FARSI CHARACTER RECOGNITION USING NEW HYBRID FEATURE EXTRACTION METHODS

Fataneh Alavipour [1] and Ali Broumandnia[2]

[1]Department of Electrical, Computer and Biomedical Engineering, Qazvin branch,
Islamic Azad University, Qazvin, Iran
[2] Islamic Azad University, South Tehran Branch, Department of Computer, Tehran, Iran

ABSTRACT

Identification of visual words and writings has long been one of the most essential and the most attractive operations in the field of image processing which has been studied since the last few decades and includes security, traffic control, fields of psychology, medicine, and engineering, etc. Previous techniques in the field of identification of visual writings are very similar to each other for the most parts of their analysis, and depending on the needs of the operational field have presented different feature extraction. Changes in style of writing and font and turns of words and other issues are challenges of characters identifying activity. In this study, a system of Persian character identification using independent orthogonal moment that is Zernike Moment and Fourier-Mellin Moment has been used as feature extraction technique. The values of Zernike Moments as characteristics independent of rotation have been used for classification issues in the past and each of their real and imaginary components have been neglected individually and with the phase coefficients, each of them will be changed by rotation. In this study, Zernike and Fourier-Mellin Moments have been investigated to detect Persian characters in noisy and noise-free images. Also, an improvement on the k-Nearest Neighbor (k-NN) classifier is proposed for character recognition. Using the results comparison of the proposed method with current salient methods such as Back Propagation (BP) and Radial Basis Function (RBF) neural networks in terms of feature extraction in words, it has been shown that on the Hoda database, the proposed method reaches an acceptable detection rate (96/5%).

KEYWORDS

Character Recognition, Zernike Invariant Moments, Fourier–Mellin moments, Back Propagation (BP), KNN Algorithm, Radial Neural Network (RBF).

1. INTRODUCTION

Automated recognition of text has been an active subject of research since the early days of computers. A 1972 survey cites nearly 130 works on the subject [1]. Despite the age of the subject, it remains one of the most challenging and exciting areas of research in computer science. In recent years it has grown into a mature discipline, producing a huge body of work. Handwriting recognition can be defined as the task of transforming text represented in the spatial form of graphical marks into its symbolic representation. Note only is this useful for making digital copies of handwritten documents, but also in many automated processing tasks, such as automatic mail sorting or cheque processing. In automated mail sorting, letters are directed to the correct location by recognition of the handwritten address. Similarly, cheque processing involves recognising the words making up the cheque amount. The field of handwriting recognition can be split into two different approaches [2].

The first of these, on-line, deals with the recognition of handwriting captured by a tablet or similar touch-sensitive device, and uses the digitised trace of the pen to recognise the symbol. In this instance the recogniser will have access to the x and y coordinates as a function of time, and thus has temporal information about how the symbol was formed [3].

The second approach concentrates on the recognition of handwriting in the form of an image, and is termed off-line. In this instance only the completed character or word is available [4].

In the character recognition system, after receiving of input images a feature vector is extracted to utilize in the classification step. In this paper, the feature extraction and classification steps of such system are studied. For feature extraction, geometric moments are employed to produce statistical features from input image. For classification step, an improvement on the k-NN classifier is introduced.

This paper is organized as follow: feature extraction step of a character recognition system is introduced in Section 2. The proposed method is presented in Section 3. Experimental results are described in Section 4 and the paper will be concluded in Section 5.

2. FEATURE EXTRACTION

Each character detection system contains three steps: pre-processing, feature extraction and classification. Feature extraction refers to the extraction of useful information from raw data so that they are suitable for the classification process. The feature extraction stage is characterized by a series of input patterns. The major problem of feature extraction is that it depends on application and feature extraction methods are not public.

Feature extraction methods can be divided into two majors: structural features and statistical features [5-6]. The first group is based on local structure of image. In other words, the structural features deals with local data.

