Discrete Cosine Transform (DCT) Based Face Recognition in Hexagonal Images

Muhammad Azam
EPE Department, PNEC
National University of Sciences and Technology (NUST)
Karachi, Pakistan
azam1100@yahoo.com

M. Almas Anjum
C.E Department, College of E&ME
National University of Sciences and Technology (NUST)
Rawalpindi, Pakistan
almasanjum@yahoo.com

M. Younus Javed
C.E Department, College of E&ME
National University of Sciences and Technology (NUST)
Rawalpindi, Pakistan
myjaved@ceme.edu.pk

***Abstract*— In this paper a new approach to face recognition is presented which is based on processing of face images in hexagonal lattice. The importance of the hexagonal representation is that it possesses special computational features that are pertinent to the Human Vision process. Few advantages of processing images on hexagonal lattice are higher degree of circular symmetry, uniform connectivity, greater angular resolution, and a reduced need of storage and computation in image processing operations. Proposed methodology is a hybrid approach to face recognition. DCT is being applied to hexagonally converted images for dimensionality reduction and feature extraction. These features are stored in a database for recognition purpose. Artificial Neural Network (ANN) is being used for recognition. Experiments and testing were conducted over ORL, Yale and FERET databases. The proposed methodology has given better results in recognition over square pixel based approaches.**

Keywords- Biometrics, Face Recognition, Hexagonal Image Processing, DCT, Feature Extraction

I. INTRODUCTION

Face Recognition by humans is a complex operation which is executed by humans naturally. Over the last 30 years researchers have investigated issues related to face recognition by humans and machines [1]. Face Recognition by machines, called Automated Face Recognition (AFR) simulating Human Vision System (HVS) has received extensive attention in the last few decades owing to its potentially wide applications in the field of security and surveillance.

For most modern image processing and display devices, the shape of pixels are square and the data is gathered and arranged in square lattices [2]. Computer Vision methodologies for Face Recognition are also based on square lattice. Square pixel based face recognition have been universally acknowledged. Though, the traditional method is accepted world wide but it has a number of limitations e.g. it requires more storage and the absence of a uniform neighbourhood introduces uneven edges in the images etc. These limitations degrade the performance of the recognition systems.

In this paper a new methodology for face recognition based on sampling of images on hexagonal lattice is proposed. Sampling on a hexagonal lattice is a promising solution which offers better efficiency and less aliasing [3]. The importance of the hexagonal representation is owing to its special computational features which are pertinent to the Human Vision System (HVS). It also offers a higher degree of circular symmetry, uniform connectivity, greater angular resolution and a reduced requirement of storage and computation in image processing operations.

The Hexagonal Image Processing (HIP) addresses problems related to square based image processing by providing a hexagonal lattice having a uniform neighbourhood throughout the image. HIP provides higher packing density and a uniform connectivity [4, 5] while ensuring minimal storage [6] and smooth feature extraction [7]. The hexagonal images have smooth features that make the recognition process smarter in conformance with human traits.

This paper is organized as follows. The proposed system architecture is presented in section II. Two basic modules of the proposed approach i.e. Training and Recognition are also discussed in section II. Sections III present the results and analysis. Conclusion is drawn in section IV.

II. SYSTEM ARCHITECTURE

In this paper a new approach for face recognition is presented. This approach gains the benefits of hexagonal image processing for face recognition. The approach is based on processing of face images in hexagonal lattice in two major modules i.e. Training and Recognition. The Training Module takes the training set of hexagonally converted face data, extracts features from the training faces and stores them in a database. The Recognition module then extracts face features from the input image and compares it with the stored features in the database for classification. The results are

978-1-4244-5586-7/10/$26.00 © 2010 IEEE

then displayed by Results/Display module. Block diagram of the approach is shown in Fig. 1.

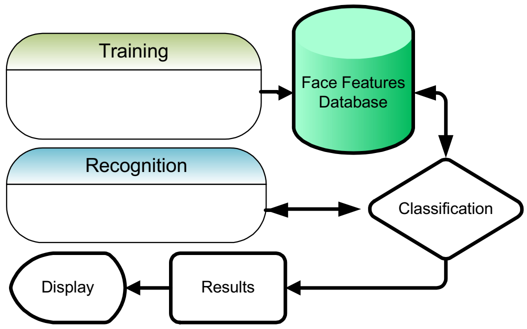

Figure 1. New approach for Face Recognition in hexagonal images

A. *Training Module*

Training Module trains the system for recognition/classification. Input from a camera or face database of square images is taken as training image. The image is first converted from square lattice to hexagonal lattice. Hexagonal image is then converted to log-space so as to apply DCT on the image for feature extraction. These features are then stored in a features database for later recognition/classification. Training module block diagram is presented in Fig. 2.

