A Graphical Model Approach for Matching Partial Signatures

Xianzhi Du[1], David Doermann[2], and Wael AbdAlmageed[3]

[1]UMIACS, University of Maryland, College Park, `xianzhi@umiacs.umd.edu`
[2]UMIACS, University of Maryland, College Park, `doermann@umiacs.umd.edu`
[3]Information Sciences Institute, University of Southern California, `wamageed@isi.edu`

Abstract

In this paper, we present a novel partial signature matching method using graphical models. Shape context features are extracted from the contour of signatures to capture local variations, and K-means clustering is used to build a visual vocabulary from a set of reference signatures. To describe the signatures, supervised latent Dirichlet allocation is used to learn the latent distributions of the salient regions over the visual vocabulary and hierarchical Dirichlet processes are implemented to infer the number of salient regions needed. Our work is evaluated on three datasets derived from the DS-I Tobacco signature dataset with clean signatures and the DS-II UMD dataset with signatures with different degradations. The results show the effectiveness of the approach for both the partial and full signature matching.

1. Introduction

In today's electronic world, signatures are still widely used as a method for authorization, identification, and in some cases, document retrieval. This is due to the fact that signatures are written in a ballistic manner and will carry discriminative features that are difficult to forge by an amateur [7] [9]. The signature matching problem aims to identify the author of a query signature, based on shape comparisons with reference signatures.

For matching we distinguish between two problems—*full signature matching*, which assumes an accurate segmentation and an author who has produced a complete and consistent signature, and *partial signature matching*, where we enroll a full reference signature, but may only have a partial signature or even initials to match against the reference signature. A lot of work has shown great promise for addressing the full signature matching problem using feature descriptors such as shape context [16] [8]. However, partial signature matching remains an open problem. When dealing with real applications, even though full signatures are collected, many query signatures are only part of the authors' full name. This is due to multiple factors including the fact that people sometimes only sign their first/last names or initials, part of a signature may be missed when performing signature extraction, or when the signature is obscured by other information, such as the machine-printed text in a signature block.

In general, state of the art methods for signature matching can be grouped into global-shape based approaches and point-level approaches. Global-shape approaches use holistic shape representations to address this problem. For example, Srihari et al. [12] used global shape structure and Lin et al. [10] used Fourier descriptors. When working with full, high quality signatures, these methods work efficiently and accurately. However, global-shape methods work poorly when degradations are present, or when the signatures become occluded or are partial.

For point-level methods, signatures are modeled as contour points or skeleton points, and features such as shape context [4], or FREAK features [2] are used to describe the relations between local points. Signatures are matched either by computing the point-wise correspondences of two signatures or by comparing high-level features built for each signature. Du et al. [8] clustered all features of contour points and then built a term-frequency histogram for each signature as the high-level feature. Nearest neighbor search was then used to find the closest sample for a query signature. This method is fast and efficient, but it still does not work on partial signatures as local information is used together to build holistic features for each signature. Zhu et al. [16] compared all pairs of points between each pair of signatures to form global weights in the training stage. Each query signature is then compared point-wise with all samples to find the closest match. This method can tolerate a low degree of variation, which makes it more robust to the partial matching problem, but it does not scale to large datasets since it is computationally expensive due to the pairwise comparisons.

To address the partial signature matching problem, we developed a method based on the combination of supervised latent Dirichlet allocation (sLDA) [5] and hierarchical Dirichlet processes (HDP) [13]. SLDA was first developed as a statistical model of labeled documents, derived from latent Dirichlet allocation (LDA) [6]. Unlike LDA, in sLDA, each document is paired with a label. Topic distributions are estimated over the vocabulary and relation between topics and labels are discovered in the training stage. For an unlabeled document, the label is regressed from its topic structure. In our formulation, a "signature" corresponds to a "document", a "salient region" corresponds to a "topic", an "observation" corresponds to a "word", and an "authorship" corresponds to a "label".

In our approach, sLDA is first used to discover the salient regions in all training signatures. A salient region is a distribution over the features in the visual vocabulary, which groups similar co-occurring observations. Each author is modeled as a combination of all salient regions with different proportions. For a query signature, classification is performed by computing the salient region proportions for the signature based on observations. Further, instead of guessing the number of salient regions empirically, HDP is used to estimate the number needed for the given dataset.

We organize this paper as follows. Section 2 describes how we use observations for each signatures to build a vocabulary over all signatures. Section 3 provides a detailed description of how to build the supervised topic models and how HDP is used to estimate the number of salient regions. Section 4 discusses the experiments and results, and conclusion and discussions are provided in Section 5.

