International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:2, No:8, 2008

Rotation Invariant Face Recognition Based on Hybrid LPT/DCT Features

Rehab F. Abdel-Kader*, Rabab M. Ramadan, and Rawya Y. Rizk

Abstract—The recognition of human faces, especially those with different orientations is a challenging and important problem in image analysis and classification. This paper proposes an effective scheme for rotation invariant face recognition using Log-Polar Transform and Discrete Cosine Transform combined features. The rotation invariant feature extraction for a given face image involves applying the log-polar transform to eliminate the rotation effect and to produce a row shifted log-polar image. The discrete cosine transform is then applied to eliminate the row shift effect and to generate the low-dimensional feature vector. A PSO-based feature selection algorithm is utilized to search the feature vector space for the optimal feature subset. Evolution is driven by a fitness function defined in terms of maximizing the between-class separation (scatter index). Experimental results, based on the ORL face database using testing data sets for images with different orientations; show that the proposed system outperforms other face recognition methods. The overall recognition rate for the rotated test images being 97%, demonstrating that the extracted feature vector is an effective rotation invariant feature set with minimal set of selected features.

Keywords—Discrete Cosine Transform, Face Recognition, Feature Extraction, Log Polar Transform, Particle Swarm Optimization.

I. INTRODUCTION

FACE Recognition (FR) has emerged as one of the most extensively studied research areas that spans multiple disciplines such as pattern recognition, signal processing and computer vision [1], [2], [3], [4]. This is due to its numerous important applications in human-computer interactions, authentication, security, and surveillance. FR is a matching process between a query face's features and target face's features. The process becomes difficult because variations in a single face can be very large, while the variations between different faces can be quite small. Furthermore, face information depends on ethnicity and registration method (i.e., capture method, lighting condition, and device). Among all the applications of biometrics identification, face recognition is most suitable for automatic visual surveillance systems.

* Corresponding Author
R. F. Abdel-Kader is with the Electrical Engineering Department, Faculty of Engineering at Port-Said, Suez-Canal University, Port-Said, Egypt; phone: (+20) 0101172820; e-mail: r.abdelkader@scuegypt.edu.eg
R. M. Ramadan is with the Electrical Engineering Department, Faculty of Engineering at Port-Said, Suez-Canal University, Port-Said, Egypt; e-mail: rabab.ramadan58@gmail.com
R. Y. Rizk is with the Electrical Engineering Department, Faculty of Engineering at Port-Said, Suez-Canal University, Port-Said, Egypt; e-mail: r.rizk@ scuegypt.edu.eg

That is the reason why a great deal of research is focusing on developing new algorithms and enhancing the capability and robustness of face recognition; an excellent survey paper on the different face recognition techniques can be found in [1]. However, most of these techniques are only capable of recognizing frontal views of faces assuming that the person was looking straight into the camera. The frontal face recognition approach is adequate in access control applications where the user is consistent from session to session. However, in surveillance applications where the user is often not aware of the task, it is important for the system to handle faces with in plane and in depth rotations. Rotation invariant face recognition is an important issue to address because of its many real-world applications, especially in surveillance. It is clear that if a robust system is created, it will have enormous applications in many different areas in commercial and military technology.

Many common FR techniques include single-template matching, eigenfaces [1] [2], Fisher's discriminant analysis [3], and neural networks [4]. These systems work well for classifying frontal views of faces. However, they are not robust against rotation variations since the whole-face approach is highly sensitive to translations and rotations. In general, there are limitations to many of the existing techniques for rotation invariant face recognition systems, which involve extracting a large number of features and/or require a high-computational complexity. In [5], [6] a rotation and size spreading associative neural network (RS-SAN net) was developed based on space and 3-D shape recognition systems in the brain. Using RS-SAN net, a new personal authentication method, which was not influenced by the rotation and size changes, was proposed. The RS-SAN net learned human faces, and correctly recognized their shape, rotation and size, regardless of their rotation and size, once they were learned. However, in real applications it is impossible to obtain recognition images that are identical to those which were learned. [7] Presents a new technique for rotation invariant face recognition based on Fourier descriptors and neural networks.