In the statistics-based feature extraction techniques, global data is employed to create a set of feature vector elements in order to perform recognition. A mixture of irrelevant data, which are usually part of a facial image, may result in an incorrect set of feature vector elements. The statistics-based feature extraction techniques are Principle Component Analysis (PCA), Legender Moment (LM) [7] and Zernike Moments (ZM) [8] and Fourier-Mellin Moment (FMM) [9]. Legendre functions are Legendre differential equation. The main advantage is that Legendre moments like Legendre basis functions are orthogonal. Legendre moments are independent of each other and are free of data redundancy. In this study, we use ZM and FMM to recognize characters that are presented in the next Section.

3. PROPOSED METHOD

The previous feature extraction methods suffer from some drawbacks: sensitivity to noise such as simple k-NN [10], to image rotation such as Principal Component Analysis (PCA) [11] and Legendre Moment (LM) [7], to image size such as k-NN [10]. As previously mentioned, both the feature extraction and classification of the most important steps to detection numbers.

The objective of this paper is improving of two above mentions steps. For this purpose, at first, we focus on the combination of two Zernike and Fourier-Mellin moments as the feature extraction step. The combined approach has overcome disadvantages of the previous methods.

For instance to solve the problem of image rotation, the features extracted by methods Zernike are statistical and so they are independent of the scaling and rotation. It should be noted that geometrical-based feature extraction methods are depend on the image size, position, big and small picture elements But statistical methods are not rely on the image size. Also, the combined moment is robust to noise because the noise is considered as the geometrical and locational information. The second contribution of this paper is on the classification step of character detection process. The proposed classification method is an improvement on k-Nearest Neighbor (k-NN).

3.1. Zernike Moment (ZM)

ZM is geometric-based moment that is a two dimensional function of orthogonal polynomials on the unit disk. The orthogonal moments of ZM are rotation and scale invariants which are suitable for pattern recognition applications [4][12][13][14]. ZM contains several orthogonal sets of complex-valued polynomials defined as

$$V_{nm}(x,y) = R_{nm}(x,y) \exp\left(jm \tan^{-1}\left(\frac{y}{x}\right)\right) \tag{1}$$

where $x^2 + y^2 \leq 1, n \quad 0, |m| \quad n$, and the radial polynomials $\{R_{nm}\}$ are defined as

$$R_{nm}(x,y) = \sum_{s=0}^{(n-|m|)/2} S_{n,|m|,s}(x^2 + y^2)^{\frac{n-2s}{2}} \tag{2}$$

where

$$S_{n,|m|,s} = (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!} \tag{3}$$

The ZM of order n and repetition m can be computed as

$$ZM_{nm} = \frac{n+1}{\pi}\sum_x \sum_y f(x,y)V_{nm}^*(x,y) \tag{4}$$

It should be noted that the PZM is computed for positive m because $V_{nm}(x,y) = V_{nm}^*(x,y)$. Center of the unit disk is located on the origin of coordinates and so ZM technique is independent of scaling and rotation. In the next Section, ZM approach will be utilized to extract feature vector elements.

3.2 Fourier-Mellin Moment (FMM)

The orthogonal Fourier-Mellin Moments (FMMs) are proposed by sheng and sheng [9] which they are suitable for scale and rotation-invariant pattern recognition. The orthogonal FMMs may be thought of as generalized Zernike moments and orthogonalized complex moments. The orthogonal FMMs are defined as:

$$\phi_{nl} = \frac{n+1}{\pi}\int_0^{2\pi}\int_0^1 Q_n(r)e^{-il\varphi} f(r,\varphi)r\, dr\, d\varphi \quad n = 0,1,2,\dots,l = 0,\pm1,\pm2,\dots \tag{5}$$

where $f(r,\varphi)$ is the image and the radial function $Q_n(r)$ is calculated as:

$$Q_n(r) = \sum_{s=0}^n (-1)^{n+s}\frac{(n+s+1)!}{(n-s)!s!(s+1)!}r^s \tag{6}$$

The first six polynomials of the moments are shown as follows:

$$Q_0(r) = 1,$$
$$Q_1(r) = -2 + 3r,$$
$$Q_2(r) = 3 - 12r + 10r^2,$$
$$Q_3(r) = -4 + 30r - 60r^2 + 35r^3,$$
$$Q_4(r) = 5 - 60r + 210r^2 - 280r^3 + 126r^4,$$
$$Q_5(r) = -6 + 105r - 500r^2 + 1260r^3 - 1260r^4 + 462r^5. \tag{7}$$

The main difference between orthogonal FMM and ZM is absence of restrictions on $n \geq l$ and also independence of the radial function $Q_n(r)$ in the repetition of factor l.