2. Observation and vocabulary building

In this section, we describe how signatures are modeled as a group of observations.

2.1. Partial shape context feature extraction

The first step of modeling signatures as a group of observations is to find a proper feature descriptor. As we are working with 2-D binary shapes, we want to find a feature descriptor that captures the relations between points in the binary shapes while preserving the local information between full and partial signatures. In this case, popular features like SIFT [11] and SURF [3] are not suitable since they use gradient information of feature points which is not informative with binary shapes. We therefore use shape context features, a feature designed for 2-D shape, which describes the relations between nearby points while tolerating slight shape distortion. This is especially important for the signature matching problem since even high-quality signatures from same author may have slight differences.

To build observations, we first extract contour points from each signature proportional to the total length of the contour. The result is that partial signature will have similar contour points as its corresponding part in the full signature. For each contour point c_i, a $r * \theta$ log-polar space is formed around it with uniform bins. A histogram s_i is built by calculating the number of nearby points that fall in each bin in a certain order, based on the relative distance and angle of the two points, as shown in Equation (1).

$$\begin{gathered} s_i = [s_i(1), s_i(2), ..., s_i(n)] \\ s_i(k) = \# \{p : p \in bin(k), p \neq c_i\} \end{gathered} \tag{1}$$

where $n = r * \theta$ is the total number of histogram bins, and p represents contour points. In order to make shape context applicable to partial signature matching, instead of normalizing the pairwise point distances within each signature, we normalize the pairwise point distances for all possible point pairs in all signatures by dividing by the mean value $\bar{D}$, as shown in Equation (2).

$$\begin{gathered} \bar{D} = (\sum_{l=1}^{N_{sig}} \sum\nolimits_{i,j \in s_l} D(p_i, p_j))/N_{pair} \\ D_{norm}(p_i, p_j) = D(p_i, p_j)/\bar{D} \end{gathered} \tag{2}$$

where $D(p_i, p_j)$ represents the distance between two points p_i and p_j in signature s_l, N_{sig} is the total number of signatures, and N_{pair} is the total number all possible pairs of points in all signatures. Since the size of log-polar space is fixed for all signatures, partial signatures have shape context features similar to a corresponding full signature, as illustrated in Figure 1.

2.2. Building the visual vocabulary

After extracting contour points and computing shape context features, the next step is to build the visual vocabulary for all signatures. One intuitive way is to cluster all contour points and treat each cluster label as one observation. The vocabulary consists of all the cluster labels. For each signature, the number of contour points being classified into one cluster is regarded as the appearance frequency of this observation.

In our method, we use K-means clustering. A histogram h_t for the t_{th} signature is formed to indicate the appearance frequencies of all observations, as shown in Equation (3).

$$\begin{gathered} h_t = [h_t(1), h_t(2), ..., h_t(K)] \\ h_t(k) = \# \{p : p \in Cluster(k), p \in Sig(t)\} \end{gathered} \tag{3}$$

Finally, signatures are represented by an observation-frequency histograms.

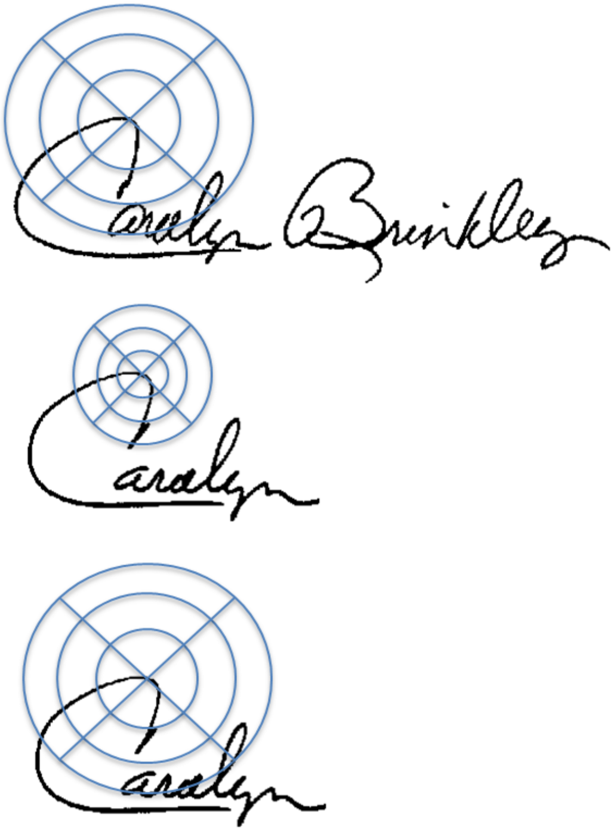

Figure 1. Shape context features for three signatures. The first row shows a full signature and the log-polar space of the shape context features. The second row shows its partial signature and the log-polar space by using the standard shape context features. The third row shows the same partial signature and the log-polar space by using our modified shape context features.

