A New Combination Method Based on Different Representation of Data

Reza Ebrahimpour[1], Mona Amini [2], Afra Vahidi Shams [2]

[1] *Department of Electrical Engineering, Shahid Rajaee Teachers Training University, Tehran, Iran, ebrahimpour@ipm.ir*

[2] *Mechatronics Engineering in Department of Mechatronics Engineering, Islamic Azad University, South Tehran Branch, Tehran, Iran, Mona.amini.mecha86@gmailcom, afra@vahidishams.ir*

Abstract

This paper proposes a new classification method for Farsi handwritten word recognition using gradient and gradient based features. The extracted feature vectors were classified using two Multi Layer Perceptron networks as basic experts, and one Radial Basis Function was applied to choose the best expert. The experiments were performed using the Iranshahr dataset. This dataset consists of 780 samples of 30 city names of Iran out of which, 600 samples were used to train the network and 180 samples to test it. A set of experiments were conducted to compare proposed method with some other combination rules. Results show that the proposed method achieved 91.11% recognition rate.

Keywords: *Farsi handwritten word recognition, Feature extraction, and Classifier fusion.*

1. Introduction

Handwritten recognition can be seen as a sub-task of more general Optical Character Recognition (OCR). Providing simple interface between man and machine is advantage of handwritten recognition systems. First, recognition of isolated handwritten digits and characters was investigated [1,2]. Later recognition of whole words [3,4] was addressed. There are many published works that deal with handwritten recognition of English [5-7], Japanese [8], Chinese [9] and Arabic scripts [10,11]. Farsi/Arabic handwritten recognition has slowly proceeded due to the special characteristics of these languages. Three important steps in handwritten recognition are pre-processing, feature extraction and classification. The aim of preprocessing is to improve the quality of the images for further processing and analysis task [12]. Handwritten recognition performance largely depends on the feature extraction approach. For extract features from word's image various approaches are proposed. The most commonly used features are: zoning [13,14], Gradient [15], Projection histogram [16], Wavelet [17], crossing points and contours [18]. Some of these methods extract high dimension feature vectors. The selection of high performing and scale invariant feature extraction method is an important but difficult task in developing Farsi handwritten word recognition (FHWR) systems. In this paper, a novel method for feature extraction presents. For extracting the features, images should be passed from the thinning stage that caused reduction of samples thickness into a single pixel [19]. Four of the 3×3 masks are applied to the thinned word images to extract horizontal, vertical, right and left-diagonal lines. Images store separately correspond to each line. Afterward, decomposed images should be partitioned to eight sectors around the center of image and the number of black pixels in each sector calculated and normalized by dividing them upon the total number of black pixels in word images for feature vector. This method solves the problem of scale invariance, has a low feature dimension and high recognition rate.

In recent years, many new classifiers have been proposed and tested on various OCR databases. Improving recognition performance in difficult classification problems [20,21] is objective of multiple classifiers combination. There are two main strategies in combining classifiers: classifier selection and classifier fusion [22]. Classifier selection attempts to choose the best classifier for a given task. This method assumes that a classifier is an expert on a subset of the feature space. In classifier fusion assume that all classifiers are trained over the whole feature space [23-25]. In this paper, a templates based classifier fusion method is proposed to work out the handwritten word recognition problem. First, for training samples the decision profile matrix is constructed. Decision templates are composed with averaging observed decision profiles for each class among training. The higher similarity between decision profiles of test samples and the decision template for each class is causing the higher support for that class [26]. The decision templates fusion uses all classifier outputs to calculate the final support for each class. This method is in sharp contrast to most other fusion methods which use only the support for that particular class to make their decision. The paper is organized as follows. Section 2 introduces new feature extraction method. Section 3 describes proposed model for fusion in details. Simulations and results are discussed in section 4 and conclusions are presented in Section 5.

2. Correlation Reduction Strategies in Multiple Classifier Systems

The The diagram in Figure 1 illustrates application of diverse classifiers which are used to build an ensemble system. There are various methods to create diversity between classifiers, which will be described in the following sections [23].