1) *Input Images*

The system is designed to take hexagonal images as input but as all the currently available image capturing devices like cameras and scanners store images in square row-column base format i.e. square images. It is therefore necessary to transform the square images into hexagonal images.

2) *Conversion to Hexagonal Image*

The square images received from normal camera/database needs conversion to hexagonal lattice for further processing. When the face image is converted to hexagonal image, the layout, arrangement and shape of pixels in the image are changed to hexagonal format. To convert a square image to a hexagonal image, first an Overlay is defined. The overlay consists of Cartesian coordinates of a hexagonal structure which gives exact location of centre of each hexagonal pixel in the space. A special indexing scheme is also employed on the hexagonal structure. This Overlay is centre aligned with the square image and the intensities are extracted from square image. These intensities are then stored in an array according to the indexing scheme [8].

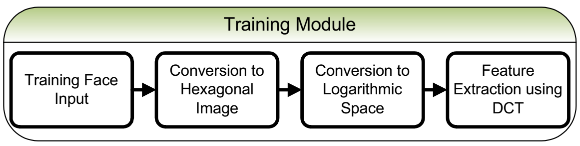

Figure 2. Training Module

3) *Defining Overlay*

Hexagonal Image indexing scheme [9] is defined from centre outwards. The centre pixel of the hexagonal structure is defined as pixel '0' and the layer 1 is defined around the '0' pixel by arrangement of 6 pixels around the centre pixel, numbered as 1, 2, 3, 4, 5 and 6. Pixel '0' will be the centre of each layer in the hexagonal structure. Layer 2 octave centers will be multiple of 10^n, where n is the number of layers. For layer 2 the six octave centre pixels will be 10, 20, 30, 40, 50 and 60. The rest of the pixels around the octaves will be numbered in sequence w.r.t. the centre pixel. For pixel 10 the surrounding pixels will be numbered as 11, 12, 13, 14, 15 and 16. Structure of 45 = 7^2 hexagons in spiral architecture is shown in Fig. 3.

After defining the hexagonal structure and its indexing scheme, the square image is then translated to a hexagonal image. For a proper translation, that a hexagon in the hexagonal lattice and a square in the rectangular grid should have the same area. For a hexagon with radius r, the area A of the hexagon would be:

$$A_h = 6*\frac{r}{2}*\sqrt{\frac{3}{4}r^2} = \sqrt{\frac{27}{4}r^4} \quad (1)$$

Therefore the side length 's' of a square should be defined as:

$$S = \sqrt[4]{\frac{27}{4}}r \quad (2)$$

Along with the scaling of the rectangular image, the pixels are split into several, equally distributed points, where each point gets the intensity of the pixel. To translate a rectangular image to the hexagonal lattice, the hexagonal structure is put on top of the intensity points. The intensity of each hexagon is evaluated as the average of the underlying points from the rectangular image. The quality of the translation depends on the number of points, in which each square is divided. Fig. 4 shows the translation of a 3 x 3 image on the rectangular structure to a structure of 49 hexagons.

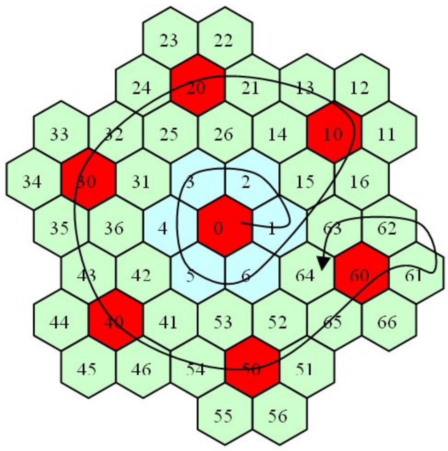

Figure 3. Structure of 49 = 7^2 hexagons in spiral architecture.

