

Available online at

<http://www.TJMCS.com>

The Journal of Mathematics and Computer Science Vol. 4 No.2 (2012) 139 - 152

## Implementation of a System for 3D Face Detection and Recognition

Taiebeh J. Askari<sup>\*1</sup> and A.H. Hadjahmadi<sup>2</sup>

<sup>1</sup> Faculty Member, Islamic Azad University, Iranshahr Branch, Iranshahr, Iran,  
t.askari@iauiranshahr.ac.ir

<sup>2</sup> Faculty Member, Department of Computer Engineering Vali-e-Asr University of Rafsanjan, Iran,  
hadjahmadi@vru.ac.ir

Received: January 2012, Revised: April 2012

Online Publication: June 2012

### ABSTRACT

This paper provides a method for detection and recognition face of man using 3D face extracting and applying 3D features. In this paper, it is used from the Bio-ID package contained 1520 images of man face. A number of these images were used for 2D face modeling; another number was used for 3D face modeling, and another number was used for test.

At first, the landmarks on the face image are determined, semi automatically. Then the shape, texture and appearance model of face images is constructed. Using these models and the fast Active Appearance Model search, the landmarks on the test image are determined. More ever, from 24 3D images, obtained by a 3D scanner, the variations of shape, texture and appearance are modeled. Using the 3D models, 2D landmarks and an 3D Initialized Active Appearance Model Search method (3D IAAMS), the 3D frame of face (this frame is described by 3D landmarks) in an image is constructed. These 3D frames with the texture are used for face recognition.

**Keywords:** Principal Component Analysis, Shape Modeling, Texture Modeling, Appearance Modeling, Face Model, 3D modeling, face detection and recognition.

### 1. INTRODUCTION

The applications in the world for face recognition such as security, surveillance, gaze-based control, video compression, animation and user interfacing, are led to develop face recognition systems commercially.

---

\* Corresponding Author

Whereas some of these systems is conditionally successful, but the face-recognition problem is challenging because faces are non-rigid and have a high degree of variability in size, shape, color, and texture (pose) [13]-[15].

The methods that were successful in face recognition commonly are used for images that show the face perfectly or with a few changes in the direction of the face. For the images with half-roc faces 2D face recognition methods with 100% performance are not represented. The 2D face recognitions are usually so. These systems cannot perform the face recognition perfectly, whereas, the 3D systems are not face into the challenges such as the position and orientation of the face. Because in a 3D face, the whole face is visible. Some of 3D systems use the 3D features as a bridge for building 2D faces with different view angles. Others used 3D geometric features, only [5], whereas, by using the 3D features with the texture features and with the 2D features, a high performance, accurate and high speed system for face recognition can build.

There are a collection of features in the 3D shapes that not in the 2D shape. Such as that the whole face with all of its pieces on a 3D shape is visible. If we have the 3D shape of a face, then we can use the all of points in this shape in the face recognition.

However the systems that are presented now for the face recognition are not able use the whole 3D features. But it has some reasons. For example the set of the 3D faces that are obtained with scanners have not coefficient quality and quantity [7]. Attention to these facts, in this paper we represent a method for the face recognition that uses the 3D features, the textures (color intensities of the pixels) and the 2D features (the range of the face in the 2D image) for face recognition.

At first the shape, texture and appearance model of face images is constructed. Using these models and the fast Active Appearance Model search, the landmarks on the test image are determined [14].

On the other hand, the 3D model of shape and texture variations and appearance model are built from 24 3D images that obtained by a Minulta laser scanner. Using these models and using the 2D landmarks and the 3D AAM search, the 3D frame of the face (3D landmarks) can be extracted from the training and test images.

The face recognition system uses the 3D frame, the landmarks on 2D image and the color intensities of the pixels for recognizing.

## **2. Data collection**

The 2D images used in this paper are taken from the Bio-ID images contained 1520 varied face images from 50 men (man and woman). The 3D images used in this paper are contained 24 3D face image. These images are taken with a laser scanner in school of dentistry, Denmark University [6].