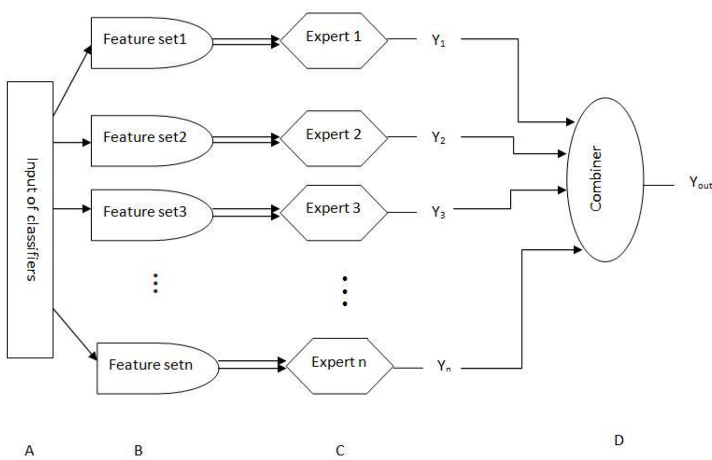

Figure 1. Shows different stages in combination method. Structure of combining methods has four levels. Level A is Data level that uses different data subsets. Level B is Feature level, in which, different feature sets should be extracted from the input data. Level C is Classifier level, where the basic classifiers are and finally Level D is Combination level, where the combiner is set.

Since classifiers are made through a training procedure, to get diverse ones, they should be trained differently. The training stage can be affected by representation of input patterns, training samples, learning procedure and supervision strategy. In addition, correlation reduction techniques are based on these mentioned items. There are many different strategies to make diverse classifiers; some of them are explained as follows:

2.1. Different representation of patterns

There are three ways to perform this task:

1. Using different feature sets or rule sets.
   This method is useful particularly when more than one source of information is available [23, 24].
2. Using decimation techniques
   Even when only one feature set exists, we can produce different feature sets [25] which can be used in training different classifiers by removing some parts of the set.
3. Feature set partitioning in a multi-net system
   This method can be used when patterns include independent parts. For example, different parts of an identification form or a different handwritten of words or digits. In this case, patterns can be divided into sub-patterns each one can be used to train one of the classifiers. A theoretical property of neural networks is the fact that they do not need special feature extraction for classification. If the number of feature sets are too high and are applied to a single network, the curse of dimensionality may occur. To avoid this problem they should be divided into independent parts and each one will be applied as the input of a sub-network [26, 27].

2.2 Different Learning Machines

There are some free parameters in any learning machine which should be set during training. The final set of these parameters depend on the training set, so even for a given structure and an identical representation of patterns, different training sets could cause different generalizations. Using identical representation of patterns has the advantage that the decision boundary of individual classifiers is located in the same axis set (space). Therefore, the effect of each sample or expert in composite classifier can be investigated. In these circumstances correlation reduction is based on partitioning the main training set into a few subsets, and using these partitions to train different experts. If partitions are not overlapped, the classifiers will be more independent, but in most practical cases due to limited number of training samples, the partitions made by perturbing the original training set, may have some overlap [28].

2.3 Different Labeling in Learning

As mentioned earlier, to classify an input pattern, it should be assigned to one of the several possible classes. Based on this fact, in supervised learning various pairs of input- target are used. [29].

3. Feature Extraction

One of the important stages in domain recognition is feature extraction. The performance of recognition largely depends on this stage. To extract features from an image of a word, a variety of approaches are proposed [48]. In this section, two more appropriate feature extraction methods for Farsi handwritten recognition are described.

3.1. Gradient

For this type of feature, Sobel filters for grey scale images are used. They are two types of masks, which are used to distinguish horizontal and vertical edges .The absolute value of the filtered image is used to derive the feature, as well as the values of horizontal and vertical Sobel masks. The gradient vector is decomposed into eight Freeman directions, if not consistent with any of them, the vector will be projected into the nearest two Freeman directions. The image is divided into 8 sub-images and for each sub-image, by applying the Gaussian mask features are extracted [1, 30, 31].

3.2. Gradient Based Feature Extraction Method

To extract the gradient based features, images should be binarized and passed through the thinning stage. This stage causes the thickness of samples be reduced into a single pixel. Thinned word images decompose into four separate images by applying four masks. These masks separate the 0, 45, 90 and 135 degree lines from word images and corresponding to each degree, images store separately. Afterwards, decomposed images should be partitioned to eight sectors around the center of image and the number of black pixels in each sector should be calculated and normalized by dividing them upon the total number of black pixels in word images and will be used as a feature vector [2].

