

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



## 3D Face Mesh Modeling for 3D Face Recognition

Ansari A-Nasser<sup>1</sup>, Mahoor Mohammad<sup>2</sup> and Abdel-Mottaleb Mohamed<sup>3</sup>

<sup>1</sup>*Ministry of Defence, QENF, Doha, Qatar*

<sup>2</sup>*University of Denver, Denver, Colorado*

<sup>3</sup>*University of Miami, Coral Gables, Florida,*

<sup>1</sup>*Qatar*

<sup>2,3</sup>*U.S.A*

### 1. Introduction

Face recognition has rapidly emerged as an important area of research within many scientific and engineering disciplines. It has attracted research institutes, commercial industries, and numerous government agencies. This fact is evident by the existence of large number of face recognition conferences such as the International Conference on Automatic Face and Gesture and the Biometric Consortium conference. Special issues of well known journals, are being dedicated to face modeling and recognition, such as the journal of Computer Vision and Image Understanding (CVIU), and the systematic empirical evaluations of face recognition techniques including the FERET (Phillips et al., 2000), XM2VTS (Messer et al., 1999), FRVT 2000 (Blackburn et al., 2000), FRVT 2002 (Phillips et al., 2002), and FRVT 2006, which evolved substantially in the last few years. There are few reasons for this trend; first the demands for machine automations, securities, and law enforcements have created a wide range of commercial applications. The second is the availability of feasible technologies developed by researchers in the areas of image processing, pattern recognition, neural network, computer vision, and computer graphics. Another reason for this growing interest is to help us better understand ourselves through the fields of psychology and cognitive science which targeted the perception of faces in the brain. Because our natural face recognition abilities are very powerful, the study of the brain system could offer important guidance in the development of automatic face recognition. Research with animals has shown that these capabilities are not unique to humans. Sheep, for example, are known to have a remarkable memory for faces (Kendrick et al., 2000). In addition, we constantly use our faces while interacting with each others in a conversation. Face gesturing helps us understand what is being said. Facial expression is an important cue in understanding a person's emotional state. In sign languages, faces also convey meanings that are essential part of the language.

A wealth of 2D image-based algorithms has been published in the last few decades (Zhaho et al., 2003). Due to the numerous limitations of 2D approaches, 3D range image-based algorithms are born. Generally, 3D facial range image or data is rich, yet making full use of its high resolution for face recognition is very challenging. It is difficult to extract

Source: State of the Art in Face Recognition, Book edited by: Dr. Mario I. Chacon M.,  
ISBN -3-902613-42-4, pp. 250, January 2009, I-Tech, Vienna, Austria

numerously reliable facial features in 3D. As a result, it becomes more challenging and computationally expensive to accurately match two sets of 3D data (e.g., matching a subject's probe data with the gallery's). Our objective in this chapter is to illustrate a model-based approach that represents the 3D facial data of a given subject by a deformed 3D mesh model useful for face recognition application (Ansari, 2007). The general block diagram of the system is shown in Fig.1, which consists of the modeling stage and the recognition stage. In the modeling stage, only three facial feature points are first extracted from the range image and then used to align the 3D generic face model to the entire range data of a given subject's face. Then each aligned triangle of the mesh model, with three vertices, is treated as a surface plane which is then fitted (deformed) to its corresponding interior 3D range data, using least squares plane fitting. Via triangular vertices subdivisions, a higher resolution model is generated from the coordinates of the aligned and fitted model. Finally the model and its triangular surfaces are fitted once again resulting in a smoother mesh model that resembles and captures the surface characteristic of the face. In the recognition stage, a 3D probe face is similarly modeled and compared to all faces in the database. Experimental application of the final deformed model in 3D face recognition, using a publicly available database, demonstrates promising recognition rates.

