Schriften des Instituts für Dokumentologie und Editorik — Band 11

Kodikologie und Paläographie im digitalen Zeitalter 4

Codicology and Palaeography in the Digital Age 4

herausgegeben von | edited by

Hannah Busch, Franz Fischer, Patrick Sahle

unter Mitarbeit von | in collaboration with

Bernhard Assmann, Philipp Hegel, Celia Krause

2017

BoD, Norderstedt

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

Digitale Parallelfassung der gedruckten Publikation zur Archivierung im Kölner Universitäts-Publikations-Server (KUPS). Stand 4. September 2017.

SPONSORED BY THE

Federal Ministry
of Education
and Research

Diese Publikation wurde im Rahmen des Projektes eCodicology (Förderkennzeichen 01UG1350A-C) mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) gefördert.
Publication realised within the project eCodicology (funding code 01UG1350A-C) with financial resources of the German Federal Ministry of Research and Education (BMBF).

2017

Herstellung und Verlag: Books on Demand GmbH, Norderstedt
ISBN: 978-3-7448-3877-1
Einbandgestaltung: Julia Sorouri, basierend auf Vorarbeiten von Johanna Puhl und Katharina Weber; Coverbild nach einer Vorlage von Swati Chandna.
Satz: LuaTEX und Bernhard Assmann

Automatic Dating of Historical Documents

Vincent Christlein, Martin Gropp, Andreas Maier

Abstract

With the growing number of digitized documents available to researchers it is becoming possible to answer scientific questions by simply analyzing the image content. In this article, a new approach for the automatic dating of historical documents is proposed. It is based on an approach only recently proposed for scribe identification. It uses local RootSIFT descriptors which are encoded using VLAD. The method is evaluated using a dataset consisting of context areas of medieval papal charters covering around 150 years from 1049 to 1198 AD. Experimental results show very promising mean absolute errors of about 17 years.

Zusammenfassung

Mit der steigenden Zahl der für Forscher zugänglichen digitalisierten Dokumente wird es möglich, wissenschaftliche Fragestellungen durch die einfache Analyse der Bilddaten zu beantworten. In diesem Beitrag wird ein neues Vorgehen für die automatische Datierung historischer Dokumente vorgestellt. Es basiert auf einem Ansatz, der erst vor kurzem für die Schreiberidentifikation entwickelt wurde und nutzt lokale RootSIFT-Deskriptoren, die mit VLAD codiert werden. Die Methode wird mit einem Datensatz evaluiert, der aus den Textbereichen mittelalterlicher Papsturkunden aus rund 150 Jahren (1049-1198) besteht. Experimentelle Ergebnisse zeigen eine sehr vielversprechende mittlere Fehlerrate von rund 17 Jahren.

1 Introduction

Dating historical documents can be a time-consuming and expensive process which typically requires the consultation of experts of history and/or paleography. While the chemical analysis of the paper through radiocarbon dating often yields reasonable accuracy, at least for the time of production of the writing medium, non-invasive methods are often preferable for a variety of reasons.

These approaches can be divided into two groups: content-based methods and image-based methods. Content-based methods relate to procedures which derive the date of production from information in the text. Either directly, e. g., an event directly referred to in the text can be related to a known date. Or indirectly, through

Kodikologie und Paläographie im Digitalen Zeitalter 4 – Codicology and Palaeography in the Digital Age 4. Hrsg. Hannah Busch, Franz Fischer und Patrick Sahle, unter Mitarbeit von Bernhard Assmann, Philipp Hegel und Celia Krause. Schriften des Instituts für Dokumentologie und Editorik 11. Norderstedt: Books on Demand, 2017. 151–164.

a linguistic analysis of the text, see for example the work of Feuerverger et al. (2008), who dated manuscripts from the 11^{th} to the 15^{th} century. This is possible if enough dated reference material exists.