3.3. Principal Component Analysis (PCA)

In some situations, the dimension of the input vector is large, and the components of the vectors are highly redundant. In this situation it is useful to reduce the dimension of the input vectors which causes less computational load. An effective procedure for performing this operation is principal component analysis. PCA allows us to compute a linear transformation that maps data from a high dimensional space to a lower dimensional space [29, 32].

3.4. Linear Discriminate Analysis (LDA)

LDA features are extracted by projecting original intensity images onto discriminate vectors. Linear discriminate analysis is a well known method for estimating a linear subspace with good discriminative properties. The idea is to find a projection of the data where the variance between the classes is large compared to the variance within the classes [26, 27].

4. Proposed New Combination Method Based on Different Representation of Data

As mentioned there are so many different ways to extract the futures. The results for some of the most applicable methods in Farsi handwritten recognition with a Multi Layer Perceptron (MLP) categorization was calculated in each class that the results have been shown in Figure 2. According to Figure 2, within the best methods Gradient and Gradient based methods complete each other, as the classes which were not properly recognized by a feature is suitably identified by the other one.

By these two features, the inputs for this problem can be divided into two categories with 15 classes. This structure is shown in Figure 3.

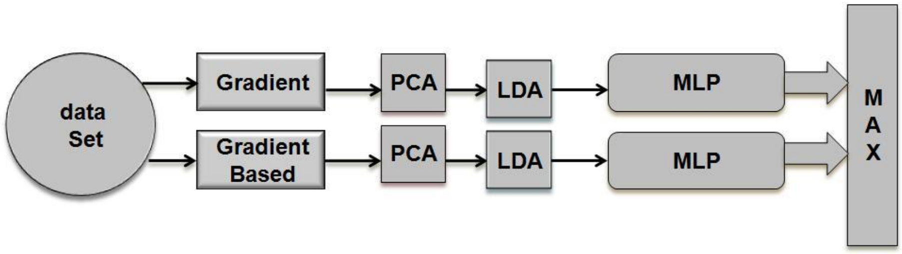

Figure 3. Primary Structure of Proposed Method

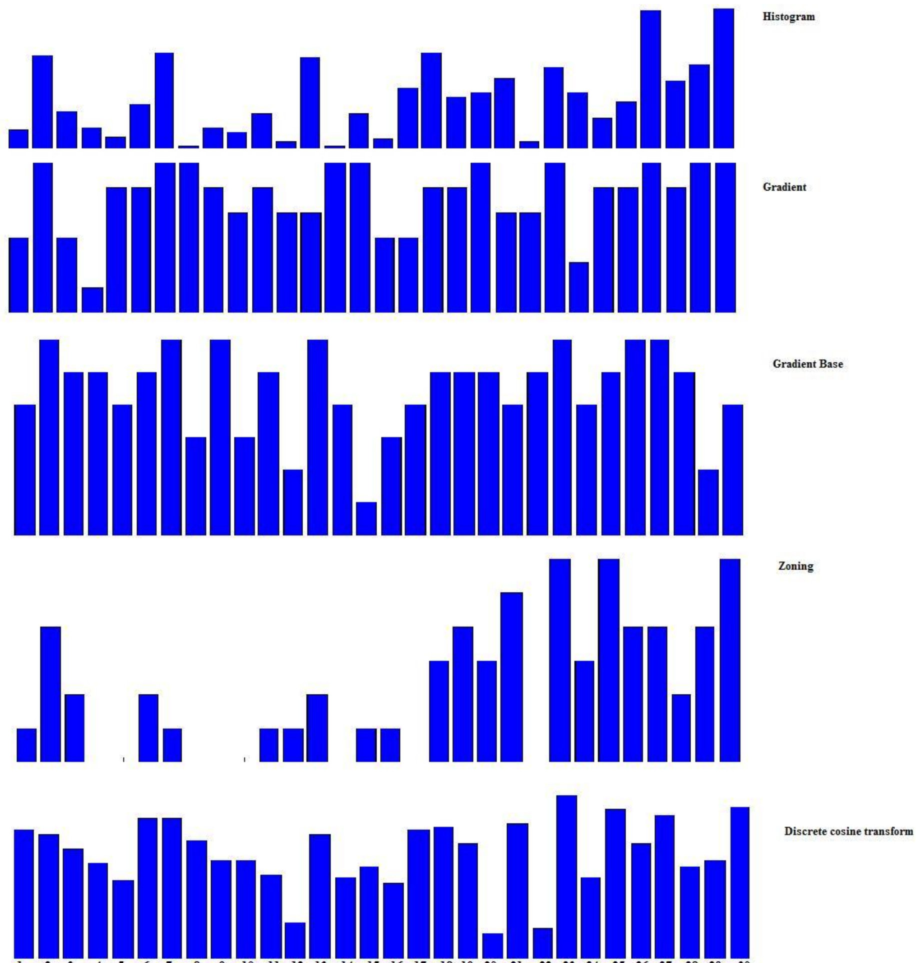

Figure 2: Results of the best feature extraction methods for 30classes of Farsi handwritten problem. Shown in the following order Histogram, Gradient, Gradient based, Zoning and discrete cosine transform.

