Detecting handwritten signatures in scanned documents

İlkhan Cüceloğlu[1,2], Hasan Oğul[1]

[1]Department of Computer Engineering, Başkent University, Ankara, Turkey
[2]DAS Document Archiving and Management Systems CO., Ankara, Turkey
`hogul@baskent.edu.tr, ilkhan.cuceloglu@das.com.tr`

Abstract *Automated localization of a handwritten signature in a scanned document is a promising facility for many banking and insurance related business activities. We describe here a discriminative framework to extract signature from a bank service application document of any type. The framework is based on the classification of segmented image regions using a set of representative features. The segmentation is done using a two-phase connected component labeling approach. We evaluate solely and combined effects of several feature representation schemes in distinguishing signature and non-signature segments over a Support Vector Machine classifier. The experiments on a real banking data set have shown that the framework can achieve a reasonably good accuracy to be used in real life applications. The results can also provide a comparative analysis of different image features on signature detection problem.*

1 Introduction

In spite of a drastic increase in the use of electronic data in bank applications, a signature in a printed document is still considered to be most reliable way of user commitment, approval and verification. Bank officers usually scan a signed copy of the document to be processed and manually inspect the signature. Manual inspection is very suspect to errors and misleading interpretations. Furthermore, it requires an additional workload, which causes either an increase in the cost of human resources for bank managerial or excessive working hours for current officers. Therefore, it has become essential to use intelligent software techniques for automated analysis of documents to perform signature-related tasks.

An important task in automated processing of scanned bank application forms is to find the position of the signature. Although the signature analysis has received a great attention of the researchers in the field of document analysis and recognition, their effort has been enriched in the direction of signature verification problem [8,16,18], where it is required to model the identity between two previously segmented signature images. There is too fewer attempt in the analysis of signature presence or position in documents.

The task of detecting signatures in scanned documents poses several challenges. First, these types of document images have usually very low resolution, which makes them difficult to enhance. Second, the background of each document is different and usually not known beforehand. Third, documents are subject to restricted processing time due to the urgency of applications. Finally, and maybe the most importantly, the documents often contain auxiliary lines and other handwritten characters that resemble or overlap with signatures.

In one of the earliest studies, Djeziri et al. [5] tackled the signature extraction problem by an intuitive approach that is supposed to mimic the human visual perception. They introduced the concept of filiformity as a criterion for the curvature characteristics of handwritten signatures. Though successful in clean bank cheques, this approach fails when other filiform objects are present in the document. Madasu et al. [12] tried to crop the image segment by estimating the area in which the signature lies using a sliding window. They then analyzed the local entropy derived from the pixel-based density of the region to decide its being signature or not. This approach disregarded the noise and therefore high-density regions are reported as signatures incorrectly. Madasu et al. [13] and Chalechale et al. [2] used geometric features including area, circularity, aspect ratio, size and position to analyze segmented regions. These features are compared using pairwise similarity metrics such as Manhattan distance. Jayavdean et al. [9] proposed another method based on variance analysis in a grid placed on the putative signature position in gray-level cheque images. Zhu et al. introduced the concept of multi-scale saliency feature to define signature characteristics [19] and also used it for signature verification [20]. A three-stage procedure was proposed by Mandal et al. [14] to extract signatures. First stage locates signature segment using word-level feature extraction. Second stage separates signature strokes that overlap with the printed text. Final stage uses skeleton analysis to classify real signature strokes. They used gradient-based features to feed a machine learning classifier. Ahmet et al. [1] used Speeded Up Robust Features (SURF) to classify segmented image blocks. Esteban et al. [6] approached the problem with the assumption that signature detection is normally followed by a verification stage and applied an evidence accumulation strategy in order to utilize a set of known signatures at hand. Although they achieved a high accuracy with this approach, their assumption is not

generally true since the bank application forms are often filled by new customers and this process is not followed by a verification but an archiving step.

In this study, we propose an automated framework to extract handwritten signatures from multi-page bank application documents assuming that the customer has no previous sign in the current database. The framework is discriminative in that a set of model parameters are learned through positive and negative signature samples. The learning model is built upon Support Vector Machines (SVMs) while the feature sets that feed SVM are selected using several local and global image feature representation schemes. The experiments on real data sets have shown that the framework can achieve a reasonably good accuracy to be used in real life applications. The study can also provide a comprehensive comparison analysis of the contributions of different image descriptors in signature extraction problem.

