3D face recognition using 2DPCA

*'** Eng .Amr Rashad ** Dr. Alaa Hamdy ** Dr. Mohamed Ali Saleh **Prof.Mohamed Eladawy

*Teaching assistant in MTI University
**Department of Electronics, Communications, and Computers Helwan University Faculty of engineering

ABSTRACT
Robustness of face recognition systems are measured by its ability to overcome the problem of changing in facial expression and rotation of individuals' face images. This paper represents a face recognition system that overcomes the problem of changes in facial expressions in three-dimensional (3D) range images. A local variation detection and restoration method based on the two-dimensional (2D) principal component analysis (PCA) is proposed. The depth map of 3D facial image is first thresholded to discard the back ground information. Then, the detected face shape is normalized to standard size 100x100 pixels and nose point is selected to be the image center. Image depth values are scaled between 0 and 255 and nose tip has the highest value 255 for translation and scaling –invariant identification. In preprocessing stage, the local variation is minimized by smoothing the face image. The (2D) principle component analysis is applied to the resultant range data for feature extraction and the corresponding principal images are used as the characteristic feature vectors of the subject to find his/her identity in the database of pre-recorded faces.
The system performance is tested against the GavabDB facial database. The facial modeling technique presented here is implemented in a PC platform with the software support of Matlab version 7.7. Experimental results show that the proposed method is able to identify subjects with different facial expression in their 3D facial images.
Key words:
3D Face, 2DPCA

1. INTRODUCTION

The need of security and fraud control applications to establish personal authentication has caused the increase of research in biometric systems [1]. Automatic face recognition has received a great deal of attention and emerged as an active research area especially over the last 20 years [2]. The major purpose of face recognition is to identify the humans from data acquired from their faces, as humans do. A good recognition system has to be fully automatic and robust enough for real life conditions such as illumination, rotations, expressions and occlusions.
Robustness of a face recognition system is measured by the ability to identify human subjects even in the presence of many variations in appearance of their faces. Face can be illuminated from a variety of light sources and surrounded by arbitrary background. Therefore, its 2D projected image appearance can vary significantly. For non-contrived recognition, faces need to be found and recognized despite these changes. The human face is not a unique rigid object. There are millions of different faces, each of which can presume different types of deformations. Variations between different faces can be due to identity, race, or genetics while variations between multiple instances of the same face can be due to deformations, aging, expression, and facial hair.
Face detection and recognition system has to attribute a unique identity to each face by matching it to a large database of persons even in the presence of such image acquisition problems as camera distortion, noise and low image resolution. Despite these rigid design specifications, one needs to maintain the usability of the system on contemporary computational devices. In other words, the processing involved should be efficient with respect to run-time and storage space.
Most research efforts have concentrated on recognizing a human face from two-dimensional (2D) images [3], recently three dimensional (3D) approaches have been receiving more attention [11, 18-26]. Even though the latest 2D face recognition systems have achieved good performance in constrained environments, they are still unable to deal with such problems as changes in head poses and illumination conditions [10]. Since the human face is a 3D object whose 2D image projection is sensitive to these changes, utilizing the 3D face information can improve the face recognition performance considerably. Further, 3D measurements help solving scale, illumination, and rotation problems encountered in 2D analysis [17].
Despite the above advantages, internal deformation problem still exists in 3D images [11, 12]. Thus, there is a need for a 3D model that deals with the aforementioned non-rigid variations.
Our objective in this research is to develop a face recognition system that overcomes the problem of changes in facial expression and gesture in 3D range images. This can be due to the differences in facial expression from one image to another, and is still the main source of errors in most of the existing 3D systems. The recognition error can be reduced by smoothing the images to compensate the changes. In order to achieve

Manuscript received December 5, 2009
Manuscript revised December 20, 2009

our goal a facial range images is subject to smoothing process. This helps avoiding any degradation in one's face caused by changes due to facial expressions. After preprocessing stage, 2DPCA (two dimensional principle component analysis) is applied for feature extraction on the smoothed range images and the resultant principal or Eigen-images are used as the feature vectors for the matching process. The Euclidian distance is used for measuring the similarity between two images.