According to the above figure the problem is splitted intot two categories with different features. As dimensions of the input are not the same, the features were initially changed by PCA=30 and then by LDA to reach to a uniform feature dimension. One of the most prominent characteristics of the latter feature is to make the patterns within the same class close to each other and apart different classes from each other. Each of these features will be given to an MLP with an individual structure from the others. As each of these MLPs was involved in reidentification of a 15-class problem, the results of each of them will be a fifteen-element vector which will be transferred to maximum operator together with the results of the other MLP, and the winner class will be selected. To complete the proposed procedure, a selective Radial Basis Function (RBF) network was allocated to help the decision making. This network solves the task as a two-class problem. In such a way that includes two neurons in the output layer, each one refers to one of the basic experts. The greatest value of these two shows the more effective MLP. This MLP is being weighed by the value of the corresponding neuron. Later on, the results of the chosen MLP will be sent to the maximum operator and the winner class is selected. The proposed structure is shown in Figure 4.

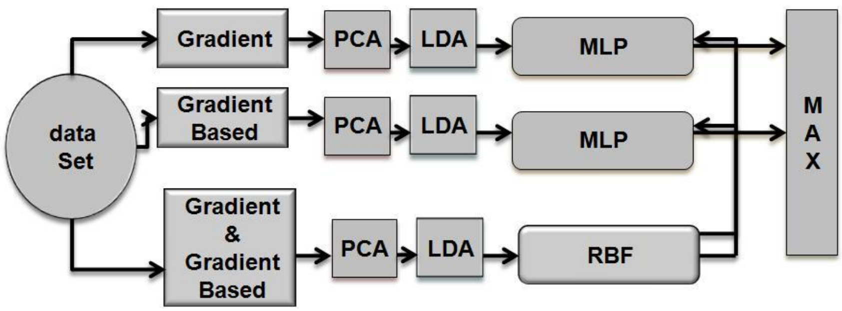

Figure 4. Final Structure of Proposed Method

5. Experimental Results

The proposed combining method which is introduced in this research was tested on a dataset of names of the cities of Iran (Iranshahr dataset). The Iranshahr dataset consists of 780 samples out of 30 names of the cities. For each name 26 samples are available which were written by 26 different persons. In this paper 20 samples had been selected for training and 6 for test, so training set consists of 600 samples and test set is made up of 180 samples. All of the samples are scanned at 96 dpi resolution in gray scale format. Some sample images are shown in Figure 5.

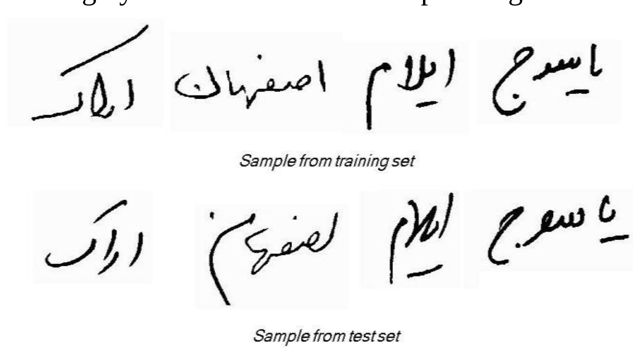

Figure 5.Some Sample Images.