This is also a prerequisite for image-based methods. In contrast to content-based methods, however, the text does not need to be transcribed first. For several manuscripts, a rough date can be estimated (manually or automatically) based on the layout of the document or the symbols/images it contains. In papal charters, for example, there typically is a *rota* symbol containing the name of the pontificate. Moreover, the handwriting can give a clue to the date since different handwriting styles were used in different periods of time . By extracting these information, a semi or fully automatic program can assist a paleographer in dating handwritten documents. It is also to be noted that large-scale dating, i. e., the dating of hundreds of manuscripts or more, might be too time-consuming for an individual. Here, automated methods suggesting a probable date might be useful for initial estimates or may also point out interesting documents to the researcher. For example, outliers in a large corpus of documents might just relate to an interesting handwriting — or the style could actually point towards a later date than the content, indicating a potential document forgery.

Wahlberg et al. (2016) showed that content- and image-based methods can also be combined for an improved automatic dating.

Automatic dating may also help to improve OCR quality as specialized classifiers can be trained for specific date ranges when they are known. Li et al. (2015) have shown great improvements in OCR when estimating the date of printed manuscripts in advance.

Algorithmically, the dating of handwritten text is closely related to the problem of (automatic) writer identification.[1] But while there are fixed classes of writers in the case of writer identification, image-based dating is typically seen as having a regression problem, i. e., we determine continuous targets (the dates) instead of fixed classes (the writers).

In this paper, we propose a new method for automatic dating. The individual parts of the approach have already been used successfully for writer identification (Christlein et al. 2014; Christlein et al. 2015). These publications draw on clean benchmark datasets, while this work relies on experiments with historical documents. Historical documents are typically digitized in high definition. Thus, we evaluate different strategies to lower the computational burden. Moreover, historical documents often contain large deficiencies such as holes or stains. We evaluate different strategies for feature sampling and study their effects on dating accuracy. An example image can be seen in figure 1.

---

1 Note: “writer” and “scribe” are used interchangeably throughout this paper.

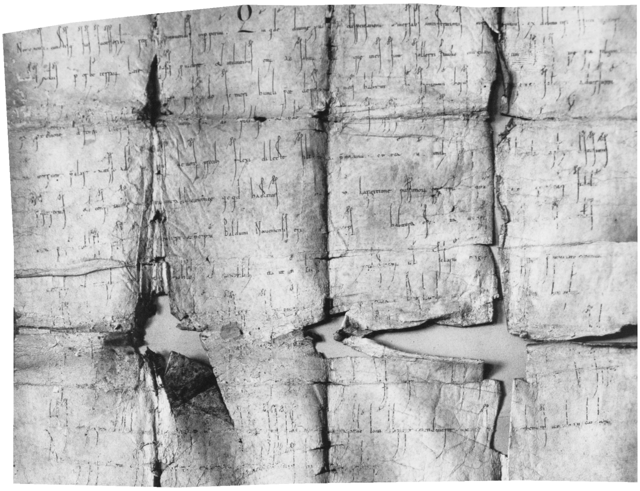

Figure 1: Image excerpt of a papal charter. Jaffé / Loewenfeld no. 4671; pontificate: Alexander II; date: January 28, 1070; image courtesy of the Göttingen Academy of Sciences and Humanities.

This paper is organized as follows: after the related work is presented in section 2, our proposed method is explained in section 3. Section 4 covers the evaluation of our experiments and results. The paper is concluded in section 5.

2 Related work

Dating of historical manuscripts

Automatic image-based dating of historical manuscripts is a relatively new discipline with virtually no visible research until only a few years ago, which was probably owed to the lack of sufficiently large digitized collections of suitable documents. In 2014, He et al. presented a new dataset and used Hinge and Fraglets features in a nested SVR approach to predict the year of a document's creation. In the following year, Wahlberg et al. (2015) proposed a method focused in particular on low-quality images, based on shape context and Stroke Width Transformation. Recently, He et al. (2016a) added a new unsupervised attribute learning step and finally treated document dating as a classification problem, an approach they continue in their later work (He et al. 2016b) with local contour fragments and stroke fragments features. While Wahlberg et al. advance to place special emphasis on incorporating language information in their 2016 paper, requiring manual transcriptions that are not easily available in many cases, they also continue to improve their solely image-based method.