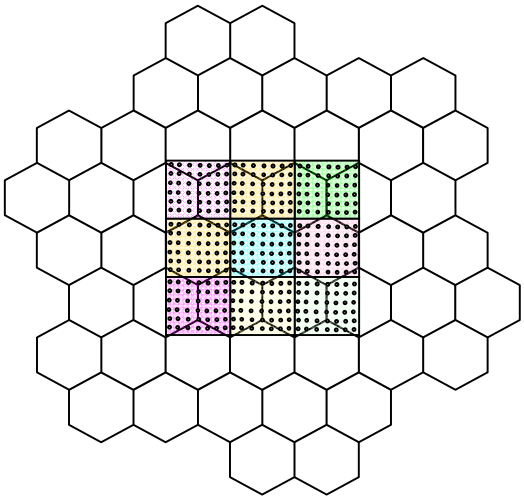

Figure 4. Translation of a 3 x 3 image on the rectangular structure

4) Conversion to Log Space

The logarithm space defines a sequence throughout the image which is found by repeated spiral addition. Thus, by repeated spiral adding address 1, the logarithm space preserves the sequence of hexagons in each column from top to bottom. Only the jump from one column to another causes discontinuity. This sequence is used for input to DCT features extraction module directly.

Special addresses in the structure are used to describe every other address, by repeated modular spiral addition and modular spiral multiplication respectively. These special addresses are identified as the generator or prime root of the discrete logarithms. Repeated spiral addition or spiral multiplication leads to a cycle through the structure as shown in Fig. 5. The logarithm of a certain address is defined as the number of repeated operations needed to reach the address within the given structure. Therefore the logarithm depends on the size of the structure. As a result, any transformation can be reduced to a normal addition of the logarithm values.

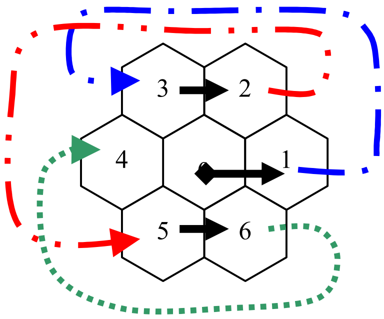

Figure 5. Cycle of log space

Address 1 is used to define the logarithm spaces for modular spiral addition. Starting with log value 0 for address 0, the algorithm iterates along every spiral address once, before it returns to address 0. Every time, the algorithm would leave the given structure, the wrap around effect maps it back to a new position in this structure. The logarithm values change with the size of the structure, thus, they are evaluated for every possible size of the structure individually. The logarithm spaces is evaluated once and stored for a later use [10].

5) Feature Extraction Using DCT

In the proposed approach DCT coefficients are being used as extracted features from the face in hexagonal images. The 2D DCT of a square image gathers its coefficients in the upper left corner, as shown in Fig. 6. Similarly if the 2D DCT is applied to hexagonal image which is stored in an array format, the resultant coefficients are collected at the start of the array, as shown in Fig. 7. Thus very few of these coefficients can be used for recognition purpose.

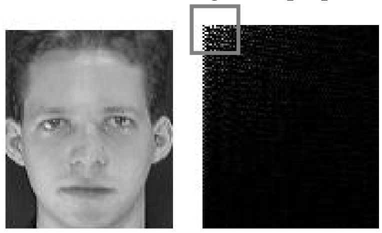

Figure 6. ORL square image and its DCT

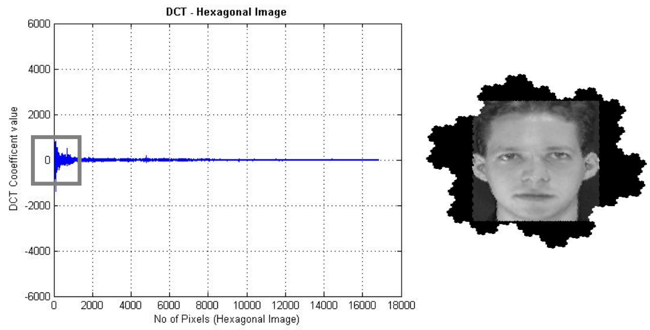

Figure 7. ORL Hexagonal image and its DCT

6) Storing features in Database

The DCT features are stored in the form of an array where the required features of the image are stored as coefficients of DCT in the start of the array. Using the initial few values delivers reasonable performance for face recognition. Increasing the number of selected coefficient, recognition rates do not improve the recognition results considerably.

