A New Method for Writer Identification of Handwritten Farsi Documents

F. Shahabi and M. Rahmati
Department of Computer Engineering & IT, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran.
fa_shahabi@yahoo.com, rahmati@aut.ac.i

Abstract

Most studies about writer identification are based on English documents and to our knowledge no research has been reported on Farsi or Arabic documents. In this paper, we have proposed a new method for off-line writer identification which is based on Farsi handwriting and text-independent. Based on the idea that has been presented in the previous studies, here we assume handwriting as texture image and a set of features which are based on multi-channel Gabor filters are extracted from preprocessed image of documents. Substantially, the property of proposed method is using of the bank of Gabor filters which is appropriate for structure of Farsi handwritten texts and vision system. Also, a new feature extraction method is proposed which is based on Gabor-energy and moments. For the first, we survey different methods for feature extraction from output of Gabor filters. These methods with co-occurrence matrix and Said method are implemented and experimental results on handwriting of 40 peoples demonstrate that the proposed method achieves better performance on Farsi handwritten documents.

1. Introduction

Handwriting as a behavioral biometric is easy to obtain and studies have shown that different people have different handwritings [1]. Therefore, writer identification has been studied in a wide variety of applications recently, such as security, financial activity, forensic and used as access control. Among of all, analysis of handwritten documents has great bearing on the criminal justice systems. Writer identification is the task of determining the writer of a document among different ones. Writer identification methods can be categorized into two types: text-dependent methods and text-independent methods. In text-dependent methods, a writer has to write the same fixed text to perform identification but in text-independent methods any text may be used to establish the identity of writer. These methods can be performed on-line, where dynamic information about the writing is available, or off-line, where only a scanned image of the writing is available.

Recently, different approaches for writer identification have been proposed. A scientific validation of individuality of handwriting is performed by Srihari et al. [1]. In this study handwriting samples of 1500 individuals, representative of the U.S. population with respect to gender, age, ethnic groups, etc., were obtained. The writer can be identified based on Macro features and Micro features that are extracted from handwritten documents. Said et al. [2] proposed a global approach based on multi-channel Gabor filtering, where each writer's handwriting is regarded as a different texture. Bensefia et al. [3] used local features based on graphemes extracted from segmentation of cursive handwriting. Then writer identification is performed by a textual based information retrieval model. Schomaker et al. [4] presented a new approach, using connected-component contours codebook and its probability-density function. Also combining connected-component contours with an independent edge-based orientation and curvature PDF yields very high correct identification rates. Schlapbach et al. [5] propose a HMM based approach for writer identification and verification. Bulacu [6] evaluated the performance of edge-based directional probability distributions as features in comparison to a number of non-angular features. Marti et al. [7] extracted a set of features from handwritten lines of text. The features extracted correspond to visible characteristics of the writing, for example, width, slant and height of the three main writing zones. In Zois et al. [8] a new feature vector is employed by means of morphologically processing the horizontal profiles of the words.

Because of the Lack of the standard database for writer identification, the comparison of the previous

978-0-7695-3725-2/09 $25.00 © 2009 IEEE
DOI 10.1109/ICDAR.2009.290

studies is not possible. Since our purpose is to introduce the automatic method and no limitation for handwriting is considered methods which need no segmentation or connected component analysis is regarded. Most previous studies are based on English documents with the assumption that the written text is fixed (text-dependent methods) and no research has been reported on Farsi or Arabic documents. In this paper, we have proposed a new method for off-line writer identification based on Farsi handwriting, which is text-independent. Based on the idea that has been presented in Said et al. [2], we assume handwriting as texture image and a set of new features are extracted from preprocessed image of documents.

In the following, in section 2 a description of our proposed method is given. Section 3 presents the experimental results. Finally in section 4, we draw conclusions from this work.

2. Proposed method

The writer identification method presented in this paper is based on the idea that has been introduced by Said [2]. We assume handwritten text as texture and extract features from document by means of texture analysis. We describe different stages of our proposed method in the following.