Handwriting classification

The problem of dating handwritten text is methodically similar to text style recognition or writer identification. Writer identification can be categorized into two groups: *textural* methods and *allograph*-based methods. In textural based methods, comprehensive statistic information are computed from the handwriting, e. g., the width of the ink stroke. A prominent example describes the handwriting by means of local binary patterns (Nicolaou et al. 2015). In comparison, allograph-based methods rely on a background model computed from local descriptors of a training set. This background model is then used to *encode* the local descriptors, i. e., to compute statistics from them. The most closely related publications belong to the latter group (Christlein et al. 2014; 2015a; 2015b). In our earliest work (Christlein et al. 2014), we used RootSIFT descriptors as local descriptors in combination with GMM supervectors for encoding. A variant of the GMM supervectors was also used in our most recent work (Christlein et al. 2015b), where they are employed to encode CNN activation features. In contrast, *vectors of locally aggregated descriptors* (VLAD) are used to encode Zernike moments which were evaluated densely at the script contour in our other work (Christlein et al. 2015a). This approach won the ICDAR 2015 competition on multi-script writer identification (Djeddi et al. 2015).

3 Methodology

Since the contour extraction involves more steps in historical documents than for clean benchmark data, we employ sparsely sampled RootSIFT descriptors (Arandjelović and Zisserman 2012) for our baseline method. For the aggregation of these local descriptors, we use multiple VLAD encodings (Jégou and Chum 2012; Jégou et al. 2012). The global descriptors of the training set are used to train a classifier for the date prediction.

The full workflow consists of three main steps: 1) local feature extraction, where we employ RootSIFT descriptors, 2) the aggregation of the local feature descriptors in the encoding step, 3) estimation of the date by means of linear regression.

3.1 Feature extraction

We make use of the Scale Invariant Feature Transform (SIFT) (Lowe 2004). SIFT descriptors encode the orientations of gradients in the neighborhood of scale and rotation invariant positions (*keypoints*) in the image. Note that we set the keypoint-angles to zero, since rotation-invariance is not necessary for the classification of handwriting (Fiel and Sablatnig 2013). Each SIFT descriptor is normalized using the Hellinger kernel (Arandjelović and Zisserman 2012), i. e., the square root is applied to each element, hence the name *RootSIFT*. This normalization reduces the effect of dominating values in the SIFT descriptor and has been shown to be very beneficial for writer identification (Christlein et al. 2014).

3.2 Encoding

The formation of a global descriptor is accomplished by the use of VLAD (Jégou et al. 2012). First, a dictionary C is computed from local descriptors using k-means. It consists of K cluster centers $\mu_k \in \mathbb{R}^D,\ k \in \{1, \dots, K\}$. For each cluster, all residuals between the cluster center and its nearest local descriptors are aggregated. Formally, given T as local descriptors $\mathrm{x_t} \in \mathbb{R}^D,\ t \in \{1, \dots, T\}$ of a single image:

$$\mathrm{v}_k = \sum_{\mathrm{x}_t :\ \mathrm{NN}(\mathrm{x}_t)=\mu_k} (\mathrm{x}_t - \mu_k)\,, \tag{1}$$

where NN(x) denotes the nearest neighbor of x. Then, the full $K \times D$ dimensional global descriptor follows by concatenation:

$$\mathrm{v} = (\mathrm{v}_1^\top, \dots, \mathrm{v}_K^\top)^\top\,. \tag{2}$$

Jégou and Chum (2012) showed that it is beneficial to use more than one dictionary resulting in multiple global descriptors. These are jointly decorrelated and dimensionality is reduced by means of PCA whitening. This has also been shown to improve the results for writer identification (Christlein et al. 2015a).

3.3 Date regression

To estimate the date the decorrelated VLAD vector v is used in a linear Support Vector Regression (SVR). The best hyper-parameters for the SVR are selected in an inner 5-fold cross-validation.

4 Evaluation

In this section, we introduce the dataset and error metrics that we use for evaluation. For the evaluation, we conduct several experiments using different preprocessing and sampling techniques for the feature extraction.