B. Recognition Module

Recognition module takes a square image as input and covert the image to hexagonal architecture by using similar conversion process as training module. Then the hexagonally converted image is converted to Log Space and DCT coefficients are extracted in similar fashion. Theses coefficients are then compared to already stored DCT coefficients in the database. ANN is used as feature matching technique and classification is done on the bases of closest match. Block diagram of the Recognition module is presented in Fig. 8.

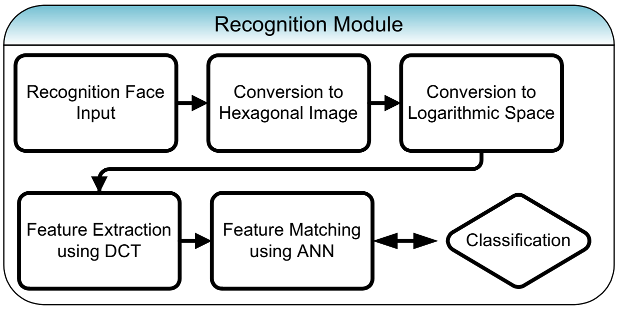

Figure 8. Recognition module block diagram

1) *Feature matching Using ANN*

After number and locations of transform coefficients are selected, the selected coefficients are arranges in one dimensional format and are fed into a classifier for recognition. The feed forward neural network classifier with only one hidden layer is used. A quick back propagation algorithm is used as the training algorithm.

Learning is improved by representing the input and output in bipolar form. Number of face subjects in the Multi Layer Perceptron (MLP) is the number of outputs. The target outputs are set to bipolar form. For an unknown image, the scaling factor obtained for the training set are applied to the retained coefficients to obtain the input vector to the MLP.

2) *Image Display in Hexagonal Images*

Displaying hexagonally converted images is yet an important task in Hexagonal Image Processing. This is especially needed to observe the results after transforming images on spiral architecture. The image viewer on spiral architecture has to consider several features. The image viewer has to display pixels in the shape of hexagons. Each of the hexagons has to have the same arrangement of the edges and has to be defined by its 6 corners; each of the corners consists of two coordinates. If the position of all the hexagons would be determined while executing the displaying function, the function would be inconvenient because of the complexity. Therefore it would be the best to store the coordinates of the corners of all the hexagons. On the other hand, the storing space is kept as low as possible.

III. RESULTS AND ANALYSIS

The methodology for face recognition in hexagonal images was tested on three databases i.e. ORL, Yale and FERET, to give variety of experimental scenarios. Based on these test results conclusive remarks can be drawn on the fact that processing of face images in hexagonal image processing is better then square lattice based image processing. From each class 5 images were selected for training and recognition results were defined using remaining 5 images from each class. The selection was also inverted and brought no significant results.

A. *Conversion Time*

Conversion time is an overhead to the face recognition in hexagonal images which can be eliminated with the availability of HIP hardware. Conversion time for a single image is dependent on image size and it remains consistent for a single image. The conversion time (in seconds) in different databases with increase in number of classes is shown in Table 3-1. The outcome of 5 experiments with increase in number of classes is presented in this Table I. The results show how the change in number of classes of different databases has effect on the conversion time.

This conversion time can also be avoided by converting the face database in advance and storing the coefficients in the coefficient database. However, new face images need to be converted to hexagonal images in real time. For the proposed methodology the conversion time cannot be ignored.

TABLE I. CONVERSION TIME – SQUARE TO HEXAGONAL

Database	**Number of Classes**				
	1	***2***	***3***	***4***	***5***
ORL	***0.45s***	***0.92s***	***1.36s***	***1.82s***	***2.25s***
YALE	***0.48s***	***0.90s***	***1.39s***	***1.85s***	***2.33s***
FERET	***0.63s***	***1.27s***	***1.89s***	***2.53s***	***3.15s***

B. *Training Time*

Time taken by training module to train the system for increasing number of classes is presented in Table II (excluding the conversion time). The training time variations are demonstrated with increase in number of classes. Total 5 number of subjects were used from each class for training. With 5 number of classes the time 0.96s signify the time taken to training 5 classes of ORL databases i.e. total of 25 face images.