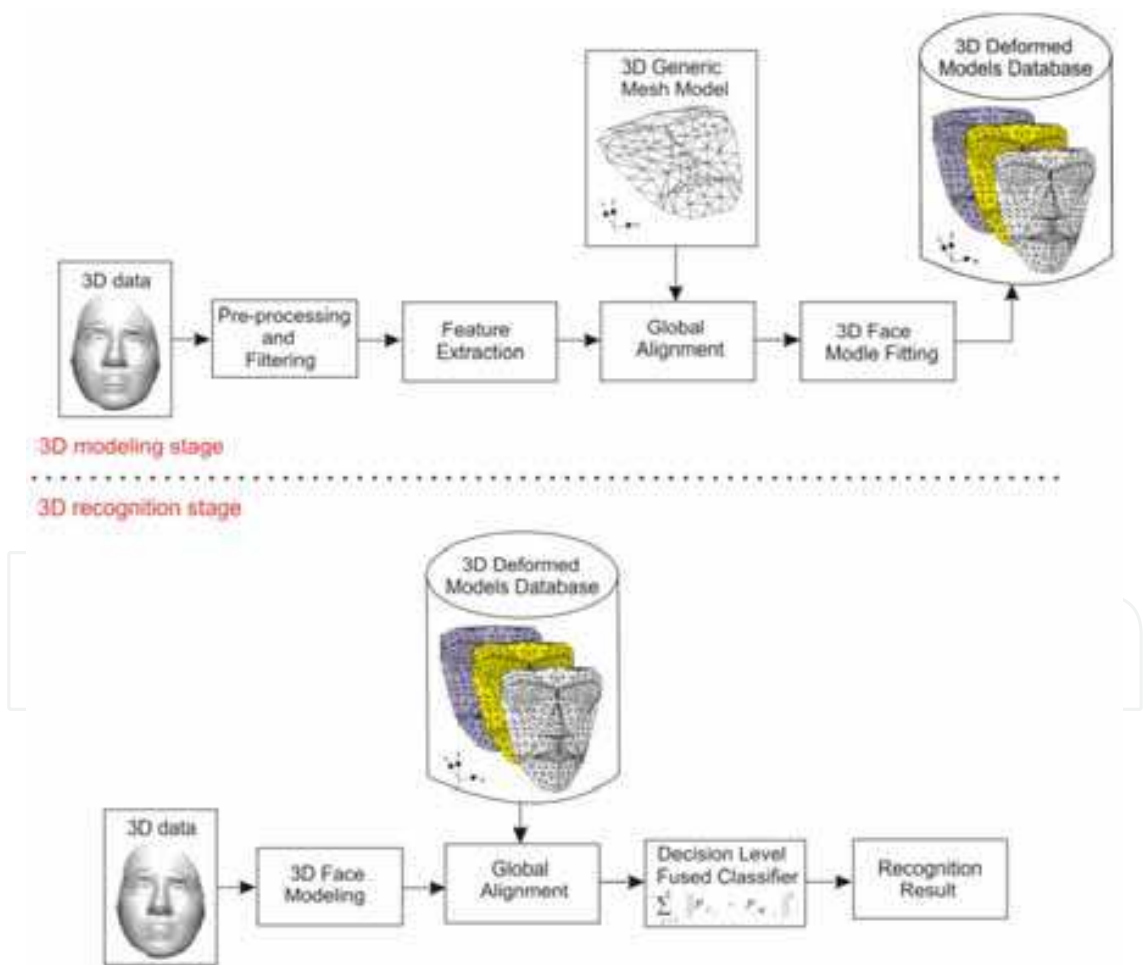


Fig. 1. 3D face modeling and recognition system.

This chapter is organized as follow: Section 2 explains the limitations and challenges of face recognition. Section 3 covers a review of related work. Section 4 describes the data pre-

processing and facial features extraction. Section 5 illustrates the process of 3D face modeling. Section 6 demonstrates experimental results. Finally, conclusion and discussion are given in section 7.

2. Limitations and challenges

Despite the great potentials and the significant advances in face recognition technology, it is still not robust and prone to high error rates, especially in unconstrained environments and in large scale applications. The process of identifying a person from facial appearance has to be performed in the presence of many often conflicting factors which alter the facial appearance and make the task difficult. In Table 1, we categorize the variations in facial appearance into two types: intrinsic and extrinsic sources of variation.

Variation in appearance	sources	Effects / possible task
Extrinsic	Viewing geometry Illumination Imaging process Other objects	Head Pose light variations, shadow, self shadow Resolution, scale, focus, sampling Occlusion, shadowing, indirect illumination, hair, make-up, surgery
Intrinsic	Identity Facial expression Age Sex speech	Identification, known-unknown Inference of emotion or intension Estimating age Decide if male or female Lip reading

Table 1. Extrinsic and intrinsic variations in facial appearance.

The intrinsic variations take place independently of any observer (camera) and are due purely to the physical nature of the face. The identity source is an important intrinsic variation in identifying people from one another, yet problems arise when combined with aging or facial expression because they are difficult to characterize analytically. The extrinsic sources of pose variations, due to the relative position of the camera, and illuminations present a major challenge in face recognition. Recognition systems are highly sensitive to the light conditions and circumstances under which the images being compared are captured. These lighting conditions can be due to the environment or the physical characteristics of the image capturing device, i.e., two cameras of the same brand may give different exposures. The pose of the face is determined by the relative three dimensional position and orientation of the capturing device. Usually, two face images of the same subject taken at different poses are more different than two images of two subjects taken at the same pose. While change in pose is considered a rigid 3D motion, a face can also undergo non-rigid motion when its 3D shape changes due to speech or facial expression. It is very difficult to model both types of motion at the same time. All these factors and conditions make the images used for training the recognition system different from the images obtained for recognition. If these factors are not incorporated and modeled properly, they dramatically degrade the accuracy and performance of a recognition system. Another challenge is the need for an evaluation standard for measuring recognition performance under different environments and conditions. As a result of this necessity, an

independent government evaluation standard was initiated, called Face Recognition Vendor Tests (FRVT) (Blackburn et al., 2000). FRVT was developed to provide evaluations of commercially available and prototype face recognition technologies. These evaluations are designed to provide the U.S. government and law enforcement agencies with information to assist them in determining where and how facial recognition technology can best be deployed. In addition, FRVT results help identify future research directions for the face recognition community. In the past, many factors have been evaluated in FRVT 2002 (Phillips et al., 2003).

A recent challenge to face recognition systems is the concern about possible privacy violations. For example, the American Civil Liberties Union (ACLU) opposes the use of face recognition systems at airports due to false identification and privacy concerns. The ACLU claims that face recognition technology poses the danger of evolving into a widespread tool for spying on citizens as they move about in public places.

### 3. Related work

Semi-automated facial recognition system dates way back to 1965. (Chan & Bledos, 1965) showed that a computer program provided with facial features extracted manually could perform recognition with satisfactory performance. In the past few years, face recognition has received great attentions. A literature survey of face recognition is given in (Zhaho et al., 2003), where most of the paper surveys 2D algorithms. In addition, the work and survey by (Bowyer et al., 2004) compare face recognition techniques based on 2D data, 3D data, and 2D+3D data fusion (also referred to as multimodal). They reported that 3D face recognition approaches outperform 2D approaches and the fusion of 2D + 3D data produces slightly better results than 3D alone. Recently, a survey by (Boyer et al., 2006) cited some algorithms with 2D recognition approaches outperforming the 3D approaches. There is a belief that it is still premature to make this judgment because current approaches do not yet make full use of 3D data either in the recognition algorithms or the rigorous experimental methodology. In this chapter, we only review relevant 3D algorithms processed on range images (3D data) alone.