4.1 Dataset

The dataset used for evaluating the date estimation consists of 697 digitized medieval papal charters with known date. The documents come from three different archives. The majority (580) were provided by the Göttingen Academy of Sciences and Humanities (papsturkunden.de), 67 charters are provided by the Collaborative Archive

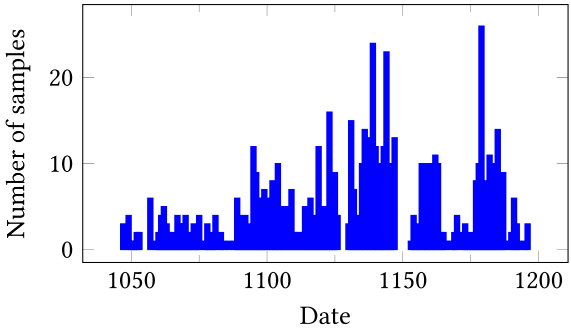

Figure 2: Distribution of documents in the dataset over the years.

Monasterium.net (Mom), and 50 stem from the Lichtbildarchiv älterer Originalurkunden - Philipps Universität Marburg (LBA). Most digitized images are retro-digitizations, i. e., digitizations from analog photos. Thus, the resolution and size of the documents vary greatly. Many documents also contain characteristics such as folds, stains, and rips, cf. figure 1. Also note that documents from the LBA contain a watermark which might have a small effect on the test accuracy (although the test set only contains two LBA-charters). The charters consist of one single document image. As a consequence, our experiments are inherently document-independent. We cannot guarantee an evaluation independent of the writer because the scribal hand is not known for the majority of the corpus. For in the time between 753 and 1197 AD, around 25 000 papal charters are handed down, about 20 000 of which are dated to the 12th century (see Hiestand 1999, 4), the chance of finding the same scribal hand in two different charters is presumably quite low. The dates of the charters of our dataset range between 1047 and 1196 AD. The year-sample distribution is depicted in figure 2.

We do not use the complete charters, but only the main context area, see figure 1 for an example. This way, it is guaranteed that graphical symbols (*rota*, *benevalete*, etc.) do not influence the results, and only the handwriting style is used for the date estimation. The main context areas were annotated during the project *Script and Signs. A computer-based analysis of high medieval papal charters. A key to Europe's cultural history* (PuhMa). We randomly split the dataset in roughly independent training (630 documents) and test (69 documents) sets.

4.2 Error metrics

We evaluate the predicted years of writing according to several error metrics. The *Mean Absolute Error* score (MAE) provides a simple indication of the average performance of the estimator:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| , \tag{3}$$

Variant	MAE	RMSE
Baseline	20.62	25.09
Baseline w.o. extr. kpts	20.81	25.21

Table 1: Evaluating the influence of extremely sized keypoints. The first row shows the results for the baseline, while the second row shows the results for the baseline method without extremely sized keypoints.

where N is the number of test documents, and y_i and $\hat{y}_i$ are the true and estimated years for document i, respectively.

In order to gain some more insight into the behavior regarding outliers, we consider another metric, the Root Mean Squared Error (RMSE), which puts more emphasis on outliers than MAE:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}\,. \tag{4}$$

Finally, the *Cumulative Score* (CS) (Geng et al. 2007) is a useful metric in cases where there is no or little value in a perfectly exact prediction. Instead, it assumes an *acceptable error* α (here given in years) and gives the percentage of the predictions that fall within this margin of error:

$$\mathrm{CS}_\alpha = \frac{N_{e\leq\alpha}}{N} \cdot 100\%\,. \tag{5}$$

4.3 Experiments

We evaluate different aspects regarding the size and sampling strategies for the RootSIFT descriptors. First, we try to limit the number of descriptors, next we experiment with reducing the image size. Finally, we evaluate the impact of different sampling strategies.

The *baseline* in our experiments denotes the pipeline as explained in section 3. Table 1 shows that the baseline approach gives an MAE of about 20 years and an RMSE of 25 years. According to the literature (see section 2), this is comparable to the state of the art in image-based dating. It follows that the transfer from a writer identification method to a date estimation method was successful.

Reducing the number of descriptors

In a first experiment, we removed keypoints varying more than twice the standard-deviation from the mean keypoint size. This way, non-standard keypoints are removed.

Variant	MAE	RMSE
Baseline (unscaled)	20.62	25.09
Scale 2048	21.97	26.53
Scale 1024	40.67	47.28
Center-crop 2048	23.52	28.65
Center-crop 1024	32.56	38.84

Table 2: Comparison of the unscaled baseline results with scaled, or cropped versions of the image.

