CURSIVE SIGNATURE EXTRACTION AND VERIFICATION

Abdolah Chalechale[1], Golshah Naghdy[1], PrashanPremaratne[1], and Alfred Mertins[2]
[1]School of Electrical, Computer, and Telecommunications Engineering
University of Wollongong, Australia
{ac82, golshah, prashan}@uow.edu.au
[2]Signal Processing Group, Institute of Physics
University of Oldenburg, Germany
alfred.mertins@uni-oldenburg.de

ABSTRACT

This paper presents a new approach for document image decomposition and verification based on connected component analysis and geometric properties of labeled regions. The database contains document images with Persian/Arabic text combined with English text, headlines, ruling lines, trademarks and cursive signatures. In particular, Persian/Arabic signature extraction is investigated using special characteristics of the signature that is fairly different from English signatures. A set of efficient, invariant and compact features is extracted for verification purposes using spatial partitioning of the signature region. Comparative results show the robustness of the proposed method.

Keywords: Document image processing; Cursive signature; Feature extraction; Invariant features; Document verification; Image retrieval.

1 INTRODUCTION

Automatic document analysis is a fundamental issue in many IT applications including Optical Character Recognition (OCR), form and bank check reading and document image storage and retrieval. Document image understanding covers a variety of documents such as facsimiles [1], bank checks [2], business letters [3], forms [4], postal mail parcels [5] and technical articles [6,7]. It has been an interesting research area for a long time [8].

Classification of a document image into text and graphics is investigated in [9]. The approach is based on the different textural properties of graphics and non-graphics components in the document. Li and Gray [10] developed a method for segmenting document images into four classes: background, photograph, text and graph. The distribution patterns of wavelet coefficients in high frequency bands are employed to extract features for the proposed classification.

Background thinning is used for page segmentation in [11]. The approach is effective but sensitive to content and computationally expensive. Wang and Chi [12] improved the approach by speeding up the process through applying a hierarchical content classification and script determination. They introduced a neural network based classifier to classify a sub-image into text or picture. They also developed an algorithm that can determine Chinese and Western scripts in the text region using a three-layer feed forward network.

One of the most important content components in official/business letters and bank checks is the signature. The effective extraction and verification of the signature play important roles in automatic processing of such documents. Moreover, paperless organizations are growing fast and the interaction with other paper-based organizations and individuals needs efficient methods for converting a paper-based document to an electronic version. Automated document and check signature processing techniques consist of two main modules: a low-level processing for signature extraction and a high-level processing for signature verification.

English signature analysis, verification and recognition have been studied extensively. They could be divided into two broad areas: on-line and off-line. A comparison between wavelet-based and function-based on-line signature verification has been reported by Da Silva and De Freitas [13] whilst Justino *et al*. [14] have focused on off-line signature classification using Hidden Markov Models. Persian/Arabic signatures however, have a different characteristic. They usually are cursive sketches, which are independent of the person's name while English signatures are often reshaped handwritten names.

Normalized Zernike moment invariants (NZMI) is adopted by MPEG-7 standard as region-based shape descriptor [15]. NZMI and projection methods are studied for off-line Chinese signature verification in [16]. Different orders of moment invariants are tested and the results show the better performance of NZMI method compared with the projection method.

A signature recognition approach based on line segment distribution (LSD) has been proposed recently [17]. The approach utilizes the derivative of the chain code to extract straight-line segments, which represent the signature. The affine invariant properties of the line segments including length and distance from the center of mass (COM) are used for the recognition task. The method is applied effectively for Persian signature recognition and outperforms chain code histogram and invariant moments approaches.

This paper proposes a new approach for signature-based document retrieval. It utilizes connected components analysis, labeling and geometric properties. The document may contain Persian/Arabic cursive signature, trademarks and ruling lines anywhere in the document in addition to headlines and normal text. The signature is detected first, features extracted, and verified by a novel spatial partitioning method. The partitioning method is based on accumulating pixels in the sectors defined adjacently in the signature region. We use the magnitude of the Fourier transform in order to achieve rotation invariance.