2.1. Preprocessing

Prior to the texture analysis stage, handwriting documents need to be preprocessed. For this purpose, documents are normalized with respect to different word spacing, line spacing, etc. The preprocessing can be done in the following steps:

(a) Projection profile has been widely used in line and word detection [9]. We use a modified version of the same algorithm extended to gray-level images [10]. First, the horizontal projection profile is computed and then smoothed with a low pass Gaussian filter (Fig. 1). In smoothed projection profile, the peaks correspond to the space between lines and the valleys correspond to the text lines. The peaks can be computed by setting the derivative of the projection profile to zero. The smoothing and the derivative operation can be combined into one step by convolving the projection profile with a Gaussian derivative as follows:

$$d / dy * G(y;\delta) * P(y) = \frac{dg(y;\delta)}{dy} * P(y) \qquad (1)$$

(b) Next, each text line that is located in previous step is binarized and its vertical projection profile is computed. In the vertical projection profile, the valleys correspond to the spacing between characters or words. Then the spacing greater than a threshold, is reduced to a predefined value. Also, blank spaces in each line are filled up by means of text padding, so that the width of text lines gives to a predefined size.

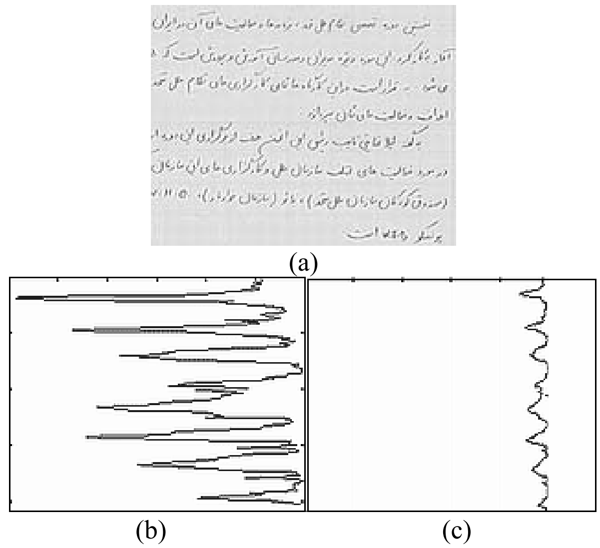

Figure 1: Extraction of lines from handwritten text. (a) image of document (b) horizontal projection profile (c) smoothed projection profile

(c) If there exist blank spaces in the bottom of image from the previous steps, then padding is applied; copying the text lines from top of the image, in order to achieve a image of text with a predefined size.

2.2. Feature extraction

In order to achieve a robust method for writer identification, we have to define features that reflect the large variability between handwriting. The preprocessed image obtained earlier, is divided into 6 non-overlapping blocks and we can employ any texture analysis technique for feature extraction from these blocks. Among different ones, the Gabor filtering has been shown to be useful for similar application [2, 11]. Therefore, we use multi-channel Gabor filtering with moment-based features. For comparison, we also investigated different feature extraction methods from output of Gabor filters, Said method and co-occurrence matrix features. Gabor filtering is inspired by multi channel filtering theory for visual information processing in the early stages of human visual system. Daugman [12] proposed the use of Gabor filter in the modeling of the receptive fields of simple cells in the visual cortex of some mammals. We use a modified parameterization according to Kruizina and Petkov [13] that takes into account restrictions found in experimental data. We employ the following family of two-dimensional Gabor function to model the properties of simple cells:

$$g_{\varepsilon,\eta,\lambda,\theta,\varphi}(x,y)=\exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\varphi\right)$$
$$x'=(x-\varepsilon)\cos\theta-(y-\eta)\sin\theta \qquad (2)$$
$$y'=(x-\varepsilon)\sin\theta+(y-\eta)\cos\theta$$

where the pair (ε,η) determines the *center of a receptive field* in image coordinates. The standard deviation σ of the Gaussian envelope specifies the *size of the receptive field*. The parameter γ, called the spatial aspect ratio, determines the *ellipticity of the receptive field*. The parameter λ is the *wavelength* and $1/\lambda$ the *spatial frequency* of the channel that is modeled by Gabor functions. The ratio δ/λ determines the spatial frequency bandwidth and is fixed to 0.56, which corresponds to a bandwidth one octave at half-response:

$$\frac{\sigma}{\lambda}=\frac{1}{\pi}\sqrt{\frac{\ln 2}{2}}\cdot\frac{2^b+1}{2^b-1} \qquad (3)$$

where the parameter b is *bandwidth* (in octaves). The angle parameter θ, determines *orientation of channel*. The parameter φ is a *phase offset* that specifies the filter is symmetric or anti-symmetric. The response $r_{\varepsilon,\eta,\lambda,\theta,\varphi}$ of a simple cell to input image *f(x,y)* is computed by convolution:

$$r_{\varepsilon,\eta,\lambda,\theta,\varphi}(x,y)=f(x,y)*g_{\varepsilon,\eta,\lambda,\theta,\varphi}(x,y) \qquad (4)$$