3. Supervised topic models for partial signature-matching

In this section, we build a supervised topic model for the partial signature matching problem and briefly review the variational inference solution based on the work of Wang et al. [14].

3.1. Building supervised topic model

For our problem, the generative process for the n_{th} observation in the t_{th} signature is given as follows:

1. For the t_{th} signature, draw salient region proportions θ_t from $Dir(\alpha)$
2. For the n_{th} observation:
   (a) Draw a salient region assignment $S_{t,n}$ from $Mult(\theta_t)$
   (b) Draw an observation $O_{t,n}$ from $Mult(\beta_{S_{t,n}})$

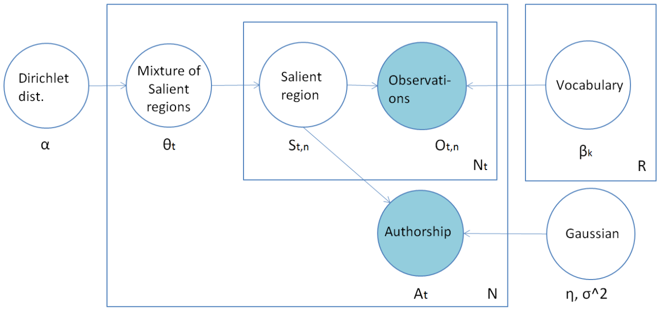

Figure 2. SLDA model for our problem.

3. Draw authorship variable A_t from $N(\eta^{\intercal}\bar{S}_t, \sigma^2)$

$Dir(\cdot)$, $Mult(\cdot)$, $N(\cdot)$ represent the Dirichlet distribution, the Multinomial distribution, and the Normal distribution respectively. α is an R-dimensional hyperparameter for the Dirichlet distribution with R being the number of salient regions. $\beta = [\beta_1, \beta_2, ..., \beta_R]$, where each β_r is the distribution of the salient region r over the vocabulary, and $\bar{S}_t$ is the mean of the salient regions of the t_{th} signature. With only observations and authorship given, we want to estimate $\alpha, \beta, \eta, \sigma^2$. The model is shown in Figure 2.

The variational EM algorithm [14] is used to solve the sLDA model. Here we give a brief description.

By introducing the free variational parameters γ and ϕ, the posterior distribution $p(\theta, S|O, A, \alpha, \beta, \eta, \sigma^2)$ is approximated by Equation (4).

$$q(\theta, S|\gamma, \phi) = q(\theta|\gamma)\prod_{n=1}^{N} q(S_n|\phi_n) \tag{4}$$

Here γ is a R-dimensional variational Dirichlet hyperparameter that governs the salient region distribution of each signature. We can regard it as an approximation for α in the original model. $\phi = [\phi_1, \phi_2, ..., \phi_{N_t}]$ is an approximation to θ, but specific for each observation.

The variational EM algorithm works as follows: In the E-step, γ and ϕ are computed to minimize the KL divergence between the true posterior and the approximation. In the M-step, each of the salient regions β_r is estimated by counting how many times each observation is assigned to this salient region among all signatures. η and σ^2 are estimated by the relationships between salient regions and the labels for all training signatures. For a new signature, the label is predicted by Equation (5).

$$\mathbb{E}[A|O, \alpha, \beta, \eta, \sigma^2] = \eta_{est}^{\intercal}\mathbb{E}[S|O, \alpha, \beta] = \eta_{est}^{\intercal}\phi_{est} \tag{5}$$

For more details about the algorithm in this section please refer to [6] [14] [5].

3.2. HDP for salient region estimation

In LDA and sLDA, the number of salient regions needs to be prefixed and it is always chosen empirically. When processing new and massive data, it is not possible to easily choose an optimal number and it is also very expensive to reprocess the massive data multiple times to find out the optimal number. To solve this problem, Teh [13] provided a new topic model structure called hierarchical Dirichlet processes, which lets the data estimate the number of salient regions needed.

