Comparison of Gabor-Based Features for Writer Identification of Farsi/Arabic Handwriting

F. Shahabi *M. Rahmati*
Department of Computer Engineering & Information Technology, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran.
{fshahabi, rahmati}@aut.ac.ir

Abstract

Writer identification recently has been studied and it has a wide variety of applications. Most 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 method for off-line writer identification based on Farsi handwriting, which is text-independent. Based on idea that has been presented in previous studies, in this paper we assume handwriting as texture image and a set of features which are based on multi-channel Gabor filtering and co-occurrence matrix features, are extracted from preprocessed image of documents. Experimental results demonstrate that the best result, 92% correct identification in a hit list with size 3 and 88% in a hit list with size 1, is acquired by using Gabor-energy features on Farsi handwritten documents from 25 peoples.

Keywords: Gabor filters, handwriting, multi-channel filtering, writer identification.

1. Introduction

Handwriting as a behavioral biometric is easy to obtain and studies have shown that different people have different handwritings [1, 2]. Therefore, writer identification recently has been studied and it has a wide variety of applications, such as security, financial activity, forensic and used as access control. Especially, 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. Writer identification task 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 to writer identification have been proposed. A scientific validation of individuality of handwriting is performed by Srihari et al. [1, 2]. 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 with a 98% correct identification, based on Macro features and Micro features that are extracted from handwritten documents. Said et al. [3] proposed a global approach based on multi-channel Gabor filtering, where each writer's handwriting is regarded as a different texture. Bensefia et al. [4] use 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. [5] present 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. [6] proposed a HMM based approach for writer identification and verification. They built an individual recognizer for each writer and train it with text lines of that writer. Bulacu et al. [7] evaluated the performance of edge-based directional probability distributions as features in comparison to a number of non-angular features. It is noted that the joint probability distribution of the angle combination of two hinged edge fragments outperforms all other individual features. Marti et al. [8] extracted a set of features from 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. Additionally, features based on the fractal behavior of the writing, are used. In Zois et al. [9] a new feature vector is employed by means of morphologically processing the horizontal profiles of the words. Efficiency of the proposed features is tested by Bayesian classifiers and neural networks.

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 presented a comparison of Gabor-based features for off-line writer identification based on Farsi handwriting, which are text-independent. Based on idea that has been presented in Said et al. [3], we assume handwriting as texture image and writer identification as texture

classification. For this purpose, in the first step the image of document is preprocessed and then we apply a set of Gabor filters with eight equidistant orientations and three spatial frequencies, resulting in 24 filtered images. Then, features are extracted from these filtered images based on four methods: symmetric Gabor filters, sigmoidal transform of Gabor output, Gabor-energy and Fourier transform of Gabor. Evaluation of these methods demonstrates that Gabor-energy features achieve better performance on Farsi handwritten documents from 25 peoples. Furthermore by means of global features, we need no segmentation or connected component analysis.

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

2. Proposed method

The proposed method is based on the idea that has been presented in Said et al. [3]. We assume handwritten text as texture and extract features from document by means of texture analysis. We describe the stages of our proposed method in the following.

2.1. Preprocessing

Prior to texture analysis, handwriting documents need to be preprocessed. For this purpose, documents are normalized with respect to different word spacing, line spacing, etc. [3, 13]. The preprocessing can be done in the following steps:

1) Projection profile has been widely used in line and word detection [10]. We use a modified version of the same algorithm extended to gray-level images [11]. First, the horizontal projection profile is computed and then smoothed with a low pass Gaussian filter. Smoothing eliminates false local maxima and reduces sensitivity to noise. Fig. 1 shows an image of document, its projection profile and smoothed version of projection profile. 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. Since convolution is a linear operation, the smoothing and the derivative operation can be combined into one step by convolving the projection profile with a Gaussian derivative as follows:

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

2) Next, each text line that is located in previous step is binarized with Otsu algorithm [12] 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 normalized by reducing it 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. The result from this step has been shown in Fig. 2a.