In this paper we propose an effective scheme for rotation invariant face recognition using Log-Polar Transform and Discrete Cosine Transform combined features. The rotation invariant feature extraction for a given face image involves applying the log-polar transform to eliminate the rotation effect and to produce a row shifted log-polar image.

International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:2, No:8, 2008

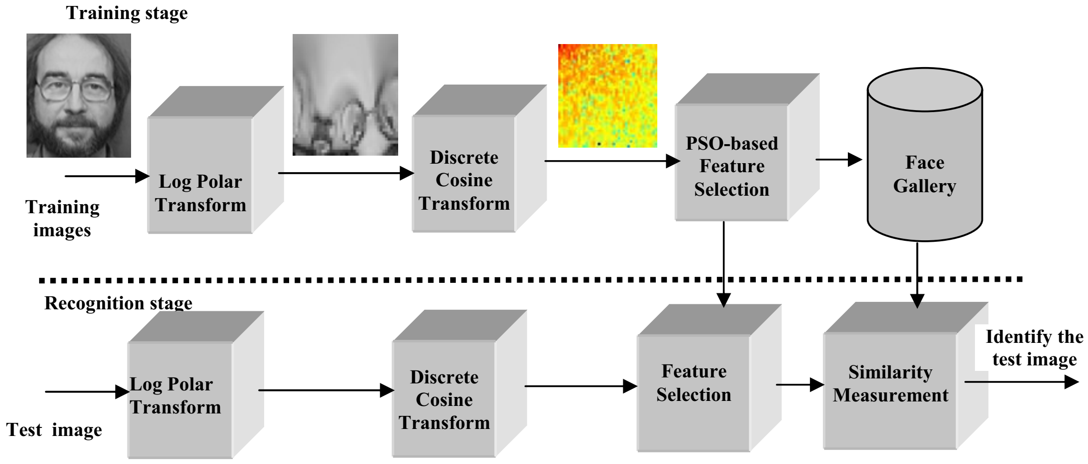


Fig. 1 Block diagram of the proposed face recognition system

The discrete cosine transform is then applied to eliminate the row shift effect and to generate the low-dimensional feature vector.

The main contribution of this work is:

- The development of a complete rotation-invariant face recognition system based on hybrid LPT/DCT feature vectors. Particle swarm optimization based
- feature selection algorithm is utilized in the system to search for the optimal feature subset to increase recognition rate and class separation.
- Evaluation of the proposed system using the ORL face database and comparing its performance with other FR methods.

The outline of this paper is organized as follows: in the next section, the scheme for extracting the rotation invariant hybrid LPT/DCT coefficients is presented. The various components used for feature extraction and feature selection are described. In section III, we discuss the experimental results and the recognition performance of the proposed system. Finally, conclusions are drawn in section IV.

II. PROPOSED FACE RECOGNITION SYSTEM

The proposed rotation invariant face recognition system using the hybrid LPT/DCT Features consists of two stages, namely, the training stage and recognition stage. The training stage represents a set of reference images using the LPT/DCT extracted features and stores them into a gallery (reference library). The recognition stage translates a probe image into a LPT/DCT feature vector, and then matches it with those referent images stored in the gallery to identify the face image. Fig. 1 shows the block diagram of the proposed FR system. The different building components in the system are explained in the following sections.

A. Feature Extraction

The success of any FR methodology depends heavily on the particular choice of the features used by the (pattern) classifier. A typical feature extraction algorithm tends to build a computational model through some linear or nonlinear transform of the data with subsequent feature selection for reducing the dimensionality of facial image so that the extracted feature is as representative as possible. It is known that a good feature extractor for a face recognition system is claimed to select as more as possible the best discriminant features which are not sensitive to arbitrary environmental variations such as variations in pose, scale, illumination, and facial expressions.