The main difference between HDP and standard LDA lies in the model structure. In HDP, each set of data (signatures in our case) has its own mixture model with random probability measure G_t. G_t's are distributed as a Dirithlet process with a global base distribution G_0. G_0 itself is also distributed as a Dirichlet process.

$$G_t \sim DP(\alpha_0, G_0) \tag{6}$$

$$G_0 \sim DP(\gamma, H) \tag{7}$$

where α_0, γ are the concentration parameters, and H is the base distribution for G_0.

A straightforward way to explain HDP is using the Chinese restaurant franchise process. We give a brief introduction as follows: Let $\theta_{t,n}$ be the salient region for the n_{th} observation in the t_{th} signature, $\psi_{t,r}$'s be the existing salient regions for signature t, $N_{t,r}$ be the number of observations in signature t under salient region r, M_t be the number of salient regions used in signature t. For a new observation, the salient region assignment is given as follows:

$$\begin{gathered}\theta_{t,n} | \theta_{t,1}, ..., \theta_{t,n-1}, \alpha_0, G_0 \sim \\ \sum_{r=1}^{M_t} \frac{N_{t,r}}{n-1+\alpha_0} \delta_{\psi_{t,r}} + \frac{\alpha_0}{n-1+\alpha_0} G_0\end{gathered} \tag{8}$$

The salient region of a new observation can be either an existing salient region within this signature, which is a draw from the first summation term on the right-hand side of Equation (8) with a probability proportional to the number of observations under that salient region, or can be a new salient region with probability given by the second term on the right-hand side of Equation (8). If a new salient region is needed for this observation, we draw one salient region $\psi_{t,r}$ from G_0 and increase M_t by one as follows:

$$\begin{gathered}\psi_{t,r} | \psi_{1,1}, ..., \psi_{t,r-1}, \gamma, H \sim \\ \sum_{r_0=1}^{R} \frac{M_{r_0}}{M+\gamma} \delta_{\phi_{r_0}} + \frac{\gamma}{M+\gamma_0} H\end{gathered} \tag{9}$$

where M_{r_0} is the number of times salient region r_0 is used in all signatures and M is the total number of times all salient regions are used in all signatures. If the first term of the right-hand side of Equation (9) is chosen, the new salient region for $\theta_{t,n}$ is picked among the existing salient regions with a probability proportional to the number of times one salient region is used in all signatures. If the second term is chosen, a new salient region is introduced and the total number of salient regions is increased by one. For more details about the algorithm, please refer to [13] [15].

4. Experiments and results

4.1. Datasets

Two popular datasets for the signature matching problem are the DS-I Tobacco signature dataset and the DS-II UMD signature dataset [1]. The DS-I Tobacco signature dataset contains 189 relatively clean and high-quality signatures from 63 authors, with three per author. All the signatures are full signatures, meaning each contains the full name of the author. The DS-II UMD signature dataset contains 26661 signatures from 887 authors. It is a more challenging and a more realistic dataset since the signatures are extracted from the Tobacco litigation corpus, containing memos and letters. It is a mixture of both full signatures and partial signatures as many signatures are partial signatures or initials. Moreover, signatures in this dataset have other different kinds of degradations such as: signatures with low-quality; machine-printed texts like address and date that appear in the signature block; occluded signatures; and signatures are too simple to be classified.

Our method is tested on two partial signature datasets and one full signature dataset. The two partial signature datasets are built based on the DS-I Tobacco signature dataset and the DS-II UMD signature dataset. We refer to them as the DS-I partial dataset and DS-II partial dataset. The full signature dataset is the DS-I Tobacco signature dataset itself.

DS-I partial dataset: As there are no partial signature datasets which contain clean and high-quality partial signatures, so we manually selected the DS-I partial set from the DS-I Tobacco signature dataset by clipping out the first names of all full signatures. Full signatures are used to train and partial signatures are used to test in the experiment. Samples from the DS-I partial set are shown in Figure 3.

