Human Face Recognition Using Radial Basis Function Neural Network

Javad Haddadnia Ph.D Student
Department of Electrical and Engineering
Amirkabir University of Technology
Hafez Avenue, Tehran, Iran, 15914
E-mail: H7423970@cic.aku.ac.ir
FAX: (+98)-21-6406469

Karim Faeze Prof.
Department of Electrical and Engineering
Amirkabir University of Technology
Hafez Avenue, Tehran, Iran, 15914
E-mail: kfaez@cic.aku.ac.ir
FAX: (+98)-21-6406469

Abstract

A neural network based face recognition system is presented in this paper. The system consists of two main procedures. The first one is face features extraction using Pseudo Zernike Moments (PZM) and the second one is face classification using Radial Basis Function (RBF) neural network. In this paper, some new results on face recognition are presented. Simulation results indicate that PZM with RBF neural network produce higher detection and lower missing rates than several existing state-of-the-art face detection systems, with an average false detection rate. Also experimental results show that high order degrees of PZM contain very useful information about face recognition process. The proposed system has been applied on face database of Olivetti Research Laboratory (ORL) with very good results.

1 Introduction

Automatic face recognition has received significant attention from the communities of computer vision, neural network and signal processing [1]. The interest is motivated by applications such as access control systems, model-based video coding, criminal identification and authentication in secure system like computer or bank teller machines. Although many face recognition by human beings and machines, it is still difficult to design an automatic system for the task because in real world, illumination, complex background, visual angle and facial expression for face images are highly variable [1,2]. Several methods have been proposed for face detection, including graph matching [3], neural networks [4,5,6], and also geometric feature based [7]. Generally in the procedure of machine face recognition two issues are central: (1) what features can be used to represent a face under environment changes?, and (2) how to classify a new face image, based on the chosen features, into one of the possibilities. In first issue, many successful face feature extraction procedures have been presented and developed. In this paper we used PZM based on central moments as face features. The advantages of considering PZM are that they are shift, rotation and scaling invariant and very robust against noise. Also simulation results indicate that these features are very robust against change of face expression. In the other hand, PZM transform the input image into very low dimensional features vector, this optimization of feature set allows the designer to focus on complex and correct classifier.

In second issue, classifier plays an essential role in the face detection process. In many face recognition systems, the Nearest Neighbor is widely used for classification. Neural network-based (NN) classifier has been proven to have many advantages for classification such as incredible generalization and good learning ability. The NN approach takes the features vector as input and training a network to learn a complex mapping for classification and using of the NN for classification avoids the need for the simplification of the classifier. In this work, face classifier is implemented via RBF neural network to take advantage of NN approaches.

The organization of this paper is as follow: section 2 describes the face feature extraction method. Face classifier technique is presented in section 3. Experimental results are shown in section 4 for comparing our system with other systems.

2 Face features extraction

The invariance properties of moments of images have received considerable attention in recent years. The term invariant denotes an image feature remains unchanged if that image undergoes one or a combination of the changes such as: change of size (scale), change of position (translation), change of orientation (rotation), and reflection. Above properties of moments, occurred that moments have been proposed as pattern sensitive features in classification and recognition applications.

In this paper Pseudo Zernike Moments (PZM) is proposed as facial features in face recognition system. PZM as statistical features are very ease of use, computing and extraction, also the advantage of PZM is that they are very robust against noise. Also simulation results indicates that PZM as face features are very robust against change of face expression.

2.1 Pseudo Zernike Polynomials and Moments

Zernike and Pseudo Zernike polynomials are well known and widely used in the analysis of optical systems. Pseudo Zernike Polynomials are an orthogonal set of polynomials of following form:

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

Where $V_{nm}(x,y)$ denotes a complete set of complex-valued polynomials, in two real variables x and y, which are orthogonal in the interior of the unit circle, n represents the degree of the polynomials, m represents its angular dependence, $R_{nm}(x,y)$ represents a real-valued set of polynomials inside the unit circle as follow:

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

$$S_{n,|m|,S} = (-1)^S \frac{(2n+1-S)!}{S!(n-|m|-S)!(n-|m|-S+1)!} \quad (3)$$

The Pseudo Zernike Moments are defined as follow:

$$PZM_{nm} = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)V^*_{nm}(x,y) \quad (4)$$

From the above equations, it is obviously determined that $PZM_{n,m} = PZM^*_{n,-m}$ and these moments only computed for positive value of m.