We can broadly classify 3D face recognition into three categories, namely, 3D surface matching, representative domain, and model-based approaches. A surface matching method, known as Iteratively Closest Point (ICP) approach (Besl & McKay, 1992), is often used as a necessary step in aligning or matching the datasets of two subjects (Lu & Jain, 2006). ICP is based on the search of pairs of nearest points in the two datasets and estimation of the rigid transformation that aligns them. Then the rigid transformation is applied to the points of one set and the procedure is iterated until convergence. Hausdorff distance is another matching approach which is often used in conjunction with ICP. Hausdorff distance attempts to match two datasets based on subset points from the datasets (Huttenlocher et al., 1993); (Russ et al., 2005). The problems with these two matching approaches are expensive computations and sometimes fail to give accurate results. The main reason for using ICP or Hausdorff is not having direct correspondences between the two compared datasets. In the presented algorithm of this chapter, the two compared datasets have direct feature correspondences, which eliminate the need for the above alignment/ matching algorithms. (Medioni & Waupotitsch, 2003) present an authentication system that acquires the 3D image of the subject using stereo images based on internally and externally calibrated cameras. They use the ICP algorithm to calculate similarity between two faces achieving



98% on a database of 100 subjects. (Lu et al., 2004) filter and stitch five multiple views of 2.5D facial scan of each subject to obtain a more complete 3D facial scan. The complete 3D facial scan model is used in the gallery for recognition and the partial 2.5D scans are used as probes. Matching is performed using ICP between a 3D scanned test face with the faces in the database. A 96% recognition rate is obtained with a database of 19 subjects. (Lu & Jain, 2005) extended their previous work using an ICP-based recognition approach by (Russ et al., 2004) to deal explicitly with variations due to the smiling expression. In their experiments, they used a 100-person dataset, with neutral-expression and smiling probes, matched to neutral-expression gallery images. The gallery entries were whole-head 3D models, whereas the probes were 2.5D scan frontal views. They report that most of the errors are after the rigid transformation resulted from smiling probes, and these errors are reduced substantially after the non-rigid deformation stage. For the total of 196 probes (98 neutral and 98 smiling), performance reached 89%. (Uchida et al., 2005) propose two sets of a passive stereo system using four cameras to capture facial images. One set contains two cameras with short baseline intended for accurate correspondence matching. The other two cameras are separated with wide baseline for accurate 3D reconstruction. ICP matching is used between the probe and the gallery faces of a database of 18 subjects each with four simultaneous images. Unfortunately, no recognition rate was reported. (Chang et al., 2005) present an Adaptive Rigid Multi-region Selection (ARMS) approach to independently match multiple facial regions and create a fused result. The ARMS is a classifier type approach in which multiple overlapping sub-regions (e.g., areas around the nose) are independently matched by ICP. Then, the results of the multiple 3D matching are fused. Their experiments on FRGC version 2.0 database resulted in a 91.9 % rank-one recognition rate for automatic Regions of Interest (ROIs) finding and 92.3 % rank-one recognition rate for manual ROIs finding. (Achermann & Bunke, 2000) used two range scanners to capture ranges image in order to overcome the holes and missing data that might result from using one scanner. In addition, they used an extension of 3D Hausdorff distance for 3D face matching. Using 10 images per each of the 24 subjects, they reported 100% recognition rate. (Lee & Shim, 2004) incorporate depth information with local facial features in 3D recognition using Hausdorff distance weighted by a function based on depth values. The weights have different values at important facial features such as the nose, eyes, mouth, and face contour. They achieved rank five recognition rate of 98%. (Russ et al., 2004) use Hausdorff distance matching for range images. In a verification experiment for 200 subjects enrolled in the gallery and the same 200 persons plus an additional 68 in the probe set, they report a verification rate of 98%. In a recognition experiment, 30 persons enrolled in the gallery and the same 30 persons imaged at a later time were used in the probe set. A 50% recognition rate is achieved at a false alarm rate of 0.

Other researchers attempted to represent the 3D data in a different domain and made recognition comparison in the representative domain. Examples of those are 3D Principle Component Analysis (PCA) (Hesher et al., 2003), shape index (Lu et al., 2006), point signature (Chua et al., 2000), spine image (Johnson & Hebert, 1999), and local shape map (Wu et al., 2004). PCA is a statistical approach commonly used in recognition. One reason for using PCA is to reduce the dimensionality of the data, while sacrificing the performance of the recognition algorithm. (Hesher et al., 2003) explore PCA techniques using different number of eigenvectors, image sizes, and different expressions. They report a high recognition rate, but their system degrades if the expression of a test face is different from

the expressions in the database. (Xu et al., 2004a) slightly improved the recognition rate by computing a feature vector from the data in the local regions of the mouth, nose, left eye, and right eye. The dimensionality of the feature vector is reduced with PCA and matching is based on minimum Euclidean distance. Experiments on 120 subjects in the dataset resulted in 72% recognition rate, and on a subset of 30 subjects resulted in a 96% recognition rate. It should be remarked that the reported performance was obtained with five images of a person used for enrollment in the gallery. Performance is generally expected to be higher with more images used to enroll a person. (Pan et al., 2005) apply PCA to the range images using a novel mapping technique. Finding the nose tip to use as a center point, and an axis of symmetry to use for alignment, the face data are mapped to a circular range image. Experimental results are reported for the FRGC version 1.0 data set with 95% rank-one recognition rate and 2.8% Equal Error Rate (EER). Another example of a representative domain approach is the use of transform or wavelet. (Cook et al., 2006) present an approach based on Log-Gabor template for providing insensitivity to expression variation in range images. They decompose the facial image into overlapping 147 sub-jets (49 sub-regions and three scales) using Log-Gabor wavelets. For face verification, they use the Mahalanobis cosine distance measure and un-weighted summation to combine the result of classifying each region. Their experiments resulted in a 92.3 % rank-one recognition rate.

Model-based approaches use a priori facial model such as graph or mesh model. Graph representation has shown to be successful (Wiskott et al., 1997); (Blome, 2003). The idea is to use a graph to model the face with nodes and edges. The edges are labeled with distance information and nodes are labeled with local wavelet responses. However, the graph models in the literature have some limitations. For example, there is no justification for defining the edges of the graph. (Mahoor et al., 2008) improved a graph model which they refer to as Attributed Relational Graphs (ARG). The ARG is a geometric graph also with nodes and edges, where the nodes represent the facial landmarks and the edges connects the nodes based on Delaunay triangulation. A set of mutual relations between the sides of the triangles are defined in the model and are used in the recognition process in addition to the nodes and edges.