Feature extraction algorithms mainly fall into two categories: geometrical features extraction and, statistical (algebraic) features extraction [1], [2], [3], [4]. The geometrical approach represents the face in terms of structural measurements and distinctive facial features that include distances and angles between the most characteristic face components such as eyes, nose, mouth or facial templates such as nose length and width, mouth position, and chin type. These features are used to recognize an unknown face by matching it to the nearest neighbor in the stored database. Statistical features extraction is usually driven by algebraic methods such as principal component analysis (PCA) [3], and independent component analysis (ICA) [4]. These methods find a mapping between the original feature spaces to a lower dimensional feature space. Alternative algebraic methods are based on transforms such as downsampling, Fourier transform (FT), discrete cosine transform (DCT), Log Polar Transformation (LPT) and the discrete wavelet transform (DWT) [8], [9], [10], [11], [12], [13].

Transformation based feature extraction methods such as the LPT and the DCT were found to generate good FR accuracies with very low computational cost. The LPT and the DCT feature extraction methods are explained in detail in the following sections.

Feature Extraction Using Log Polar Transformation (LPT)

The first step in the proposed rotation invariant FR system is the application of the LPT to eliminate the rotation effects in the input image by converting the image into the corresponding log-polar image. The log-polar transform (LPT) is a well-known space-variant image encoding scheme used in computer vision inspired by the process of optical nerves visual image projection in the cerebellar cortex in humans [8] [9], [10], [11].

If I x, y is a rectangular image in the Cartesian coordinates, then the log polar transform ($I^*(\rho,\phi)$) with origin x_0, y_0 is given by:

$$I^*(\rho,\phi) = \{I(x,y)(x_0,y_0)\} \tag{1}$$

Where $\rho = M.log(r)$, (2)

$$r = \sqrt{(x-x_0)^2 + (y-y_0)^2} \tag{3}$$

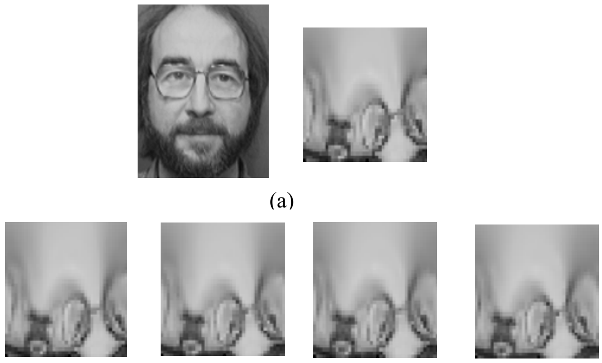
$$\phi = \tan^{-1}\left(\frac{(y-y_0)}{(x-x_0)}\right) \tag{4}$$

Where, $I^*(\rho,\phi)$) is the surface of a cylinder of unit radius. The ρ-coordinate of a point in I^* is proportional to the logarithm of the radial distance from the origin (plus an offset) of the corresponding point in *I*. The ϕ-coordinate is equal to the angular distance of the corresponding point from the *x*-axis in *I*. The parameter *M* is determined by the selected height of the transformed image. The angle resolution is given by the selected width of the transformed image. The LPT allows a significant reduction of the visual data to be processed. This data diminution is produced by the logarithmic sampling of the input signal in the radial direction and by the constant sampling (the same number of points is taken) in each angular sector to be transformed. Additionally, this mapping provides an invariant representation of the objects, because rotations and scaling of the input signal are transformed into translations [8], which can be easily compensated. Fig. 2 shows an original face image and its LPT and the LPT of the rotated versions of the input image with rotation angles $5^\circ, 10^\circ, 15^\circ, 20^\circ$.