The Gabor energy is related to a model of complex cells which combines the responses of a pair of simple cells with a phase difference of $\pi/2$ **:**

$$E_{\varepsilon,\eta,\lambda,\theta}=\sqrt{r^2_{\varepsilon,\eta,\lambda,\theta,0}+r^2_{\varepsilon,\eta,\lambda,\theta,(-\pi/2)}} \qquad (5)$$

2.2.1. Our proposed features. Hubel and Wiesel [17] deduced that simple cells are sensitive to specific orientations with approximate bandwidths of 30º. Therefore, we use a bank of Gabor filters with eight orientations ($\theta=k(\pi/8), k=1,2,...8$) and three frequencies ($\lambda=2.8$, $\lambda=4.2$ and $\lambda=5.6$). The frequencies and orientations are selected such that appropriate coverage of the spatial-frequency domain is achieved. For each channel, the response of Gabor filters represents the regions which are in direction of frequency and orientation of the channel. The useful criterion for distinction of different texture is the feature extraction from shape of these regions. In this paper, the features extraction is performed by moments [16]. For this purpose, the following procedure is applied:

First, we use the bank of Gabor filters to obtain Gabor energy images. Let the filtered image for the *i*th filter be $E_i(x,y)$. Second, for each filtered image, the moments are computed within small local window. For this purpose, the convolution of the filtered image $E_i(x,y)$ is computed by a set of masks that are given as follows:

$$m_{00}=\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}\quad m_{10}=\begin{bmatrix}-1&-1&-1\\0&0&0\\1&1&1\end{bmatrix}\quad m_{01}=\begin{bmatrix}-1&0&1\\-1&0&1\\-1&0&1\end{bmatrix}$$

$$m_{20}=\begin{bmatrix}1&1&1\\0&0&0\\1&1&1\end{bmatrix}\quad m_{11}=\begin{bmatrix}1&0&-1\\0&0&0\\-1&0&1\end{bmatrix}$$

Let the M_{ik} be the *i*th filtered image that are convolved with *k*th mask (*k=1, 2, 3, 4, 5*). Third, feature F_{ik} corresponding to the moment image M_{ik} with mean $\overline{M}$ is obtained using following non-linear transform:

$$F_{ik}=\frac{1}{m\times n}\sum|\tanh(\varepsilon(M_{ik}-\overline{M}))| \qquad (7)$$

where F_{ik} is the *k*th feature which is extracted from the *i*th filtered image of size m×n pixels. The parameter ε controls the shape of logistic function which is set to 0.25. Indeed, by using non-linear transform, the moment images with *identical second-order statistics* same average energy are distinguished. In this way a total of 120 (=5×24) features are obtained for each input image.

2.2.2. Gabor-energy features. The quantity of Gabor energy is computed based on the previous bank of Gabor filters and a set of 24 filtered images are acquired. Since the histograms of filtered images are often close to a Gaussian shape [11], only the mean and standard deviation of 24 filtered images are calculated. So a total of 48 features are achieved for each input image.

2.2.3. Fourier transform of Gabor output. Tan [19] has proposed a set of rotation invariant features which are based on Fourier transform of Gabor energy. These features are computed on predefined bank of Gabor filters and for each frequency, a set of five coefficients is acquired. In fact, a total of 15 (=3×5) feature are obtained for each input image.

2.2.4. Symmetric Gabor filters. Symmetric Gabor filters are defined by using Eq. 2 with a phase parameter φ=0. The mean and standard deviation of 24 filtered images are calculated and a total of 48 features are achieved for each input image.

2.2.5. Sigmoidal transform of Gabor. Jain and Farrokhnia [18] applied a bank of symmetric filters to input image and then, filtered images subjected to a nonlinear sigmoidal function that behaves like a blob detector. An energy measure was defined on the transform image in order to compute different texture feature for different blobs. We use this method with the previous bank of Gabor filters and the mean and standard deviation of 24 feature images are computed. So, a total of 48 features are obtained.