More formally, a keypoint k is removed when:

$$s(k) \gtrless \mu \pm 2 \cdot \sigma\,, \tag{6}$$

where $s(k)$ is the size of the keypoint k, and μ and σ are the mean and standard deviation of all keypoint sizes in the image respectively. See for example figure 3b, where the orange keypoints denote the extreme keypoints, i. e., those omitted for this experiment. Interestingly, table 1 shows that this step is not advisable in comparison to the baseline raised by the RMSE and MAE. It seems that larger keypoints, which result in descriptors covering a larger image portion, are beneficial. Thus, we do not remove extremely sized keypoints in the following experiments.

Influence of image scaling

Next, we evaluate the impact of image scaling. Since the images are quite large (in average 2603 × 2021 pixels), this would decrease the computational load. Thus, we scale down the images in such a way that the larger dimension consists of 2048 (1024) pixels by retaining the aspect ratio of the image. In two subsequent experiments, we take the center-crop of 2048 × 2048 pixels (1024 × 1024). If one image dimension is smaller we take this dimension, i. e., min(2048, width) × min(2048, height), proceeding similarly for center-crops with 1024 pixels in each dimension.

Table 2 shows that any scaling harms the date estimation. However, results for the unscaled baseline are only slightly better than a moderate scaling of 2048 pixels. A scaling to 1024 pixels worsens the results drastically. A possible reason might be the lower number of detected keypoints, and, thus, extracted descriptors in the image. Using the center-crop of 2048 pixels is slightly worse than rescaling to 2048 pixels. Interestingly, the center-crop of 1024 pixels is much better than its scaling counterpart.

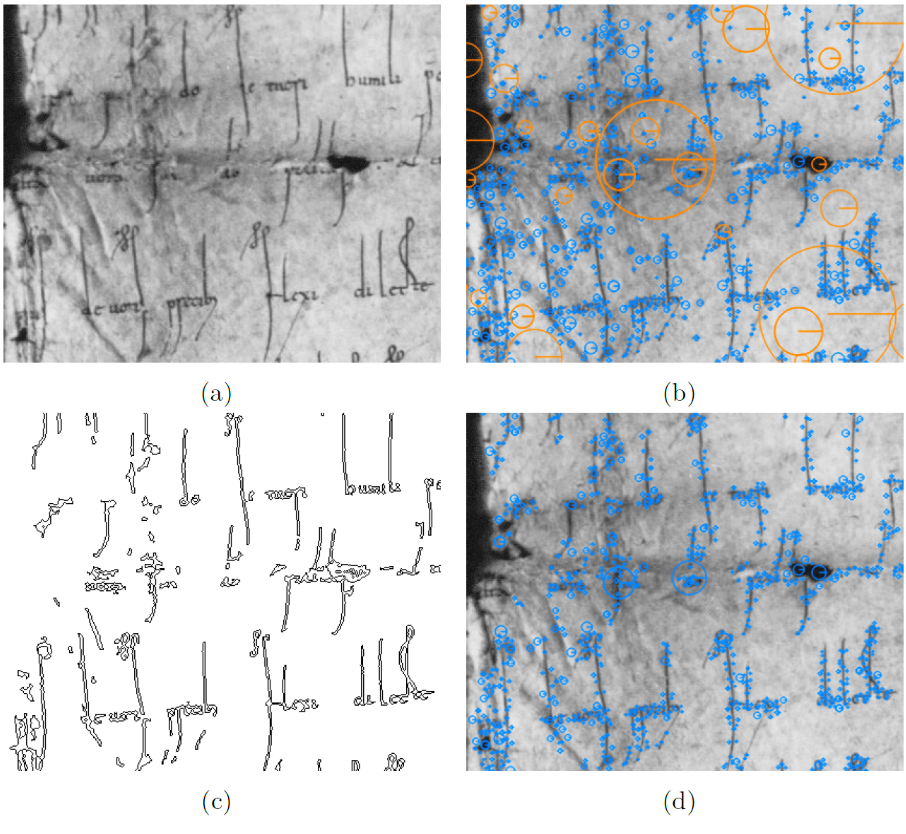

Figure 3: a) excerpt of figure 1; b) SIFT keypoints (orange: keypoints with extreme size), for the baseline results all keypoints are taken; c) contour sampling; d) masked SIFT keypoints.

