PZMI AND WAVELET TRANSFORM FEATURES IN FACE RECOGNITION SYSTEM USING A NEW LOCALIZATION METHOD

Hamidreza Rashidy Kanan
Electrical Engineering Department
Amirkabir University of Technology
Tehran, Iran, 15914
rashidykanan@aut.ac.ir

Karim Faez
Electrical Engineering Department
Amirkabir University of Technology
Tehran, Iran, 15914
kfaez@aut.ac.ir

Abstract – **This paper compares performances of the Pseudo Zernike Moment Invariant (PZMI) and Wavelet Transform features in the application of face recognition. In this study, after preprocessing and face localization of an image, we optimize the exact location of oval shape of face in the image with genetic algorithm which improves the recognition rate. High order PZMI and discrete wavelet transform (Haar wavelet) is utilized to produce feature vectors. In the wavelet transform step, we used Mallat pyramid algorithm for finding approximation of the image in lower resolution and decomposed each image in 4 resolution level. Also RBF Neural Network with HLA learning algorithm has been used as a classifier. Simulation results on ORL database show that approximately the same results are obtained for both PZMI and wavelet features. But feature extraction using wavelet transform has a rate of 0.078 image/Sec that is about 11 times faster than the rate of PZMI feature.**

I. INTRODUCTION

Face recognition is a challenge task in computer vision. It has been an active research area in the past few years because of its wide range of application such as identity authentication, access control, surveillance and intelligent human computer interaction [1]. Face region segmentation or face localization is a fundamental step in the process of face recognition that is shown in Fig. 1 [2]. The accuracy of the localized face center coordinates and orientation has a heavy influence on the recognition performance.

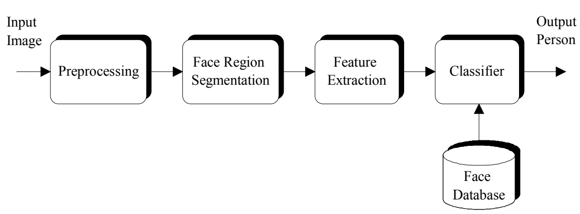

Figure 1. Block Diagram of face recognition system

Pseudo Zernike Moment Invariant is a moment that has been frequently utilized for a number of image processing tasks [2]. The nature of rotation invariance makes the Pseudo Zernike moment descriptors very valuable and it has been demonstrated theoretically and supported by a number of experimental studies. Also, Wavelet analysis and its applications have become the fastest growing research area in recent years. Advanced researches in wavelet analysis have found numerous applications in such area as signal processing, image processing and pattern recognition [3]. The wavelet transform is a powerful technique for representing data at different scales and frequencies.

RBF neural network has been found to be very attractive for many engineering problems due to its simple topological structure, its locally tuned neurons and its ability to have a fast learning algorithm in comparison with the multi-layer feed forward neural networks [2, 4].

In this paper, high order Pseudo Zernike Moment Invariants and wavelet transform with multiresolution analysis concepts have been used for feature extraction of the preprocessed face images. We used the Mallat pyramid algorithm [3] for finding the approximation of the image in lower resolution which contains the global properties of image. Also RBF Neural Network is employed in this system.

The organization of this paper is as follows: Section 2 presents preprocessing and face localization. Feature extraction and Classifier design are described in section 3 and 4 respectively and finally, section 5 and 6 attain the experimental results and conclusion.

II. PREPROCESSING AND FACE LOCALIZATION

One of the key problems in building automated systems in face recognition task is face localization. Many algorithms have been proposed for face localization and detection, which are based on using shape, color information, motion etc. See [5] and the references cited therein.

After preprocessing (histogram equalization) of the facial images, we extract connected components by applying a region growing algorithm. Then the ellipse that is a good approximation of connected components is selected by the shape information method [6].

We calculate the major and minor axes of best-fit ellipse by the approach that is described in [6] and its exact orientation and center coordinates which uses genetic algorithm is described in following.