## **3. 2D Face Detection**

In this section, the full range of man face is extracted. This range contains the face round from the top of an eyebrow to the bottom of chin and from the left ear to the right ear (no contains two ears). For this purpose, at first, the landmarks in the 2d face image are obtained, semi automatically, Then the variations of shape, texture and appearance be modeled. Finally, using these models and the active appearance model (AAM) search algorithm, the landmarks in a new face (test) image are obtained. These landmarks are used in extracting the 3D face frame.

### 3-1. Obtaining the Landmarks

The obtaining of the landmarks in 2D images can be done by placing the points manually. However, the methods are presented that can do it automatically. But, these methods are limitedly applied. Namely, these are used for certain applications, and, for all applications not working properly. Because of the landmarks' number in 2D images is low, in some cases lower than 100 landmarks, it can be obtained manually. In this paper, a program is implemented in Matlab that the landmarks are obtained by clicking on the landmarks and storing the coordinate of the points in a vector. Then the landmarks of all 2D images are stored in a database. As an example, figure1 shows a face image and the 81 landmarks on it. Thus, a shape is represented by a vector of points  $[x, y]$ :

$$x = [x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_n]^T \quad (1)$$

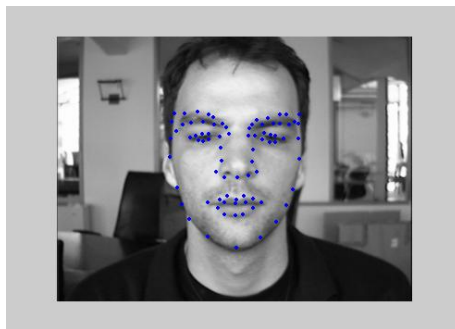


Figure 1: 81 specified landmarks in a face.

### 3-2. Shape Alignment

To obtain a true shape representation (according to our definition) location, scale and rotational effects need to be filtered out (The location, scale and rotation commonly known as pose). This is carried out by establishing a coordinate reference to which all shapes are aligned. One can use from the Procrustes analysis procedure for shape alignment. [1]-[14].

### 3-3. Modeling Shape Variation

After shape alignment, the shape model is created using principal component analysis (PCA) of the covariance matrix of aligned shapes. PCA is defined by H. Hotelling (1933) [4]. By performing PCA on covariance matrix, the eigenvectors,  $\Phi_s$ , and eigenvalues are obtained. Then a shape instance can be generated by deforming the mean shape by a linear combination of eigenvectors.

Essentially the point or nodal representation of shape has now been transformed into a modal representation where modes are ordered according to their deformation energy – i.e. the percentage of variation that they explain. Notice that an eigenvector is a set of displacement vectors, along which the mean shape is deformed.

Figure 2 shows shapes generated using change in the first four modes of shape model ( $b_{s(1)}$  to  $b_{s(4)}$ ).

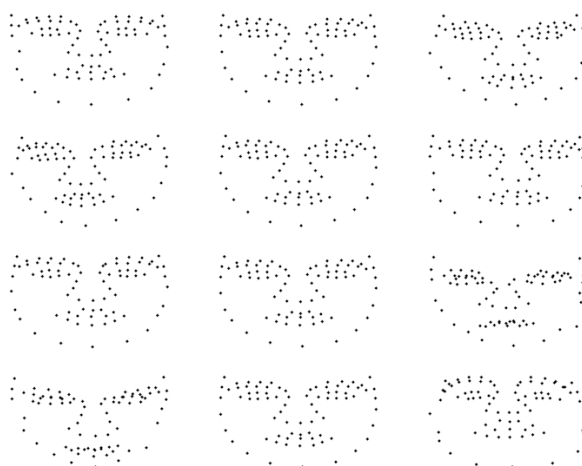


Figure 2: Generated shapes by changing the first four modes of shape model  $b_{s(1)}$  to  $b_{s(4)}$ .

### 3-4. Texture Modeling

To form a complete model of appearance one must not only consider shape, must also consider the information constituted by pixels themselves.

In the shape case, the data acquisition is straightforward because the landmarks in the shape vector constitute the data itself. In the texture case, one needs a consistent method for collecting the texture information between the landmarks, i.e. an image warping function needs to be established. This can be done in several ways. Here, a piece-wise affine warp based on the Delaunay triangulation of the mean shape is used. Thus to obtain texture information from the training set, each shape is warped to a reference shape (here the mean shape is used as reference shape) and sampled.