The basic components of the new algorithm are discussed in the next section. Section 3 exhibits comparative results and Section 4 concludes the paper.

2 SIGNATURE EXTRACTION AND VERIFICATION

We consider official and business letters containing Persian/Arabic text combined with English characters, headlines, ruling lines, trademarks (logos) and cursive signatures (see Figure 1 for an example). The main objective is to extract the signature region from the original image and transform it into a new compact feature vector that supports measuring the similarity among signatures for validation purposes.

Figure 1: A document image example

Let *I* be the binary image of the original document. The connected components labeling operation [18] that performs the unit change from pixel to region is employed to label *I* as follows: all pixels that have value binary 1 and are connected to each other in an 8-connectivity neighborhood are given the same identification label. The label is a unique index of the region to which the pixels belong. For efficient implementation of the labeling algorithm we first run-length encode the input image *I* and then assign preliminary labels while scanning the runs and recording label equivalences in a local table. The equivalence classes are resolved next and the runs are relabeled based on the resolved equivalence classes.

Based on a combination of heuristic geometric properties including area, circularity, eccentricity, size and position of the labeled regions in *I* the signature region is determined. The bounding box of the signature region is then normalized to $J*J$ pixels using bilinear interpolation. The signature enclosed within the normalized region *B* might overlap with other materials such as text or ruling lines. The unwanted extra parts in the signature region are partially eliminated by applying the algorithm given in [19]. The adverse effect of remaining unwanted regions is overcome by applying angular-radial partitioning (ARP) process. The ARP algorithm will also ensure scaling and rotation invariance properties.

The ARP algorithm uses the surrounding circle of *B* for partitioning it to $M*N$ sectors, where *M* is the number of radial partitions and *N* is the number of angular partitions. The angle between adjacent angular partitions is $\theta=2\pi/N$ and the radius of successive concentric circles is $\rho=R/M$ where *R* is the radius of the surrounding circle of the image *B* (see Figure 2).

The number of edge points in each sector of *B* is chosen to represent the sector feature. The scale invariant signature feature is then $\{f(k,i)\}$ where

$$f(k,i) = \sum_{\rho=\frac{kR}{M}}^{\frac{(k+1)R}{M}} \sum_{\theta=\frac{i2\pi}{N}}^{\frac{(i+1)2\pi}{N}} B(\rho,\theta) \tag{1}$$

for $k = 0,1,2\ldots M\text{-}1$ and $i = 0,1,2\ldots N\text{-}1$. The feature extracted above will be circularly shifted when the signature image B is rotated by $\tau = l2\pi/N$ radians ($l = 0,1,2\ldots$). To show this, let B_τ denote the image B after rotation by τ radians in counterclockwise direction:

$$B_\tau(\rho,\theta) = B(\rho,\theta-\tau). \tag{2}$$

Then

$$f_\tau(k,i) = \sum_{\rho=\frac{kR}{M}}^{\frac{(k+1)R}{M}} \sum_{\theta=\frac{i2\pi}{N}}^{\frac{(i+1)2\pi}{N}} B_\tau(\rho,\theta) \tag{3}$$

are the signature feature elements for B_τ for the same k and i. We can express f_τ as

$$\begin{aligned} f_\tau(k,i) &= \sum_{\rho=\frac{kR}{M}}^{\frac{(k+1)R}{M}} \sum_{\theta=\frac{i2\pi}{N}}^{\frac{(i+1)2\pi}{N}} B(\rho,\theta-\tau) \\ &= \sum_{\rho=\frac{kR}{M}}^{\frac{(k+1)R}{M}} \sum_{\theta=\frac{(i-l)2\pi}{N}}^{\frac{(i-l+1)2\pi}{N}} B(\rho,\theta) \\ &= f(k,i-l) \end{aligned} \tag{4}$$

where $i\text{-}l$ is a modulo M subtraction. It means that there is a circular shift (for individual k's) in the signature feature $\{f_\tau(k,i)\}$ compared to the signature feature $\{f(k,i)\}$ representing I_τ and I respectively.

