3D Face Recognition Method Using 2DPCA-Euclidean Distance Classification

Yashar Taghizadegan[1], Hassan Ghassemian[2], and Mohammad Naser-Moghaddasi[3]
[1,3] Department of Electrical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
Email: [1]yashar.taghizadegan@gmail.com, [3]mn.moghaddasi@srbiau.ac.ir
[2] Department of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran.
Email: ghassemi@modares.ac.ir

***Abstract*— In this paper, we present a 3d face recognition method that is robust to changes in facial expressions. Instead of locating many feature points, we just need to locate the nose tip as a reference point. After finding this reference point, pictures are converted to a standard size. Two dimensional principle component analysis (2DPCA) is employed to obtain features Matrix vectors. Finally Euclidean distance method is employed for classifying and comparison of the features. Experimental results implemented on CASIA 3D face database which including 123 individuals in total, demonstrate that our proposed method achieves up to 98% recognition accuracy with respect to pose variation.**

***Index Terms*— 3D Face recognition, depth information, features vectors, two dimensional principle component analysis (2DPCA), Euclidean distance**

I. Introduction

Face recognition attempts, return to last century when Galton [2] did first research to recognise face. Face recognition and identification have so many applications such as in police department and security system to identity guilty people. 2D face recognition system has some serious issues with light environment sensitivity, face expressions (such as happy, unhappy, wondering ...) and also turning of head. To remove these issues, using 3D pictures is suggested. The 3D systems are not sensitive to transformation, turning and light condition [3-6]. 3D pictures are the pictures that instead of light condition level of pixels, the deep information of pixels exist.

A. Previous Works

Researchers have employed different methods for automatic 3D face recognition. Some methods are based on the face curve analysis. Gordon [7] has presented a method based on algorithm using 3D curve face features. In Gordon's method face is divided to three subdivision named: ridge and valley lines and then the location of nose, mouth, eyes, and other features which are used for recognition, specified. Lee et al [8] presented a method based on locating features of eight points on face and using supportive vector machine [SVM], get the 96% accuracy in face recognition. Database used by Lee, contains of 100 different pictures. In Lee's method face features points are chosen manually and this can be one of the issues of this method. Moreno et al. [9] employed a set of eighty six features using a database of 420 3D range images, 7 images per each one of a set of 60 individuals. After the feature discriminating power analysis, the first 35 features of the ordered list of features according to the Fisher coefficients were used to represent faces in their face recognition experiments. The features offering better recognition results were angles and distances measurements.

Chang et al. [10] proposed to use the principal component analysis (PCA) for extracting the 3D facial appearance features, which achieved a promising recognition performance. After that, many appearance based features are applied to improve the 3D face recognition results.

Khalid et al. [11] presented a method using 53 features extraction from 12 reference points. The local geometric features are calculated basically using Euclidean distance and angle measurement.

B. General overview of recommended method

In this paper a novel algorithm for face recognition that is robust to changes in facial expressions (such as unhappy, wondering and…) will be presented. This method provides the face recognition with high level of speed and accuracy and deals with facial expressions. The experimental results are performed on CASIA 3D face database. The block diagram of suggested method is shown in Fig. 1. 3D sample pictures of CASIA 3D face database relates to different facial expressions are shown in Fig. 2

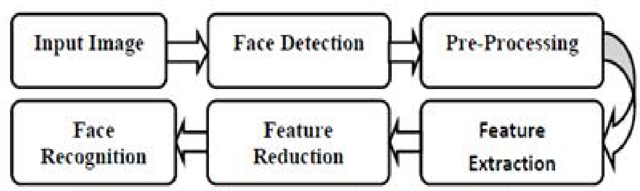


Fig. 1: The block diagram of suggested method

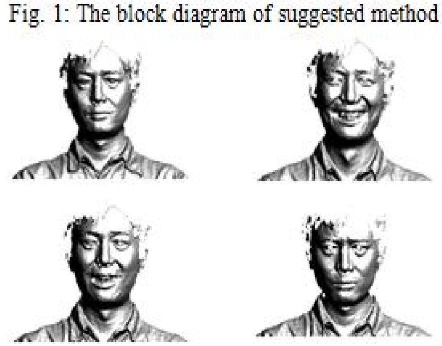


Fig. 2: normal (up-left)- Happy (up-right)- wondering (down-left)- unhappy (down-right)

DOI: 01.IJCSI.03.01.70

II. Our Method

A. Face detection

As shown in Fig. 3 unprocessed information of data contains some additional information like neck, ear, hair, and some parts of clothes which are not usable and should be deleted. In this paper thresholding method over the pictures depth coordination (*Z*) is used.