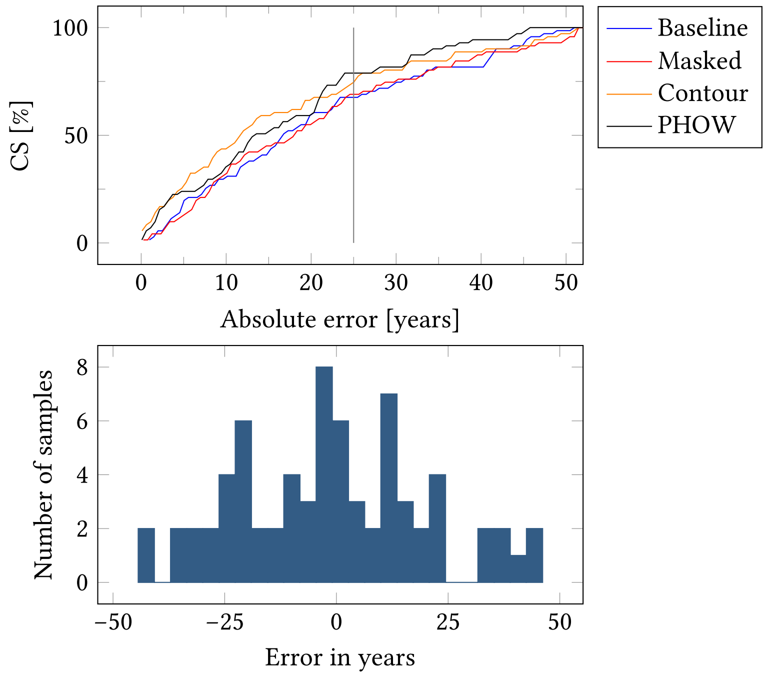

Figure 4: Top: Cumulative absolute error distributions of the different sampling strategies. Bottom: Histogram (with 25 bins) of errors using the PHOW method.

Influence of feature sampling

As a last experiment we evaluated different sampling strategies, i. e., we compare different positions (keypoints) at which feature descriptors are computed. The baseline uses the original SIFT keypoint detection proposed by Lowe (2004), see for example figure 3b. At these keypoints, the RootSIFT descriptors are extracted. We compare it with three different variants:

1.) We compute the keypoints as before but use only those which are close to the handwriting script (denoted as Masked RootSIFT). Therefore, we compute a mask which mainly consists of handwriting. To segment the handwriting in background and text, we apply the binarization technique proposed by Su et al. (2010). Remaining noise is reduced by removing connected components thare are too small (less than 20 points) or too large (larger than 3000 points). The mask is dilated by a 5×5 circular shape to allow keypoints at the border of the handwriting. Figure 3d shows an example of masked keypoints.

Variant	MAE	RMSE	CS_{25}
Baseline	20.62	25.09	67.61
Masked RootSIFT	21.45	27.19	69.01
Contour RootSIFT	17.17	23.01	73.24
PHOW	**16.95**	**21.04**	**78.87**

Table 3: Evaluation of different sampling strategies.

2.) We use the points situated at the contour of the handwriting (denoted as Contour RootSIFT). Therefore, we use the same strategy as before without the dilation step, see for example figure 3c for the extracted contour. At each contour point we evaluate the RootSIFT descriptor.
3.) Finally, a fast and dense variant of SIFT, known as Pyramid Histogram of Visual Words (PHOW) (Bosch et al. 2007) is computed. We extract the PHOW descriptor from the slightly downscaled version where the larger image dimension was resized to 2048 pixels. Descriptors having a norm lower than 0.05 were discarded since they stem from homogeneous areas. We used multiple bin sizes (4, 10, 16) and a step size of 10. The descriptors are Hellinger-normalized, similarly to the RootSIFT descriptors.

Table 3 shows the results for the four different strategies. It reveals that the masked variant of RootSIFT slightly worsens the results. This might be related to parts where the segmentation for the mask creation fails. In contrast, the contour-based RootSIFT and the densely sampled RootSIFT descriptors both surpass the baseline results by a significant margin. Both achieve similar results of about 17 years MAE. Regarding the RMSE, PHOW is in favor. Note, however, more keypoints are computed for these two methods and an order of magnitude more than for the baseline. This effects the computational costs for the feature extraction (especially for the contour-based method) and for the encoding step since more descriptors need to be accumulated.