2. Background and review of past work

Despite that 2D intensity image has been the most popular and commonly used image for face recognition and easy to use but it has intrinsic problem that it cannot handle the change of illumination, facial expression and changing in pose. Recently with new technology and using 3D scanners, 3D face recognition has attracted a lot of researches because it is a more reliable system and able to face facial expression and illumination problem. A review of (3D) face recognition researches is as follows.

Most of the 3D approaches for face recognition rely on both the depth and color information extracted from the face [15]. M. Turke and A. pentland used Eigen faces for recognition [13]. Gordon [23] extracted the facial curvature features in two stages: high level features and low level features. This kind of method usually has to depend on the high-quality 3D data, which could characterize delicate features. The research work in [17] represented a new attempt to face recognition based on 3D point clouds by constructing 3D eigenfaces. The work in [39] represented a survey on 2D and 3D face recognition. In [6, 17, 19, 20], size invariant PCA-based approaches are represented. In [32], a detailed comparison between the existing approaches of 3D face recognition is given. Chang et al. [33] report a 92.8% recognition rate by performing PCA on the range-images of 277 people. However, they perform manual normalization which is not desirable in a real system. Lee and Milios [22] create an extended Gaussian image for each convex region in the image. Some of the existing 3D face recognition systems, except the deformations, are from some form of known gesture. The presented work deals mainly with changes in facial expression. It cannot handle the changes in target pose by training each individual using 2DPCA then minimum Euclidian distance is used for matching. Individual images are smoothed to overcome internal deformation. However, the focus of this paper is to handle changes in facial expression ignoring pose changes. Moreno and Sanchez [37] report 78% rank-one recognition on the same data set GavaDB. The presented method achieves 80.3% recognition rate on the same dataset with 8 principle component vectors and median filter with window size 9.

3. Main Work

This paper represents a modeling technique for 3D face recognition with the ability to overcome the problem of changing in facial expression. The proposed method is illustrated by a block diagram representation in Figure 1.

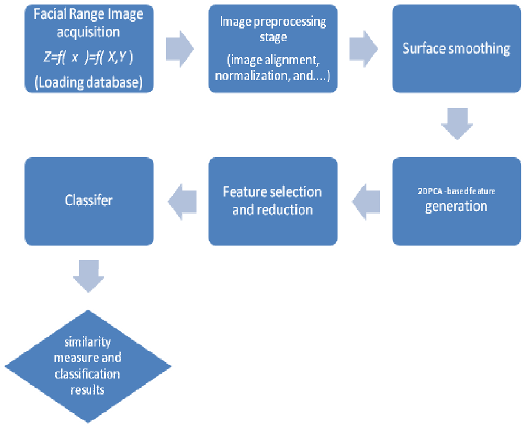

Fig.1

A facial image is first subjected to preprocessing stage. First depth map of all database facial images is thresholded to discard their background using Otsu's method. Then the resultant facial images are normalized to standard size 100x100 then aligned so that the nose tip is located in the center with the highest depth value 255 and all images depth values are between 0 and 255. Then, images are smoothed using Median filter with different window size to find best recognition rate and to prevent variation false matching results. After the smoothing stage, we train each smoothed individual image using 2DPCA where each image per individual is described by its principle component. The resultant principal components are used as feature vectors in the classification stage to calculate the similarity between two facial images. The nearest neighbor classifier is used in the matching process.

To eliminate background, Otsu [35] suggested a criterion by which the best threshold for images with bimodal histogram can be determined. The criterion states that the threshold should be chosen in such a way that minimizes the weighted sum of within group variances for the two groups that result from separating the gray levels at the threshold value (T).