Hereafter a photometric normalization of the obtained textures is done to remove influence from global linear changes in pixel intensities. Hence, a compact PCA representation is derived to deform the texture in a manner similar to what is observed in previous section [14].

Figure 3 illustrates the mean shape of the normalized texture vectors for face images.



Figure 3: The mean shape of the normalized texture vectors for face images.

#### 3-4-1. Modeling Texture Variation

For texture modeling, we used PCA procedure as in shape modeling is used.

Figure 4 illustrates changing in the first four modes of texture modeling and resulted textures.



Figure 4: Generated textures by changing the first four modes of texture model  $b_{g(1)}$  to  $b_{g(4)}$ .

### 3-5. Combined Model of Shape and Texture

How to unify the presented shape and texture models into one complete compact model, called appearance model is the topic of this section. After this, the appearance model is called the complete model.

### 3-6. Combining of Shape and Texture Models

From the previous chapter, it was seen that an object instance can be constructed using the two set of model parameters of shape,  $b_s$ , and texture,  $b_g$ . The appearance model is then generated by combining two shape and models. Therefore, a 3rd PCA is performed on the concatenated shape and texture parameters,  $b$ , of the training set to obtain the combined model parameters,  $c$ :

$$b = \Phi_c c$$

where  $\Phi_c$  denotes a set of eigenvectors.

In figure 5 changes in the first four modes of appearance model parameters ( $c(1)$  to  $c(4)$ ) and the produced instances are shown.



Figure 5: Changes in the first four modes of appearance model parameters ( $c(1)$  to  $c(4)$ ) and the produced instances.

### 3-7. Active Appearance Model (AAM)

One of the pleasant properties of the statistical models of shape and texture is that it is possible to use these to search images for new instances of the class of objects that they represent. This is called the Active Appearance Models.

The foundation of the AAM search is to treat the search as an optimization problem in which the difference between the synthesized object delivered by the AAM and an actual image is to be minimized. Formally this can be written as the difference vector  $\delta I$ :

$$\delta I = I_{Image} - I_{Model} \quad (2)$$

In this way the fit can be enhanced by adjusting the model and pose parameters to fit the image in the best possible way. Since the following will be based on normalized texture vectors,  $\delta I$  will be denoted as  $\delta g$ .

### 3-8. Solving Parameter Optimization Off-line

It is proposed that the spatial pattern in  $\delta g$  can predict the needed adjustments in the model and pose parameters to minimize  $\delta g$ . The simplest model we can arrive at constitutes a linear relationship:

$$\delta c = R \delta g \quad (3)$$

To determine a suitable  $R$  in equation (4) a set of experiments are conducted which is fed into a multivariate linear regression framework.

The multivariate linear regression algorithm is fully described in [12]. Each experiment consists of displacing a set of ground truth parameters by a known amount and measuring the difference between the model and the part of the image below the model. The value of these changes and the manner of performing these experiments is fully described in [14].

#### 3-8-1. Iterative Model Optimization

The previous sections have established a firm foundation for the optimization process in AAMs. This section will focus on the details in using the linear regression predictor in what is coined the AAM search.

Suppose that a test image contain a face is available. The purpose of this algorithm is extracting the face round from this image. At first, it is considered a shape and texture as initial. Then extract the area of image under the initial shape and then compares it with the texture of model. According to this difference, an adjustment is performed on model parameters ( $\delta c$ ). This is repeated until no improvement is observed in the model fit. A pseudo code of this algorithm is available in [14].

Figure 6 shows the AAM search (using texture) on an unknown image that converges after 12 iterations. In comparison with the original image, the AAM is somewhat blurrier because of the texture approximation.



Figure 6. Search result: From left to right: The initial model; The AAM after 2 iterations; The AAM after 10 iterations; The converged AAM (12 iterations).