Mesh model approaches use a priori defined facial mesh which is usually morphed or deformed to a given face. A detailed example of this approach is illustrated in this chapter, which has the advantages of eliminating some of the previously stated problems of both the surface matching and the representative domain algorithms. Firstly, by representing the huge facial range data by a mesh model with smaller number of vertices, we reduce the amount of data points for facial processing, data storage, and recognition comparisons. Secondly, having a predefined and labeled-vertices in the deformed mesh model, establishes direct features correspondences between compared probe's and gallery's facial data. Hence faster recognition comparisons are achieved. Both the labeling of the model's vertices and the data reduction, resulting from representing the face by the vertices of the model, are vital in reducing the complexity of the face recognition system. The presented method in this chapter is similar to work of (Xu et al., 2004b) but differs in the followings: (a) The method in this chapter uses a generic face mesh model and (Xu et al., 2004b) use a general mesh grid model, (b) here, the aligned model's mesh triangles coordinate are deformed to the data and (Xu et al., 2004b) simply align the grid mesh coordinates to the range data then copy the  $z$  coordinate at each  $x$  and  $y$  coordinates, hence in their way the pose of the  $z$  coordinate is not considered, (c) the presented system establishes direct correspondences

with other models in the database, hence direct comparison is achieved in recognition, while the method of (Xu et al., 2004b) has no correspondences and would require facial surface alignment and matching.

(Vetter & Blanz, 1999) proposed a face recognition algorithm based on computer graphics techniques, where they synthesize the 3D model of a face from a single 2D image of known orientation and illumination. However, their algorithm is computationally expensive and initially requires manual user assistance and a database of 200 different real scans of faces obtained from a 3D scanner. Correspondences across these 3D scans are pre-computed. The input face image is estimated as a linear combination of the projected 3D scans in the database; subsequently, the output 3D model is a linear combination of the 3D scans. Similar approach is proposed by (Jiang et al., 2004) which they referred to as analysis-by-synthesis 2D to 3D face reconstruction, in which they use a single frontal 2D image of the face with a database of 100 3D faces captured by 3D scanner. In this approach frontal face detection and alignment are utilized to locate a frontal face and the facial feature points within an image, such as the contour points of the face, left and right eyes, mouth, and nose. Then, the 3D face shape is reconstructed according to the feature points and a 3D face database. Next, the face model is textured-mapped by projecting the input 2D onto the 3D face shape. Finally, based on the resulting 3D model, virtual samples of 3D models are synthesized with pose and expression variations and are projected to 2D for recognition. (Hsu & Jain, 2001) adapts a generic face model to the facial features extracted from both registered range and color images. The deformation iteratively moves the vertices of the mesh model using vertices displacement propagation. (Ansari & Abdel-Mottaleb., 2005) deformed a generic model to few 3D facial features obtained from one frontal and one profile view calibrated stereo images. The additional profile view complements and provides additional information not available in the frontal view. For 29 subjects, a recognition rate of 96.2 % is reported. In (Ansari et al., 2006) an improved modeling and recognition accuracy is presenting using dense range data obtained from two frontal and one profile view stereo images for 50 subjects attaining 98% recognition rate.

#### 4. Data pre-processing and facial features extraction

This section explains the pre-processing of the data, localization of the facial region, and the facial features extraction. Further details are given in (Mahoor et al., 2007). Range images, captured by laser scanners, have some artifacts, noise, and gaps. In the pre-processing step, we first apply median filtering to remove sharp spikes and noise, that occur during the scanning of the face, followed by interpolation to fill up the gaps, and low pass filtering to smooth the final surface. This is followed by face localization using facial template matching to discard the neck, hair, and the background areas of the range image. The facial range image template is correlated with the range images of a given face using normalized cross-correlation. We start by roughly detecting the location of the nose tip and then translate the template such that the detected tip of the nose is placed on the location of the nose tip of the range image under test. Afterward, we iteratively apply a rigid transformation to the template and cross-correlate the result with the subject's range image to find the best pose. Finally, the area underneath the template with the maximum correlation is considered as the localized facial region. Subsequently, we use Gaussian curvature to extract the two inner corners of the eyes and the tip of the nose. The surface that either has a peak or a pit shape has a positive Gaussian curvature value  $K > 0$  (Dorai & Jain, 1997). Each of the two inner



corners of the eyes has a pit surface type and the tip of the nose has a peak surface type that is detectable based on the Gaussian curvature. These points have the highest positive Gaussian curvature values among the points on the face surface. Fig.2.a shows the result of calculating the Gaussian curvature for one of the sample range images in the gallery.

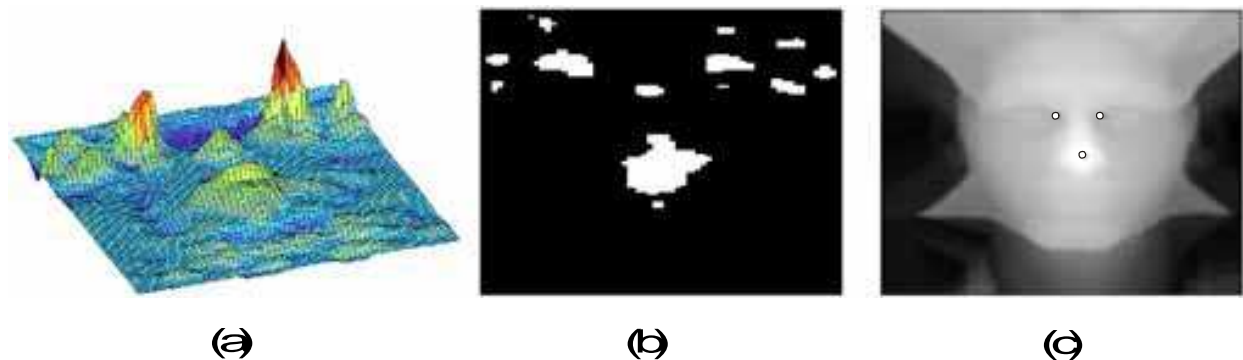


Fig. 2. Features extraction process (a) Gaussian curvature showing high values at the nose tip and eyes corners (b) Result of thresholding Fig.2.a (c) Final result of feature extraction.