This method is named Otsu [12]. The *Z* coordination of pictures is divided to the facial and non-facial information, and then Otsu algorithm specified the best estimation of threshold border. All the information of threshold will be kept and unused information will be eliminated.

Assumed $p(i); i = 1,2,..., I$ .are values of normalized histogram, including *I* levels. Assumed the value of *I* (coordination info) are divided to C_1 and C_2 classes. The value of C_1 is $i=1$ to k and C_2 is $i=k+1$ to I.

In this case, probabilities of happening for each class are shown in Equation (1) and (2).

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

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

And also average μ_i and variance σ_i^2 of C_1 and C_2 classes are calculated in Equation (3) through (6).

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

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

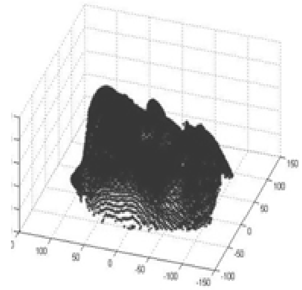
$$\sigma_1^2(k) = \sum_{i=1}^{k} [i - \mu_1(k)]^2 .P(i) / c_1(k) \quad (5)$$

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

The best value of threshold *T* with recursive method is specified in a way, for all *K* values from 1 to *I* the variance of σ_w^2 minimized.

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

Fig. 3 shows the data of a picture used in Cartesian coordination presentation and Fig.4 shows the final picture with elimination of unused points.



FiFig. 4: The Fig. 3 after elimination of unused points

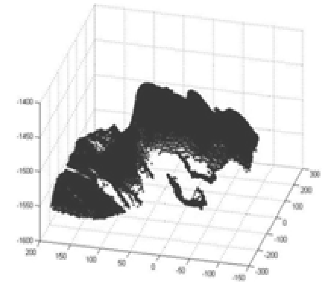


Fig. 3: The data of a picture before elimination of unused points

B. Pre-processing

3D pictures are taken by laser camera, have spark noise and also some holes in some areas of pictures. These holes and noises reduce the recognition system accuracy. Median filter is used for eliminating spark noise, and interpolation is used for filling the probable holes. In accordance of Fig. 3, *X* variable change from 150 to 200 and *Y* from 100 to 200. These limits for different pictures even for a person are not the same, there for it is not possible to compare the pictures. For a meaningful comparison of pictures, the pictures should be normalized to a 100×100 network. The limit of *Z* coordinate as the depth of the pictures should be mapped in [0 255]. In the following, the complete explanation of this method will be presented.

After face detection, the difference between max and min values of *X* & *Y* has been obtained and with 1/99 steps, is sampled, the obtained pictures are mapped to a 100×100 net and Fig. 5 is the result. In regarding of inequality in pictures depth information and different changes, it is necessary to map the pictures depth info to a space between the 0 to 255 levels. The results have been shown in Fig. 6.

C. Face borders and nose tip detection

To obtain the face borders, minimize a 100×100 square such as the corners of square touch the face corners. Result of this process has been shown in Fig.5 Final obtained 100×100 square with the nose tip will be mapped in a 100×100 net. After detecting of face borders, regarding of this fact that the nose tip has the highest height in the pictures, a simple method will provide the coordination of nose tip by this way: pictures will be scanned with a 3×3 window and sum of all points inside and below the 3×3 window will be obtained, the largest number in these data is the nose tip. For some pictures the depth of chin is more than nose, so in this kind of pictures, to prevent wrong nose info, the method will accept just the points in the central areas of pictures. In other word, if an obtained point as nose centre point, is locates up or down of the picture, it means that this point is not the nose tip and the next maximum point should be regarded as the nose tip. It is necessary to continue this procedure to approach as much as possible to the centre of picture. Finally the nose tip should be placed in centre of picture as shown in Fig.6.