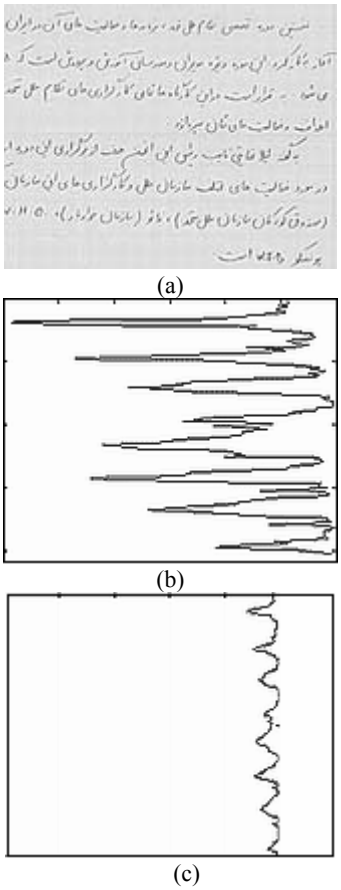

Figure 1. Extraction of lines from handwritten text. (a) Image of documents, (b) Horizontal projection profile and (c) Smoothed projection profile

Figure 2. (a) Space between words and lines is normalized and the padding is done for each line of text, (b) the final preprocessed image

3) If blank spaces exist in the bottom of image from previous steps, then padding is applied, by copying the text lines from top of the image, in order to achieve a uniform image of the text with a predefined size. In Fig. 2b, the final preprocessed image of text is shown.

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 4 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 [3, 13]. Therefore, we use features which are based on multi-channel Gabor filtering. For comparison, we also investigated a classical method which is based on gray-level 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 [14] 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 [15] 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. Based on experiments, the frequency bandwidth of simple cells in the visual cortex is about one octave [16]. Therefore, the ratio δ/λ that determines the spatial frequency bandwidth 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:

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

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 $\pi/2$. The results of a pair of symmetric and anti-symmetric filters are combined into the Gabor energy as follows:

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

Indeed in multi-channel Gabor filtering, each channel is modeled by a pair of Gabor filter which are symmetric and anti-symmetric.

For feature extraction, different methods can be applied to the output of Gabor filters. Some of these methods were investigated for writer identification and comparison of Gabor features with co-occurrence matrix features was made.

2.2.1. Symmetric Gabor filters

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 three frequencies ($\lambda = 2.7,\ 4.1,$ and 5.4) and eight equidistant orientations ($\theta = k(\pi/8), k = 1,2,...8$). The frequencies and orientations are selected such that appropriate coverage of the spatial-frequency domain is achieved as shown in Fig. 3. In fact, by this method the input image is decomposed into a number of filtered images, each of which contains intensity variation over a narrow band of frequency and orientation.

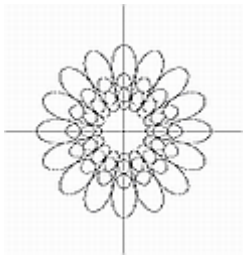

Figure 3. Coverage of spatial-frequency domain by using Gabor filters [15]

Therefore, a total of 24 filtered images which are based on symmetric filters, are obtained. These filters are defined by using Eq. 2 with a phase parameter φ=0. Since the histograms of filtered images are often close to a Gaussian shape [13], only the mean and standard deviation of 24 filtered images are calculated. So, a total of 48 features are achieved for each input image.

2.2.2. 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. The feature image $f_i(x,y)$, corresponding to filtered image $r_i(x,y)$ was computed by:

$$f_i(x,y) = \frac{1}{M^2}\sum_{(a,b)\in W}|\tanh(\alpha r_i(x,y))| \qquad (6)$$

where α is set to a constant value α= 0.25 and *M* is the size of averaging window. We use this method with the previous bank of Gabor filters with three frequencies and eight orientations to obtain feature images. The mean and standard deviation of 24 feature images are computed and a total of 48 features are obtained for each input image.