Different research groups have employed the LPT in face recognition applications. Chien and Choi have used the LPT to locate landmarks in faces [9]. Minut *et al.* have used the LPT together with Hidden Markov Models for the recognition of faces [10]. Tistarelli and Grosso [11] have implemented an active face recognition system that uses the LPT together with Principal Component Analysis.

Discrete Cosine Transform (DCT)

The dimension of image feature vector after the log polar transform is still high. The 2-dimensional DCT is used for further dimensionality reduction. DCT has emerged as a popular transformation technique widely used in signal and image processing [12], [13], [14], [15], [16].



(a) (b)

Fig. 2 Face image and the LPT (a) The original face image and the corresponding LPT. (b) The LPT of the face image with rotation angles $5^\circ, 10^\circ, 15^\circ, 20^\circ$

This is due to its strong "energy compaction" property: most of the signal information tends to be concentrated in a few low-frequency DCT coefficients. DCT exploits inter-pixel redundancies to render excellent decorrelation for most natural images. After decorrelation each transform coefficient can be encoded independently without losing compression efficiency. The DCT helps separate the image into parts (or spectral sub-bands) of differing importance (with respect to the image's visual quality).

The general equation for the *2*- dimensional DCT of an *NxM* image $f(x,y)$ is defined by the following equation:

$$F(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} f(x,y)\cos[\frac{\pi.u}{2.N}(2x+1)]\cos[\frac{\pi.u}{2.M}(2y+1)] \tag{5}$$

Where $f(x,y)$ is the intensity of the pixel in row *x* and column *y*; *u= 0, 1, …., N-1* and *v=0, 1, …., M-1* and the functions $\alpha(u)$, $\alpha(v)$ are defined as

$$\alpha(u), \alpha(v) = \begin{cases} \sqrt{\frac{1}{N}} & for \quad u,v=0 \\ \sqrt{\frac{2}{N}} & for \quad u,v \neq 0 \end{cases} \tag{6}$$

An important characteristic of facial images is its frequency properties. Facial images generally include facial components such as eyes, eyebrows and mouth, which have strong horizontal and vertical frequency properties. For most facial images, much of the signal energy lies at low frequencies (corresponding to large DCT coefficient magnitudes); these are relocated to the upper-left corner of the DCT array. Conversely, the lower-right values of the DCT array represent higher frequencies, and turn out to be small enough to be truncated or removed with little visible distortion, especially as *u* and *v* approach the sub-image width and height, respectively. This means that the DCT is an effective tool that can pack the most effective features of the input image into the fewest coefficients. This makes the choice of the number of DCT coefficient initially used in the feature vector for the face recognition system very critical. Energy is one of the most important properties in image processing [17]. Given a result of DCT transform on a facial image, $F(u, v)$ with the size $N\times N$, the energy using DCT can be calculated as:

$$Energy_F = \sum_{u=1}^{N}\sum_{v=1}^{N} \left|F(u,v)\right|^2 \quad (7)$$

Where $Energy_F$ is total energy content for the facial image. From the $Energy_F$ and the square value of each coefficient, the energy Probability $EP(u, v)$ is defined as follows:

$$EP(u,v) = \frac{\left|F(u,v)\right|^2}{Energy_F} \quad (8)$$

The magnitude of $EP(u, v)$ can be used as a criterion for selecting valid coefficients. It represents the relative energy magnitude of the coefficient at position (u, v). Therefore, coefficients of high *EP* value hold more valid information than coefficients with low *EP* value. In this research the 10x10 coefficients in the top left corner of the DCT array are selected as the feature vector and were found to contain 96.7% of the total energy in the signal.

B. Feature Selection

Feature selection (FS) in pattern recognition involves the derivation of the feature subset from the raw input data to reduce the amount of data used for classification and simultaneously provide enhanced discriminatory power [15], [16], [18], [19], [20], [21] In the proposed FR system we utilized an evolutionary feature selection algorithm based on swarm intelligence called the PSO-based feature selection algorithm [21].