#### 4. 3D Face Detection

The systems that recognize objects, in particular, the objects that are non-rigid, using 2D features are failed in the most cases. An example of these objects is the face image. As we know, many images can be taken from the face of a man, in different states of eye, lip, mouth and other pieces of the face and the different states of placing relative to the camera and in the various lighting conditions. If we want to use the 2D systems for the face recognition, if this system is not strong, consequently, the two images of a man in different poses it may recognize as two men. A 3D image has many important features. For example, the position of the man relative to the camera isn't effective on the recognition. On the other hand, a 3D image contains the all of pieces of the face in all different views [11].

For these reasons, the recognition systems are ventured use of the 3D features. Thus, it must be used from a method that extracts the 3D frame from 2D image. In the previous sections, the extraction of the face round from a 2D image is described. In this section, we use the 2D extracted face round and the 2D landmarks to extracting the 3D frame of face, and then will use this frame for face recognition.

The various but low methods are proposed for this purpose. Because the AAM works well with non-rigid objects such as faces, in this paper to extracting the 3D frame of face, the 2D AAM is generalized into 3D AAM and with a robust initialization method. This initialization caused the algorithm will be accurate and quick.

##### 4-1. 3D Face Modeling

As described above, to detecting the 3D frame of face in an unknown image, at first, we must model the variations of shape, texture and appearance of 3D faces, and then, using these models and an initialized 3D AAM search, extract the 3D frame of face.

The 3D shape is described with the landmarks same as 2D shape. In 2D case, obtaining 2D landmarks is very easy, because the number of them is about hundreds, in most cases such as face the shape is described using less than one hundred. In case of face, about 80 landmarks are sufficient to describing 2D shape. Therefore the landmarks can be obtained manually. But in 3D case, thousands of landmarks are required to capture the complex surface of a human face. Obviously, it is not feasible to annotate these by hand. Instead, some sort of automated process is necessary. A semi-automatic algorithm is used here, which constructs a dense distribution of corresponding points from a sparse set of manually placed landmarks [8]-[9]-[10].

The unregistered face data is unordered. This means that a certain 3-D point can be placed anywhere on a face. On one face the point can be part of the ear, while on another it is part of the nose. If each point represented the same position on every face, they would be in correspondence. The idea of the registration algorithm is to change the order of the points to give all objects the same representation. If all objects have an equal amount of points and the same point ordering, every point will act as a landmark.

The algorithm uses one of the objects as a template shape. The idea is to pick a suitable face as the template, and then change the extent and point ordering of the other faces to match the template [14].

##### 4-2. Defining a Sparse Set of Landmarks

As described earlier, a limited number landmark must manually be determined for using in registration algorithm. Here nine landmarks were used as shown in figure 7 and were listed in table 1. These landmarks are obtained in Matlab, manually.



Figure 7: Nine landmarks on the face.

Table 1: The list of nine landmarks on the face.

Number	Description
1	The tip of the chin
2	The left corner of the mouth
3	The right corner of the mouth
4	The tip of the nose
5	The left corner of the left eye
6	The right corner of the left eye
7	The minimum of the curve defined by the adjoining of the nose and the forehead
8	The left corner of the right eye
9	The right corner of the right eye

After the forming the Dense Correspondence, the shape alignment and removing the location and orientation effects, the achieved mean is viewable.

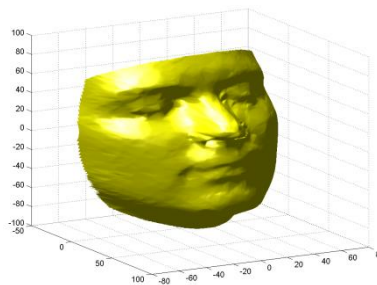


Figure 8: The mean of the 3D shapes after removing the location and orientation effects.

### 4-3. Building a Shape Model

A shape model produces linear combinations of the training shapes. This makes it possible to produce shapes not present in the training set. The shape space created by the model spans all the training shapes, all in between shapes as well as extrapolations, depending on the coefficients. As long as the choice of coefficients is reasonable, plausible, i.e face-like, shapes are generated. The 3D shape modeling is fully similar to 2D shape modeling. Therefore, we don't repeat it and only we have shown the results of changing in the first four modes in the shape model.





Figure 9: The shapes resulted from the changing the first four modes in the shape model.

The choice of the suitable values for parameters and dimension reduction is similar to the 2D shape modeling.