2 Materials and Methods

The general handwritten signature detection framework is shown in Figure 1. The process starts with the pre-processing stage to acquire and get a better view of input image. In the second stage, an image segmentation procedure is applied to obtain signature candidates. The third stage is where the candidate signatures are represented by a set of numerical features extracted from image content. The feature vectors are fed into machine learning classifiers in the final stage. The details of the proposed framework are given as follows.
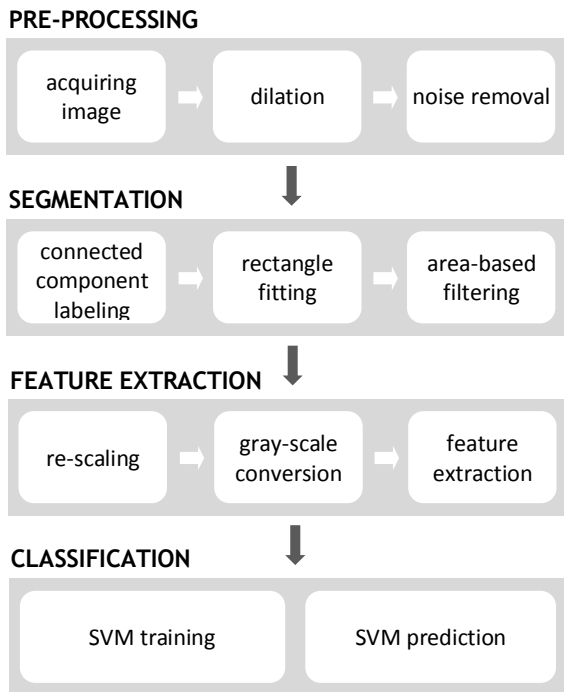
Figure 1: A brief outline of proposed framework for signature extraction

2.1 Pre-processing

Pre-processing stage involves acquiring the input image and extracting single page samples from multi-page input. To enhance the input image, we consider applying a simple dilation operator to make the lines more visible and a noise removal step by Median filtering to smooth the image. No other pre-processing step is applied to keep as much as possible information for segmentation phase.

2.2 Segmentation

Segmentation is a crucial step in signature detection. It is desired to catch all true positive (signature-containing) segments while allowing false negatives to some degree as they are expected to be removed in following steps. For an effective segmentation, we follow a two-scan connected component labelling approach described in He et al. [7]. This approach involves three processes: (1) assigning each pixel a provisional label by a 4-neigbor mask and finding equivalent labels, (2) recording equivalent labels and finding a representative label for each equivalent provisional label, (3) replacing each provisional label by its representative label. For a more efficient version, the image pixels and lines are processed two by two as opposed to conventional methods that perform these processes one by one. After connected component labelling, corresponding image regions are fit into a rectangle to record candidate signature segments. The segments involving less than 350 pixels are removed since the signature regions are often larger.

2.3 Feature Extraction

Since we follow a discriminative approach for signature detection, each segment should be vectorized to be fed into a machine learning classifier. This vectorization is made by selecting and extracting a set of content-based features to represent the segment as being a signature or not. Before feature extraction, a number of pre-processing steps are performed. First, the segments are re-scaled into 126x126 pixels. Then, the bitonal image is converted into a gray-scale image by applying 2x2 median filtering 5 times. We evaluate several feature representation schemes to distinguish signatures from other connected components.

Gradient-based features: First feature set is based on the local pixel representatives that are evaluated by gradient vectors. To extract this feature set, the re-scaled segment image is partitioned into 9x9 blocks. The arctangent (strength) of the gradient is quantized into 16 directions and the strength is accumulated with each of the quantized direction. This is implemented by a Robert's operator. The histograms of the values of 16 quantized directions are computed in each of 9x9 blocks. After a down-sampling from 9x9 to 5x5 by a Gaussian filter, the resulting feature vector has 400 dimensions.

HOG: Another popular gradient-based feature called HOG (Histogram of Oriented Gradients) [4] is also considered with its default parameters. Each HOG vector is composed of 8100 features, which includes 9-bin gradients on 15x15 blocks with 4 cells each.