Because of the background effect, the ellipse that is obtained in previous part is not the best-fit ellipse. So, we must optimize its orientation and center coordinates. In the optimization section, the population was initialized with random points. In order to select the individuals for the next generation, GA's roulette wheel selection method was used. Further genetic parameters are:

Population size: 30

Chromosome length: 23

Probability of crossover: 0.8

0-7803-9252-3/05/$20.00 ©2005 IEEE

Probability of mutation: 0.003

We use two different coding methods (gray and binary) for genotype coding and the symmetric of localized image with respect to major axis of obtained ellipse for fitness function. With this optimization method, the position of ellipse is found precisely while the Tilt Error and Translation Error which are described with the equations (1) and (2) are lower before optimization.

$$Error_{tilt} = \left|tilt_r - tilt_o\right| \tag{1}$$

$$Error_{trans.} = \sqrt{(x_r - x_o)^2 + (y_r - y_o)^2} \tag{2}$$

Where indices of "r" and "o" refer to "real" and "obtained" words. Also (x_r, y_r) and (x_o, y_o) refer to real and obtained center's coordinates of localized face respectively.

Fig. 2 shows the localized faces before and after using the genetic algorithm optimization.

(a) (b) (c)

Figure 2. Localized Face: (a) Original image (b) Before optimization (c) After optimization

III. FEATURE EXTRACTION

A. PZMI Feature Extraction

The advantages of considering orthogonal moments are that they are shift, rotation and scale invariant and very robust in the presence of noise. The invariant properties of moments are utilized as pattern sensitive features in classification and recognition applications [4]. Pseudo Zernike polynomials are well known and widely used in the analysis of optical systems.

Pseudo Zernike polynomials are orthogonal sets of complex-valued polynomials defined as:

$$V_{nm}(x, y) = R_{nm}(x, y)\exp(jm\tan^{-1}(\frac{y}{x})) \tag{3}$$

Where $x^2 + y^2 \le 1$, $n \ge 0$, $|m| \le n$ and $n - |m|$ is even and Radial polynomials $\{R_{nm}\}$ are defined as:

$$R_{nm}(x, y) = \sum_{s=0}^{n-|m|} D_{n,|m|,s}(x^2 + y^2)^{\frac{n-s}{2}} \tag{4}$$

Where

$$D_{n,|m|,s} = (-1)^S \frac{(2n+1-s)!}{s!(n-|m|-s)!(n-|m|-s+1)!} \tag{5}$$

The PZMI of order n and repetition m can be computed using the scale invariant central moments $CM_{p,q}$ and the radial geometric moments $RM_{p,q}$ as follow:

$$\begin{aligned} PZM_{nm} = {} & \frac{n+1}{\pi} \sum_{(n-m-s)even,\, s=0}^{n-|m|} D_{n,|m|,s} \sum_{a=0}^{k} \sum_{b=0}^{m} \binom{k}{a}\binom{m}{b} \\ & (-j)^b S_{n,|m|,s} CM_{2k+m-2a-b,2a+b} \\ & + \frac{n+1}{\pi} \sum_{(n-m-s)odd,\, s=0}^{n-|m|} D_{n,|m|,s} \sum_{a=0}^{d} \sum_{b=0}^{m} \binom{d}{a}\binom{m}{b} \\ & (-j)^b S_{n,|m|,s} RM_{2d+m-2a-b,2a+b} \end{aligned} \tag{6}$$

Where $k = (n-s-m)/2$, $d = (n-s-m-1)/2$ and also $CM_{p,q}$ and $RM_{p,q}$ are as follow:

$$CM_{p,q} = \frac{\mu_{p,q}}{M_{00}^{(p+q+2)/2}} \tag{7}$$

$$RM_{p,q} = \frac{\sum_x \sum_y f(x,y)(\hat{x}^2 + \hat{y}^2)^{1/2} \hat{x}^p \hat{y}^q}{M_{00}^{(p+q+2)/2}} \tag{8}$$

Where $\hat{x} = x - x_0$, $\hat{y} = y - y_0$ and x_0, y_0, M_{pq} and μ_{pq} have been defined in part 2.