Particle Swarm Optimization (PSO)

PSO proposed by Dr. Eberhart and Dr. Kennedy in 1995 is a computational paradigm based on the idea of collaborative behavior and swarming in biological populations inspired by the social behavior of bird flocking or fish schooling [22], [23]. When PSO is used to solve an optimization problem, a swarm of computational elements, called particles, is used to explore the solution space for an optimum solution. Each particle represents a candidate solution and is identified with specific coordinates in the D-dimensional search space. The position of the i -th particle is represented as $X_i = (x_{i1}, x_{i2} ..., x_{iD})$. The velocity of a particle (rate of the position change between the current position and the next) is denoted as $V_i = (v_{i1}, v_{i2}, ..., v_{iD})$. The fitness function is evaluated for each particle in the swarm and is compared to the fitness of the best previous result for that particle and to the fitness of the best particle among all particles in the swarm. After finding the two best values, the particles evolve by updating their velocities and positions according to the following equations:

$$V_i^{t+1} = \omega * V_i^t + c_1 * rand_1 * (p_{i\_best} - X_i^t) + c_2 * rand_2 * (g_{best} - X_i^t) \quad (9)$$

$$X_i^{t+1} \quad X_i^t \quad V_i^{t+1} \quad (10)$$

Where $i = 1, 2, ... N$, and N is the size of the swarm; p_{i_best} is the particle best reached solution and g_{best} is the global best solution in the swarm. c_1 and c_2 are cognitive and social parameters that are bounded between 0 and 2. $rand_1$ and $rand_2$ are two random numbers, with uniform distribution $U(0,1)$. $-V_{max} \le V_i^{t+1} \le V_{max}$ (V_{max} is the maximum velocity). In equation (9) the first component represents the inertia of pervious velocity. The inertia weight ω, is a factor used to control the balance of the search algorithm between exploration and exploitation; the second component is the "cognitive" component representing the private experience of the particle itself; the third component is the "social" component, representing the cooperation among the particles. The recursive steps will go on until we reach the termination condition (maximum number of iterations K).

PSO-Based Feature Selection Algorithm

A binary version of the PSO algorithm has been developed in [23]. In the binary version, the particle position is coded as a binary string that imitates the chromosome in a genetic algorithm. The particle velocity function is used as the probability distribution for the position equation. That is, the particle position in a dimension is randomly generated using that distribution.

The equation that updates the particle position becomes the following:

$$\text{if } rand_3 < \frac{1}{1+e^{-V_i^{t+1}}} \quad \text{then } X_i^{t+1} = 1; \text{ else } X_i^{t+1} = 0 \quad (11)$$

A bit value of {1} in any dimension in the position vector indicates that this feature is selected as a required feature for the next generation, whereas a bit value of {0} indicates that this feature is not selected as a required feature for the next generation.

The initial coding for each particle is randomly produced where each particle is coded to imitate a chromosome in a genetic algorithm; each particle was coded to a binary alphabetic string $P = F_1F_2 ...F_n$, $n = 1, 2, ..., m$; where m is the length of the feature vector extracted by the hybrid LPT/DCT feature vector. Each gene in the m-length chromosome represents the feature selection, “1” denotes that the corresponding feature is selected, otherwise denotes rejection. The m-genes in the particle represent the parameters to be iteratively evolved by PSO. In each generation, each particle (or individual) is evaluated, and a value of *goodness* or *fitness* is returned by a fitness function. This evolution is driven by the fitness function F that evaluates the quality of evolved particles in terms of their ability to maximize the class separation term indicated by the scatter index among the different classes.

