SIGNATURE-BASED DOCUMENT RETRIEVAL

Abdolah Chalechale, Golshah Naghdy

University of Wollongomg
Wollongong, Australia

Alfred Mertins

University of Oldenburg
Oldenburg, Germany

ABSTRACT

This paper presents a new approach for document image decomposition and retrieval based on connected component analysis and geometric properties of the labelled regions. The database contains document images with Arabic/Persian text combined with English text, headlines, ruling lines, trade mark and signature. In particular, Arabic/Persian 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 validation purposes using angular-radial partitioning of the signature region. Experimental results show the robustness of the proposed method.

1. INTRODUCTION

Automatic document analysis is a fundamental issue in many applications including Optical Character Recognition (OCR), form and check reading and document image storage and retrieval. Document image understanding covers a variety of documents such as facsimiles [1], checks [2], business letters [3], forms [4], postal mail parcels [5] and technical articles [6, 7] and it has been studied for a long period [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 parts in the document. J. Li and R. M. 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 [11]. The approach is effective but sensitive to content and computationally expensive. Q. Wang and Z. Chi [12] improved the approach for speeding up the process by 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 checks is the signature. The effective extraction and verification of the signature play an important role 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 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. Arabian/Persian 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.

A signature recognition approach based on line segment distribution has been proposed recently [15]. The approach utilizes the derivative of the chain code to extract straight line segments which are representing the signature. The affine invariant properties of the line segments including length and distance from the center of mass 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 document retrieval based on signature. It utilizes connected components analysis, labelling and geometric properties. The document may contain Arabian/Persian signature, trade mark and ruling lines anywhere in the document in addition to headlines and normal text. The signature is then feature extracted and verified by a novel angular-radial partitioning (ARP) method. The ARP method is based on accumulating of the signature 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 proposed algorithm are discussed in the next section. Section 3 exhibits comparative results and Section 4 gives concluding remarks.

2. SIGNATURE EXTRACTION AND VERIFICATION

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

Logo Example

This is a test document image

در حالی که ورود پاسداران صلح و وسائل نقلیه نظامی به پایتخت لیبریا
ادامه دارد، شورشیان و نیروهای دولتی این کشور از خطوط مقدم
جبهه در مونروویا گذشته و یکدیگر را در آغوش گرفته اند.

براساس تخمین ها حدود 250 هزار نفر از ساکنان شهر با شرایط
دشواری روبرو هستند و فورا به آذوقه، آب و تدارکات پزشکی نیاز
دارند.

تامو امبکی، رئیس جمهور آفریقای جنوبی گفته است که چارلز تیلور،

قرار است نیروهای سازمان ملل در روز اول اکتبر جایگزین سه هزار
نیروی نیجریه ای شوند.

لیبریا از سال 1989 زمانی که چارلز تیلور دست به شورش مسلحانه
زد اسیر درگیری های کهگلی بوده است.

هم اکنون دو کشتی آمریکایی در سواحل لیبریا مستقر هستند و کشتی
دیگری در راه است. این کشتی ها در مجموع سه هزار نیروی
آمریکایی را حمل می کنند.

یک سخنگوی جرج بوش، رئیس جمهور آمریکا گفت که استقرار اولیه
نیروهای آفریقای غربی چند روز وقت می برد و آمریکا فعالانه به این
استقرار کمک خواهد کرد.

اما این سخنگو بار دیگر تاکید کرد تا زمانی که چارلز تیلور قدرت را
ترک نکند نیروهای آمریکایی وارد این کشور نخواهند شد.

Cursive Signature

Figure 1: A document image example.

Let I be the binary image of the original document. The connected components labelling operation [16] 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 identifying label. The label is a unique index of the region to which the pixels belong. For efficient implementation of the labelling algorithm we first run-length encode the input image I and then assign preliminary labels with scanning the runs and recording label equivalences in a local table. The equivalence classes are resolved next and the runs are relabelled based on the resolved equivalence classes.