2.2.6. Said method. In this paper, Said method which are proposed for English documents, are investigated. In this method, the Gabor filters which are introduced in [12] are used. So, the Gabor energy is calculated for 4 frequencies and four orientations and the mean and standard deviation of 16 images are obtained as futures.

2.2.7. Co-occurrence matrix features. In our experiments, a set of 16 co-occurrence matrices is calculated by four distances d=1, 2, 3, 4 and four directions α= 0, 45, 90, 135. Since preprocessed image is binary, these matrices are of dimension 2×2. Also due to the diagonal symmetry, from each of the matrix three values are considered. Therefore, a total of 48 features are obtained for each input image.

2.3. Writer identification

Since the number of writers in a realistic problem is very large, use of techniques such as the support-vector machine (SVM) or multilayer perceptron (MLP) is not trivial in the writer identification problem. Therefore, we investigated two simple classifiers with low computational cost: weighted Euclidean distance and χ^2 distance. Since the results of χ^2 distance was better than weighted Euclidean distance, in this paper only these results are presented. The features that are extracted for unknown input text are compared with the features of known writers. The writer that his/her features have minimum distance from features of unknown input text is considered as identity of unknown input text. The χ^2 distance measure is defined as follow:

$$\chi_{ij}{}^2 = \sum_{k=1}^{n} \frac{(f_{ki} - m_{kj})^2}{(f_{ki} + m_{kj})} \tag{9}$$

where f_{ki} is the kth feature of the unknown input text i and m_{kj} is the mean value of the kth feature of writer j, that are computed from training blocks of writer j. The advantage of using the χ^2 distance measure is that differences for small features are weighed more importantly than weighted Euclidean distance.

3. Experimental results

For evaluation of our proposed method, we selected 40 participants with respect to age, gender, education, etc. and asked from each participant to copy out a desired text in his/her natural handwriting on an A4 page. Handwritten documents are digitally scanned and preprocessing is performed for those documents. Each preprocessed image is divided into 6 non-overlapping blocks, and 4 blocks is considered for training and two blocks for the testing.

Our proposed features were extracted from preprocessing documents and the effect of different combination frequencies in these features was tested. These results of identification rate for five hit list sizes are summarized in Table 1. Experiments show that mean features in comparison with the standard deviation features, contain more information about characteristic of handwriting. Also, identification rate for high frequencies (lower wavelength) is better than low frequencies.

Table 1: The identification rate (%) for our proposed features based on χ^2 distance

Feature type	Number of features	Size of hit list				
		1	2	3	4	5
All features	120	**75**	**86.25**	**93.75**	**95**	96.25
$\lambda = 2.8$	40	63.75	81.25	90	93.75	95
$\lambda = 4.2$	40	62.5	81.25	90	93.75	95
$\lambda = 5.6$	40	56.25	78.75	81.25	87.5	93.75
$\lambda = 2.8, 4.2$	80	72.5	83.75	92.5	95	**97.5**
$\lambda = 4.2, 5.6$	80	70	85	91.25	95	97.5

Other features were evaluated by similar method and for each method the best results are given in Table 2. The performance of Fourier transform of Gabor outputs is close to Gabor-energy. Although, these results are poorer than Gabor energy and proposed features, but these features are significant with considering their low dimensions and rotation invariant characteristic. Also, by comparison symmetric Gabor with sigmoidal transform (see Table 2), it can be seen that sigmoidal features have a better performance. Indeed, the combination of multi-channel filtering and the nonlinear stages can be viewed as performing a multi-scale blob detection, since writer identification is associated with differences in the attributes of these blobs in different handwritings. It is noted that classical method such as co-occurrence matrix features are much poorer than methods which are based on multi-channel filtering such as Gabor filters.