3.3 Creating feature vector

After resizing and denoising of input image, the ZM and Fourier-Mellin moments are employed separately to extract feature vector. The feature vector elements are defined according to ZM orders as

$$FV\_ZM_j = \{ZM_{km}| \, k = j, j+1, \dots, N\} \tag{8}$$

where j is interval [1,N−1] and so, FV_j contains all the ZM from order j to N. Samples of feature vector elements will be demonstrated in Table 1 for j = 3, 5 and 9, and N = 10. As Table 1 shows, increasing of j decreases the number of elements in each feature vector (FV_j).

Table 1. Feature vector elements based on the ZM

j value	FV$_j$ feature elements (ZM_{km})		Number of feature element
	K	M	
4	4	0,2,4	30
	5	1,3,5	
	6	0,2,4,6	
	7	1,3,5,7	
	8	0,2,4,6,8	
	9	1,3,5,7,9	
	10	0,2,4,6,8,10	
6	6	0,2,4,6	24
	7	1,3,5,7	
	8	0,2,4,6,8	
	9	1,3,5,7,9	
	10	0,2,4,6,8,10	
9	9	1,3,5,7,9	11
	10	0,2,4,6,8,10	

For FMM, the feature vector is obtained as:

$$FV\_FMM = \{\phi_{nl}| \, n = 0,1,2, \dots, l = 0, \pm1, \pm2, \dots\} \tag{9}$$

The final feature vector is the combination of FV_ZM_j and FV_FMM which contains all necessary information for detection of characters.

3.4 Classification Using Improved *k*-Nearest Neighbor (*k*-NN)

After the construction of feature vector, classification step of input image will be performed. In this step, the specific character shown in input image will be detected. For this purpose, a new classification approach based on *k*-NN classifier is proposed to improve the accuracy of the classifier called Improved *k*-NN.

In the classic *k*-NN, *k* determines number of samples as the neighbors of test sample where for the variable, only one value is set. After that, the Euclidean distance between the test sample and its neighbors are calculated as the follows

$$d = \sqrt{\sum_{i=1}^{m}(x_i - y_i)^2} \tag{10}$$

and finally, the test sample is classified to the class which covers most of the neighbors of the test sample. The idea of one value for *k* decreases the accuracy of *k*-NN method. Also, it causes that the simple *k*-NN is sensitive to noisy data. In the other words, in *k*-NN, decision on the class of the test sample with just one value may be provided in random result and so, the efficiency of *k*-NN classifier is reduced.

In the Improved *k*-NN, an interval of integer values is considered for variable *k* to determine the class of test sample. In this manner, an interval is set for the variable *k* and so a class is assigned to the test sample per each value of the interval. Finally, the class with highest frequency of selection is chosen as the main class of the test sample. For instance, suppose four values 3,5,7 and 9 for the *k*. To classify test sample *A*, the distance between the *A* and its neighbors is computed and for *k*=3 three nearest of its neighbors are selected in order to determine the class of *A* for this value of *k*. After that this process is repeated for *k*=5, 7 and 9, and the class with maximum selection is considered as the class of sample test *A*. This idea increases the accuracy of *k*-NN classifier. Unlike the simple *k*-NN, the results of Improved *k*-NN is produced without randomness and the most reliable results are obtained using Improved *k*-NN. Also, the Improved *k*-NN is robust to noisy data. The other advantage of Improved *k*-NN is that the runtime of character detection process is reduced.in the proposed method and so the proposed method will converge faster to optimal solution. In comparison to simple *k*-NN.