TABLE II. TRAINING TIME

Databases	**Number of Classes**				
	1	***2***	***3***	***4***	***5***
Yale	***0.18s***	***0.37s***	***0.57s***	***0.76s***	***0.96s***
ORL	***0.20s***	***0.38s***	***0.58s***	***0.76***	***0.94s***
FERET	***0.17s***	***0.34s***	***0.52s***	***0.77s***	***0.97s***

C. *Recognition Time*

Recognition time in presented approach is also effected by conversion process. However, overall recognition time without the conversion time was found to be better for face recognition in HIP. The Recognition time is reduced with system getting mature due to increase in recognized images, which also become part of the training database.

D. *Recognition Rate*

Face recognition methodology on hexagonal images were then tested on ORL and Yale databases to full length to determine the recognition rate on full databases. Randomly selected classes were chosen from FERET database which included male and female subjects. Misclassified images are shown in Fig. 9. Moreover, recognition rate on complete database is presented in Fig. 10. For lesser number of classes the recognition rate remained high. It gradually reduced with increase in number of classes. Average recognition rate on ORL database remained 98.01%.

Figure 9. Misclassified images – ORL Database

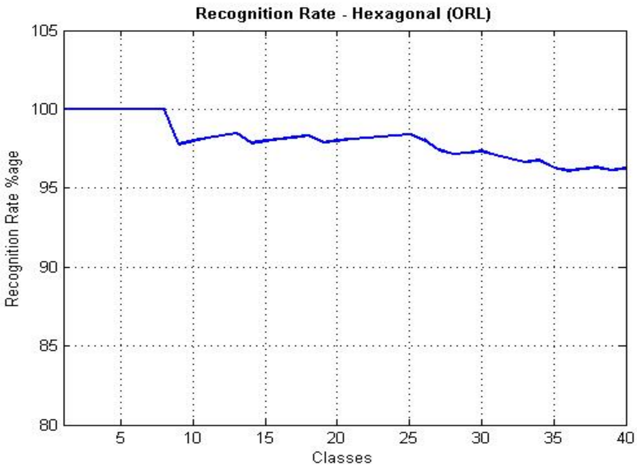

Figure 10. Recognition rate – ORL Database

Recognition rate on Yale database remained 92.77%. Misclassified images are shown in Fig. 11. Moreover, recognition rate on complete database is presented in Fig. 12. In multiple tests on Yale database the misclassified images remained the same.

Recognition rate on FERET database remained 83.31%. Misclassified images are shown in Fig. 13. Moreover, recognition rate on randomly selected 40 classes is presented in Fig. 14.

The recognition rate in FERET database is not significant due to the fact that this methodology is not very robust in order to cater for pose and illumination variations.

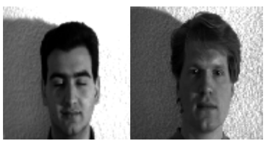

Figure 11. Misclassified Images Yale

Figure 12. Recognition rate – Yale database

Figure 13. Misclassified Images in FERET database

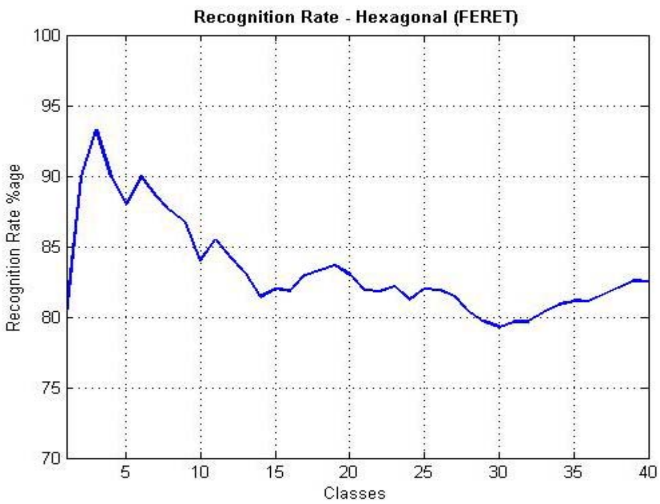

Figure 14. Recognition rate FERET

E. *Comparison with Square Lattice Based Approaches to Face Recognition*

The proposed methodology of face recognition in hexagonal images using DCT and ANN was compared with popular research in face recognition and its out come are compared. As shown in Table III the different tests carried out on ORL database shows that the proposed methodology is better in recognition rate. However, the methodology is costly with respect to Training and Recognition times.

