Ergebnisse des Grazer Pilotprojekts DAMalS. In *Internationales Jahrbuch für Editionswissenschaft*, volume 22, pages 90–117, 2008.
10. H. Kansal, S. Sanyal, and D. Gupta. Dewarping and deskewing of a document using affine transformation. In A. Ranchordas and H. Araújo, editors, *VISAPP (2)*, pages 73–78. INSTICC Press, 2009.
11. G. Kim, V. Govindaraju, and S. N. Srihari. An architecture for handwritten text recognition systems. *IJDAR*, 2(1):37–44, 1999.
12. T. Obafemi-Ajayi, G. Agam, and O. Frieder. Historical document enhancement using LUT classification. *IJDAR*, 13(1):1–17, 2010.
13. V. Romero, A. H. Toselli, L. Rodríguez, and E. Vidal. Computer assisted transcription for ancient text images. In M. S. Kamel and A. C. Campilho, editors, *ICIAR*, volume 4633 of *LNCS*, pages 1182–1193. Springer, 2007.
14. N. Serrano, A. Gimnez, J. Civera, A. Sanchis, and A. Juan. Interactive handwriting recognition with limited user effort. *IJDAR*, pages 1–13, 2013.
15. E. H. Barney Smith, J. Darbon, and L. Likforman-Sulem. A mask-based enhancement method for historical documents. In G. Agam and C. Viard-Gaudin, editors, *DRR*, volume 7874 of *SPIE Proceedings*, pages 1–10. SPIE, 2011.
16. J.-H. Worch, M. Lawo, and B. Gottfried. Glyph spotting for mediaeval handwritings by template matching. In C. Concolato and P. Schmitz, editors, *ACM Symposium on Document Engineering*, pages 213–216. ACM, 2012.