4. EXPERIMENTAL RESULTS

In this Section, we evaluate the proposed character detection system. In this regard, the proposed hybrid Zernike-Mellin Moment is compared with both of the ZM and FMM on the detection rate, speed and the size of feature vector. Also, the Improved k-NN classifier is compared with two neural networks such as Radial Basis Function (RBF) , Back Propagation (BP) on the accuracy and runtime.

The used dataset is Hoda dataset [15] which contains the number from "0" to "9". For each number, there are 100 samples image in different handwriting forms in size of 13×13.

At the first step, the ZM, FMM and Zernike-Mellin Moments as the feature extractors are compared on the runtime which is demonstrated in Table 2. In this table, the mentioned moments are compared in different sizes of feature vector. The runtime is increased when the size of

feature vector is increased. The results show the success of hybrid Zernike-Mellin because both of the ZM and FMM extract the feature vector in a parallel form and so an informative feature vector is extracted in the reduced time.

In Table 3, the results of BP neural network and the Improved k-NN are analysed in terms of character identification speed. In this evaluation, both of the classifiers are tested in different size of feature vector and the results prove the superiority of the Improved k-NN to the BP network because the feature extraction step of Improved k-NN is performed using hybrid Zernike-Mellin Moment which has significant effect on the performance of the proposed classifier. The graphical results of Table 3 are shown in Figure 1. It should be noted that the results of BP and RBF neural network in this test are very similar to each other that the only results of BP is presented.

Next, the impact of each individual moment, ZM, FMM and hybrid Zernike-Mellin, on the classification is examined. As can be seen in Table 4, the feature vector extracted by the ZM has high impact on the Improved k-NN algorithm. In the second place, the best results are obtained using BP network, and RBF neural network is located on the third level. The reason of the success of Improved k-NN is that it employs multi-level neighborhood rather than a value.

In the next step, the previous testing phase is performed using the feature vector extracted using FMM method which is shown in Table 5. In this table, the best results are produced by proposed classifier. It should be noted that the results using ZM is higher than the FMM ones that shows the superiority of ZM to FMM in creation of informative feature vector.

Table 6 presents the results of the above mentioned classifiers on the feature vector extracted by the proposed hybrid Zernike-Mellin Moment. In this table, the Improved k-NN shows the high results in comparison to BP and RBF neural networks. As can be seen in Tables 4, 5 and 6, the results of Improved k-NN using hybrid Zernike-Mellin Moment have priority to the other classifiers with ZM and FMM feature extractors. The first reason refers to the feature extracted by the hybrid Zernike-Mellin Moment because the produced feature vector includes all necessary and important information for detection of character. The second reason of this success is that the Improved k-NN selects the class of test samples with high confidence and the accuracy of classification is enhanced.

Table 2. Comparison of ZM, FMM and Zernike-Mellin Moment on the runtime. Suppose that N=17. The bold values show the best obtained results. FE means Feature Element.

No. of FE	ZM	FMM	Zernike-Mellin Moment
3	0.003	0.003	**0.002**
4	0.005	0.006	**0.004**
5	0.007	0.006	**0.005**
6	0.01	0.011	**0.006**
7	0.024	0.013	**0.012**
10	0.04	0.014	**0.017**
13	0.07	0.05	**0.024**
15	0.11	0.08	**0.032**
17	0.13	0.08	**0.041**
25	0.19	0.17	**0.14**

Table 3. Comparison of BP and Improved *k*-NN classifiers on the detection of characters on the runtime. Suppose *N*=17. The bold values show the best results. FE means Feature Element.