Using 1-D discrete Fourier transform of $f(k,i)$ and $f_\tau(k,i)$ for each k we obtain

$$F(k,u) = \frac{1}{N}\sum_{i=0}^{N-1} f(k,i)e^{-j2\pi ui/N} \tag{5}$$

and

$$\begin{aligned} F_\tau(k,u) &= \frac{1}{N}\sum_{i=0}^{N-1} f_\tau(k,i)e^{-j2\pi ui/N} \\ &= \frac{1}{N}\sum_{i=0}^{N-1} f(k,i-l)e^{-j2\pi ui/N} \\ &= \frac{1}{N}\sum_{i=-l}^{N-1-l} f(k,i)e^{-j2\pi u(i+l)/N} \\ &= e^{-j2\pi ul/N}F(k,u). \end{aligned} \tag{6}$$

Because of the property $|F(k,u)| = |F_\tau(k,u)|$, the scale and rotation invariant signature features are chosen as $\{|F(k,u)|\}$ for $k = 0,1,2\ldots M\text{-}1$ and $u = 0,1,2\ldots N\text{-}1$.

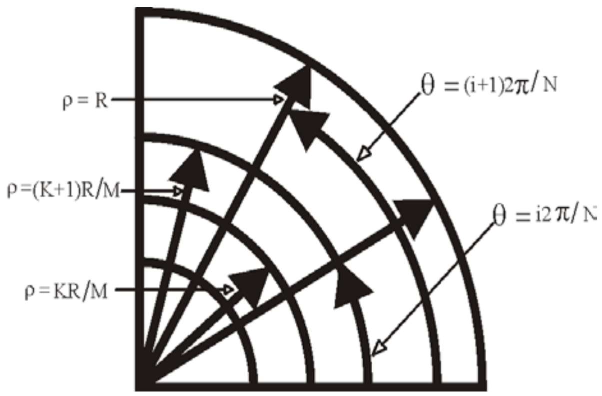

Figure 2: Angular Radial Partitioning of a region to N angular and M radial sectors where $k = 0,1,2\ldots M\text{-}1$ and $i = 0,1,2\ldots N\text{-}1$

Experimental results (Section 3) confirm the robustness and efficacy of the extracted features.

2.1 SIMILARITY MEASUREMENT AND EVALUATION METHOD

The similarity between the signature features is measured by the L_1 (Manhattan) distance between the two feature vectors. That is

$$d(S_i,S_j) = \sum_{k,u} |F_{s_i}(k,u) - F_{s_j}(k,u)| \tag{7}$$

where S_i and S_j are two different signatures.

Average Normalized Modified Retrieval Rank (ANMRR) is used as retrieval performance measure. The ANMRR considers not only the recall and precision information but also the rank information among the retrieved images. It is defined in MPEG-7 standard as follows:

$$AVR(q) = \sum_{g=1}^{NG(q)} \frac{Rank(g)}{NG(q)} \tag{8}$$

$$MRR(q) = AVR(q) - 0.5 - \frac{NG(q)}{2} \tag{9}$$

$$NMRR(q) = \frac{MRR(q)}{K + 0.5 - 0.5 * NG(q)} \tag{10}$$

$$ANMRR = \frac{1}{Q}\sum_{q=1}^{Q} NMRR(q) \tag{11}$$

where *NG(q)* is the number of ground truth images for a query *q*. *K* = min*(4*NG(q),2*GTM)* where *GTM* is max *{NG(q)}* for all *q*'s of a data set. *Rank(g)* is the rank of the found ground truth images, counting the rank of the first retrieved image as one. A *Rank* of *K+1* is assigned to each of the ground truth images, which are not in the first *K* retrievals. For example, suppose a given query q_i has 10 similar images in an image database (*NG* = 10). If an algorithm finds 6 of them in the top 20 retrievals (*K* = 20) in the ranks of 1,5,8,13,14 and 18, then the $AVR(q_i)$ = 14.3, $MRR(q_i)$ = 8.8 and $NMRR(q_i)$ = 0.5677.