As described in the previous section, the MLP is used as the base classifier with one hidden layer that it is trained using the back–propagation algorithm minimizing the squared error criterion.

Each MLP is being trained by different learning sets, one of them being trained by the result of Gradient feature and the other by Gradient based feature. Since these two feature sets have different dimensions, PCA algorithm had to be taken before they could be used as a base classifier input. Results of different PCAs are shown in Table1.

Table1. Different PCAs Calculated for Equalizing the Dimensions of Features and Choosing the Best PCA

	10	20	30	Average Accuracy	Rank
Gradient	51.41	63.2	64.5	59.7	2
Proposed method	67.52	74.74	75.73	72.66	1
Rank	3	2	1		

As the above table shows, PCA=30 is the best one. At the next step the MLPs of two base experts had 30 input nodes due to dimension of feature vector, and 15 output nodes corresponding to 15 classes. The selector section has 60 input nods and 2 output nods.

The MLP has different learning parameters, such as; number of epochs, learning rate and number of neurons in hidden layer. The required number of epochs to reach the highest recognition rate was estimated by four-fold cross validation on the training set. Figure 6 shows the diagram of selected epochs and in Table 2 the best topologies for base classifiers are shown. Note that experiment has been repeated 10 times and the best structure for basic experts have been chosen.

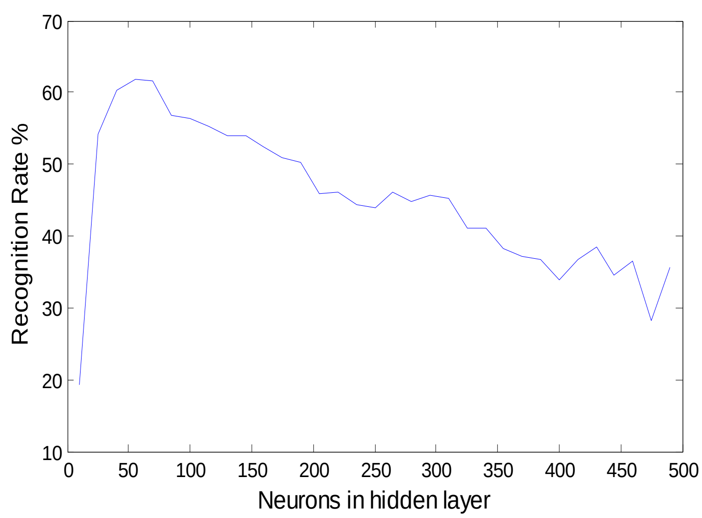

Figure 6. Diagram of Selected Epochs

Table 2- Shows the best topologies for base classifiers to have the highest recognition rate. Note that experiment has been repeated 10 times to get the recognition rates.

	Structure	Performance
First Expert	30:40:15 Eta:0.1 Epoch: 400	74.05
Second Expert	30:40:15 Eta:0.1 Epoch: 400	76.27

To find the best structure for proposed method, different structures were examined and the best form was chosen.

Table 3. Different Structures and the Associated Results for the Proposed Procedure.

Structure	Expert1	Expert2	Selector	Performance
1	MLP	MLP	MLP	80.5
	91.3	93.17	82.34	
2	KNN	KNN	KNN	82.22
	76.12	76.12	87	
3	RBF	RBF	RBF	88.61
	87.65	89	93.7	
4	MLP	MLP	RBF	91.11
	93.45	95	96.6	

According to the above table, three structures were tried out for basic and selector experts, Multi Layer Perceptron, Radial Basis Function and K-Nearest Neighbor are assessed. According to the results, the MLPs as basic expert and RBF as selector had higher percentage which resulted in a recognition rate of 91.11 percent.

To show the robustness of the proposed method, the result was compared with some other combining methods. Each of these methods applies four basic experts. The results of recognition domain for different methods on Iranshahr dataset are shown in Table 4.

Table 4. The Results of Recognition Domain for Different Methods on Iranshahr Dataset

Fusion methods	Min	Max	Product	Sum	Averaging	Weighted Averaging	Genetic Algorithm	Stack Generalization
Gradient	75.78	79.22	81.1	68.33	80.34	83.22	87.78	87.9
Gradient base	82.94	85.22	88.5	87.38	88.22	88.77	88.97	89.12

After comparing the best performances in both Tables 3 and 4 robustness, of the proposed method became obvious. Proposed method has %1.99% improvement in recognition rate.