DS-II partial dataset: This dataset contains 495 full signatures for training from 495 authors and 1732 signatures contain both partial and full signatures for testing. To build the DS-II partial dataset, we use authors in the DS-II UMD dataset with at least one partial signature. Here we define a partial signature as follows: For each signature, when compared to a full sample, if any part that should belong to the full signature is missing, we consider it to be a partial signature. We didn't test on the whole dataset because the number of partial signatures is relatively too small compared to the number of full signatures. Due to the limited samples in this dataset, we only pick one full signature per author to train (which is the usual case in real life ap-

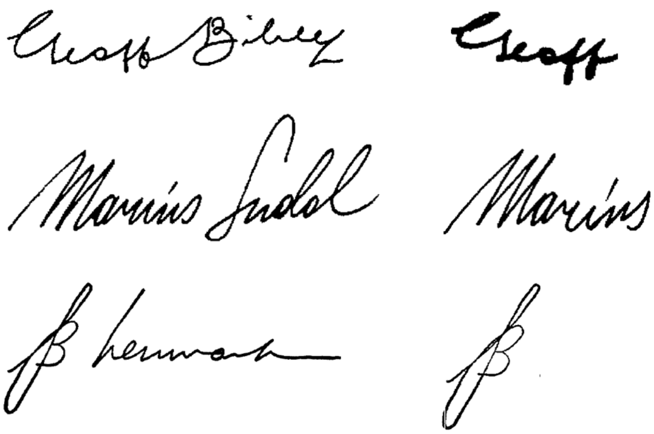

Figure 3. Sample signatures in DS-I partial dataset. Left column shows three full signatures. Right column shows their partial signatures.

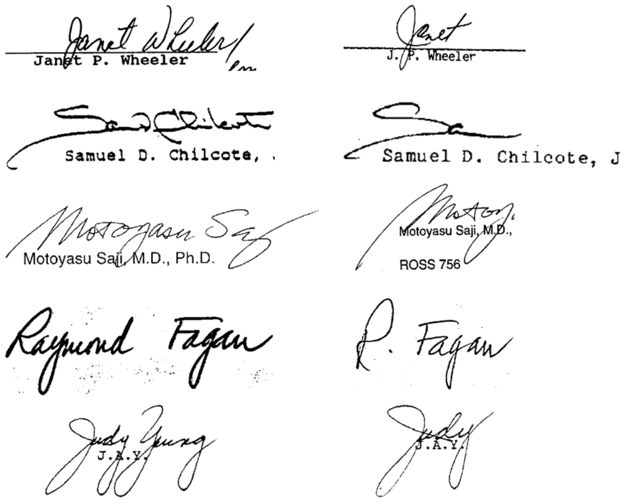

Figure 4. Sample signatures in DS-II partial dataset. Left column shows five full signatures. Right column shows their partial signatures with different kinds of degradations.

plications) and use other signatures (may contain both full and partial signatures) to test. Each author has at least one and at most 30 signatures for testing. Samples from DS-II partial dataset are given in Figure 4.

4.2. Evaluation protocol

In our experiments, the top-N rank accuracy is used for evaluation. Top-N rank strategy means for each query signature, in the prediction stage, if the full signature sample of the true author appears in any position of the first N_{th} ranks, the top-N rank accuracy for that sample is set to 1. The total top-N rank accuracy is the mean of all query signatures.

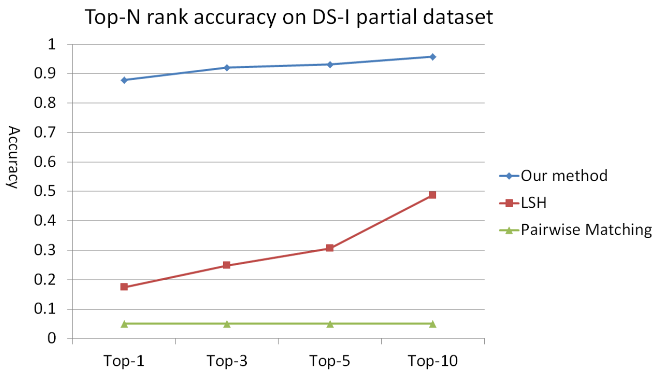

Figure 5. Performance curves on DS-I partial dataset.

	Top-1	Top-3	Top-5	Top-10
LSH	0.175	0.249	0.307	0.487
Pairwise matching	<0.05	<0.05	<0.05	<0.05
Our method	**0.878**	**0.921**	**0.931**	**0.958**

Table 1. Results on DS-I partial dataset

4.3. Results on DS-I partial dataset

We normalize the height of each signature to the mean height of all training signatures and set the widths accordingly to preserve the aspect ratio. In this experiment, the heights of all signatures are fixed to 182. For shape context features, $r = 5$ and $\theta = 12$ are used. As we have three signatures from each author in this dataset, each time we pick two full signatures to train and the partial signature of the rest one to test. We run three rounds to cover all the signatures in both training and testing stages. The average accuracy is reported.