Let P (i); i=1,…,I be the histogram probabilities of an image with I gray levels. Suppose that the pixels are divided into two classes; C1 with gray values [1,...,k] and

C2 with gray values [k+1,...,I]. Then the probabilities of class occurrence are given by

$$c_1(k) = \sum_{i=1}^{k} P(i) \qquad (1)$$

$$c_2(k) = \sum_{i=k+1}^{I} P(i) \qquad (2)$$

Means and variances for the classes C1 and C2 are computed using

$$\mu_1(k) = \sum_{i=1}^{k} i.P(i)/c_1(k) \qquad (3)$$

$$\mu_2(k) = \sum_{i=k+1}^{I} i.P(i)/c_2(k) \qquad (4)$$

$$\sigma_1^2(k) = \sum_{i=1}^{k} [i-\mu_1(k)]^2 .P(i)/c_1(k) \qquad (5)$$

$$\sigma_2^2(k) = \sum_{i=k+1}^{I} [i-\mu_2(k)]^2 .P(i)/c_2(k) \qquad (6)$$

The best threshold T can be calculated [36] by a simple sequential search through all the values of k (1,..,I) to determine the threshold that minimizes the within-group variance σ_w that is defined by

$$\sigma_w^2(k) = c_1(k).\sigma_1^2(k) + c_2(k).\sigma_2^2(k) \qquad (7)$$

Figure 2 represents original image from GavabDB database [37] and Figure 3 shows the histogram of the facial image in Figure 2 taken from the GavabDB [37] data-base and the resultant Otsu threshold. The depth value of each pixel, measured in millimeters (mm), represents the distance between that point and the 3D scanner. Figure 4 illustrates the histogram of the same image after thresholding is performed. The thresholded image is given in Figure 2.

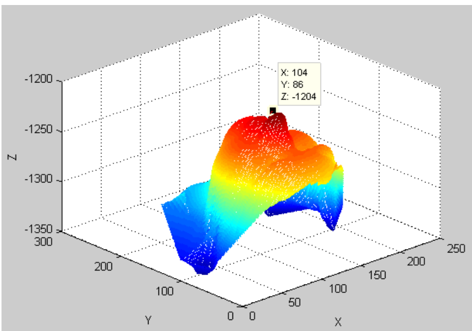

Fig.2 (original image from GavabDB)

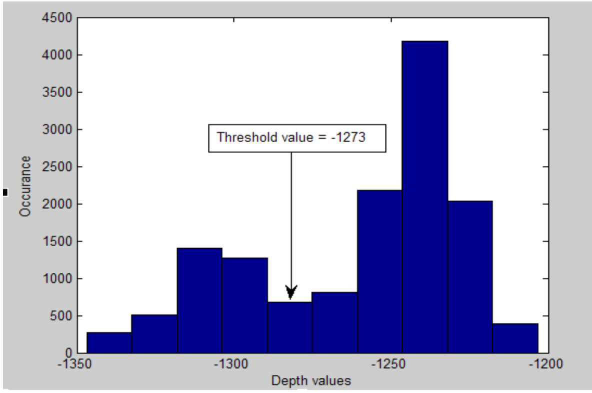

Fig.3 (Depth histogram of facial image) Threshold value = -1273

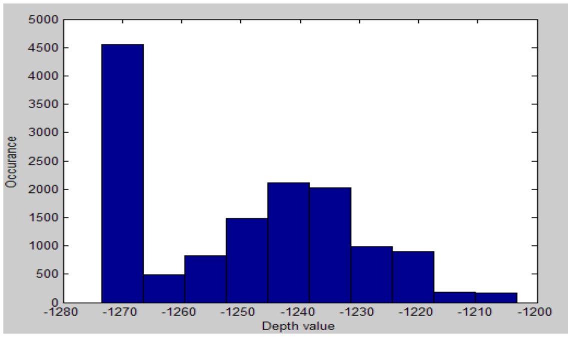

Fig.4 (Depth histogram of the thresholded facial image)

Thresholding is used to crop the images so that only the front of the face is used in the model. The sides of the head, neck, and ears are not used for recognition. These features are not very discriminative and it is difficult to deal with their irregularities.

After the thresholding stage, we locate the location of nose tip to align all the images. Furthermore, it is highly rigid in the sense that its shape does not change with facial expression. In most images, the nose is the closest part of the face to the 3D scanner; i.e., it has the highest depth value among all the facial points. By using a 3x3 window that calculates the sum of the depth values of its corresponding pixels, the nose is detected as the coordinates of the central pixel of the window with the maximum value. After detecting the nose, all images in the database are normalized to a standard 100x100 pixels in size and then aligned so that the nose lies exactly in the center of each image at the (50,50) x-y coordinate as shown in Figure 5. After normalization and alignment, all images are smoothed using median filter with different window size to achieve highest recognition rate and to overcome internal deformation. Figure 6 represents a facial image before and after smoothing.