Let $w_1, w_2, ..., w_L$ and $N_1, N_2, ..., N_L$ denote the classes and number of images within each class, respectively. Let $M_1, M_2, ..., M_L$ and M_o be the means of corresponding classes and the grand mean in the feature space, M_i can be calculated as:

$$M_i = \frac{1}{N_i}\sum_{j=1}^{N_i} W_j^{(i)}, \qquad i = 1, 2, ..., L. \quad (12)$$

Where $W_j^{(i)}$, $j = 1, 2, ..., N_i$, represents the sample images from class w_i, and the grand mean M_o is:

$$M_o = \frac{1}{n}\sum_{i=1}^{L} N_i M_i, \quad (13)$$

Where n is the total number of images for all the classes. Thus, the between class scatter fitness function F is computed as follows:

$$F = \sqrt{\sum_{i=1}^{L} (M_i - M_o)'(M_i - M_o)} \quad (14)$$

The main steps of PSO-based feature selection algorithm are as follows [21]:

1) Initialization
- initialize the maximum number of iterations (K)
- initialize parameters V_{max}, c_1, c_2 and ω

2) Particles generation
- Generate N particles and initialize each particle with random position and velocity.

3) Particles evaluation
- The selected feature subsets represented by the particles are evaluated using the fitness function given in (14).
- If the fitness of the particle X_i^t is greater than the fitness of the particle's best P_{i_best} update $P_{i\_best} = X_i^t$
- If the fitness of X_i^t is greater than the fitness of g_{best} update $g_{best} = X_i^t$

4) Check the stop criterion
- If the number of iterations exceeds the maximum allowed iteration exit, otherwise continue.

5) Particles evolution
- Update the velocity and the position of particles using (9) and (11) respectively.

6) Go to 3 and continue.

C. Euclidian Distance Classifier

A typical and popular Euclidean distance is employed to measure the similarity between the test vector and the reference vectors in the gallery. Euclidean distance is defined as the straight-line distance between two points. For example, in a plane with p_1 at (x_1, y_1) and p_2 at (x_2, y_2), the Euclidean distance is given by:

$$D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (15)$$

For N-dimensional space, the Euclidean distance between two any points' p_i and q_i is given by:

$$D = \sqrt{\sum_{i=1}^{N} (p_i - q_i)^2} \quad (16)$$

Where p_i (or q_i) is the coordinate of p (or q) in dimension i.

In the application of this approach for face recognition, distances in the feature space from a query image to every image in the database are calculated. The index of the image which has the smallest distance with the image under test is considered to be the required index.

III. EXPERIMENTAL RESULTS

The performance of the proposed system is evaluated using the standard Cambridge ORL gray-scale face database. The ORL database of faces contains a set of face images taken between April 1992 and April 1994 at the AT&T Laboratories (by the Oliver Research Laboratory in Cambridge, UK) [15] [16], [21]. The database is composed of 400 images corresponding to 40 distinct persons. The original size of each image is 92x112 pixels, with 256 grey levels per pixel. Each subject has 10 different images taken in various sessions varying the lighting, facial expressions (open/ closed eyes, smiling/ not smiling) and facial details (glasses/ no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement).

A. Experiment 1

In this experiment, 4 pictures for each of the 40 subjects in the ORL database were selected as the training set, for a total of 160 face images. The training images were rotated with rotation angles $\pm 5^\circ, \pm 10^\circ, \pm 15^\circ, \pm 20^\circ, \pm 25^\circ, \pm 30^\circ$ and used as our test set (160 X 12 images). We compared the recognition rate and the number of selected features for four different approaches which are the LPT combined with the PSO feature selection algorithm (LPT + PSO), the proposed system using the hybrid LPT/DCT features and PSO feature selection (LPT + DCT+ PSO), the DCT, and DWT followed by the DCT (DWT + DCT). Fig. 3 shows the recognition rates of the different approaches versus the rotation angles. The LPT+PSO and the proposed system yields similar results of a recognition rate of approximately 100% for all rotated images rotated with angles less than 20°. The LPT+PSO approach yields better results for rotation angles greater than 20°.