B. Wavelet Feature Extraction

The wavelet first mentioned by Haar in 1909, had a compact support which means that it vanishes outside of the finite interval. Mallet gave a boost to digital signal processing in inventing the pyramidal algorithms, and orthonormal wavelet bases. Later Daubechies used Mallet's work to construct a set of wavelet orthonormal basis functions that are the cornerstone of wavelet application today.

A set of wavelet basis functions $\{\psi_{a,b}\}$, can be generated by translation and scaling the basis wavelet, $\psi(x)$, as

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi(\frac{x-b}{a}) \tag{9}$$

Where a>0 and b are real numbers. The parameter "a" reflects the scale of particular basis function, while "b" specifies its translated position along the x-axis.

The wavelet transform coefficients are given as inner products of the function being transformed with each of the basis functions. The discrete presentation of an orthonormal compactly supported wavelet basis of $L^2(R)$ is formed by dilation and translation of basic wavelet. It is assumed that the dilation parameters "a" and "b" take only the discrete values:

$a = a_0^{-j}$, $b = kb_0 a_0^{-j}$ where $k, j \in Z, a_0 > 1, b_0 > 0$.

The wavelet functions may be written as

$$\psi_{j,k}(x) = a_0^{j/2}\psi(a_0^j x - kb_0) \tag{10}$$

For dynamic wavelet transform we have $a_0 = 2, b_0 = 1$, and discrete wavelet transform becomes

$$DWT(f) = \langle f, \psi_{j,k} \rangle = \int_{-\infty}^{\infty} f(x) 2^{j/2} \psi(2^j x - k) dx \qquad (11)$$

Mallet introduced an efficient algorithm to perform the DWT known as the Multiresolution Analysis (MRA). The MRA of $L^2(R)$ consists of successive approximation of the space V_j of $L^2(R)$. A scaling function $\phi(x) \in V_0$ exists such that

$$\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k) \quad j,k \in Z \qquad (12)$$

For the scaling function $\phi(x) \in V_0 \subset V_1$, there is a sequence, $\{h_k\}$,

$$\phi(x) = \sum_k h_k \sqrt{2} \phi(2x - k) \qquad (13)$$

Furthermore, let us define W_j as a complementary space of V_j in V_{j+1} such that $V_{j+1} = V_j \oplus W_j$ and $\bigoplus_{j=-\infty}^{\infty} W_j = L^2(R)$. Since the $\psi(x)$ is a wavelet and it is also an element of V_0, a sequence $\{g_k\}$, exists as

$$\psi(x) = \sum_k g_k \sqrt{2} \phi(2x - k) \qquad (14)$$

The projection of signal $f(x) \in V_0$ on V_j defined by $P_V^j f$ is given by

$$P_V^J f(x) = \sum_{k \in Z} c_{j,k} \phi_{j,k}(x) \qquad (15)$$

Here, $c_{j,k} = \langle f, \phi_{j,k}(x) \rangle$ (as approximation coefficients). Similarly, the projection of function $f(x)$ on subspace W_j is also defined by

$$P_W^J f(x) = \sum_{k \in Z} d_{j,k} \psi_{j,k}(x) \qquad (16)$$

Where $d_{j,k} = \langle f, \psi_{j,k}(x) \rangle$ (as detail coefficients). Because $V_j = V_{j-1} \oplus W_{j-1}$, the original function in the scale "J" can be written as

$$f(x) = \sum_k c_{j_0,k} \phi_{j_0,k}(x) + \sum_{j_0}^{J-1} \sum_k d_{j,k} \psi_{j,k}(x) \quad J > j_0 \qquad (17)$$

The coefficients $c_{j,k}$ and $d_{j,k}$ are given by

$$c_{j,k} = \sum_m h_{m-2k} c_{j+1,m} \qquad (18)$$

and

$$d_{j,k} = \sum_m g_{m-2k} d_{j+1,m} \qquad (19)$$

The two-dimensional wavelet transform can be computed with a pyramidal algorithm similar to one dimensional algorithm. The two-dimensional wavelets transform can be seen as one-dimensional wavelets transform along the x and y axis. At each step we decompose $A_{2^{j+1}}^d f$ (the approximation of the signal f(x,y) at a resolution 2^{j+1}) into $A_{2^j}^d f$, $D_{2^j}^1$ (horizontal edges), $D_{2^j}^2$ (vertical edges), $D_{2^j}^3$ (the corners). We first convolve the rows of $A_{2^{j+1}}^d f$ with one-dimensional filter (Equation 16), retain every other row, convolve the columns of resulting signals with another one-dimensional filter (Equation 19) and retain every other column [3].