The highest points in Fig.2.a correspond to the points with pit/ peak shape. We threshold the Gaussian curvature to find the areas that have positive values greater than a threshold, producing a binary image. See Fig.2.b. The threshold is calculated based on a small training data set different from the images used in the recognition experiments. Finally, the three regions with the largest average value of the Gaussian curvature are the candidate regions that include the feature points. The locations of the points with maximum Gaussian curvature in these regions are labeled as feature points. Fig.2.c shows a final result of the three feature extraction points. These features are used in the 3D model alignment as we show next.

## 5. 3D face modeling

This section deals with modeling the human face using its extracted features and a generic 3D mesh model. The idea is to align the 3D model to a given face using the extracted 3D features then proceed with fitting the aligned triangles of the mesh to the range data, using least square plane fitting. Next, the aligned triangles of the model are subdivided to higher resolution triangles, before applying a second round of plane fitting, to obtain a more realistic and a smoother fitted surface resembling the actual surface of the face. Fig.3.a shows our neutral 3D model with a total of 109 labeled feature vertices and 188 defined polygonal meshes. In addition, the model is designed such that the left and right sides of the jaw fall within but not on the edges of the face boundary. This approach avoids incorporating inaccurate data at the facial edges of the captured range images. We explain next the process of aligning the mesh model to the range data.

### 5.1 Global alignment

In the global alignment step, we rigidly align the 3D model using the three 3D feature points,  $P_I$ , obtained from the range image, and their corresponding feature vertices,  $P_M$ , in the model. Subscripts  $I$  and  $M$  indicate image features and model vertices, respectively. To achieve this goal, the model must be rotated, translated, and scaled. Eq.1 gives the sum squared error between  $P_I$  and  $P_M$  in terms of scale  $S$ , rotation  $R$ , and translation  $T$  for  $n = 3$  points.

$$\text{Min } E(S, R, T) = \sum_{j=1}^n \|P_{I_j} - P_{M_j}\|^2 \quad (1)$$

An example of the aligned 3D model to the range data is demonstrated in Fig.3.b and Fig.3.c for 2D view and 3D view, respectively. As shown in the figures, the triangles of the model are buried either totally or partially above or below the 3D data. We show next how to segment the 3D data points within the aligned 3D model.

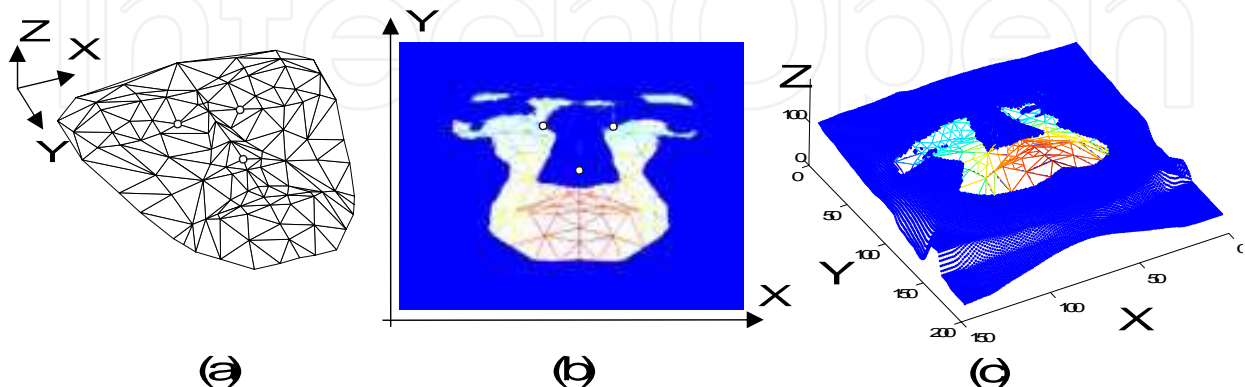


Fig. 3. (a) 3D mesh model (b) 2D view of aligned 3D model vertices to the range data (c) 3D view of model to the range data.

### 5.2 3D facial points segmentation

The first step prior to deforming the model is to segment and extract the 3D data points facing (above, below, or within) each mesh triangle using a computer graphic technique referred to as *Barycentric Coordinate* (Coxeter, 1969). A barycentric combination of three point vertices  $P_1$ ,  $P_2$ , and  $P_3$ , forming a triangular plane is shown in Fig.4

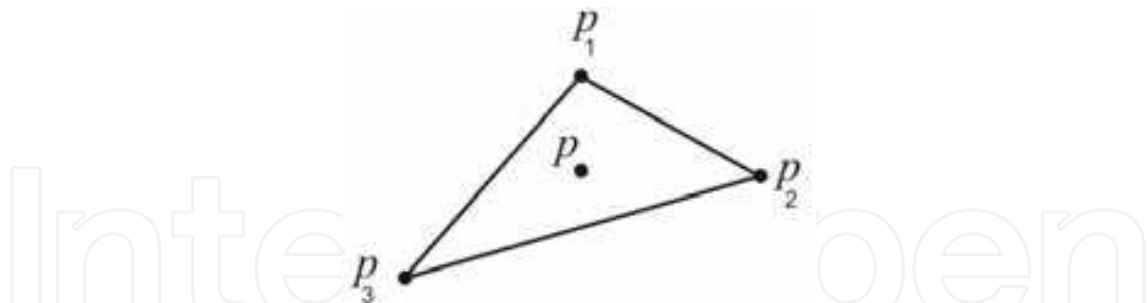


Fig. 4. The barycentric coordinates of a point  $p$  with respect to the triangle vertices.