DOI: 01.IJCSI.03.01.70

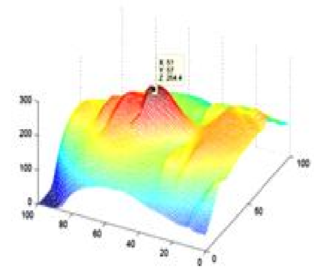


Fig. 6: The final picture after pre-processing

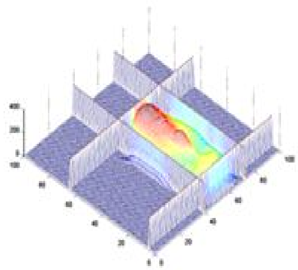


Fig. 5: The picture showing obtained borders of the face

D. Face Smoothing

Face smoothing converts face curves to a smooth and soft surface and removes the facial expressions. By this way, before processing the information, the facial expressions will be eliminated. In this paper, minimum square error method is used for minimizing the errors between input pictures and final smoothed picture. For this case variance of input picture is calculated and then entire picture matrix is scanned by using a window with $p\times q$ size. Scanning of matrix is started from up and left side and element by element. In each scan the element of window centre will be changed by Equation (8). For each window, local average value and variance is calculated.

$$s(i,j) = \mu_{p,q} + \frac{\sigma^2_{p,q} - \sigma^2_n}{\sigma^2_{p,q}}(I(i,j) - \mu_{p,q})\,. \qquad (8)$$

In Equation (8), $\mu_{p,q}$, is average value and $\sigma^2_{p,q}$ is depth value variance matched with p×q size window. σ^2_n Is the variance of noise, $I(i,j)$ and $S(i,j)$ are the primary and the normalized pictures.

The scanned window size specifies the level of smoothing. A bigger window will causes more smoothing level, but the level of smoothing should not be in a way that the features and important information of picture be removed. Fig. 7 shows an unsmooth picture from CASIA 3D face database and Fig. 8 shows the same picture after smoothing. Fig. 7 shows a smiley face but in Fig. 8 the smile is removed from picture.

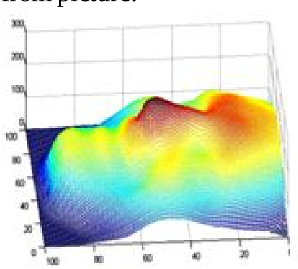


Fig. 8: The Fig. 7 after smoothing

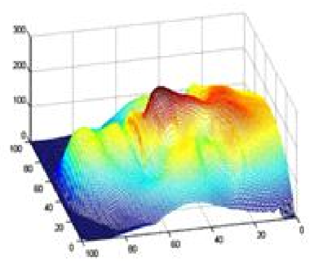


Fig. 7: An unsmooth picture from CASIA 3D face database

E. Feature extraction

The features should be in such a way that for two different persons, their pictures are clarified from each other but for one person, different pictures contain same information. By using Two Dimensional Principle Component Analysis (2DPCA), the features of pictures have been obtained.

The normalized pictures are named $A_1, A_2, ..., A_M$, the size of each A_i assumed $n\times n$. The average of pictures A_1 through A_M is defined in Equation (9).

$$\overline{A} = \frac{1}{M}\sum_{i=1}^{M} A_i\,. \qquad (9)$$

The covariance matrix of M picture will be gotten from Equation (10).

$$C = \frac{1}{M}\sum_{i=1}^{M}(A_i - \overline{A})^T(A_i - \overline{A})\,. \qquad (10)$$