Note that NMRR and its average (ANMRR) will always be in the range of [0,1]. Based on the definition of ANMRR, the smaller the ANMRR, the better the retrieval performance.

3 EXPERIMENTAL RESULTS

We used a document image database containing 350 documents signed by 70 different persons who have Persian or Arabic cursive signatures. The content of the images include a variety of mixed text of Arabic, Persian and English alphanumeric with different fonts and sizes, a company logo, some horizontal and vertical lines and a cursive signature. All documents are processed to generate a feature vector for the signature within the image. The processing involves: a) Signature region detection and b) Generation of a compact feature vector for the signature region in the document and also for the query presented by the user as a handwritten signature.

For the detection phase, we applied the connected component analysis with a set of heuristic conditions based on cursive signature geometry (Section 2). The signature region was found correctly in 346 cases (98.86%) and the signature extracted completely in 342 cases (97.71%). This is due to the fact that some cursive signatures have several disjoint parts while the algorithm focuses on neighboring connected parts.

For the generation of feature vector, we examined three different approaches. The first approach is based on line segment distribution (LSD) method, which uses an 80-entry feature vector for the signature description [17]. The second approach is based on normalized Zernike moment invariants (NZMI) [15,16]. A 36-entry feature vector is obtained using the algorithm given in [15]. We applied J=100, M=5 and N=12 in proposed ARP method resulting in a 60-entry feature vector $\{|F(k,u)|\}$ as explained in Section 2. The vectors are used to describe the signature region. Automatic links were established between feature vector of the extracted signature and the corresponding document image using file names. This will enable the retrieval to be conducted automatically.

We used ANMRR as retrieval performance measure. In our experiments the $NG(q)$ = 5 for all q's, K = 10 and Q = 70.

Table 1 exhibits ANMRR for normalized Zernike moment invariants (NZMI), line segment distribution (LSD), and the proposed angular-radial partitioning (ARP) methods. The retrieval performance of the proposed ARP approach is the best (0.2612). LSD method's performance is better than NZMI method (0.3018<0.3305). In time-based comparison, NZMI approach is found to be the fastest due to minimum vector-length (35 entries). On the other hand, it has the longest extraction time resulting from higher computational cost arising in handling complex numbers.

The LSD method is more sensitive to extra and eroded parts in the signature than the ARP method. The latter is more robust to such effects since it looks at larger areas (sectors) in the signature region and compares the spatial distribution of the pixels in the region of interest.

Table 1: Retrieval performance of different methods

Method	NZMI	LSD	ARP
ANMRR	0.3305	0.3018	0.2612

4 CONCLUDING REMARKS

A new approach for signature-based decomposition and retrieval of document images has been introduced. Cursive Persian/Arabic signature recognition and verification are investigated as a case study. Connected component analysis and labeling along with geometric properties are used to recognize the signature region. A novel angular-radial partitioning scheme is also introduced for description of spatial distribution of pixels in the region of interest. The approach utilizes magnitude of Fourier transform resulting in rotation invariant characteristics. The approach is also scale and translation invariant.

Experimental findings confirm that the proposed approach generates better retrieval performance compared with two other approaches known from the literature. ANMRR was employed as the performance indicator.

The region of interest detection and feature extraction methods introduced in this paper can be utilized for other applications including sketch-based image retrieval.

Acknowledgments. The authors would like to appreciate all participants who gave permission to use their signatures in this study. The first author is financially supported by the Ministry of Science, Research and Technology of I.R. Iran.