The CS_{25} draws a picture similar to the MAE and RMSE values. However, figure 4 (top) shows that in ranges below 20 years, the cumulative score of the contour-based sampling strategy is in favor. Figure 4 (bottom) depicts the error histogram of the PHOW method. While there are fewer documents outside errors of ±25 years, there is a clear peak around 0 showing that several documents could be dated very exactly. Note that the results show a significant (Pearson-)correlation of 92% between the regression output and the true date (significance level 0.001).

5 Conclusion

In this work, we have shown that a method originally developed for writer identification can be transferred to fulfil the task of dating historical manuscripts. The historical manuscripts we used are not comparable to clean benchmark data, they are typically digitized in high resolution but contain deficiencies such as holes or stains. For this reason, we evaluated different strategies to lower the computational burden by reducing the image size. The results show that, while moderate scaling is acceptable, the results drop drastically in case of excessive scaling.

We also showed that sampling strategies other than SIFT keypoints improve the results. Both a dense SIFT variant (PHOW) as well as contour-based sampling surpass the baseline achieving an MAE of about 17 years and an RMSE of 21 years. However, the increase in keypoints comes at the cost of an increased computational complexity.

For future research, we would like to expand our studies regarding the feature sampling. Maybe other keypoint strategies, such as a sparse contour sampling could decrease the computational cost. Given enough training data, deep learning techniques could also be used for dating handwritten text similar to the work of Li et al. (2015).

Acknowledgments

Many thanks to Benedikt Hotz, Benjamin Schönfeld, Thorsten Schlauwitz, and Viktoria Trenkle for the annotation of the papal charters.

Bibliography

Arandjelović, Relja, and Andrew Zisserman. “Three things Everyone Should Know to Improve Object Retrieval.” *Computer Vision and Pattern Recognition* (CVPR). IEEE Conference. 2012. 2911–2918.

Bosch, Anna, Andrew Zisserman, Xavier Mu, and Xavier Munoz. “Image Classification Using Random Forests and Ferns.” *IEEE 11th International Conference onComputer Vision* (ICCV). 2007. 1–8.

Christlein, Vincent, David Bernecker, and Elli Angelopoulou. “Writer identification using VLAD encoded contour-Zernike moments.” *Document Analysis and Recognition* (ICDAR). 13th International Conference. 2015. 906–910.

Christlein, Vincent, David Bernecker, Florian Hönig, and Elli Angelopoulou. “Writer Identification and Verification Using GMM Supervectors.” *2014 IEEE Winter Conference on Applications of Computer Vision* (WACV). 2014. 998–1005.

Christlein, Vincent, David Bernecker, Andreas Maier, and Elli Angelopoulou. “Offline Writer Identification Using Convolutional Neural Network Activation Features.” In Gall, Juergen, Peter Gehler, and Bastian Leibe (eds.) *37th German Conference on Pattern Recognition.* GCPR 2015. Aachen, Germany, October 7-10, 2015. Proceedings. London: Springer International Publishing, 2015. 540–552.

Djeddi, Chawki, Somaya Al-Maadeed, Abdeljalil Gattal et al."ICDAR2015 Competition on Multi-script Writer Identification and Gender Classification using 'QUWI' Database." *13th International Conference on Document Analysis and Recognition* (ICDAR). 2015. 1191–1195.

Feuerverger, Andrey, Peter G. Hall, Gelila Tilahun, and Michael Gervers "Using Statistical Smoothing to Date Medieval Manuscripts." In *Beyond parametrics in interdisciplinary research: Festschrift in honour of Professor Pranab K. Sen.* 2008. 321–331. <http://arxiv.org/abs/0805.2490>.

Fiel, Stefan and Robert Sablatnig. "Writer Identification and Writer Retrieval using the Fisher Vector on Visual Vocabularies." *12th International Conference on Document Analysis and Recognition* (ICDAR). 2013. 545–549.

Geng, Xin, Zhi-Hua Zhou, and Kate Smith-Miles. "Automatic Age Estimation Based on Facial Aging Patterns." *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 29.12 (2007). 2234–2240.