2.2.3. Gabor-energy features

The quantity of Gabor energy (Eq. 5) is computed based on the previous bank of Gabor filters and a set of 24 filtered images are acquired. Then, features are extracted from these images based on the mean and standard deviation values.

2.2.4. Fourier transform of Gabor output

Tan [19] has proposed a set of rotation invariant Gabor features which are based on Fourier transform. Based on this method, the average value $Q(\lambda,\theta)$ of energy output $E_{\varepsilon,\eta,\lambda,\theta}$ is computed by:

$$Q(\lambda,\theta)=\frac{1}{\Omega}\iint_{\Omega}E_{\varepsilon,\eta,\lambda,\theta}(x,y)dxdy \tag{7}$$

where Ω is area of $E_{\varepsilon,\eta,\lambda,\theta}$. The features $Q(\lambda,\theta)$ are rotation dependent, so for a certain frequency 1/λ, Fourier transform of $Q(\lambda,\theta)$,or simply $Q(\theta)$ is computed and coefficients of Fourier transform are used as features. Since $Q(\theta)$ is a discrete function, the maximum number of independent features is *M*/2+1, where *M* is the total number of orientations at the certain frequency. These features are computed on predefined bank of Gabor filters and for each frequency 1/λ, a set of five coefficients is acquired. In fact, a total of 15 (=3×5) feature are obtained for each input image.

2.2.5. Co-occurrence matrix features

Gray-level co-occurrence matrix, is a classical method for texture analysis [20]. This matrix reveals certain properties about the spatial distribution of the gray values in the texture image. 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 acquired 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 *k*th feature of the unknown input text *i* and m_{kj} is the mean value of the *k*th 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 25 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 at 300 dpi resolution and preprocessing is performed for those documents. Each preprocessed image is divided into 4 non-overlapping blocks, and 3 blocks is considered for training and one block for the testing. Fig. 4 shows samples of handwriting from different participants. The performance of the five described methods for feature extraction was investigated and results are given in the following.

Figure 4. Samples of handwriting from different participants

The features of symmetric Gabor filters were extracted from preprocessing documents, as described in section 2.2.1 and identification was performed using χ^2 distance classifier. The effect of different combination frequencies in these features was tested and results of identification rate for two hit list sizes are summarized in

Table 1. By comparing these results with the results of sigmoidal transform (see Table 2), it can be seen that sigmoidal features have a better performance. Using the symmetric component requires smoothing to achieve reasonable results, so this method solely is not preferable. 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. The best result for sigmoidal features, identification rate 84% in hit list of size 3 and 80% in hit list of size 1, was obtained using all features. 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 symmetric Gabor filters

Feature type	Number of features	Identification rate (%)	
		Hit list of size 1	Hit list of size 3
All features	48	36	68
All at $\lambda = 2.7$	16	40	60
All at $\lambda = 4.1$	16	32	56
All at $\lambda = 5.4$	16	12	28
All at $\lambda = 2.7, \lambda = 4.1$	32	**52**	**76**
All at $\lambda = 4.1, \lambda = 5.4$	32	24	60
Only mean	24	28	52
Only standard deviation	24	20	44
Average		**30.5**	**55.5**

Gabor-energy features were evaluated by similar method and results are given in Table 3. It is clear these features give better performance than two previous features. Table 3 shows that best identification rate that is 92% in hit list of size 3 and 88% in hit list of size 1, was achieved using all features.

Table 2. The identification rate for sigmoidal transform of Gabor

Feature type	Number of features	Identification rate (%)	
		Hit list of size 1	Hit list of size 3
All features	48	**80**	**84**
All at $\lambda = 2.7$	16	76	80
All at $\lambda = 4.1$	16	68	80
All at $\lambda = 5.4$	16	48	76
All at $\lambda = 2.7, \lambda = 4.1$	32	80	80
All at $\lambda = 4.1, \lambda = 5.4$	32	72	84
Only mean	24	76	84
Only standard deviation	24	32	60
Average		**66.5**	**78.5**

The results from Fourier transform of Gabor outputs are presented in Table 4 and show that the performance of these features is close to Gabor-energy. The best result for these features, identification rate 92% in hit list of size 3 and 80% in hit list of size 1, was acquired using all features. Although, these results are poorer than Gabor energy but, these features are significant with considering their low dimensions and rotation invariant characteristic.