First we run HDP on the DS-I parital dataset to estimate the number of salient regions. $R = 90$ is suggested after 1000 iterations. For K in K-means algorithm, empirically we choose $K = 1300$. Due to the random initialization of K-means algorithm, we rerun the whole experiment 10 times. The term most frequently appears on the 10 i_{th}'s rank is taken. Our method is compared with Du et al. [8] and Zhu et al. [16]. We refer to them as LSH method and pairwise matching method. Results are given in Figure 5 and Table 1.

Our method achieves 87.8% top-1 accuracy on this partial dataset, which significantly outperforms the LSH method by 17.5% and the pairwise matching method. The pairwise matching method reported a less than 5% top-1 rank accuracy on this dataset due to incapability of dealing with partial signature-matching problem and insufficient training samples.

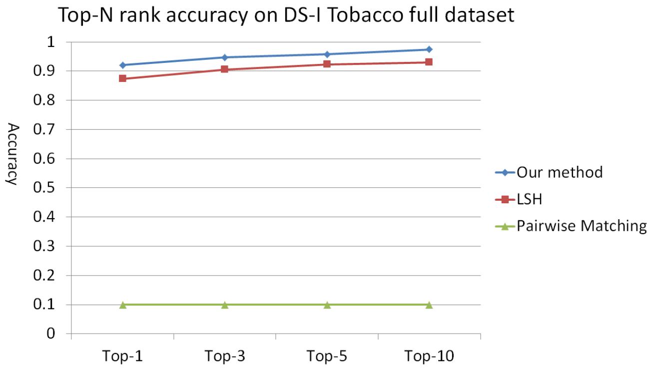

Figure 6. Performance curves on DS-I Tobacco full dataset.

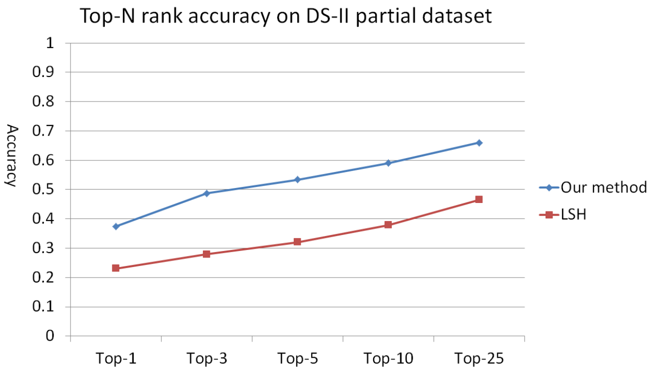

Figure 7. Performance curves on DS-II partial dataset.

	Top-1	Top-3	Top-5	Top-10
LSH	0.874	0.906	0.923	0.928
Pairwise matching	<0.1	<0.1	<0.1	<0.1
Our method	**0.921**	**0.947**	**0.958**	**0.974**

Table 2. Results on DS-I Tobacco full dataset

4.4. Results on DS-I Tobacco full signature dataset

In order to show that our method also works well on a full signature dataset, we test on the DS-I Tobacco full signature dataset. We run this experiment three times and report the average accuracy as we did in the previous experiment except that the full signatures are used to test. We use the same setups as the previous experiment, since the training stages are identical. The results are compared to the LSH method and the pairwise matching method reported in [8]. Figure 6 and Table 2 show the results.

Our 92.1% top-1 rank accuracy outperforms both of the previous methods on this dataset. This proves that our method is not only specifically designed for partial signature matching, but also works well on full signature dataset.

4.5. Results on DS-II partial dataset

This dataset is extremely challenging since we only have one training sample per class to perform the 495-class classification task. To setup the experiment, the heights of all signatures are fixed to 166, and $r = 5$ and $\theta = 12$ shape context features are used. 230 salient regions are estimated from HDP after 1000 iterations, and empirically $K = 1500$ is chosen. Figure 7 and Table 3 show the results from our method and the LSH method. Our algorithm achieve 37.8% top-1 rank accuracy, which outperforms the LSH method by 23.1%. We didn't compare to the pairwise matching method since it is not scalable to large datasets.

4.6. Effects of parameters

The two parameters that have impact on the performance are the number of centers K in K-means algorithm and the number of salient regions R in sLDA. We test on different values of K and R to see how they change the performance.

Number of clusters: The choice of K depends on the data and the size of the dataset. A larger value of K tends to over-classify the data and make the classification more computationally intensive. A smaller value of K tends group incorrect data together. The results of K versus top-1 rank accuracy graphs on DS-I partial dataset and DS-II partial dataset are given in Figure 8 and Table 4, and Figure 9 and Table 5, at fixed $R = 90$ and $R = 230$ respectively.