IV. RBF NEURAL NETWORK CLASIIFIER

Radial Basis Function neural networks have been found to be very attractive for many engineering problem because: (1) they are universal approximators, (2) they have a very compact topology and (3) their learning speed is very fast because of their locally tuned neurons.

An RBF neural network structure is shown in Fig. (3) The construction of the RBF neural network involves three different layers with feed forward architecture. The input layer of this network is a set of "n" units that are fully connected to the hidden layer with "r" hidden units. Connections between the input and hidden layers have unit weights and, as a result, do not have to be trained. In this structure, hidden layer is named RBF units. The RBF units are also fully connected to the output layer.

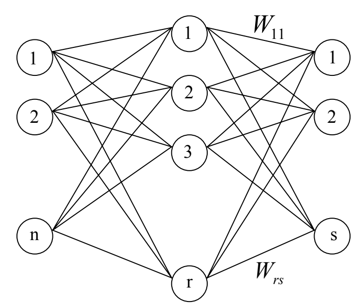

Figure 3. RBF neural network structure

The RBF neural network is a class of neural networks, where the activation function of the hidden units is determined by the distance between the input vector and a prototype vector. The activation function of the RBF units is expressed as follow [2]:

$$R_i(x) = R_i\left(\frac{\|x - c_i\|}{\sigma_i}\right) \quad , \quad i = 1,2,...,r \qquad (20)$$

It should be noted that x is an n-dimensional input feature vector, c_i is an n-dimensional vector called the center of the RBF unit, σ_i is the width of RBF unit and r is the number of the RBF units.

For designing a classifier based on RBF neural network, we have set the number of input nodes in the input layer of neural network equal to the number of feature vector elements. The number of nodes in the output layer is set to the number of image classes. The RBF units are selected using the clustering procedure [2].

Training of the RBF neural network involves estimating output connection weights, centers and widths of the RBF units. We use the HLA method, which combines the gradient method and the linear least squared method for training RBF neural network [2].

V. EXPRIMENTAL RESULTS

For experimental studies, we have considered ORL gray scale face image data sets. This database contains 400 facial images from 40 individuals in different states. The total number of images for each person is 10.

After preprocessing and face localization that is described in part 2, we have face images in which the redundant information such as background and hair are omitted. So, according to simulation results, the performance of the classifier is improved.

Table 1 shows the mean of Tilt Error and Translation Error of localized face images on the ORL database before and after optimization step.

Table 1. Tilt and Translation Errors of Localized images

Method	Tilt Error (degree) ($Error_{tilt}$)	Translation Error ($Error_{trans.}$)
Shape Information Based Method (before optimization)	17.5	26.7
Proposed Algorithm	4.2	5.4

For feature extraction, since the face images are not continuous, we used Haar wavelet which is also discrete. We applied pyramid algorithm to each preprocessed image for decomposing it into 4 resolution levels that is shown in Fig. 4. Then we used the approximation of images at level 4 and converted them into vectors by concatenating the columns.

Dimensions of ORL database images are 92×112, so after decomposing them, the length of wavelet feature for each image is 42. Also, because of PZMI advantages, we extract PZMI of face images with different order in equal length of wavelet feature. Simulation results (Fig. 5) show that the error rate for order of 9 and higher is lower than the other orders.