SIFT: Third feature set is based on interest-point descriptors. SIFT (Scale Invariant Feature Transform) generates features for gray-level images called keypoints which is invariant to image scaling, rotation and partially to change in illumination [11]. The statistics of local gradient directions of image intensities are accumulated to give a summary description of the local structures in a neighbourhood around each keypoint. The feature set is highly distinctive if a sufficient number of keypoints is found.

LTP: LTP (Local Ternary Patterns) is a spatial method for modeling texture in an image. It was recently introduced by Suruliandi and Ramar [17] as an extension of Local Binary Patterns [15]. LBP is based on recognizing certain local binary texture patterns termed 'uniform'. The central pixel is compared with P pixels at the radius R of a circular neighborhood. The binary level comparisons computed along the boundary of the circular neighborhood are used to find a uniformity measure. It indeed refers to the number of spatial transitions. A pattern with uniformity level less than a predefined threshold is assigned with a label in the range 0 to P. Non uniform patterns are assigned to a single label, e.g. P+1. The LBP feature representation consists of a vector of the discrete occurrence histogram of these uniform patterns computed over the area under consideration. LTP allows operating on ternary pattern instead of binary pattern. It permits to detect the number of transition or discontinuities in the circular presentation of the patterns. The uniformity of a pattern is evaluated by such transitions that are found to follow a rhythmic pattern. The occurrence frequency of these patterns over the larger region then represents the LTP feature representation.

Global features: Last feature set is related to global description of segments. This set includes entropy, aspect ratio, and energy. Given a block of image i, and pixel density of P_i, the *entropy* is given by $E_i = -P_i log P_i$. The entropy is a measure of global information contained in that region. The *energy* is calculated by adding up the squares of all pixel intensities and dividing by the area of the segment. The *aspect ratio* is another global feature that is given by the ratio of width to height of the segment.

2.4 Classification

The classification of a segment is done using a popular machine learning method called Support Vector Machines (SVMs). SVM is a binary classifier that works based on the structural risk minimization principle. The inputs of an SVM in training phase are *n*-dimensional feature vectors which represent the predefined properties of the training samples. The SVM non-linearly maps its *n*-dimensional input space into a high dimensional feature space. In this high dimensional feature space a linear classifier is constructed. In the prediction phase, this linear classifier provides a discriminant score corresponding to the sample in question. In a binary classification task, a positive value of this score indicates that the test sample is belonging to that class. In our study, we use SVMs having a linear, polynomial and Gaussian kernels with their default input parameters in LIBSVM [3].

3 Results

3.1 Datasets

We used two datasets for performance evaluation. First dataset is an extension of a benchmark set, called Tobacco-800 [19]. It contains 755 segments, 353 of which are signatures and the others are not. It is available at www.baskent.edu.tr/~hogul/signds.rar. Second dataset contains real document images obtained from a currently operating local bank with a bilateral privacy agreement. In this dataset, there are 2670 multi-page documents where the total number of pages is 9943. In 4082 of the pages, there exists at least one handwritten signature. 5861 of the pages have no signature on them.

3.2 Experimental setup

The task in first dataset is to detect if given segment is a signature or not. For the second dataset, the task is to identify whether a given document contains a signature or not. Basically, we set out a rule that; if any of the page yields at least a signature segment as a classifier output, then this document is labelled as a signature-containing page. The actual position of the signature segment in the document is reported as the location of the signature. A 5-fold cross-validation is performed in first dataset for evaluation. To discern the practical ability of the framework in signature extraction, the system is trained using all segments, including positive and negative samples, in first data sets, and run through the documents in second independent dataset. Two datasets have no common documents. The performance evaluation is done by measuring the accuracy (the proportion of correctly predicted labels), the sensitivity (the proportion of positive samples which are correctly identified as such) and the specificity (the proportion of negative samples which are correctly identified as such). Here, a sample corresponds to an individual segment and a document page in the first and second datasets, respectively. The experiments are conducted for several classifier models alternating for feature representation schemes, noise removal application and SVM kernel used. For each model, only the results for the kernel that leads to the best accuracy are reported.