Table 3. The identification rate for Gabor-energy features

Feature type	Number of features	Identification rate (%)	
		Hit list of size 1	Hit list of size 3
All features	48	**88**	**92**
All at $\lambda = 2.7$	16	76	92
All at $\lambda = 4.1$	16	72	92
All at $\lambda = 5.4$	16	68	88
All at $\lambda = 2.7, \lambda = 4.1$	32	80	92
All at $\lambda = 4.1, \lambda = 5.4$	32	80	88
Only mean	24	84	92
Only standard deviation	24	44	68
Average		**74**	**88**

In fact, Gabor-energy models the response of complex cells in the human visual system (HVS) and performs better than others Gabor features.

Table 4. The identification rate for Fourier transform of Gabor output

Feature type	Number of features	Identification rate (%)	
		Hit list of size 1	Hit list of size 3
Five amplitude of Fourier	15	**80**	**92**
Five amplitude at $\lambda = 2.7$	5	76	84
Five amplitude at $\lambda = 4.1$	5	64	88
Five amplitude at $\lambda = 5.4$	5	52	80
Four amplitude at $\lambda = 2.7$	4	56	84
Four amplitude at $\lambda = 4.1$	4	52	84
Four amplitude at $\lambda = 5.4$	4	52	76
Three amplitude at $\lambda = 2.7$	3	40	80
Average		**59**	**83.5**

Finally, the same experiment was carried out for co-occurrence matrix features and the results from these features are shown in Table 5. Different distances was computed for these features and demonstrated the nearest distances have better performance. The best identification rate that is 76% in hit list of size 3 and 64% in hit list of size 1, was obtained using two distances *d= 1, 2*.

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.

Since, the diversity of orientation for Farsi/Arabic handwritings is powerless in comparison to English handwritings, considering only symmetric filters is inadequate for feature extraction and more information was acquired by using other methods such as Gabor-energy and Fourier transform.

Table 5. The identification rate for co-occurrence matrix features

Feature type	Number of features	Identification rate (%)	
		Hit list of size 1	Hit list of size 3
All features	48	64	72
d= 1, 2, 3	36	64	72
d= 1, 2	24	**64**	**76**
d= 1	12	52	72
d= 2	12	64	72
d= 4	12	40	64
Average		**58**	**71.3**

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. We assumed handwriting as texture image and after a preprocessing stage, a set of features which are based on Gabor features and co-occurrence matrix features, were extracted from image of documents. Experimental results demonstrate that Gabor-energy and Fourier transform of Gabor output perform much better than others 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. Therefore, we can obtain a higher identification rate by using of features which contain details of texture. Since no research has been reported on Farsi/Arabic documents for comparison, our method is promising and we can achieve better performance with improvement in preprocessing stages and combination of local features with texture features.

Acknowledgments

The authors would like to thank the Iran Telecommunication Research Center (ITRC), for their financial support.

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] S. N. Srihari, H. Arora, S. H. Cha and Sangjik Lee, "Individuality of handwriting: a validation study," *Proc. Of 6th IEEE Conf. on Document Analysis and Recognition*, pp. 106-109, 2001.

[3] 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.

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

[5] 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.

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

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

[8] U.V. Marti, R. Messerli and H. Bunke, "Writer identification using text line based features," *IEEE Proc. of 6th Int. conf. on Document Analysis and Recognition*, pp. 101-105, 2001.

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

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

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

[12] N. Otsu, "A threshold selection method for gray-level histogram," *IEEE Trans. On Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.

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

[14] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Amer.*, vol. 2, pp. 1160–1169, 1985.

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

[16] 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.

[17] 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.

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

[19] 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.

[20] R. Haralick, K. Shanmugam and I. Dinstein, "Textural features for image classification," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610-621, 1973.