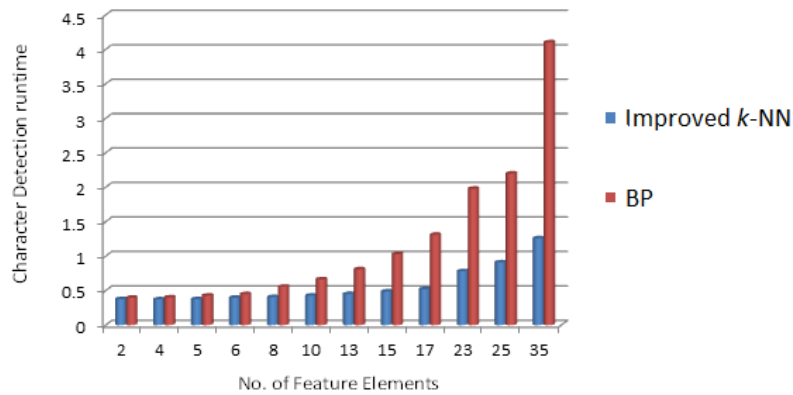| No. of FE	BP	Improved *k*-NN
2	0.399	**0.378**
3	0.339	**0/380**
4	0.402	**0.374**
5	0.429	**0.378**
6	0.450	**0.399**
8	0.557	**0.407**
10	0.664	**0.428**
13	0.811	**0.452**
15	1.03	**0.459**
17	1.311	**0.523**
23	1.978	**0.782**
25	2.20	**0.91**

Figure 1. Comparison of BP and Improved *k*-NN classifiers on Character detection runtime

Table 4. Comparison of BP, RBF and Improved *k*-NN classifiers on the accuracy based feature extracted by ZM. Suppose N=17. The bold values show the best results. FE means Feature Element.

No. of FE	BP	RBF	Improved k-NN
3	86.2	61.3	**88.4**
4	86.7	61.6	**88.5**
5	86.1	62.2	**88.5**
6	86.7	64.6	**88.7**
8	96.2	69.2	**99.7**
10	86.2	72.8	**88.8**
13	86.3	81.7	**89**
15	86.7	85.1	**89.5**
17	86.9	85.9	**89.7**
23	87.7	86.1	**90**
25	88.2	86.3	**90.1**

Table 5. Comparison of BP, RBF and Improved k-NN classifiers on the accuracy based feature extracted by FMM. Suppose that N=17. The bold values show the best obtained results. FE means Feature Element.

No. of FE	BP	RBF	Improved k-NN
3	87.4	80.1	**89.1**
4	87.3	80.2	**89.5**
5	88.0	81.9	**89.9**
6	88.1	83.4	**90.5**
8	88.5	85.4	**91.3**
10	88.8	86.6	**91.5**
13	89.6	87.2	**92.7**
15	89.7	87.9	**93.2**
17	89.9	88.1	**93.8**
23	90.4	89.1	**96.2**
25	91.5	90.4	**96.8**

Table 6. Comparison of BP, RBF and Improved k-NN classifiers on the accuracy based feature extracted by hybrid Zenike-Mellin Moment. Suppose that N=17. The bold values show the best obtained results. FE means Feature Element.

No. of FE	BP	RBF	Improved k-NN
3	87.5	73.9	**89.5**
4	87.9	84.8	**90.7**
5	88.6	85.5	**91.3**
6	89.1	84.7	**91.5**
8	89.7	85.6	**93.2**
10	89.9	66.5	**93.4**
13	90.1	75.2	**94.1**
15	90.3	72.2	**94.3**
17	90.3	70.4	**94.4**
23	90.4	80.2	**95.8**
25	91.3	82.1	**96.9**