REFRENCES

[1] D. G. Elliman, "Document recognition for facsimile transmission," *Proc. IEE Colloquium on Document Image Processing and Multimedia*, 1999, pp. 3/1–3/5.

[2] S. Djeziri, F. Nouboud, and R. Plamondon, "Extraction of signatures from check background based on a filiformity riterion," *IEEE Trans. Image Processing*, vol. 7, no. 10, pp. 1425–1438, 1998.

[3] A. Dengel, "Initial learning of document structure," *Proc. Second Int. Conf. Document analysis and recognition*, 1993, pp. 86–90.

[4] B. Yu and A. K. Jain, "A generic system for form dropout," *IEEE Trans. Patt. Anal. and Mach. Intell.*, vol. 18, no. 11, pp. 1127–1134, 1996.

[5] B. Yu, A. K. Jain, and M. Mohiuddin, "Address block location on complex mail pieces," *Proc. Fourth Int. Conf. Document Analysis and Recognition*, 1997, vol. 2, pp. 897–901.

[6] A. K. Jain and B. Yu, "Document representation and its application to page decomposition," *IEEE Trans. Patt. Anal. and Mach. Intell.*, vol. 20, no. 3, pp. 294–308, 1998.

[7] G. Nagy, S. Seth, and M. Viswanathan, "A prototype document image analysis system for technical journals," *IEEE Computer*, vol. 25, no. 7, pp. 10–22, 1992.

[8] G. Nagy, "Twenty years od document image analysis in PAMI," *IEEE Trans. Patt. Anal. and Mach. Intell.*, vol. 22, no. 1, pp. 38–62, 2000.

[9] M. Acharyya and M. K. Kundu, "document image segmentation using wavelet scale-space features," *IEEE Trans. Circ. and Syst. for Video Tech.*, vol. 12, no. 12, pp. 1117–1127, 2002.

[10] J. Li R. M. Gray, "Context-based multiscale classification of document images using wavelet coefficient distributions," *IEEE Trans. Image Processing*, vol. 9, no. 9, pp. 1604–1616, 2000.

[11] K. Kise, O. Yanagida, and S. Takamatsu, "Page segmentation based on thinning of background," *Proc. IEEE 13th Int. Conf. Patt. Recog.*, 1996, vol. 3, pp. 788–792.

[12] Q.Wang, Z. Chi, and R. Zhao, "Hierarchical content classification and script determination for automatic document image processing," *Proc. IEEE 16th Int. Conf. Patt. Recog.*, 2002, vol. 3, pp. 77–80.

[13] A. V. Da Silva and D. S. De Freitas, "Wavelet-based compared to function-based on-line signature verification," *Proc. IEEE Int. Conf. Comput. Graphics and Image Processing*, 2002, pp. 218–225.

[14] E. J. R. Justino, F. Bortolozzi, and R. Sabourin, "The interpersonal and intrapersonal variability influences on offline signature verification using hmm," *Proc. IEEE Int. Conf. Comput. Graphics and Image Processing*, 2002, pp. 197–202.

[15] ISO/IEC JTC1/SC29/WG11/N3321, "MPEG-7 visual part of experimentation model version 5," Nordwijkerhout, March 2000.

[16] H. Lin and H.-Z. Li, "Chinese signature verification with moment invariants," *IEEE Int. Conf. Sys., Man, and Cyber.*, 1996, vol. 4, pp. 2963–2968.

[17] A. Chalechale and A. Mertins, "Persian signature recognition using line segment distribution," *IEEE TENCON'03*, 2003, vol. 1, pp. 11–15.

[18] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, Addison-Wesley, 1992.

[19] J. D. Hobby, "Page decomposition and signature finding via shape classification and geometric layout," *Proc. IEEE 5th Int. Conf. Document Analysis and Recognition*, 1999, pp. 555–558.

[20] ISO/IEC JTC1/SC29/WG11-MPEG2000/M5984, "Core experiments on MPEG-7 edge histogram descriptors," Geneva, May 2000.