3.3 Empirical results

Table 1 shows the results for signature prediction task in the first dataset. The table demonstrates the results obtained with single feature representation schemes and some their combinations as well. It is obvious that SIFT feature set can not contribute so much in signature detection; all performance metrics with solely SIFT feature is too low in comparison to the others. HOG feature set provides the highest specificity with a terrible sensitivity and lower accuracy. The gradient-based and LTP feature sets can achieve the highest accuracy levels with a reasonable balance between specificity and sensitivity, whilst the LTP feature set performs slightly better. On the other hand, the performance achieved by solely use of LTP can not be enhanced with the addition of other features in an integrated feature set. Although

global feature set looks useless in its single use, it can improve the overall accuracy when it is combined with gradient-based features. This combination indeed attains the highest performance for all metrics. An interesting result is that a noise removal step is not useful even harmful for signature detection performance. This is probably because of the fact that connected lines inside the signatures, which are considered to be major indicators of being signature, are slightly removed by the noise reduction procedure. Knowing this fact, the noise removal is not applied in the models of integrated feature sets and the experiments for second dataset.

The experimental results for the task of detecting whether a whole document page contains a signature are given in Table 2. This experiment is performed on the second dataset containing complete document pages with a classifier model trained by all signature samples in the first data set. The results still suggest that the use of gradient-based feature sets with global features can serve the most reliable way of detecting signatures in scanned documents.

Manual inspection of individual results reveals some difficulties in signature detection problem. It evidently appears that there are two major reasons for false positives: (1) machine printed texts with highly connected fonts, and (2) handwritten initials, marks or other signs to check some choices on the application documents (Figure 2). Missing true positives is usually caused by the intervening signatures with other texts or figures of the document, especially when the signature is much smaller than the other intervened part or the faint fonts due to low scan quality or resolution (Figure 3).

Features	Noise removal applied	SVM Kernel*	Sensitivity	Specificity	Accuracy
Gradient	yes	linear	87,0	92,8	90,1
	no	linear	90,1	93,0	91,7
HOG	yes	linear	39,9	94,0	68,7
	no	linear	35,1	93,8	66,4
SIFT	yes	linear	70,5	76,9	73,9
	no	linear	77,9	75,9	76,8
LTP	yes	polynomial	93,8	90,0	91,8
	no	polynomial	92,9	91,0	91,9
Global Features	yes	polynomial	58,9	40,8	49,3
	no	polynomial	71,7	66,9	69,1
Gradient + Global	no	linear	94,9	95,0	95,0
LTP + Global	no	linear	92,9	93,0	92,9
Gradient + LTP+ Global	no	polynomial	94,1	91,8	92,8

Table 1: Results with first dataset for signature prediction task in image segments.

Features	Sensitivity	Specificity	Accuracy
Gradient	93,8	47,7	66,6
LTP	97,6	29,7	57,5
SIFT	98,5	16,3	51,2
Global Features	99,8	14,3	49,3
Gradient + Global	95,0	54,3	71,0
Gradient + Global + LTP	94,9	44,8	65,4
LTP + Global	97,8	29,1	57,3

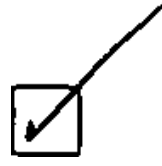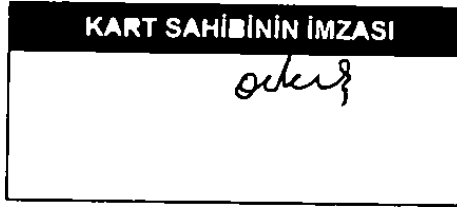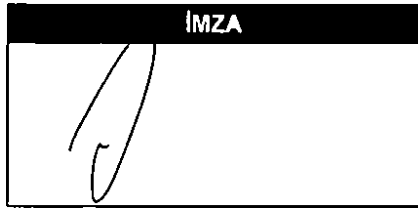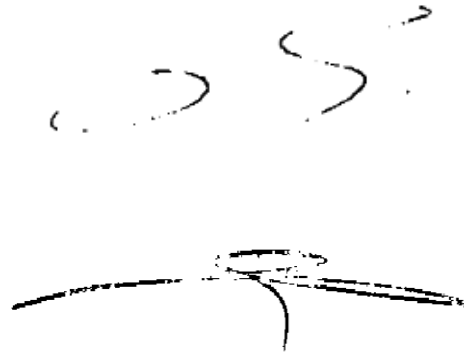**Table 2:** Results with second dataset for signature detection task in whole documents.

Figure 2: Most common false positives: examples for (a) machine printed texts, (b) handwritten initials or check marks.