#### 4-4. Building a Texture Model

For building the texture model, at first, the effect of the color intensity is removed, and then the textures must be normalized. The removing the effect of the color intensity and the normalization and the building texture model in 3D is similar to the 2D. For this reason, we only pay for showing the results. Figure 10 shows the mean of the textures after normalization. Figure 11 shows the textures resulted from the changing in the first four modes of the texture model.



Figure 10: The mean of the textures after the normalization.

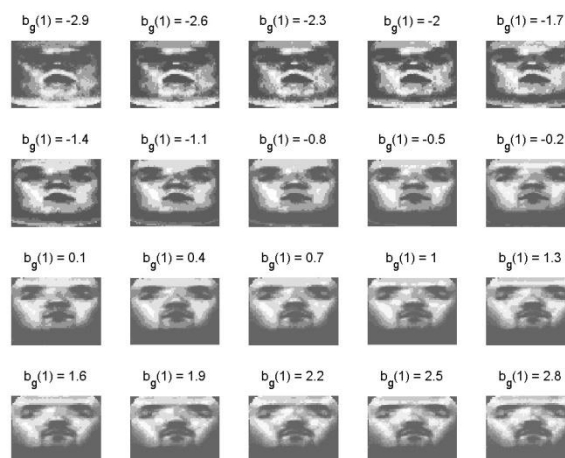


Figure 11: The textures achieved by changing in the first four modes in the texture model.

#### 4-5. Building an Appearance Model

The combination of the shape and the texture and constructing a complete model in 3D case is similar to the 2D case. Here we only pay for showing the results of changing in the first four modes in the appearance model. Figure 12 shows these results.



Figure 12: The results of the changing in the first four in the appearance model.

#### 4-6. Extracting the 3D Frame of an Unknown Face

The principle of this paper is the extraction of the 3D frame of a test face that is not in training set and the recognition of it. Namely, the 3D frame of a face is extracted and this frame is compared with the frames stored in the database and then the most similar frame with it is selected. On the other hands, the man who is the most similar to the test face is determined.

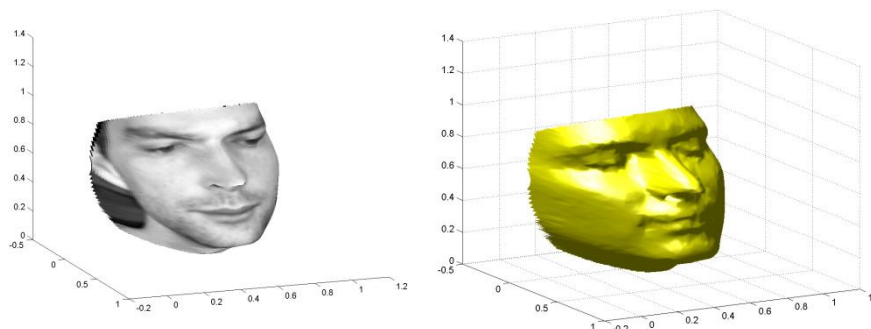
For the constructing the 3D database and, too, for the determining an unknown face, the extracting of the 3D frame is a principle stage. As it is clear, it must use from the 3D face model that is constructed. The approach is used here is the 3D AAM that has not some difference with 2D AAM.

At first, a number of 2D images, that the landmarks on them are obtained, are selected (it is tried to select at least one image from a particular man as possible as), and then, the 3D frame of them is extracted (using 3D AAM). The 3D shape and the texture of them is stored in a database for using in recognition.