6. Conclusion

This paper presented a new combination method for Farsi handwritten word recognition. The proposed method consists of two stages, the first stage applies two basic experts and the second one includes only one selector. For the basic experts, Multi Layer Perceptron (MLP) and for the selector, Radial Basis Function (RBF) was used. In order to evaluate the proposed method, the result was compared with some other combining methods. The experiments were conducted on Iranshahr dataset and the proposed algorithm was validated with 91.11% of recognition rate.

References

[1] Sh. Abdleazeem, and E. EL-sherif, "Arabic handwritten digit recognition," IJDAR, Vol.11, pp.127-141, 2008.
[2] R. Ebrahimpour, R. D. Vahid and B.M. Nezhad, "Decision Templates with Gradient based Features for Farsi Handwritten Word Recognition", International Journal of Hybrid Information Technology, 2010.
[3] H. Choi, S.J. Cho, and J. H. Kim "Generation of Handwritten Characters with Bayesian network based On-line Handwriting Recognizers," Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003) 0-7695-1960-1/03 $17.00 © 2003 IEEE
[4] F.ch Li and F. Guan: 'heuristic Model Research on Decision Tree Algorithm'. Intelligent Interaction and Affective Computing, 2009. ASIA '09. International Asia Symposium on pp. 149 – 152, 2009.
[5] DS. Lee, and SN. Srihari, "A theory of classifier combination: the neural network approach." Proceedings of the third international conference on document analysis and recognition, Montreal, Canada, pp. 42–5, 1995.
[6] G. Giacinto, F. Roli, and L, Bruzzone, "Combination of neural and statistical algorithms for supervised classification of remote-sensing images," Pattern Recognition Letters, Vol. 21, pp. 385-397, 2000.
[7] CJC. Burges, "A tutorial on support vector machines for pattern recognition," Knowl Disc Data Min, Vol. 2, pp.1–43, 1998.
[8] G. M. FUNG, "Multi category Proximal Support Vector Machine Classifiers," Machine Learning, Vol. 59, pp. 77–97, 2005.
[9] T.K. Ho, "Multiple classifier combination: Lessons and the next steps. In A. Kandel and H. Bunke editors, Hybrid Methods," Pattern Recognition. World Scientific Publishing, pp. 171–198, 2002.
[10] L. I. Kuncheva, "Combining Pattern Classifiers, Methods and Algorithms," INC., PUBLICATION,2004.
[11] R. Polikar, "Ensemble Based sysems in Decision Making". IEEE CIRCUITS AND SYSTEMS MAGAZINE, Vol. 06, pp. 1531-1636, 2006.
[12] S.J. Soltysiak, "Visual information in word recognition: word shape or letter identities," in: Proceeding Workshop Integration of Natural Language and Vision Processing, pp. , 1994.
[13] M. Leung, and A.M. Peterson, "Scale and Rotation Invariant Texture Classification," Proceeding International Conference Acoustics, Speech, and Signal Processing, pp. 461–165, 1991.
[14] K. Woods, W.P. Kegelmeyer, and K. Bowyer, "Combination of multiple classifiers using local accuracy estimates," IEEE Trans. Pattern Anal. Mach. Intell, Vol.19 , pp.405-410, 1997.
[15] L. Xu, A. Krzyzak, and C.Y. Suen, "Methods of combining multiple classifiers and their application to handwriting recognition,' IEEE Trans. Systems Man Cybernet. Vol.22, pp. 418-435, 1992.
[16] K.-C. Ng, and B. Abramson, "Consensus diagnosis: a simulation study,' IEEE Trans. Systems Man Cybernet. Vol.22, pp. 916-928, 1992.
[17] L. I. Kuncheva, James C. Bezdek, and Robert P.W. Duin, "Decision Templates for Multiple Classifier Fusion: An Experimental Comparison," Pattern Recognition, vol. 34, pp. 299-314, 2001.
[18] L. Xu, A. Krzyzak, and C.Y. Suen, "Methods of combining multiple classifiers and their application to handwriting recognition," IEEE Trans. Systems Man Cybernet. Vol.22, pp. 418-435, 1992.
[19] K.-C. Ng, B. Abramson, "Consensus diagnosis: a simulation study," IEEE Trans. Systems Man Cybernet. Vol.22, pp. 916-928, 1992.
[20] R.A. Jacobs, M.I. Jordan, S.J. Nowlan, and G.E. Hinton, "Adaptive mixtures of local experts," Neural Comput. Vol.3, pp. 79-87, 1991.
[21] L.A. Rastrigin, and R.H. Erenstein, "Method of Collective Recognition," Energoizdat, Moscow, 1982 .
[22] E. Alpaydin, and M.I. Jordan, "Local linear perceptrons for classification," IEEE Trans. Neural Networks, Vol.7, pp. 788-792, 1996.