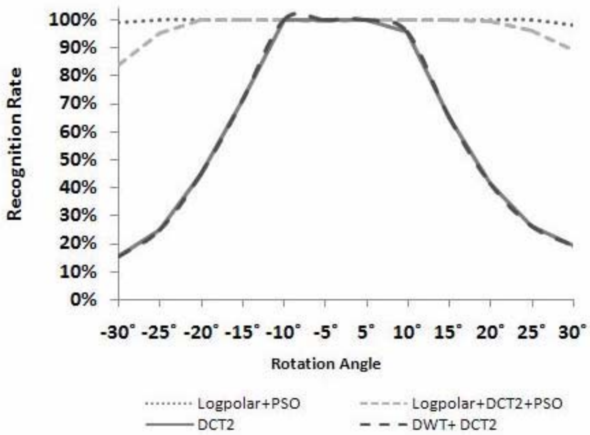


Fig. 3 Recognition rate versus the rotation angles.

Table I presents the average recognition rates for the different approaches using the complete test set with various rotation angles. The table shows the number of selected features in each approach. The first two approaches outperforms the DCT and DWT+DCT with approximately equal recognition rates using only 54 selected features in the proposed system versus 803 features in the LPT+PSO approach.

TABLE I
PERFORMANCE COMPARISON USING THE ORL DATABASE

Approach	No. of Features	Recognition Rate
LPT + PSO	803	99.7%
LPT + DCT+ PSO	54	96.9%
DCT	100	58.6%
DWT + DCT	100	58.7%

B. Experiment 2

In this experiment, 4 rotated images of each of the trained images were added to the training set. The used rotation angles were randomly selected from the set $\pm 12^{\circ}, \pm 18^{\circ}, \pm 24^{\circ}, \pm 28^{\circ}$ (a total of 160 + 160*4 = 800 training images). The same test set was used as in the first experiment. Table II shows the recognition rate of our proposed system with and without the rotated images in the training set. The addition of the rotated images increases the recognition rate from 96.9% to 100%.

TABLE II
PERFORMANCE OF THE PROPSED SYSTEM USING DIFFERENT TRAINING SETS

Training set	Recognition Rate
Train images do not include rotated images	96.9%
Train images include rotated images	100%

IV. CONCLUSION

The paper presents a rotation invariant face recognition system method that combines the log polar transform and the discrete cosine transform to generate a rotation invariant feature vector for face recognition systems. A PSO-based feature selection algorithm is effectively utilized in the system to search the feature vector space for the optimal feature subset for further feature dimensionality reduction. Experimental results based on the ORL face database using testing data sets for images with different orientations, facial expression, small occlusions and different types of illuminations show that the proposed system outperforms other face recognition methods for rotated face images and requires short time for both the training and querying process. The overall recognition rate of the rotated test images being up to 97% with only 54 features. This recognition rate can be improved to 100% if random rotated images are included in the training stage. The experimental results indicate that the concept of hybrid LPT/DCT features is a compact, rotation invariant, and meaningful global representation of facial information obtained by projection the low-frequency components of the DCT coefficients.