Table 2: The best identification rate (%) for other features based on χ^2 distance

Method type	Size of hit list				
	1	2	3	4	5
Gabor energy	56.25	71.25	80	86.25	90
Fourier of Gabor	48.75	61.25	71.25	72.5	76.25
Symmetric Gabor	36.25	48.75	58.75	65	67.5
sigmoidal transform	52.5	60	70	73.75	78.75
Said method	26.25	33.75	46.25	57.5	62.5
Co-occurrence	31.25	47.5	55	62.5	66.25

Experiments indicate that the use of proposed features in writer identification gives better performance. We think that the reason for this resides in fact that, the diversity of dominant orientation for Farsi/Arabic handwritings is powerless in comparison of English handwritings and the mean and standard deviation is not enough for feature extraction. Therefore, the result of Gabor-energy features based on weighted Euclidean distance, are poorer with respect to results which were obtained for English documents by Said [2]. In our method practically, more information was extracted by computing the moments on filtered images and therefore, a better performance was achieved in comparison with Gabor-energy and other features.

4. Conclusion

In this paper, the problem of writer identification was studied and a method for off-line writer identification based on Farsi/Arabic handwriting was presented which is text-independent. Experimental results demonstrate that proposed features perform much better than other investigated methods. Also, the comparison of results shows that, a better performance was achieved in frequency domain for higher frequencies (details) and in spatial domain for small distances. Due to no research has been reported on Farsi/Arabic documents, our method is very promising and can achieve better performance with improvement in preprocessing stages and combination local features with texture features.

5. References

[1] S.N. Srihari, H. Arora, S.H. Cha, and Sangjik Lee, "Individuality of handwriting," *Journal of Forensic Sciences*, vol. 47, no. 4, pp. 1-17, 2002.

[2] A. Bensefia, T. Paquet, and L. Heutte, "A writer identification and verification system," *Pattern Recognition Letters*, vol. 26, issue 13, pp. 2080-2092, 2005.

[3] L. Schomaker and M. Bulacu, "Automatic writer identification using connected-component contours and edge-based features of upper-case western script," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 787-798, 2004.

[4] M. Bulacu, L. Schomaker, and L. Vuurpijl, "Writer identification using edge-based directional features," *Seventh Int. Conf. on Document Analysis and Recognition*, pp. 937-941, 2003.

[5] E.N. Zois and V. Anastassopoulos, "Morphological waveform coding for writer identification," *Pattern Recognition*, vol. 33, pp. 385-398, 2000.

[6] A. Schlapbach and H. Bunke, "Using HMM based recognizers for writer identification and verification," *9th Int. Workshop on Frontiers in Handwriting Recognition*, pp. 167-172, 2004.

[7] H.E. Said, T.N. Tan, and K.D. Baker, "Personal identification based on handwriting," *Pattern Recognition*, vol. 33, No. 1, pp. 149-160, 2000.

[8] J. Ha, R.M. Haralick and I.T. Phillips, "Document page decomposition by the bounding-box projection techniques," *Proc. of Int. Conf. on Document Analysis and Recognition*, pp. 1119-1122, 1995.

[9] R. Manmatha and L. Rothfeder, "A scale space approach for automatically segmentation words from historical handwritten documents," *IEEE Trans. On PAMI*, vol. 27, no. 8, pp. 1212-1225, 2005.

[10] G.S. Peake and T.N. Tan, "Script and language identification from document images," *In Proc. of the BMVC '97*, vol. 2, pp. 610-619, 1997.

[11] Y. Zhu, T.N. Tan, and Y. Wang, "Font recognition based on global texture analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, No. 1, pp. 1192-1200, 2001.

[12] P. Kruizinga and N. Petkov, "Nonlinear operator for oriented texture," *IEEE Trans. on Image Processing*, vol. 8, no. 10, pp. 1395-1407, 1999.

[13] D.A. Pollen and S.F. Ronner, "Visual cortical neurons as localized spatial frequency filters," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 13, no. 5, pp. 907-916, 1983.

[14] D.H. Hubel and T.N. Wiesel, "Receptive fields and functional architecture in two nonstriate visual areas 18 and 19 of the cat," *J. of Neurophysiology*, vol. 28, pp. 229-289, 1956.

[15] A.K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp. 1167-1186, 1991.

[16] T.N. Tan, "Rotation invariant texture features and their use in automatic script identification," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 7, pp. 751-756, 1998.

[17] T.N. Tan, "Texture feature extraction via visual cortical channel modelling," *Proc. of 11th IAPR Int. conf. on Pattern Recognition*, vol. 3, pp. 607-610, 1992.