The coordinate of point  $p$  inside the triangle is defined by

$$p = up_1 + vp_2 + wp_3 \quad \text{where} \quad u + v + w = 1 \quad (2)$$

Therefore,  $p$  lies inside the triangle and we say  $[u, v, w]$  are the barycentric coordinates of  $p$  with respect to  $p_1$ ,  $p_2$ , and  $p_3$  respectively. Equivalently, we may write

$$p = up_1 + vp_2 + (1 - u - v)p_3 \quad (3)$$

Eq.3 represents three equations and thus we can form a linear system given by

$$\begin{bmatrix} p_1 & p_2 & p_3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = p \quad (4)$$

which can be solved for the unknowns in  $[u, v, w]$ . Points inside a triangle have positive  $u, v$ , and  $w$ . On the other hand, points outside a triangle have at least one negative coordinate. Eq.4 is computationally expensive because each of the 188 triangles of the mesh model has to check all the range data coordinates to determine whether or not the coordinate points, if any, fall within its interior. A practical implementation is to window the data enclosed by the triangle coordinates as shown in Fig.5. Only the point coordinates within the rectangle are applied in Eq.4.

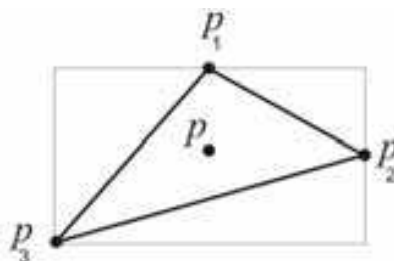


Fig. 5. Windowed 3D range data points.

Figure 6 shows a 2D view of an actual 3D mesh model, superimposed on the range data points. The figure shows an example of segmented 3D data points within one triangle of the eyebrow meshes. We show next how to fit and deform the model's triangles to be as closely as possible to the 3D data.

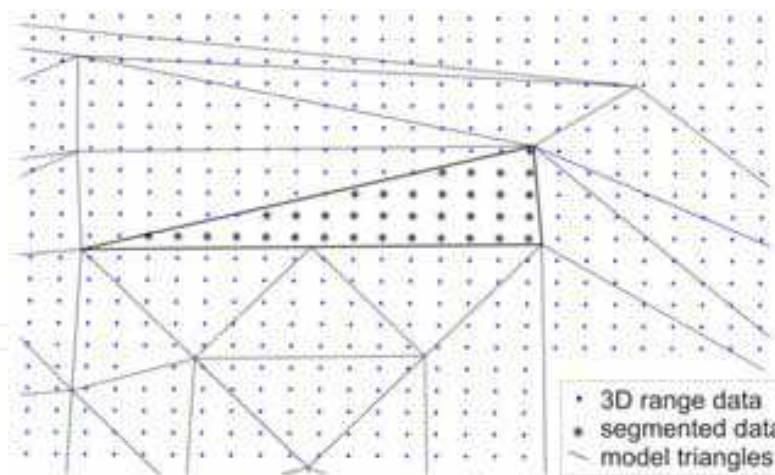


Fig. 6. Segmentation example of the 3D data points within one triangle using barycentric coordinates.

### 5.3 3D face model deformations

Once the cloud of the 3D data points is segmented by the barycentric coordinate, they are represented by a plane using least square fitting. The general equation of a plane, with non-zero normal vector  $N$ , is defined in 3D as

$$aX + bY + cZ + d = 0, \text{ where } N = (a, b, c) \quad (5)$$

For  $n$  number of points, Eq.5 can be written in least square form as

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 \\ X_2 & Y_2 & Z_2 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ X_n & Y_n & Z_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = AB = 0 \quad (6)$$

where the coordinates  $(X_i, Y_i, Z_i)$ 's are those of all the data points segmented by the barycentric coordinate. Eq.6 can be solved for the plane equation parameters,  $B = [a, b, c, d]$ , which is then substituted in Eq.5, leading to a plane representing the 3D data points. Fig.7.a illustrates a concept example of a triangle with 3D data points in 3D space. Fig.7.b shows the segmented data within the triangle which are represented by a plane using Eq.5. From the mathematical geometry of a plane, having the parameters of  $B$ , any point on the plane can be evaluated. In this work, we deform each corresponding mesh triangle to the 3D data points, by first discarding the three vertices  $Z$  coordinates, evaluating the  $X$  and  $Y$  coordinates, and solving for the new  $Z$  coordinate (given the parameters in  $B$  from Eq.6). This produces a mesh triangle, with new depth coordinates, lying on the plane that is approximated by the dense 3D data points. Fig.7.c shows the concept of deforming the mesh triangle to the plane representing the data. Essentially, the pose of the triangle is changed to match that of the plane.