In Equation (10), T is the transpose matrix. The covariance matrix has n Eigen value and correspond n Eigen vectors. These n Eigen values stored descending and the numbers of Eigen vector (d) correspond with d as the largest Eigen value in $n\times d$ matrix. A is assumed as one of the input pictures. In this case the features vector of picture will be define as Equation (11)

$$Y = A.X\,. \qquad (11)$$

In Equation (11), X is a $n\times d$ matrix and the first column of this matrix relates to the biggest Eigen value, the second column corresponds to the second biggest number and so on. Finally the d[th] column corresponds to the largest Eigen value of covariance matrix. In Equation. (11) the number of d should be calculated such a way that the face picture is recoverable from Eigen value matrix of Y. and vector of X, and also d should be obtained such a way that the recognition rate is a desirable value. In this identification system with $d=14$, the identification rate is maximum. Although, the d value can be from 1 to 100, it is clear that with $d=14$ the size of features matrix will be reduced from 100×100 to 100×14. So small matrix creates a high processing speed and a good combination between accuracy and speed will be approach with $d=14$. Improved value of d can be obtained using Equation (12).

$$e = \sqrt{\frac{1}{1000}\sum_{i=1}^{100}\sum_{j=1}^{100}[A(i,j) - \tilde{A}_d(i,j)]^2} \le 0.02\left[\frac{1}{10000}\sum_{i=1}^{100}\sum_{j=1}^{100}A(i,j)\right]. \qquad (12)$$

In this Equation $\tilde{A}_d$ is the obtained processed picture using d to Eigen values corresponding with d to the largest Eigen values. The lowest value of d in Equation (12) can be a good value for a low rate fault and good improved picture. Calculation of d for CASIA 3D face database is 14 and the processed picture will be obtained from Equation (13).

$$\tilde{A}_d = YxTranspose(X)\,. \qquad (13)$$

Fig.9 shows the Eigen value of covariance matrix versus its index for CASIA 3D face database.

In Fig. 7 3D facial mesh of a sample picture from CASIA 3D face database is shown. Fig. 10 shows reconstructed picture of Fig. 7 using some different Eigen vector. The Figs.

DOI: 01.IJCSI.03.01.70

10-a to 10-d show reconstructed picture using 4, 8, 12 and 14 Eigen vectors respectively. It is clear that when we use 14 Eigen vectors, Fig. 7 will be reconstructed completely.

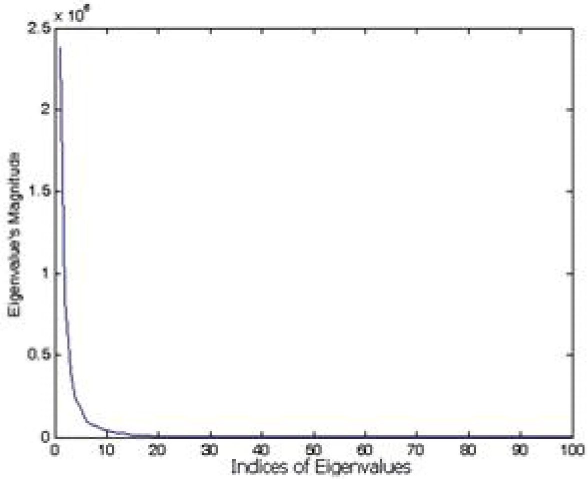


Fig. 9 : The Eigen value of covariance matrix versus its index

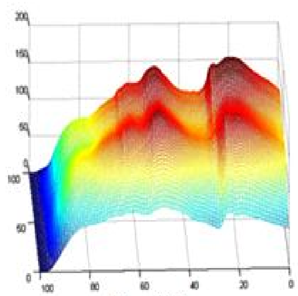


Fig. 10-b

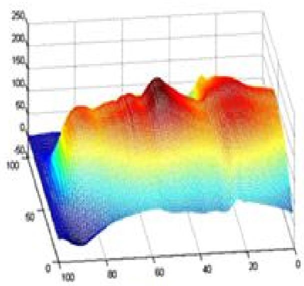


Fig. 10-a

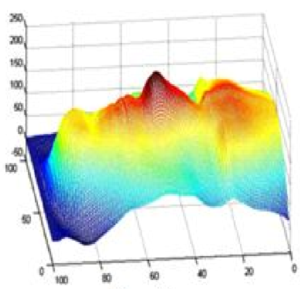


Fig. 10-c

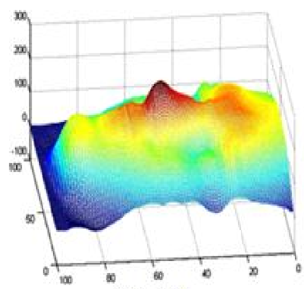


Fig. 10-d

E. Classification

Euclidean distance is used for classification and obtaining the level of similarity. The Euclidean distance between two vectors can be obtained from Equation (14).