2.2 PZM based on Central Moments

For remaining shift invariant property of moments, we used central and radial moments for computing PZM. This done as follows:

$$PZM_{nm} = \frac{n+1}{\pi}\sum_{(n-m-s)even, s=0}^{n-|m|} S_{n,|m|,s}\sum_{a=0}^{k}\sum_{b=0}^{m}\binom{K}{a}\binom{m}{b}(-j)^b CM_{2k+m-2a-b,2a+b}$$
$$+\frac{n+1}{\pi}\sum_{(n-m-s)odd, s=0}^{n-|m|} S_{n,|m|,s}\sum_{a=0}^{d}\sum_{b=0}^{m}\binom{d}{a}\binom{m}{b}\binom{K}{a}\binom{m}{b}(-j)^b RM_{2d+m-2a-b,2a+b} \quad (5)$$

Where $k=(n-s-m)/2$ and $d=(n-s-m+1)/2$, $CM_{i,j}$ is the central moments and $RM_{i,j}$ is the radial moments.

For each face image these moments are computed as face features.

3 RBF NN-based Classifier

RBF neural networks have recently attracted extensive research interests in community of neural networks because: (1) they are universal approximations, (2) they have very compact topology, (3) their learning speed is very fast because of local-tuned neurons, (4) they possess the best approximation property. In this paper, RBF neural network is used as classifier in face detection system.

3.1 Structure of RBF Neural Network

Figure 1 is showed the basic structure of RBF neural networks.

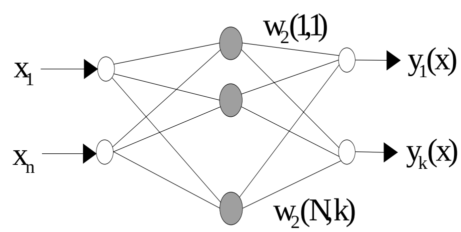

Figure1: RBF neural network

The output of the *i*th RBF unit is as follow:

$$R_i(x) = R_i\left(\frac{\| x - c_i \|}{\sigma_i}\right),\ i = 1,2,...,n \quad (6)$$

where x is an input feature vector with r dimensional, c_i is a r-dimensional vector named center of RBF node, n is the number of hidden node . Typically, $R(x)$ is chosen as a Gaussian function as follow:

$$R_i(x) = \exp\left[-\frac{\| x - c_i \|^2}{\sigma^2}\right] \quad (7)$$

The jth output of RBF neural network is:

$$y_j(x) = b(j) + \sum_{i=1}^{N} R_i(x) \times w_2(j,i) \quad (8)$$

Where $w_2(j,i)$ is the weight of the *i*th RBF node to the *j*th output node and $b(j)$ is the bias of the *j*th output. The bias is not considered in this in order to reduce network complexity. Henc:

$$y_j(x) = \sum_{i=1}^{N} R_i(x) \times w_2(j,i) \quad (9)$$

3.2 Classifier Design

For designing classifier based on RBF neural network, we set the number of input nodes in input layer of NN equal to the number of features that determined based on number of PZM. The number of nodes in output layer is set to the number of image classes. The selecting RBF nodes we do following steps:

1) We initially set the number of RBF nodes equal to outputs.

2) For each class s, s=1,2,…,k, the center of RBF nodes is selected as the mean value of the sample feature belonging to the class, i.e.

$$C^k = \frac{\sum_{i=1}^{N^k} p^k(r,i)}{N^k} \quad (10)$$

where $p^k(r,i)$ is the *i*th sample with r-dimention(number of PZM is r) belonging to class *k* and N^k is the number of image in class *k*.

3) For any class *k*, compute the distance d_k from the mean to the furthest point p_f^k belonging to class *k:*

$$d_k = | p_f^k - C^k | \quad (11)$$

4) For any class *k*, compute the distance $dc(k, j)$ between the mean of class *k* and the mean of other classes as follow:

$dc(k, j) = | C^k - C^j |$, *j*= 1, 2….s, $j \neq k$

Then find $d_{\min}(k,l) = \min(dc(k, j))$ and check the relationship between $d_{\min}(k,l)$ and d_k, d_l. If $d_k + d_l \leq d_{\min}(k,l)$ then class *k* has no overlapping with other classes, otherwise class *k* has ovrlapping with other classes and misclassifications may occur in this case.

5) For all the training data, check how data are classified.

6) Repeat step 2 to 5 untill all the training sample patterns are classified satisfactorily.

7) The mean values of the classes are selected as the centers of RBF nodes.

In this paper, a hybrid learning algorithm, which combines the gradient method and Linear Least Squared (LLS) method to adjust the parameters is used that presented in [8].

4 Simulation Results

Experimental Studies are carried out on the ORL (Olivetti Research Laboratory) database image of Cambridge University. In this database the total number of images for each person is 10. None of the 10 samples are identical to each other. They vary in position, rotation, scale and expression. The change in orientation has been accomplished by rotating the person a few degrees (maximum 20 degree) in the same plane, and also each person has changed his face expression in each 10 samples. The change in scale has been achieved by changing the distance between the person and the video camera. Each image was digitized and presented by 112*92 pixel array whose gray levels ranged between 0 and 255. One sample of these images is shown in figure 2. Like the experiment of [9] and [10] we also use a database of 400 images of 40 individuals. A total of 200 images are used to train and another 200 are used to test, where each person has 5 images.