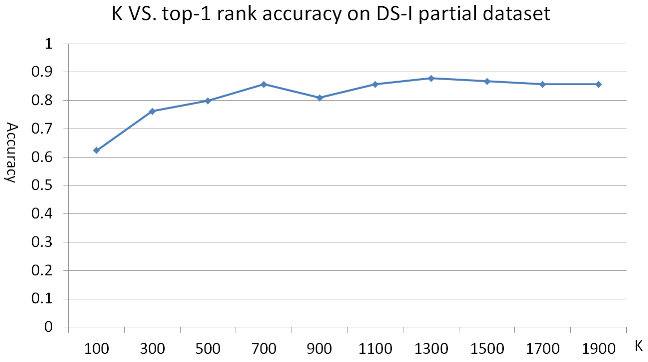

Figure 8. The impact of K on DS-I partial dataset.

Number of salient regions: The choice of R depends on the size of vocabulary and the observations. HDP provides us an alternative way to choose R. The top-1 rank accuracy comparisons between R predicted by HDP and other R values on DS-I partial dataset and DS-II partial dataset are given in Figure 10 and Table 6, and Figure 11 and Table 7, at $K = 1300$ and $K = 1500$ respectively.

	Top-1	Top-3	Top-5	Top-10	Top-25
LSH	0.231	0.279	0.321	0.379	0.465
Our method	**0.378**	**0.487**	**0.534**	**0.590**	**0.660**

Table 3. Results on DS-II partial dataset

$\#K$	100	300	500	700	900	1100	1300	1500	1700	1900
Accuracy	0.624	0.762	0.799	0.857	0.809	0.857	**0.878**	0.867	0.857	0.857

Table 4. Results from different K at $R = 90$ on DS-I partial set

$\#K$	100	300	500	700	900	1100	1300	1500	1700	1900	2100	2300	2500
Accuracy	0.290	0.334	0.335	0.327	0.337	0.303	0.358	**0.378**	0.335	0.335	0.314	0.310	0.301

Table 5. Results from different K at $R = 230$ on DS-II partial set

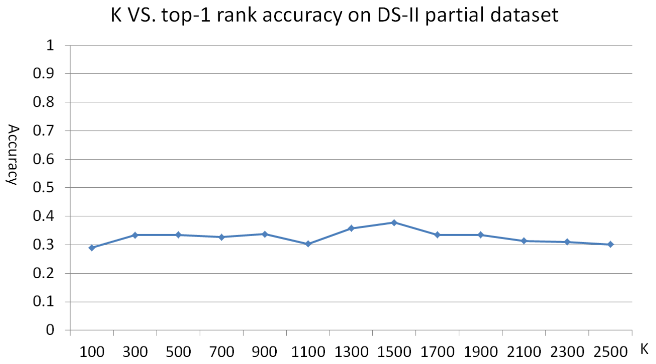

Figure 9. The impact of K on DS-II partial dataset.

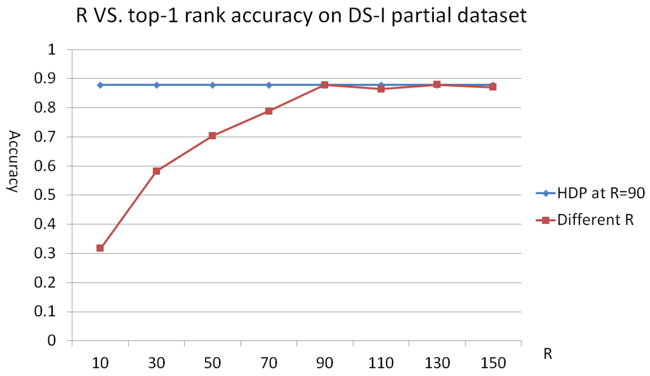

Figure 10. The impact of R and comparisons between different R and R predicted by HDP on DS-I partial dataset.

From Figure 11 we see HDP overfit the training samples of DS-II partial dataset with a large value of R. This is due to the variety of the training samples.

4.7. Failure examples

There are several common cases in the DS-II partial dataset that will lead to a mismatching, such as the full signature of a query partial signature is close to another signature, the partial signature is more similar to the full signature of another author, signatures are not informative because of too many handwritten texts, and signatures are so simple that limited information can be collected. Figure 12 shows four failure cases.