Figure 3: Most common false negatives: examples for (a) intervened signatures, (b) faint fonts.

4 Conclusion

A framework is introduced for detecting handwritten signatures in scanned documents. The framework involves a robust and reliable segmentation stage. A number of descriptive image features are studied to discern their performances on distinguishing signature and non-signature images. Based on the empirical results, it is shown that gradient-based and LTP features are more useful in classifying signature segments in their single uses. In some cases, combining feature representation schemes can enhance the reliability of the predictions. In that sense, global features such as the aspect ratio, energy and entropy of the candidate segments can serve as lucrative complementary properties. The experimental results on real daily bank operation documents motivate the use of the framework in real life applications. The framework is quite extensible with the use of additional domain-specific local image features. Our feature work involves adding various rule-bases into different stages of the current framework to filter out improbably signature segments to obtain higher accuracy in predictions. We also anticipate that integrating the predicted document type from entire page content as a latent variable into feature sets will probably improve the detection performance.

Acknowledgement

References

[1] S. Ahmed, M.I. Malik, M. Liwicki, A. Dengel. Signature Segmentation from Document Images. *International Conference on Frontiers in Handwriting Recognition.* 2012.

[2] A. Chalechale, G. Naghdy, P. Premaratne, A. Mertins. Document Image Analysis and Verification Using Cursive Signature. *IEEE International Conference on Multimedia and Expo.* 2004.

[3] C.C. Chang and C.J. Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1--27:27, 2011.

[4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 2005.

[5] S. Djeziri, F. Nouboud, R. Plamondon. Extraction of signatures from check background based on a filiformity criterion. *IEEE Trans. Image Process.* 7(10), 1425–1438, 1998.

[6] J.L. Esteban, J.F. Vélez, Á. Sánchez. Off-line handwritten signature detection by analysis of evidence accumulation. *IJDAR*, 15:359–368. 2012.

[7] L. He, Y. Chao, K. Suzuki. A New Two-Scan Algorithm for Labeling Connected Components in Binary Images. *Proceedings of the World Congress on Engineering.* 2012.

[8] D. Impedovo and G. Pirlo. Automatic Signature Verification: The State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews*, 38:5. 2008.

[9] R. Jayadevan et al. Variance based extraction and hidden Markov model based verification of signatures present on bank cheques. *International Conference on Computational Intelligence and Multimedia Applications.* 2007.

[10] J. Li, N.M. Allinson. A comprehensive review of current local features for computer vision. *Neurocomputing*, 7:1771– 1787. 2008.

[11] D.G. Lowe. (1999). Object recognition from local scale-invariant features. *7th International Conference on Computer Vision.* 1999.

[12] V.K. Madasu, B.C. Lovell. Automatic segmentation and recognition of bank cheque fields. *Digit. Image Comput. Tech. Appl.*, 80(1): 33–40. 2005.

[13] V.K. Madasu, M.H.M. Yusof, M. Hanmandlu, K. Kubik. Automatic extraction of signatures from bank cheques and other documents, *DICTA'03*. 2003.

[14] R. Mandal, P.P. Roy, U.Pal. Signature Segmentation from Machine Printed Documents using Conditional Random Field. *International Conference on Document Analysis and Recognition.* 2011.

[15] T. Ojala, M. Pietikäinen, D. Harwood. A Comparative Study of Texture Measures with Classification Based on Feature Distributions. *Pattern Recognition*, 29:51-59. 1996.

[16] R. Plamondon, S.N. Srihari. On-line and off-line handwriting recognition: a comprehensive survey. *IEEE Trans. Pattern Anal.Mach. Intell.*, 22(1), 63–84. 2000.

[17] A. Suruliandi, K.Ramar. Local Texture Patterns –A Univariate Texture Model for Classification of Images. *16th International Conference on Advanced Computing and Communications.* 2008.

[18] W.K.H. Weiping, Y. Xiufen. A survey of off-line signature verification. *International Conference on Intellingence Mechatronics and Automation.* 2004.

[19] G. Zhu, Y. Zheng, D. Doermann, S. Jeager. Multi-scale Structural Saliency for Signature Detection. *IEEE Conference on Computer Vision and Pattern Recognition.* 2007.

[20] G. Zhu, Y. Zheng, D. Doermann, S. Jaeger. Signature detection and matching for document image retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31, 2015–2031. 2009.