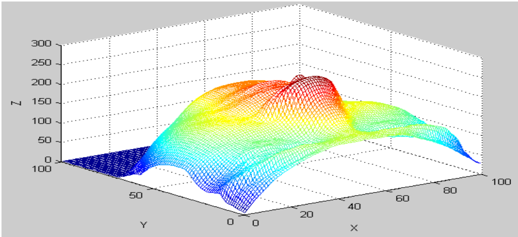

Fig.5 (facial image after normalization and alignment)

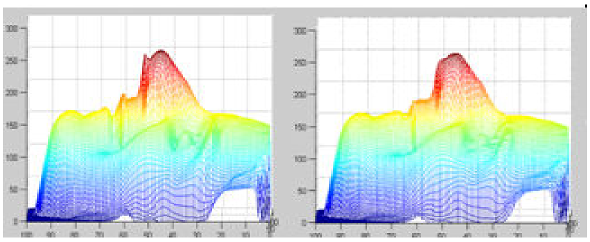

Fig.6 (facial image before & after Smoothing using median filter (5x5))

After preprocessing stage (image thresholding, normalization, alignment and Smoothing), we apply 2DPCA for feature extraction.

Let the training set of 3D facial images be A1, A2...... AM, each of which is of size nxn. The average face of the set is defined as

$$\overline{A} = \frac{1}{M}\sum_{i=1}^{M} A_i \qquad (8)$$

The image covariance matrix is given by

$$C = \frac{1}{M}\sum_{i=1}^{M}(A_i - \overline{A})^T (A_i - \overline{A}) \qquad (9)$$

The optimal number (d) of projection set of an image A, denoted by X, is defined through [38].

$$Y_k = A \cdot X_k \ (k=1, 2, \ldots, d) \qquad (10)$$

To achieve the maximum scatter of the projected feature vectors, the projection axes (X1, X2… Xd) have to be the orthonormal eigenvectors of the covariance matrix C corresponding to the first *d* largest eigenvalues. The projected vectors (Y1, Y2… Yd) are used as feature vectors in our implementation.

After feature extraction stage, we measure similarity by comparing two lists of numbers (i.e., pattern or feature vectors), and computes a scalar value which evaluates the patterns' similarity. The basis of many measures of similarity and dissimilarity is the Euclidean distance d, which is given between two vectors X and Y as

$$d(X,Y) = \sqrt{\sum_{i}^{n}(X(i) - Y(i))^2} \qquad (11)$$

Where n is the number of elements of X and Y. The Euclidean distance is appropriate for the data measured on the same scale. Let the collection of the feature vectors F= (Y1, Y2… Yd) is called the feature matrix. The similarity between two matrices F^i and F^j is measured by

$$d(F^i, F^j) = \sum_{k=1}^{d} dist(Y_k^i, Y_k^j) \qquad (12)$$

Where $dist(Y_k^i, Y_k^j)$ is the Euclidian distance between Y_k^i *and* Y_k^j. The nearest neighbor classifier is used in the classification process. In our implementation, the size of the matrix F is 100 by 6 since we select to use only 6 principle component vectors.

4. Implementation and Results

4.1 Data base

Performance of the proposed method is evaluated using the GavaDB [16] database, which contains 427 three-dimensional facial surface images corresponding to 61 individuals. There are 9 different images per each person. There are systematic variations over the pose and facial expression of each individual. In particular, there are 2 frontal and 4 rotated images without any facial expressions, and 3 frontal images in which the subject presents different facial expressions. The non frontal images are discarded and only the frontal ones are used to focus on the facial expression problem. Each image is specified in a three dimensional mesh, and every mesh is composed of points of the facial surface. The coordinates (x, y, and z) of a 3D point are referred to as a coordinate origin located in the scanner.