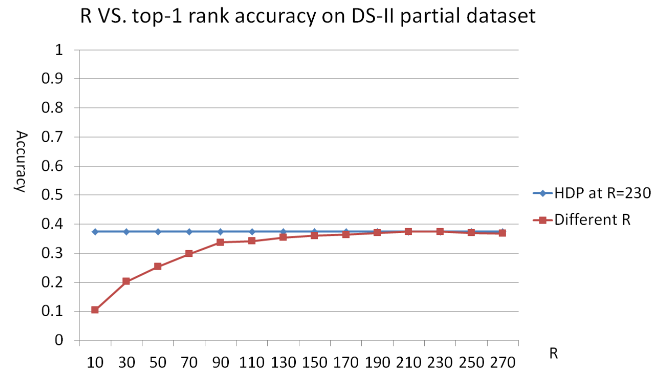

Figure 11. The impact of R and comparisons between different R and R predicted by HDP on DS-II partial dataset.

Figure 12. Failure examples for four query signatures. For each box, the first signature is a query signature; the middle signature is the incorrect match; the last signature is the true full signature.

5. Conclusion and discussions

We presented an effective and scalable partial signature matching method based on modified shape context features and topic modeling. The motivation is to deal with real sig-

$\#R$	10	30	50	70	90	110	130	150
Accuracy	0.317	0.582	0.704	0.788	**0.878**	0.864	0.878	0.870

Table 6. Results from different R at $K = 1300$ on DS-I partial dataset

$\#R$	10	30	50	70	90	110	130	150	170	190	210	230	250
Accuracy	0.105	0.203	0.254	0.298	0.337	0.342	0.354	0.360	0.364	0.370	0.378	**0.378**	0.370

Table 7. Results from different R at $K = 1500$ on DS-II partial dataset

nature matching applications when partial signatures are encountered. We modified shape context features to fit the partial matching problem better. K-means clustering algorithm is used to build the observations and vocabulary, sLDA is used to model each author as proportions of the hidden salient regions, and HDP is used to indicate the proper number of salient regions needed for each dataset.

Three experiments were conducted on two partial signature datasets and one full signature dataset. Our method outperformed previous methods on all datasets, which demonstrated that it works well on both partial and full signature datasets.

For further work, a better way to select the number of clusters to build the vocabulary is needed. We are also exploring effective methods to address other kinds of degradations.

References

[1] lamp.cfar.umd.edu.

[2] A. Alahi, R. Ortiz, and P. Vandergheynst. Freak: fast retina keypoint. *In Proc. CVPR*, pages 510–517, 2012.

[3] H. Bay, A. Ess, T. Tuytelaars, and L. Gool. Surf: Speeded up robust features. *In CVIU*, 110(3):346–359, 2008.

[4] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Trans. PAMI*, 24(4):509–522, 2002.

[5] D. Blei and J. McAuliffe. Supervised topic models. *In NIPS*, 2008.

[6] D. Blei, A. Ng, and M. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, pages 993–1022, 2003.

[7] D. S. Doermann and A. Rosenfeld. Recovery of temporal information from static images of handwriting. *International Journal of Computer Vision*, 52:143–164, 1994.

[8] X. Du, W. AbdAlmageed, and D. Doermann. Large-scale signature matching using multi-stage hashing. *In ICDAR*, pages 976–980, 2013.

[9] D. D. K. Guo and A. Rosenfeld. Forgery detection by local correspondence. *International Journal on Pattern Recognition and Artificial Intelligence*, 15:579–641, 2001.

[10] C. Lin and R. Chellappa. Classification of partial 2d shapes using fourier descriptors. *IEEE Trans. PAMI*, 9(5), 1987.

[11] D. Lowe. Distinctive image features from scale-invariant key points. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[12] S. Srihari, S. Shetty, S. Chen, H. Srinivasan, C. Huang, G. Agam, and O. Frieder. Document image retrieval using signatures as queries. *International Conference on Document Image Analysis for Libraries*, pages 198–203, 2006.

[13] Y. W. Teh, M. Jordan, M. Beal, and D. Blei. Hierarchical dirichlet processes. *Journal of the American Statistical Association*, pages 1566–1581, 2006.

[14] C. Wang, D. Blei, and F. Li. Simultaneous image classification and annotation. *In CVPR*, 2009.

[15] C. Wang, J. Paisley, and D. Blei. Online variational inference for the hierarchical dirichlet process. *In AISTATS*, 2011.

[16] G. Zhu, Y. Zheng, and D. Doermann. Signature detection and matching for document image retrieval. *IEEE Trans. PAMI*, 31(11):2015–2031, 2009.