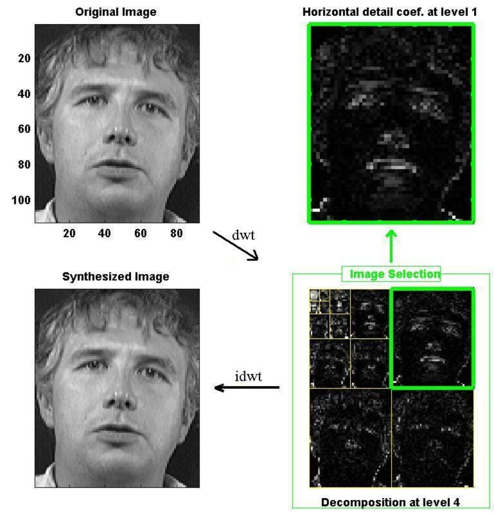

Figure 4. Decomposing of image in to 4 resolution levels

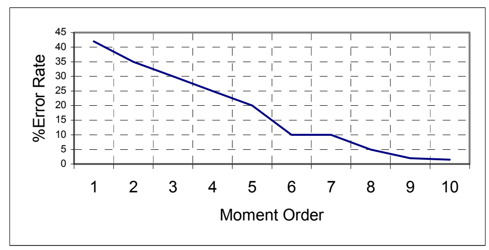

Figure 5. Error rate for different order of PZM

For scale invariancy of extracted features, we normalized them and then apply to RBF Neural Network.

In classifier step, we use RBF neural network with HLA learning algorithm. We use 5 images from each class randomly for training and the rest for the test. There is no overlap between the training and test sets. The RBF classifier is trained using the HLA method with the training sets, and finally the classifier error rate is computed using the test images. In this study, the classifier error rate is computed as the number of misclassifications in the test phase over the total number of test images.

Simulation results on the ORL database show that the PZMI and wavelet are suitable features for face recognition. Using the same train and test sets, approximately the same results are obtained for both PZMI and wavelet features. But feature extraction with wavelet transform has a rate of 0.078 image/Sec that is about 11 times faster than the rate of PZMI feature. These results are summarized in Table 2.

Table 2. Experimental results

	Recognition Rate		**Feature Extraction Time (Image /S)**
	Train Set	Test Set	
PZMI	100%	93%	0.858
Wavelet	100%	88.4%	0.078

VI. CONCLUSION

In this research, we have evaluated two kinds of features (PZMI and discrete wavelet transform) for face recognition. After preprocessing, we find the exact location of oval shape of face (best fit ellipse) in database images with genetic algorithm. Then the high order PZMI and Haar wavelet coefficient are used as a feature vectors which are fed to classifier stage. We have used RBF neural network classifier with HLA learning algorithm. Experimental results on ORL database indicate that approximately the same results are obtained for both PZMI and wavelet features. But feature extraction with wavelet transform has a rate of 0.078 image/Sec that is about 11 times faster than the rate of PZMI feature.

VII. ACKNOWLEDGMENT

This research was supported by the Ministry of Information & Communications Technology (ICT).

VIII. REFERENCES

[1] Chellappa R., Wilson C.L., Sirohry S., "Human and Machine Recognition of faces: A survey", Proceedings of the IEEE, Vol. 83, No. 5, 1995, P.P. 705-740.

[2] Hddadnia, J. Faez K. and Ahmadi M. "AN EFFICIENT HUMAN FACE RECOGNITION SYSTEM USING PSEUDO ZERNIKE MOMENT INVARIANT AND RADIAL BASIS FUNCTION NEURAL NETWORK", Int. Journal of Patt. Recog. and Artificial Intell, Vol. 17, No. 1, 2003 , P.P. 41-62.

[3] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 11, No. 7, 1989, pp. 674-693.

[4] J. Haddadnia, K. Faez and P. Moallem, "Neural Network based Face Recognition with Moment Invariants", Proceedings of ICIP2001, Thessaloniki, Greece, pp. 1018-1021, 2001.

[5] J. Daugman, "Face detection: a survey", Comput. Vis. Imag. Underst. 83, 3, 2001, P.P. 236-274.

[6] K. Sobotta and I. Pitas, "Face localization and facial feature extraction based on shape and color information", IEEE Int. Conf. on Image Processing, Vol. 3, Lausanne, Switzerland, 16-19 September, 1996, P.P. 483-486.