$$d(X,Y)=\sqrt{\sum_{i}^{n}(X(i)-Y(i))^{2}}\,. \qquad (14)$$

In this Equation n is the numbers of *X* and *Y* vector component. While all the pictures exist in data base convert to *100×100* and finally approach to *100×14* matrix. Therefore the entire features vectors have the same size and it is possible to use Euclidean distance method. Assume the $F=(Y_1,\ldots,Y_d)$ is the features vectors and the features matrix is F. The similarity and matching between Fi and Fj can be gotten from Equation (15).

$$d(F^i,F^j)=dist(Y_k^i-Y_k^j)\,. \qquad (15)$$

The $dist\,(Y_k^i-Y_k^j)$ is the Euclidean distance between two vectors of Y_k^i, Y_k^j

F. Experimental results

The Experimental results are implemented on CASIA 3D which including 123 individuals in total. Each individual contains the variations of illuminations, expressions and poses, which are the main problems for depth modalities. The Experimental results shows that *7×7* window provides the highest rate of identification. Fig. 11 shows identification rate diagram in order of pictures numbers for different windows sizes. As shown in this graph for windows bigger than *7×7* the identification rate is decreased. It means that more smoothing removes more features and information from pictures. In process with 12 training pictures the rate of identification respect to pose variation is 98%.

As shown in chapter (II), the $(Y_1,\ldots,\ Y_d)$ are the features vectors and the graph in Fig. 9 shows that the most probable information can be approach using limited features vectors. In process with d=14 the rate of identification respect to pose variation is 98%. As the features of pictures have same type of principle component, the best classification is the Euclidean distance. So, for testing and training of pictures this method was used. The database used in this method is CASIA 3D face database, containing of 4674 3d pictures. In this data base the pictures of 123 persons exist and for each person 38 pictures have been taken. All of these pictures have been divided to two training and testing categories. Table I shows the accuracy of system versus different numbers of training pictures. We validated our proposed method and compared it with existing methods using the CASIA 3D face database [13, 14]. Table II shows the recognition rate of our system in comparison with other methods.

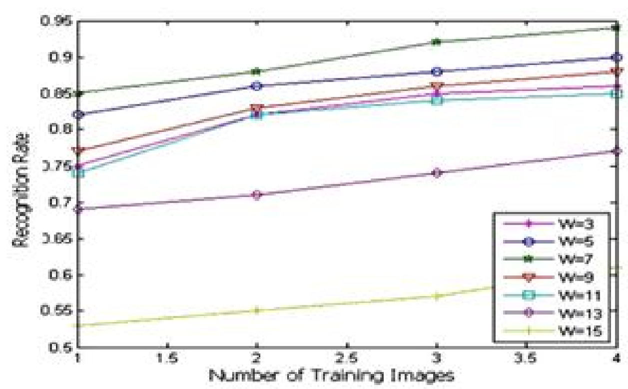


Fig. 9: identification rate diagram in order of training pictures numbers for different windows sizes

TABLE I: THE SYSTEM ACCURACY VERSUS DIFFERENT NUMBERS OF TRAINING PICTURES

Number of pictures for training	9	10	11	12
The rate of identification	92%	95%	97%	98%