Based on the geometric properties including area, contour shape, size and position of the labelled regions in the image I the signature region is determined. The bounding box of the signature region is then normalized to $J \times J$ pixels using nearest neighbor interpolation. The signature enclosed within the normalized region B might overlap other materials such as text. The unwanted extra parts in the signature region are partially eliminated by applying the algorithm given in [17]. 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 $M \times N$ 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 Fig. 2).
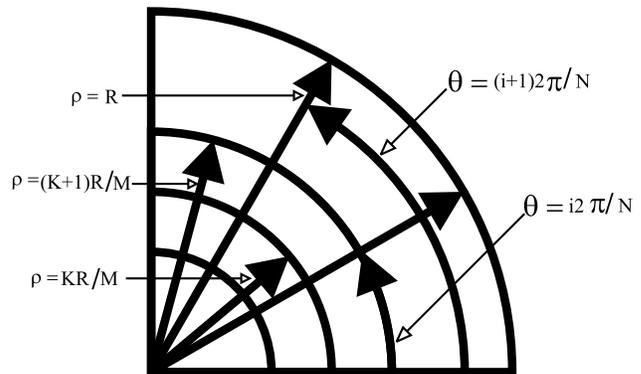
Figure 2: Angular Radial Partitioning of a region to N angular and M radial sectors where $k = 0, 1, 2 \ldots M - 1$ and $i = 0, 1, 2 \ldots N - 1$.

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

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

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

$$B_\tau(\rho,\theta) = B(\rho, \theta - \tau).$$

Then,

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

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

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

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

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

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 $k = 0,1,2\ldots M-1$ and $u = 0,1,2\ldots N-1$.

The similarity between the signature features is measured by the ℓ_1 (Manhattan) distance between the two feature vectors. Experimental results (Section 3) confirm the robustness and efficiency of the method.

3. EXPERIMENTAL RESULTS

Experimental results are presented in this section. We used a document image database containing 310 documents signed by 62 different persons. The content of the images include a variety of mixed text of Arabic/Persian and English alphanumerics with different fonts and sizes, a company logo, some horizontal and vertical lines and an Arabic/Persian 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 Arabic/Persian cursive signature geometry (Section 2). The signature region was found correctly in 303 cases (98.71%) and the signature extracted completely in 298 cases (96.13%). 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 phase, we examined two different approaches. The first approach is based on line segment distribution method that uses an 80-entry feature vector for the signature description [15]. The second approach is based on the proposed angular-radial partitioning (ARP) explained in Section 2. We applied $J = 100$, $M = 5$ and $N = 12$ in ARP method resulting in a 60-entry feature vector $\{\|F(k,u)\|\}$. The vector is 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.

Average Normalized Retrieval Rate (ANMRR) was used for measuring the retrieval performance. 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 [18] as:

$$AVR(q) = \sum_{\kappa=1}^{NG(q)} \frac{Rank(\kappa)}{NG(q)}$$

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

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

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

$NG(q)$ is the number of ground truth images for a query signature q and $Rank(\kappa)$ is the rank of the retrieved signature image in the ground truth. $K = \min(4 * NG(q), 2 * GTM)$, where GTM is max $\{NG(q)\}$ for all q's of a data set. 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. In our experiments the $NG(q) = 5$ for all q's, $K = 10$ and $Q = 62$.

The ANMRR for the line segment distribution approach is 0.32 while the ANMRR for the proposed angular-radial approach is 0.27. It indicates better performance for the new proposed approach in the retrieval of Arabic/Persian

cursive signatures in comparison with the line segment distribution method. The reason is that the line segment distribution method is more sensitive to extra and eroded parts in the signature than the angular-radial partitioning. 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 interested region.

4. CONCLUSION

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

Experimental results show the supremacy of the retrieval performance of the proposed approach over the line segment distribution method using ANMRR measure.

The discipline of the detection and the feature extraction phases can be generalized for other applications including sketch-based image retrieval.

5. REFERENCES

[1] D. G. Elliman, "Document recognition for facsimile transmission," in *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 criterion," *IEEE Trans. Image Processing*, vol. 7, no. 10, pp. 1425–1438, 1998.

[3] A. Dengel, "Initial learning of document structure," in *proc. second int. Conf. document anal. and recog.*, 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," in *proc. Fourth Int. Conf. Document Anal. and Recog.*, 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," in *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," in *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," in *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 off- line signature verification using hmm," in *proc. IEEE Int. conf. Comput. Graphics and Image Processing*, 2002, pp. 197–202.

[15] A. Chalechale and A. Mertins, "Persian signature recognition using line segment distribution," in *IEEE TENCON'03 Conf. convergent technologies for Asia-Pacific*, Bangalore, India, 2003.

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

[17] J. D. Hobby, "Page decomposition and signature finding via shape classification and geometric layout," in *proc. IEEE 5th Int. Conf. Document Anal. and Recog.*, 1999, pp. 555–558.

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