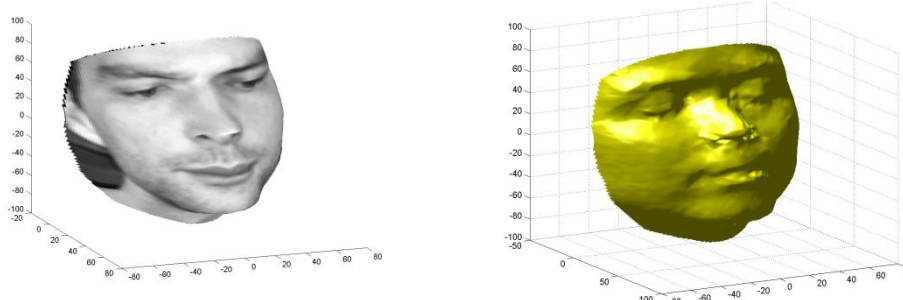
The major difference between the 2D AAM and the 3D AAM, used in this section, is the using of an initializing algorithm in the 3D AAM. Namely, for initializing, it is used from the nine landmarks for determining an initial 3D shape. This initial 3D shape is selected in a manner that the eight landmarks of the 2D shape are matched with eight landmarks of the 3D mean shape. The selection of an initial shape is resulted to convergence in the algorithm.

The remaining of the 3D AAM algorithm is similar to the 2D AAM algorithm. Here we pay to represent an example is shown in Figure 13.

The first epoch



The 10 epoch



The 11 epoch

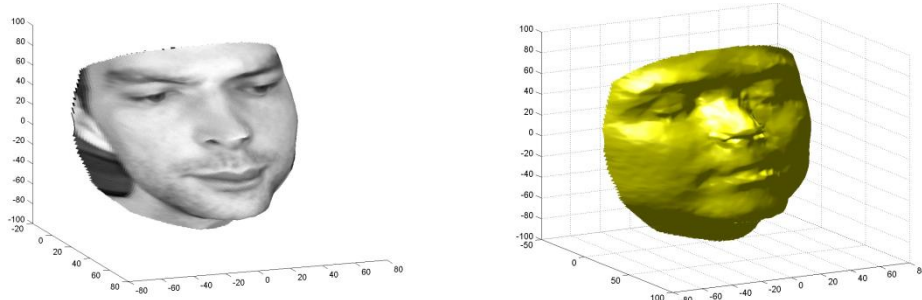


Figure 13: An example of the 3D AAM (the right is the 3D shape, and the left is the texture is mapped on the shape).

As mentioned in the introduction, in this paper, both the 2D features and the 3D features are used in face recognition. The 3D shape has several good features for using in the recognition process. Generally, it can be used from the whole points of the 3D shape as feature for recognition. The simplest approach for recognition is use of the Euclidean distance between the two shapes [2]-[3].

Because the points are landmarks, can use from all of them. In addition, the 2D shape and the texture can be used, too. A suitable combination of these features (weighted combination) can raise the recognition quality.

The performance of the recognition is relative to the rate of goodness of the feature, as is relative to the selection of the suitable features. Namely, for example, in this paper, the 3D shape must be extracted fully and precisely. Also, the place of the 2D landmarks must be determined precisely. It requires that a good algorithm and program is implemented for the 3D frame extracting and the obtaining of the landmarks. On the other hand, the performance of the recognition is a function of the performance of the obtaining algorithm of the landmarks and the 3D frame extracting algorithm.

As we know, the selection of the suitable database is effective in the performance of the face recognition. In this paper, three databases are required. First, 2D image database, second the 3D scanned faces, and third the 3D frame database. The more performance in each of them, the more performance in the face recognition.

## 5. Results

All of the algorithms and programs used in this paper are implemented in Matlab.

### 5-1. results in the section of face recognition

The performance of the recognition for 50 test image and 50 training image is obtained equal to 100%. The 3D landmarks of 50 images and their texture and their 2D landmarks are stored in a database.

Then, using the Mahalanobis distance between the training faces and an unknown face, the most similar man to the unknown face is determined. For the faces that there was an image of it in the database, the most similar man to it is recognized. If for a particular man, there were several images in the database, the most similar image to it is selected. If the man in the test image had no image in the database, then, the most similar man to him was selected (Figure 14).



Figure 14: The result of executing the algorithm on a test image, that there were several image from him in database (top-right : the test image, top-left: the selected image as output of the algorithm, bottom: two other images from him).

## 6. Conclusion and Summary

This paper has paid to representation of methods for modeling, extraction and recognition of 3D faces.

At first, a database of 2D images containing 50 various men image is used and the model of shape variations, texture variations and the model of appearance are built. These models make the detection of the landmarks (range) of the face in the image as possible. On the other hand, the 3D model of shape and texture variations and appearance model are built from 24 3D images that obtained by a Minulta laser scanner. Using these models and using the 2D landmarks, the 3D frame of the face (3D landmarks) can be extracted from an unknown image.