DOI: 01.IJCSI.03.01.70

TABLE II: THE RECOGNITION RATE OF THREE METHODS

Methods	Li [13]	Ming [14]	our method
Recognition Rate	91.1%	94.17%	98%

CONCLUSION

In this paper a novel method was presented for automatic face recognition using three dimensional pictures. At the first step, 3D pictures were pre-processed and then pre-processed pictures were detected. Pictures of database contained some additional areas such as neck and cloths. These parts didn't have useful information and they had to be eliminated. After face detection, pictures were post-processed. In this stage some effective processing that had more effects on system identification accuracy were done. Also some processes were done that made identification system immune against face expressions. By using a simple method, nose tips were obtained and they used as references. In the next stage each 3D pictures were normalized to a 100×100 matrix and smoothing of face picture was done. In features extraction section for obtaining features two dimensional principle component analysis (2DPCA) was used and in classification stage Euclidean distance method was used. Finally experimental results implemented on CASIA 3D face database showed our system t had a good immunity against face expressions. In this paper the best recognition rate is obtained 98% and it is accepted rather to other similarity methods implemented on CASIA 3D face database. Combination of 2D and 3D systems improve efficiency and the rate of face detection systems, therefore combination of these two systems can be one of the futures works. Another future works that can be continued is working over the face occlusion concept. To be able to work and study over the face coverage concept, another case for future work can be a presentation for exact and accurate estimation of face rotation angle.

REFERENCES

[1] CASIA 3D face database, National Laboratory of Pattern Recognition, CASIA, P. O. Box 2728, Beijing, 100190, China

[2] S. F. Galton, Numeralizaed profiles for classification and recognition, Nature 83, pp. 127-130, 31 March 1910.

[3] Chenghua Xu, Stan Li, Tieniu Tan, Long Quan, *Automatic 3D face recognition from depth and intensity Gabor features*, Science Direct, Pattern Recognition Journal, VOL 42 , doi:10.1016/j.patcog, 2009.01.001.

[4] K. W. Bowyer, K. I. Chang, and P. J. Flynn, *A survey of approaches to three-dimensional face recognition*, International Conference on Pattern Recognition, 1:358361, 2004.

[5] B. Moreno, A. Sanchez, J. F. Velez, F. J. Diaz, *Face recognition using 3D local geometrical features PCA vs. SVM*, Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, 2005.

[6] Thomas David Heseltine BSc. Hons, face recognition/ T*wo-dimensional and three-dimensional techniques*, Ph.D. Thesis, The University of York, Department of Computer Science, September 2005.

[7] G. Gordon, *Face recognition based on depth maps and surface curvature*. Geometric Methods in Computer Vision, SPIE 1570:1–12, July 1991.

[8] Y. Lee, H. Song, U. Yang, H. Shin, K. Sohn, *Local feature based 3D face recognition*, International Conference on Audio- and Video based Biometric Person Authentication (AVBPA 2005), LNCS, vol. 3546, July 2005, pp. 909–918.

[9] A. B. Moreno, A. Sanchez, J. F. Velez, and F. J. Diaz., *Face Recognition using 3D surface-extracted descriptors*, In Proceedings of the Irish Machine Vision & Image Processing Conference 2003 (IMVIP'03), Ulster, 3-5 September 2003.

[10] K. I. Chang, K.W. Bowyer, and P. J. Flynn., *An evaluation of multi-modal 2d+3d face biometrics*, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 27(4):619–624, 2005.

[11] Fatimah Khalid, Tengku Mohd. Tengku Sembok, Khairuddin Omar, *Face Recognition Using Local Geometrical Features – PCA with Euclidean Classifier*, 978-1-4244-2328-6/08/$25.00 © 2008 IEEE.

[12] 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.

[13] Yong-An Li, Yong-Jun Shen, Gui-Dong Zhang, Taohong Yuan, Xiu-Ji Xiao, *An Efficient 3D Face Recognition Method Using Geometric Features*, 978-1-4244-5874-5/10/$26.00 ©2010 IEEE.

[14] Yue Ming, Qiuqi Ruan, Xueqiao Wang, Meiru Mu, *Robust 3D Face Recognition using Learn Correlative Features*, ICSP2010 Proceedings, 978-1-4244-5900-1/10/$26.00 ©2010 IEEE.

DOI: 01.IJCSI.03.01.70