[23] R. ghaderi, "Arranging simple neural networks to solve complex classification problems," PHD thesis from University of Surrey, 2000.
[24] JH. Holland, "Adaptation in natural and artificial systems," University of Michigan Press; 1975.
[25] J. Kittler, A. Hojjatoleslami, and T. Windeatt, "Weighting factors in multiple expert fusions," In Proc. of British Machine Vision Conference BMVC97, Essex University, Essex U.K, pp. 42-50 1997.
[26] K. Tumer and J. Ghosh, "Error correlation and error reduction in ensemble classifiers," Connection Science, special issue on combining arti_cial neural ne tworks: ensemble approaches, Vol. 8, pp 385-404. 1996.
[27] H.A. Rowley, "Neural Network-Based Face Detection," PhD thesis, School of computer science, Computer scince Department, Carnegie Mellon universit, Pittsburg, PA 15213, 1999.
[28] T. Windeatt and R. Ghaderi, "Dynamic weighting factors for decision combining," International Conf. on Data Fusion, Great Malven, U.K., Vol. pp. 123-130, 1998.
[29] R. Ghaderi, "Arranging simple neural networks to solve complex classification problems," Submitted for the Degree of Doctor of Philosophy from the University of Surrey.
[30] N. Ahmad, T. Natarjan, and K,Roa, "Discrete Cosine Transform," IEEE. Trans Compute., Vol C-23, PP 90-93, 1974.
[31] C.L. Liu., Y.J. Liu., and R.W. Dai, "Preprocessing and statistical/structural feature extraction for handwritten numeral recognition," A.C. Downton, S. Impedovo (Eds.), Progress of Handwriting Recognition, World Scientific, Singapore, Vol. , pp. 161–168,1997.
[32] Sh. Abdleazeem, and E. EL-sherif, "Arabic handwritten digit recognition," IJDAR, Vol.11, pp.127-141, 2008.

Authors

Reza Ebrahimpour is an Assistant Professor at the Department of Electrical and Computer Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran. He is also a research scientist at School of Cognitive Sciences, Institute for Research in Fundamental Sciences, Tehran, Iran. He received the B.S. degree in Electronics Engineering from Mazandaran University, Mazandaran, Iran and the M.S. degree in Biomedical Engineering from Tarbiat Modares University, Tehran, Iran, in 1999 and 2001, respectively. He obtained Ph.D. degree in the field of Computational Neuroscience from the School of Cognitive Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran in July 2007. Dr. Ebrahimpour is the author or co-author of more than 60 international journal and conference publications in his research areas, which include Human and Machine Vision, Computational Neuroscience, Neural Networks, Pattern Recognition and Multiple Classifier Systems.

.

Mona Amini was born in Tehran Iran, in August 1984. She received the BS degree in electronic engineering from Azad University, Qazvin, Iran, in 2007. She is student in MS degree of Mechatronic Engineering in Azad University, Tehran, Iran. Her research interests include human and machine vision, neural networks and pattern recognition.

Afra Vahidi Shams was born in Tehran, Iran (Octobr 1984). He received the B.Sc. degree in electronics engineering from Mazandaran University(Noshirvani University of Technology), Babol, Iran. in sep. 2007 and his M.Sc. degree in mechatronics engineering from Islamic Azad University of Tehran (south branch), Tehran, Iran in sep. 2010. His research interests include Intelligent Robotic Systems and Humanoids,Human and Machine Vision, Biologically Motivated Object Recognition, Computational Neuroscience.