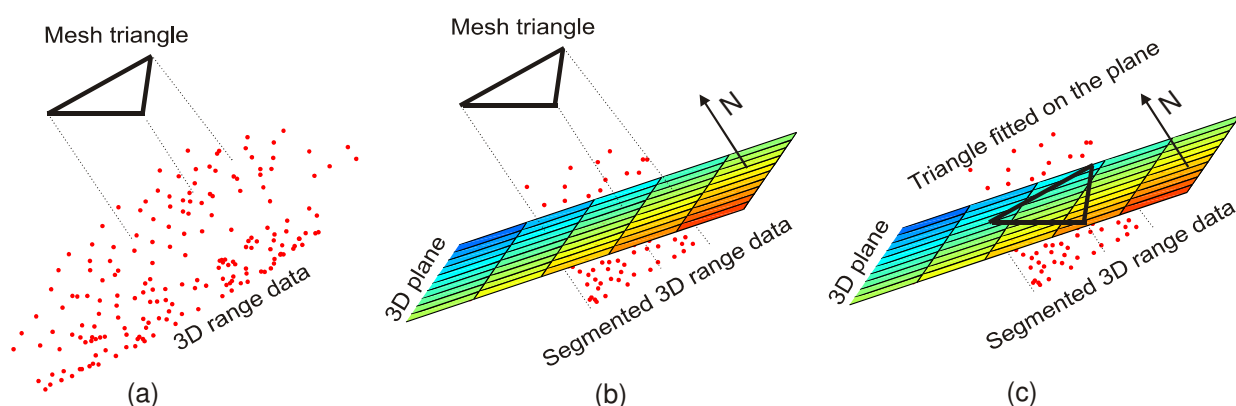


Fig. 7. The process of deforming the triangles of the 3D mesh model, (a) Given cloud of 3D data and a mesh triangle (b) Segmenting the 3D data and plane fitting (c) Deforming the mesh triangle to the plane representing the 3D data.

Subsequently, we repeat the deformation process to all the triangles of the mesh model. Fig.8. shows an example of a complete deformed model superimposed on the data in 2D and 3D views. Comparing Fig.8.a-b with the initially aligned model of Fig.3.c-d, we see that the deformation and fitting of the model to the range data are clearly observed. The triangles of the mesh model have come closer to the data.

The deformed model of Fig.8 is a good representation of the data, yet it's not smooth enough to represent the high resolution and curvatures of the 3D data. In the next step, we subdivide the triangles of the model to a higher resolution in a manner shown in Fig.9.a. New vertices are computed based on the locations of the deformed vertices. Fig. 9.b shows the result of subdividing the deformed model of Fig.8. This process increases the number of vertices and triangles (meshes) of the original model from 109 and 188, respectively, to 401 vertices and 752 polygonal meshes. Finally, because the new triangles do not reflect actual

deformation to the data, we deform them once again using the same deformation process explained above.

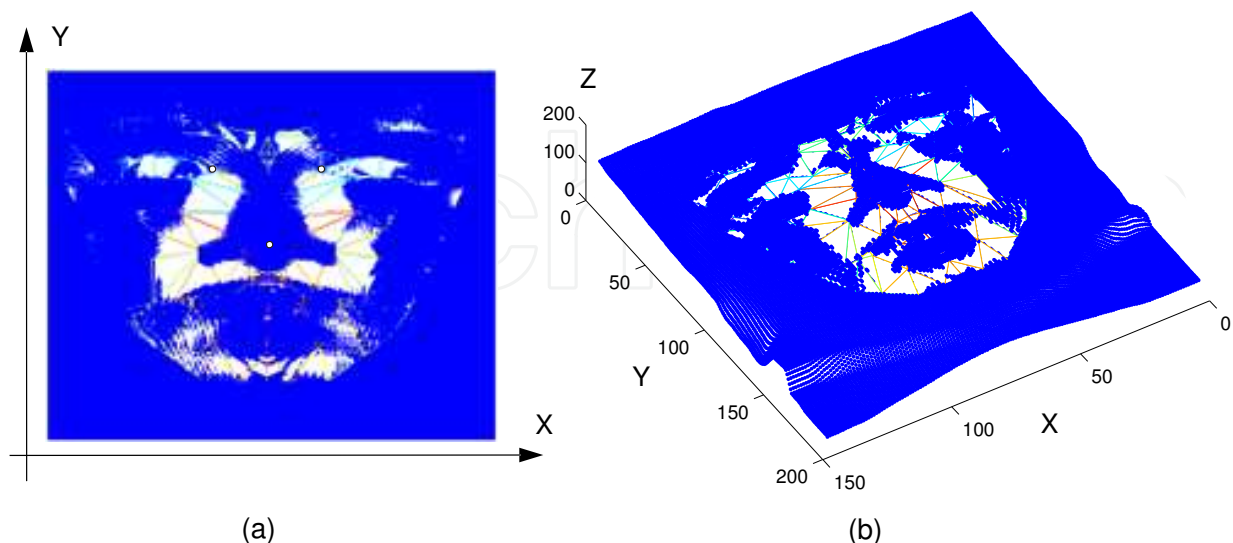


Fig. 8. Deformed model superimposed on the range data. (a) 2D view (b) 3D view.

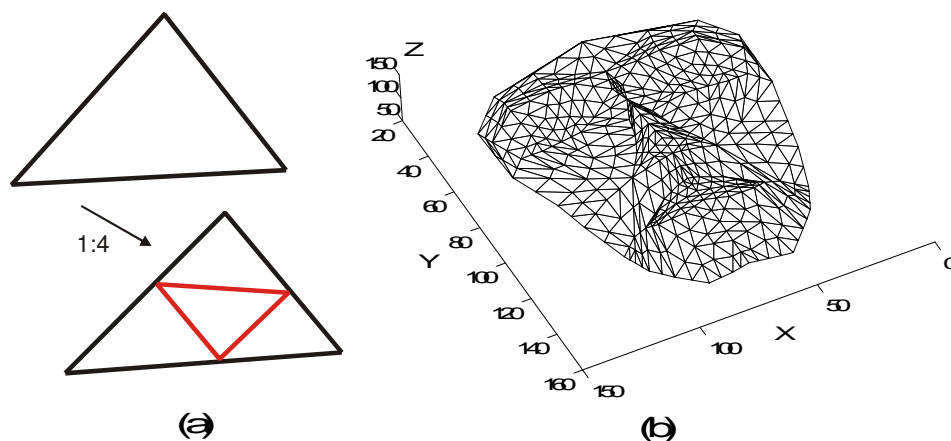


Fig. 9. (a) 1 to 4 triangle subdivision (b) Result after applying triangle subdivisions to the deformed model of Fig.5.