In feature extraction step with Pseudo Zernike Moments, simulation has been done in three step based on degree of the Pseudo Zernike Polynomials (n). Experimental results are shown in Table 1 and Table 2 respectively. In Table 1 training data used as training images and testing data as testing images and in Table 2 testing data used as training images and training data as testing images.

We also define the average error rate as follow:

$$E_{ave} = \frac{\sum_{i=1}^{m} N_m^i}{mN_t} \qquad (12)$$

Where m is the number of experimental runs, each being performed on random partition of the database into sets, N_m^i is the number of misclassification for the *i*th run, and N_t is the number of total testing images for each runs. The comparison with Convolutional Neural Network (CNN) approach [9] and Nearest Feature Line (NFL) approach [10] using the same ORL database in terms of average error rate is shown in Table 3.

Table 3: Error rate in different methods

Methods	E_{ave} %
CNN	3.83
NFL	3.125
Our Method	1.2

The lowest error rate achieved by our method is based on these condition: m=3, Number of PZM is 21. The way to partition the training set and query set is the same as that of [9] and [10]. For NFL method, the best error rate is the average of the error rates obtained on the condition: m=4, number of features is 40, and query set is the same as that of [9] and [10]. For NFL method, the best error rate is the average of the error rates obtained on the condition: m=4, number of features is 40, whereas the average error rate obtained by CNN in Table 3 is based on m=3.

5- References

[1] R. Chellappa, C.L.Wilson, and S.Sirohey,"Human and Machine Recognition of Faces: A Survey." Proceeding of the IEEE, 83(5): 705-740,May 1995.

[2] J. Daugman, "Face and Gesture Recognition: Overview." IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 675-676, July 1997.

[3] L. Wiskott, J. Fellous, N. Kruger and C.Malsburg, "Face Recognition by Elastic Bunch Graph Matching." IEEE Trans. on Pattern Recognition and Machine Intelligence, Vol 19, No. 7, pp. 775-779, July 1997.

[4] S.Z. Li and J. Lu, "Face Detection by Neural Learning." *Proceeding of ICICS-99*, Singapore, Dec. 7-10, 1999.

[5] E. Osuna, R. Freund and F. Girosi, "Training Support Vector Machines: An Application to Face Detection." *In CVPR*, pp. 130-136, 1997.

[6] H. A. Rowley, S. Baluja and T. Kanade, "Neural Network based Face Detection." *IEEE Trans. On Pattern Recognition and Machine Intelligence*, Vol. 20, No. 1, pp. 23-28, 1998.

[7] G.Z. Yang and T. S. Huang, "Human Face Detection in Complex Background.", *Pattern Recognition*, 27: 53-63, 1994.

[8] J-S. R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System," *IEEE Trans. Syst. Man. Cybern.*, Vol. 23, No. 3, pp. 665-684, 1993.

[9] S. Lawrence, C. L. Giles, A. C. Tsoi and A. D. Back, "Face Recognition: A Convolutional Neural Networks Approach," *IEEE Trans. on Neural Networks*, Special Issue on Neural Networks and Pattern Recognition, Vol. 8, No. 1, pp. 98-113, 1997.

[10] S. Z. Li and J. Lu. "Face Recognition Using the Nearest Feature Line Method," *IEEE Trans. Neural Networks*, Vol. 10, pp. 439-443, 1999.

Table 1: Error rate and parameter

Features		Training		Test	
No. of PZM	PZM Order	No. of Epochs	RMSE(1)	No. of Misclassification	Error rate(2)
25	$0 \le n \le 6$	30 ~ 50	0.03 ~ 0.02	2	1%
24	$6 \le n \le 8$	20 ~ 35	0.02 ~ 0.01	1	0.5%
21	$9 \le n \le 10$	15 ~ 22	0.04 ~ 0.03	0	0

(1)RMSE : Root Mean Squared Error

(2) Error rate = Number of misclassification / Number of total testing pattern

Table 2: Error rate and parameters

Features		Training		Test	
No. of PZM	PZM Order	No. of Epochs	RMSE(1)	No. of Misclassification	Error rate
25	$0 \le n \le 6$	50 ~ 70	0.04 ~ 0.03	6	3%
24	$6 \le n \le 8$	45 ~ 50	0.02 ~ 0.01	5	2.5%
21	$9 \le n \le 10$	30 ~ 40	0.04 ~ 0.03	3	1.5%

Figure2: Sample of Face images on ORL database