TABLE III. COMPARISON ON ORL DATABASE

S.No.	**Approach**	**Average Recognition Rate**
1.	Hexagonal + DCT+MLP (Proposed Approach)	98.01%
2.	DCT+MLP[11]	92.9%
3.	MLP [11]	77.2%
4.	Convolutional NN [12]	96.2%
5.	NN[13]	94.64%
6..	DCT+RBF NN[14]	97.68%

Table IV shows similar kind of results on the Yale database. The different tests carried out on Yale database shows that the proposed methodology is better in recognition rate or at least comparable with square pixel based approaches. The proposed methodology proves one of the better recognition techniques available. However, the approach is costly with respect to Training and Recognition times.

TABLE IV. COMPARISON ON YALE DATABASE

S.No.	Approach	Average Recognition Rate
1.	Hexagonal + DCT+MLP (Proposed Approach)	92.77%
2.	NN[13]	83.51%
3.	PCA[13]	81.13%
4.	LDA[13]	98.69%
5.	DCT+RBF NN[14]	98.95%

IV. CONCLUSION

In this paper new approach to face recognition is proposed. The approach is a paradigm shift from square based approach to a hexagonal based approach for face recognition for the reasons mentioned in section I. The proposed methodology is better or comparable in recognition rate over traditional approach based on square lattice due to the fact that lines at acute angles is better represented in faces processed on hexagonal lattice. As the proposed method is a holistic approach to face recognition, presenting the faces in hexagonal lattice has improved the DCT-ANN based face recognition approach to face recognition. A recognition time of processing images on hexagonal lattice also has improved over traditional approaches based on square lattice. Additional time for conversion of square images to hexagonal images is one drawback which can be eliminated with development of hexagonal input and output devices. The recognition rate of the proposed system came out to be 98.01% on ORL database. Better results were also observed while the same methodology was applied to Yale and FERET databases.

REFERENCES

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. "Face recognition: A literature survey". ACM Computing Surveys, 35(4), 2003.

[2] J. Foly, A. Van Dam and J. Hughes, Computer Graphics, Principles and Practice (Second Edition). Sydney: Addison Wesley, 1990.

[3] Mersereau, R.M., "The processing of Hexagonally Sampled Two-Dimensional Signals'. Proceedings of the IEEE, 67: p. 930-949, 1979.

[4] Staunton, R.C., "An Analysis of Hexagonal "Thinning Algorithms and Skeletal Shape Representation". Pattern Recognition, 29(7): p. 1131-1146, 1996.

[5] Staunton, R.C., "A one pass parallel hexagonal thinning algorithm. Image Processing and Its Applications", 1999. Seventh International Conference on (Conf. Publ. No. 465), 2: p.841-845 vol.2, 1999.

[6] Vitulli, R., et al., "Aliasing Effects Mitigation by Optimized Sampling Grids and Impact on Image Acquisition Chains". Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International, p. 979-981 vol.2, 2002. 2:

[7] Golay, M., "Hexagonal parallel pattern transformation". IEEE Transactions on computers, 18(8): p. 733-740, 1969.

[8] Q. Wu, X. He and T. Hintz, "Virtual Spiral Architecture", Proc. Int. Conference on Parallel and Distributed Processing Techniques and Applications, Vol. 1, pp 399 – 405, 2004.

[9] P. Sheridan, "Spiral Architecture for Machine Vision", Ph.D. thesis, University of Technology, Sydney, Australia, 1996.

[10] P. Sheridan, T. Hintz and D. Alexander, "Pseudo- invariant image transformations on a hexagonal lattice", Image and Vision Computing, Vol. 18, pp. 907 – 917, 2000.

[11] Z. Pan, R. Adams, and H. Bolouri, High speed face recognition based on discrete cosine transform and neural networks, submitted to IEEE Trans. On Pattren Analysis & Machine Intellegence. 1999.

[12] Lawrence, C. Lee Giles, A. Tosi and A. Back, "Face Recognition, A Convolution neural network approach", IEEE Transections on Neural Networks, vol 8, no.1, pp.98-113, 1997.

[13] Huang, R. and Pavlovic, V. and Metaxas, D.N: "A hybrid face recognition method using markov random fields", ICPR04, pp. 157-160, 2004.

[14] C. Yuehui and Z. Yaou, "Face recognition using DCT and hierarchical RBF model", E. Corchado et al. (Eds.): IDEAL 2006, LNCS 4224, pp. 355–362, 2006.