4.2 Experiment results and matching

An image of the GavabDB is processed through all the computational stages of the system in Figure 1. The nearest neighbor classifier is chosen with the Euclidian distance between the feature vector of the query and the database as being the similarity measure. Rotated images of GavabDB are discarded so only the frontal ones are used with facial expression. The facial modeling technique presented here is implemented in a PC platform with the software support of Matlab version 7.7.

Median filter window size	Recognition rate						
	Number of principle component vector used (features)						
	6	8	10	20	50	90	100
3	0.6	0.606	0.6885	0.8852	0.8689	0.9344	0.9508
5	0.7377	0.7705	0.8525	0.8689	0.9180	0.9344	0.9344
7	0.6885	0.7705	0.8197	0.9016	0.9344	0.9672	0.9508
9	0.7869	0.8033	0.8197	0.9016	0.9508	0.9836	0.9344
11	0.8033	0.8197	0.8689	0.8689	0.9180	0.9344	0.9508
13	0.7869	0.7869	0.8361	0.9016	0.9508	0.9672	0.9508
15	0.8361	0.8852	0.8852	0.9016	0.9508	0.9508	0.9672

Table 1 (recognition rate for different window size of Median filter and with different Number of principle component vector used)

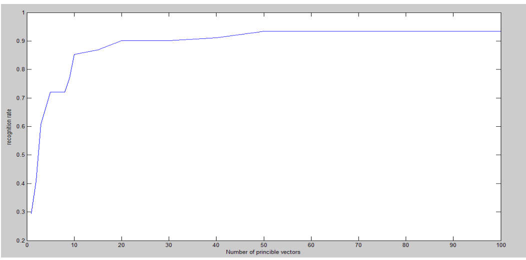

Fig.7 (Recognition rate versus number of principle component vector use with Median filter window size 5)

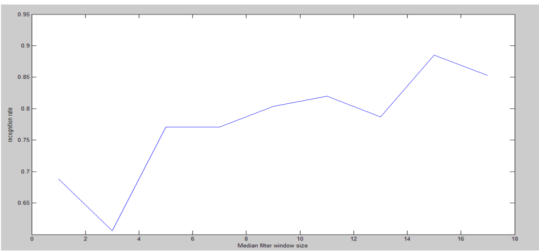

Fig.8 (Recognition rate with different window size and 8 principle vectors)

5. Conclusion

We construct a 3D face recognition system to solve the effect of facial expression on the recognition rate. A facial image is first subjected to pre processing stage in which the background is eliminated by thresholding the depth value using Otsu's thresholding. Then, we locate the nose tip to align and normalize all images to standard size (100x100) and location of nose tip lies exactly in the center of the image. It is then smoothed to eliminate local variations. The two-dimensional principal component analysis is applied to the smoothed range image and the corresponding principal or Eigen-images are extracted as the feature vectors for matching process. The proposed method is tested on the GavaDB. Classification is carried out by calculating the Euclidian distance between the feature vectors. The nearest neighbor classifier is used in choosing the closest match. Moreno and Sanchez [37] report 78% rank-one recognition on the GavaDB. Our method achieves 80.3% recognition rate on the same dataset with 8 principle component vectors and median filter with window size 9. Since 3D face images are invariant to illumination, it is expected that 3D face recognition rates are more stable than those of 2D methods in outdoor lighting.

6. Future work

A face recognition system can be built to overcome the problem of changing in the pose of individuals (rotated images).

References

[1] The Biometric Consortium, http://www.biometrics.org.

[2] A. Pentland and T. Choudhury, "Face Recognition for Smart Environments", Computer IEEE, February 2000,50-55.

[3] R. Chellappa, C.L. Wilson, and S. Sirohey, Human and Machine Recognition on Faces: A Survey. Technical Report CAR-TR-731, Computer Vision Laboratory, University of Maryland, 1994.

[4] T. Kanade, Picture Processing System by Computer Complex and Recognition of Human Faces, Ph.D. Thesis. Kyoto University, Japan, 1973.

[5] A. Yuille, D. Cohen, and P. Hallinan, "Feature extraction from faces using deformable templates," Proceedings of IEEE Computer Society Conference on Computer Vision and Templates, pp.104-109, 1989.