The face recognition is performed using the 3D features (3D landmarks), 2D landmarks and the texture of the face. The performance of the recognition for 50 images that were in the training images is obtained equal to 100%.

Each section of this paper can be used in many applications, for example, the object analysis, the object recognition and, etc. The face recognition has many applications. However, there are challenges in this case. One challenge is that the researches in this field are not used in real applications for discovering the real errors in these systems. Another of these challenges is that the speed and the accuracy of the present computers

don't permit the better use of these systems. The third challenge is that the lack of a strong and varied database of 3D scanned image is causes the performance reduced. For obtaining 3D images, the expensive 3D scanners must be used that are made in a limited number of countries.

Despite these challenges, the relatively strong, accurate and efficient algorithms can be implemented. As mentioned earlier, in this paper, the 3D features are used with the 2D features for recognition. In addition, for speeding up the 3D extracting algorithm, it is used from a method for initializing using the 2D and 3D landmarks. This lead to that the algorithm doesn't fall into local minima for 3D frame searching and lead to raise the quality and the speed of the 3D extracting algorithm.

## REFERENCES

- [1] C. Goodall., Procrustes Methods in the Statistical Analysis of Shape, 53(2):285–339, 1991.
- [2] C. McCool, V. Chandran, S. Sridharan, and C. Fookes., 3d face verification using a free-parts approach. *Pattern Recognition Letters* (In Press), 2008.
- [3] F. Khalid, N. A. Lili, 3D Face Recognition Using Multiple Features for LocalDepth Information, *IJCSNS International Journal of Computer Science and Network Security*, VOL.9 No.1, January 2009.
- [4] H. Hotelling., Analysis of Complex Statistical Variables into Principal Components, *Journal of Educational Psychology*, 24:417–441, 1933.
- [5] I. Kakadiaris, G. Passalis, G. Toderici, N. Murtuza, Y. Lu, N. Karampatziakis, and T. Theoharis., 3d face recognition in the presence of facial expressions: an annotated deformable model approach. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 29(4):640-649, 2007.
- [6] K. Skoglund., Three-Dimensional Face Modeling & Analysis, *Informatics and Mathematical Modelling, Technical University of Denmark*, 2003.
- [7] K. W. Bowyer, K. I. Chang, and P.J. Flynn., A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. *Computer Vision and Image Understanding*, 101(1):1–15, January 2006.
- [8] M. B. Stegmann and D. D. Gomez., A Brief Introduction To Statistical Shape Analysis, *Informatics and Mathematical Modeling, Technical University of Denmark* , Mars 2002.
- [9] M. B. Stegmann, R. Fisker, and B. K. Ersbøll., On Properties of Active Shape Models, *Informatics and Mathematical Modeling, Technical University of Denmark*, 2000.
- [10] M. B. Stegmann, R. Fisker, B. K. Ersbøll, H. H. Thodberg, And L. Hyldstrup., Active Appearance Models: Theory and Cases, *9th Danish Conference on Pattern Recognition and Image Analysis, Aalborg, Denmark*, Volume 1, Pages 49–57. AUC, 2000.
- [11] R. Kimmel A. M. Bronstein, M. M. Bronstein., Three-dimensional face recognition. *Intl. Journal of Computer Vision*, 64(1):5–30, August 2005.
- [12] S. C. Sharma., Applied Multivariate Techniques, *John Wiley & Sons*, 1996.
- [13] S. Gupta, M. K. Markey, & A. C. Bovik., Advances and challenges in 3D and 2D+3D human face recognition. , *Pattern Recognition Research Horizons* , F. Columbus (Editors) , Hauppauge, New York: Nova Science Publishers, 2007.

- [14] T. J. Askari, & H. J. Eghbali., A New Technique in Image Segmentation for Cardiac MR Images, Department of *Computer Science and engineering, University of Shiraz*, 2005.
- [15] X. Lu & A. K. Jain, Deformation Modeling for Robust 3D Face Matching. Proc. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2006)*, pp. 1377-1383, 2006.