The last phase of testing is evaluation of the proposed method against the "Salt & Pepper" and "Gaussian" noise distributions. The results of previous classifiers for detection of characters on the noisy images with "Salt & Pepper" and "Gaussian" distributions are reported in Tables 7 and 8, respectively. Note that for both of the noise distributions. The utilized feature extraction method is proposed hybrid Zernike-Mellin Moment. As can be seen in Tables 7 and 8, the Improved k-NN shows the best results. These tables show that the proposed classifier has significant improvement on the behaviour of simple k-NN because the simple k-NN is extremely sensitive to the noisy data. The graphical results of classifiers on detection of character on the Gaussian noisy image are demonstrated in Figure 2. In the last phase, the results of Improved k-NN on the both of the "Salt & Pepper" and "Gaussian" noise distributions are presented in Figure 3. According to this figure, the "Salt & Pepper" noise has more impact on decreasing of detection rate of Improved k-NN.
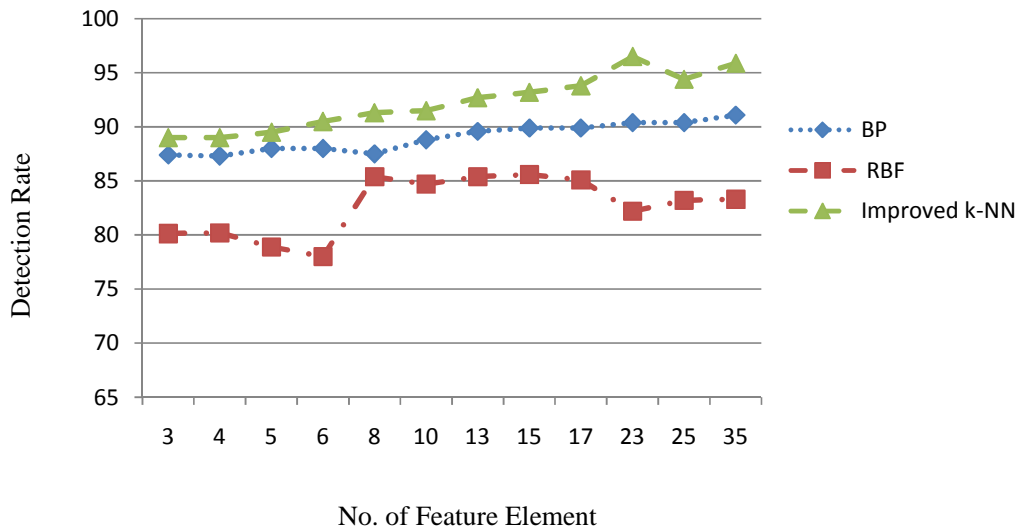
Figure 2. Comparison of BP, RBF and Improved *k*-NN classifiers on Character detection runtime

Table 7. Comparison on the detection rate of BP, RBF and Improved k-NN classifiers on the "Salt & Pepper" noisy images. Suppose N=17 and bold values show the best results.

Noise frequency (%)	BP	RBF	Improved *k*-NN
0.01	79.4	80.3	**84.4**
0.02	80.8	81.5	**83.9**
0.03	80.1	81.8	**83.6**
0.04	81.2	82.7	**83.2**
0.05	79.8	80.3	**82.4**
0.06	78.4	80.9	**81.9**

Table 8. Comparison on the detection rate of BP, RBF and Improved k-NN classifiers on the "Gaussian" noisy images. Suppose N=17 and bold values show the best results.

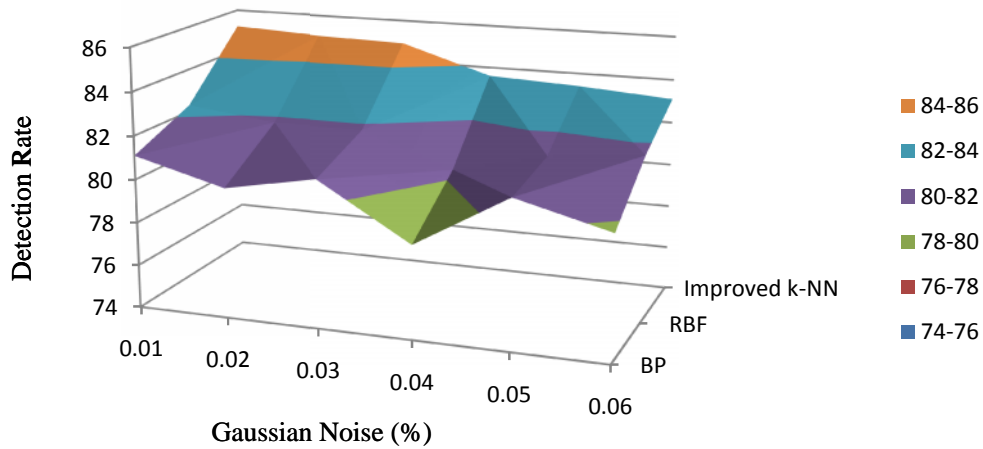| Noise frequency (%)	BP	RBF	Improved k-NN
0.01	81.1	82.3	**85.6**
0.02	80	81.8	**85.2**
0.03	80.8	81.5	**84.9**
0.04	78.3	81.7	**83.6**
0.05	80.8	81.3	**83.4**
0.06	79.7	80.7	**83.1**