REFERENCES

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," *ACM Computing Surveys,* vol. 35, no. 4, pp. 399-458, 2003.
[2] R. Brunelli and T. Poggio, "Face Recognition: Features versus Templates," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042-1052, 1993.
[3] M. A. Turk and A. P. Pentland, "Face Recognition using Eigenfaces," *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586-591, June 1991.
[4] X. Yi-qiong, L. Bi-cheng and W. Bo, "Face Recognition by Fast Independent Component Analysis and Genetic Algorithm," *Proc. of the 4th International Conference on Computer and Information Technology (CIT'04)*, pp. 194-198, Sept. 2004.
[5] K. Nakamura, and S. Miyamoto, "Rotation, size and shape recognition by a spreading associative neural network," *IEICE Trance. on Information and Systems*, vol.E-84-D, no.8, pp.1075-1084, 2001.
[6] K. Nakamura, K. Arimura, and T. Yoshikawa, "Recognition of Object Orientation and Shape by a Rotation Spreading Associative Neural Network," *Proc. IEEE-INNS Int. Joint Conf. on Neural Networks (JCNN2001)*, pp.565-570, 2001.
[7] H. El-Bakry, "A Rotation Invariant Algorithm for Recognition," Fuzzy Days 2001, LNCS 2206, pp. 284–290, 2001. Springer-Verlag Berlin Heidelberg 2001.
[8] H. R. Wilson, D. Levi, L. Maffei, J. Rovamo, and R. DeValois, "The Perception of Form: Retina to Striate Cortex", *Visual Perception: The Neurophisiologcal Foundations*, Academic Press, 1990.
[9] S. Chien, and I. Choi, "Face and Facial Landmarks location based onLog-Polar Mapping", *Lecture Notes in Computer Science – LNCS* 1811, pp. 379-386, 2000.
[10] S. Minut, S. Mahadevan, J. Henderson, and F. Dyer, "Face Recognition using Foveal Vision", *Lecture Notes in Computer Science – LNCS* 1811, pp. 424-433, 2000.
[11] M. Tistarelli, and E. Grosso, "Active Vision-Based Face Authentication", *Image and Vision Computing*, no. 18, pp. 299-314, 2000.
[12] A. S. Samra, S. E. Gad Allah, R. M. Ibrahim, "Face Recognition Using Wavelet Transform, Fast Fourier Transform and Discrete Cosine Transform," *Proc.* 46th *IEEE International Midwest Symp. Circuits and Systems (MWSCAS'03)*, vol. 1, pp. 272- 275, 2003.
[13] Z. Yankun and L. Chongqing, "Efficient Face Recognition Method based on DCT and LDA," *Journal of Systems Engineering and Electronics*, vol. 15, no. 2, pp. 211-216, 2004.
[14] Z. M. Hafed and M. D. Levine, "Face Recognition Using Discrete Cosine Transform, " *International Journal of Computer Vision,* vol. 43, no. 3, pp. 167-188. 2001.
[15] F. M. Matos, L. V. Batista, and J. Poel, "Face Recognition Using DCT Coefficients Selection," *Proc. of the 2008 ACM Symposium on Applied Computing, (SAC'08)*,pp. 1753-1757, March 2008.
[16] M. Yu, G. Yan, and Q.-W. Zhu, "New Face Recognition Method Based on DWT/DCT Combined Feature Selection," *Proc. 5th International Conference on Machine Learning and Cybernetics*, pp. 3233-3236, August 2006.
[17] K. Hyun Kim, Y.-S. Chung, J.-H. Yoo, and Y. Man Ro, "Facial Feature Extraction Based on Private Energy Map in DCT Domain," ETRI Journal, Volume 29, Number 2, pp. 243-245, April 2007
[18] E. Kokiopoulou and P. Frossard, "Classification-Specific Feature Sampling for Face Recognition," *Proc. IEEE 8th Workshop on Multimedia Signal Processing*, pp. 20-23, 2006.
[19] X. Fan and B. Verma, "Face recognition: A New Feature Selection and Classification Technique," *Proc. 7th Asia-Pacific Conference on Complex Systems*, December 2004.
[20] A. Y. Yang, J. Wright,Y. Ma, and S. S. Sastry, " Feature Selection in Face Recognition: A Sparse Representation Perspective," submitted for publication, 2007.
[21] R. M. Ramadan, and R. F. Abdel-Kader, "Face Recognition Using Particle Swarm Optimization-Based Selected Features," In Press, 2008.
[22] J. Kennedy and R. Eberhart, "Particle swarm optimization," *Proc. IEEE International Conference on Neural Networks*, pp. 1942-1948, 1995.
[23] J. Kennedy and R. C. Eberhart, "A Discrete Binary Version of the Particle Swarm Algorithm", *Proc. IEEE International Conference on Systems, Man, and Cybernetics*, vol. 5, pp. 4104-4108, Oct. 1997.