[6] I. Craw, D. Tock, and A. Bennett, "Finding face features," Proceedings of the Second European Conference on Computer Vision, pp.92-96, 1992.

[7] M. Celenk and I. Al-Jarrah, "Color face recognition by auto-regressive moving averaging," First European Conference on Color in Graphics, Image and Vision CGIV'2000, April 22-26, 2002, Poitiers, France.

[8] M. Celenk and I. Al-Jarrah, "Multiresolution ARMA modeling of facial color images," Image Processing: Algorithms and Systems, Proc. SPIE Vol.4667, 21- 23 January 2002, San Jose, California, pp.74-81.

[9] M. Celenk and I. Aljarrah, "ARMA modeling of facial color images," Journal of Electronic Imaging (submitted on June 11, 2005; in revision for publication).

[10] Z. Li and K. Jain, Handbook of Face Recognition. Springer, 2005.

[11] W. Zhao, S. Romdhani, and T. Vetter, "Face recognition and modeling," European Conference on Computer Vision, T4 Tutorial, May 2004.

[12] R. Kimmel and G. Sapiro, "The mathematics of face recognition," SIAM News, Vol.36, No.3, April 2003.

[13] M. Turk and A. Pentland, "Eigenfaces for recognition," Journal of Cognitive Neuroscience, Vol.3, No.1, pp.71-86, 1991.

[14] M. Turk and A. Pentland, "Face recognition using eigenfaces," Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.586-591, 1991.

[15] M. Lee and S. Ranganath, "Pose-invariant face recognition using a 3D deformable model," Pattern Recognition 36, pp.1835-1846, 2003.

[16] A. Bronstein, M. Bronstein, E. Gordon, and R. Kimmel, "Expression-invariant 3D face recognition," Audio and Video based Person Authentication (AVBPA 2003), LCNS 2688, J. Kittler and M. Nixon, eds., pp.62-70, 2003.

[17] C. Xu, Y. Wang, T. Tan, and L. Quan, "A new attempt to face recognition using 3D eigenfaces," the 6th Asian Conference on Computer Vision (ACCV), Vol.2, pp.884-889, 2004.

[18] S. Malasiotis and M.G. Strintzis, "Pose and illumination compensation for 3D face recognition," IEEE International Conference on Image Processing (ICIP), Singapore, October 24-27, 2004.

[19] C. Hesher, A. Srivastava, and G. Erlebacher, "Principal component analysis of range images for facial recognition," Proceedings of International Conference on Imaging Science, Systems, and Technology, Las Vegas, June 2002.

[20] F. Tsalakanidou, D. Tzovaras and M. Strintzis, "Use of depth and color eigenfaces for face recognition," Pattern Recognition Letters, Vol. 24, pp. 1427-1435, 2003.

[21] I. Fisher and J. Lee, "Correlation coefficients for random variables on a unit sphere or hypersphere," Biometrica, No. 73, pp. 159-164, 1986.

[22] J. Lee and E. Milios, "Matching range images for human faces," Proceedings of the International Conference on Computer Vision, pp. 722-726, 1990.

[23] G. Gordon, "Face recognition based on depth maps and surface curvature," Geometric Methods in Computer Vision: SPIE Proceedings, Vol. 1570, pp. 234- 247, 1991.

[24] C. Chua and R. Jarvis, "Point Signatures: A new representation for 3D object recognition," International Journal Computer Vision, Vol. 25, No. 1, pp. 63-85, 1997.

[25] C. Chua, F. Han, and Y. Ho, "3D human face recognition using point signature," Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition, pp. 233-239, 2000.

[26] A. Elad and R. Kimmel, "On bending invariant signatures for surfaces," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 10, pp.1285-1295, October 2003.

[27] A. Nefian, A Hidden Markov Model-Based Approach for Face Detection and Recognition, Ph.D. thesis. Georgia Institute of Technology, 1999.

[28] B. Weyrauch, B. Heisele, J. Huang, and V. Blanz. "Component-based face recognition with 3D morphable models," Conference on Computer Vision and Pattern Recognition Workshop, Vol. 5, No. 5, p. 85, 2004.

[29] A. Hyvarinen, J. Karhunen, and E. Oja, Independent Component Analysis. Wiley, New-York, 2001.