Figure 3. Comparison of BP, RBF and Improved *k*-NN classifiers on Character detection runtime

5. CONCLUSION

In this paper, a new method is proposed to identify the characters such as numbers. the proposed method has two innovations. The first one is combination of ZM and FMM which called hybrid Zernike-Mellin Moment in order to extract feature vector from input image. The second idea is on the classification step where an improvement on the simple *k*-NN is introduced which called Improved *k*-NN. In the proposed classifier, multi-level neighborhood rather than a value for variable *k* is presented. The proposed system is evaluated on HODA dataset which contains non-noisy and noisy images. The results show that the proposed method is able to detect the characters with 96.5% detection rate in the reduced time.

REFERENCES

[1] Khosavi, H., Kabir, E., (2007) "Introducing a very large dataset of handwritten Farsi digits and a study on their varieties", Pattern Recognition Letters, Vol. 28, No. 10, pp. 1133-1141.

[2] de Campos, T., Babu, B. R., & Varma, M. (2009) "Character recognition in natural images".

[3] Harmon, L. D. (1972) "Automatic recognition of print and script", Proceedings of the IEEE, Vol. 60, No.10, pp. 1165-1176.

[4] Broumandnia, A., Shanbehzadeh, J. (2007) "Fast Zernike wavelet moments for Farsi character recognition", Image Vision Computing, Vol. 25, No. 5, pp. 717-726.

[5] Broumandnia, A., Shanbehzadeh, J., Nourani, M. (2007) "Handwritten Farsi/Arabic Word Recognition", AICCSA 2007, pp. 767-771.

[6] Broumandnia, A., Shanbehzadeh, J., Rezakhah Varnoosfaderani, M. (2008) "Persian/arabic handwritten word recognition using M-band packet wavelet transform", Image Vision Computing, Vol. 26, No. 6, pp. 829-842.

[7] Liao, S. & Pawlak, M. (1996) "On image analysis by moments", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 3, pp. 254-266.

[8] Zernike, v.F. (1934), "Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode", Physica, Vol. 1, No. 7, pp. 689-704.

[9] Sheng, Y., & Shen, L. (1994) "Orthogonal Fourier-Mellin moments for invariant pattern recognition", JOSA A, Vol. 11, No. 6, pp. 1748-1757.

[10] Altman, N. S. (1992) "An introduction to kernel and nearest-neighbor nonparametric regression", The American Statistician, Vol. 46, No. 3, pp. 175–185.

[11] Abdi. H., & Williams, L.J. (2010) "Principal component analysis", Wiley Interdisciplinary Reviews: Computational Statistics, Vol. 2, pp. 433–459.

[12] Belhumeur, P., Hespanha, J. & Kriegman, D. (1997) "Eigenfaces vs. fisherfaces recognition using class specific linear projection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 711-720.

[13] Belkasim, S., Shridhar, M. & Ahmadi, M. (1991) "Pattern recognition with moment invariants: A comparative study and new results", Pattern Recognition, Vol. 24, No.12, pp. 1117-1138, http://www.sciencedirect.com/science/article/pii/003132039190140Z.

[14] Teh, C.H. & Chin, R. (1988) "On image analysis by the methods of moments". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 10, No. 4, pp. 496-513.

[15] Free and Commercial Resources for Farsi OCR, http://www.farsiocr.ir/