Gervers, Michael. "The DEEDS project and the development of a computerised methodology for dating undated English private charters of the twelfth and thirteenth centuries." *Dating Undated Medieval Charters.* 2000. 13–35.

He, Sheng, Petros Samara, Jan Burgers, and Lambert Schomaker. [2014.] "Towards Style-Based Dating of Historical Documents." *International Conference on Frontiers in Handwriting Recognition* (ICFHR). 2014. 265–270.

He, Sheng, Petros Samara, Jan Burgers, and Lambert Schomaker. [2016a.] "Historical Document Dating using Unsupervised Attribute Learning." *12th IAPR Workshop on Document Analysis Systems* (DAS). 2016. 36–41.

He, Sheng, Petros Samara, Jan Burgers, and Lambert Schomaker. [2016b.]"Image-based Historical Manuscript Dating Using Contour and Stroke Fragments." *Pattern Recognition 58.* 2016. 159-171. DOI: 10.1016/j.patcog.2016.03.032.

Hiestand, Rudolf. "Die Leistungsfähigkeit der päpstlichen Kanzlei im 12. Jahrhundert mit einem Blick auf den lateinischen Osten." In Herde, Peter, and Hermann Jakobs (eds.). *Papsturkunde und europäisches Urkundenwesen. Studien zu ihrer formalen und rechtlichen Kohärenz vom 11. bis 15. Jahrhundert.* Vol. 7. Archiv für Diplomatik, Schriftgeschichte, Siegel- und Wappenkunde. Köln: Böhlau, 1999. 1–26.

Jégou, Hervé, and Ondřej Chum. "Negative Evidences and Co-occurences in Image Retrieval: e Benefit of PCA and Whitening." In Fitzgibbon, Andrew et al. (eds.). *Computer Vision – ECCV 2012: 12th European Conference on Computer Vision, Florence, October 7-13 2012 Proceedings Part II.* Berlin: Springer, 2012. 774–787.

Jégou, Hervé, Florent Perronnin, Matthijs Douze, Jorge Sanchez et al. "Aggregating Local Image Descriptors into Compact Codes." *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 34.9 (2012). 1704–1716.

Li, Yuanpeng, Dmitriy Genzel, Yasuhisa Fujii, and Ashok C. Popat. "Publication Date Estimation for Printed Historical Documents Using Convolutional Neural Networks." *3rd International*

Workshop on Historical Document Imaging and Processing. Gammarth: ACM, 2015. 99–106.

LBA: *Lichtbildarchiv älterer Originalurkunden.* Philipps Universität Marburg. 2008. <http://lba.hist.uni-marburg.de/lba/>.

Lowe, David G. "Distinctive Image Features from Scale-Invariant Keypoints." *International Journal of Computer Vision* 60.2 (2004). 91–110.

Mom: *Collaborative Archive Monasterium.net.* 2005 –. <http://monasterium.net/mom>.

Nicolaou, Anguelos, Andrew D. Bagdanov, Marcus Liwicki, and Dimosthenis Karatzas. "Sparse Radial Sampling LBP for Writer Identification."*3th International Conference on Document Analysis and Recognition* (ICDAR). 2015. 716–720.

Papsturkunden.de. Göttingen Academy of Sciences and Humanities. 2015. <http://www.papsturkunden.de.

PuhMa: *Script and Signs. A computer-based analysis of high medieval papal charters. A key to Europe's cultural history.* 2012-2015. <http://www5.cs.fau.de/puhma>.

Su, Bolan, Shijian Lu, and Chew Lim Tan. "Binarization of Historical Document Images Using the Local Maximum and Minimum." *9th IAPR International Workshop on Document Analysis Systems.* 2010. 159–165.

Wahlberg, Fredrik, Lasse Mårtensson, and Anders Brun. "Large Scale Continuous Dating of Medieval Scribes Using a Combined Image and Language Model." *12th IAPR Workshop on Document Analysis Systems (DAS).* 2016. 48–53.

Wahlberg, Fredrik, Lasse Mårtensson, and Anders Brun. "Large Scale Style Based Dating of Medieval Manuscripts." *Proceedings of the 3rd International Workshop on Historical Document Imaging and Processing.* (HIP'15). 2015.