[30] L. Ma, D. Chelberg, and M. Celenk, "Spatio-Temporal Modeling of Facial Expressions Using Gabor-Wavelets and Hierarchical Hidden Markov Models," IEEE International Conference on Image Processing (ICIP), Vol. 2, pp. 57-60, 2005.

[31] H. Tanaka, M. Ikeda, and H. Chiaki, "Curvature-based face surface recognition using spherical correlation principal directions for curved object recognition," Third International Conference on Automated Face and Gesture Recognition, pp.372-377, 1998.

[32] A. Rajwade, Facial Pose Estimation and Face Recognition from Three- Dimensional Data, MS Thesis. McGill University, August 2004.

[33] K. Chang, K. Bowyer, and P. Flynn, "Face recognition using 2D and 3D facial data," Multimodal User Authentication Workshop, pp.25-32, December 2003.

[34] K. W. Bowyer, K. Chang, and P. Flynn, A Survey Of 3D and Multi-Modal 3D+2D Face Recognition, Notre Dame Department of Computer Science and Engineering Technical Report, January 2004.

[35] N. Otsu, "A threshold selection method from gray-level histogram," IEEE Transactions on Systems, Man and Cybernetics, Vol.8, No.1, pp.62-66, 1979.

[36] K. Lin, "On improvement of the computation speed of Otsu's image thresholding," Journal of Electronic Imaging, Vol. 14, No. 2, 2005.

[37] A.B. Moreno and A. Sanchez, “GavabDB: A 3D face database,” Proc. 2nd COST275 Workshop on Biometrics on the Internet, Vigo, Spain, March 25-26, 2004.
[38] J. Yang, D. Zhang, A. F. Frangi, and J. Y. Yang, “Two-dimensional PCA: A new approach to appearance-based face representation and recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.26, No.1, pp.131-137, January 2004.
[39] Andrea F. Abate, Michele Nappi *, Daniel Riccio, Gabriele Sabatino “2D and 3D face recognition: A survey” January 2007

M. E. El-Adawy was born in Cairo in Egypt, on May 09, 1951. B. Sc in communication and electronics, Assuit University, Assiut, Egypt, May 1974 with grade excellent. M. Sc in electronics, Cairo University, Egypt, June 1978 "New binary sequence generator and filter". Ph. D. in medical electronics, University of Connecticut, USA, December 1983, "Application of thermodynamic theory of generalized fields to the investigation of biological cell membrane voltage".
Currently he is working as Vice Dean for student affairs, Helwan University.

A. M. Hamdy was born in Cairo in Egypt, on August 17, 1966. He graduated from the Telecommunications and Electronics Department, Faculty of Engineering and Technology of Helwan University, Cairo, Egypt in 1989. He received the M.Sc. degree in computer engineering from the same university in 1996 and the Ph.D. degree from the Faculty of Electrical Engineering, Poznan University of Technology, Poland in 2004.
Currently he is working as a lecturer in Faculty of Engineering of Helwan University. His special fields of interest include image processing, pattern analysis, and machine vision.

M. A. Saleh was born in Cairo in Egypt, on??? 00, 0000. B.Sc. in Telecommunications and Electronics, Faculty of Engineering, Helwan University, Cairo, Egypt, June 1988. M.Sc. in Computer Science, Rochester Institute of Technology, Rochester, NY, USA, May 1995. Ph.D. in Computer Science and Engineering, University of Connecticut, CT, USA, August 2003.

Currently he is working as

- Director, Scientific Computing Center, Helwan University, Dec. 2004- Till now.
- Director, Helwan University Campus Network, Dec. 2004 – Till now.
- IT Consultant for the Ministry of State for Administrative Development for the e-Governmen Projects, 2005 – Till now.

A. R. Ahmed was born in Cairo in Egypt, on April 09, 1984. B.Sc. in Electronics and communication, Faculty of engineering, The Arab Academy for Science and Technology Maritime Transport, Cairo, Egypt in June 2006.
Currently he is working as a communication teaching assistant in communication and electronics department, MTI University. His special fields of interest include image processing, pattern analysis.