The introduction of smaller triangles gives more effective triangle fitting of the data especially at areas of high curvatures. Fig. 10.a-b-c show the final result of the deformed model, superimposed on the data in 2D view, 3D view, and a profile 2D view, respectively. In Fig.10.a-b-c, because most of the models' vertices are embedded within the data, we use the “\*” symbol to clearly show their locations. Fig.10.d shows a profile (YZ-axis) view of the model in Fig.10.c without the data. This deformed model, containing 401 vertices points, is the final representation of the facial data, which originally contained about 19,000 points (based on an average range image size of 150 by 130). This is nearly a 98 % data reduction.

We summarize below the 3D mesh model deformation algorithm:

- Given an aligned 3D mesh model to the facial range data, extract the 3D points within each triangle of the mesh model using the barycentric coordinate approach.
- For each triangle, fit a plane to the extracted 3D data points and solve for the  $B$  parameters in Eq.6.



- c. For each of the three vertices of the mesh triangle, solve for the unknown  $Z$  coordinate by evaluating the coordinates of  $X$ ,  $Y$ , and  $B$  parameters in Eq.5. This fits the triangle on the plane.
- d. Repeat steps (b) to (c) for all the mesh triangles of the model.
- e. Subdivide the resulting model and repeat steps (a) to (d).
- f. Further subdivision is possible depending on the resolution, quality, or accuracy of the captured range data points.

We show next the application of the deformed model in 3D face recognition.

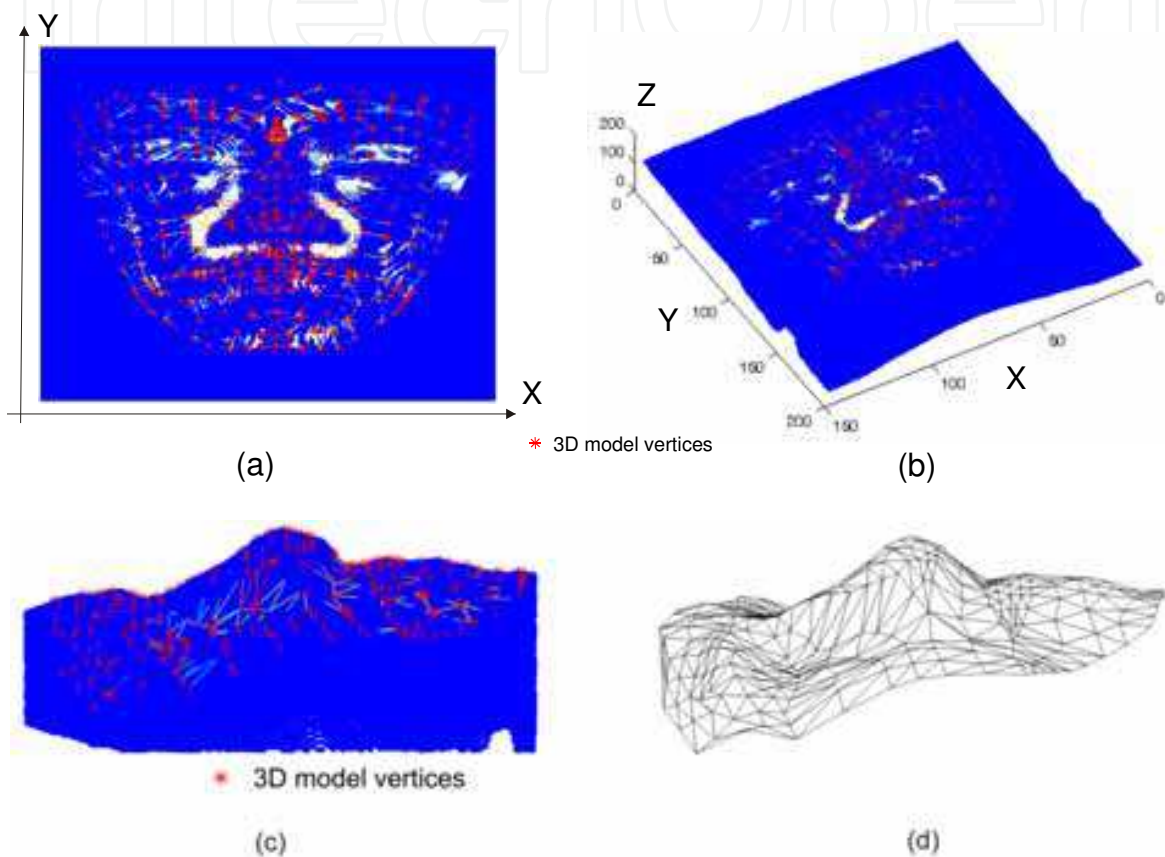


Fig. 10. Final deformed model. (a) 2D view (b) 3D view (c) profile view superimposed on the data (d) mesh model without showing the data.

## 6. 3D face recognition

Face recognition has received great attentions in the past few years. A recent literature survey for 3D face recognition is given by (Bowyer et al., 2006). The final result of Fig.10 gives a model with 401 deformed vertices specific to a given subject's 3D range data. In this section we explore for different subjects the use of the deformed final models in 3D face recognition. The recognition score is based on a decision level classifier applied to the deformed models obtained from ranges images of a public database.

### 6.1 Range image database

The range images we use in this chapter are obtained from the publicly available GAVAB database captured by a 3D scanner (Moreno & Sanchez, 2004). This database contains seven

facial range images of 61 subjects: two frontal images with normal expression, two images looking up and down, and three images with facial expression. Many subjects contain instances of dark regions in the face which do not reflect successful 3D scanning, producing in these cases incomplete facial surfaces. As a result, range image pre-processing and filtering are necessary preliminary steps. In this chapter, we are only concerned with modeling and recognizing the frontal images of the database under neutral expressions. Figure 11 contains an example of two views of the texture and range images of one subject. The texture images are not publicly available. For both sets of the frontal range images we obtain the 3D face models as outlined in previous section. One model is used as a query (probe) and the other model is used in the gallery (database). We explain next the recognition technique.

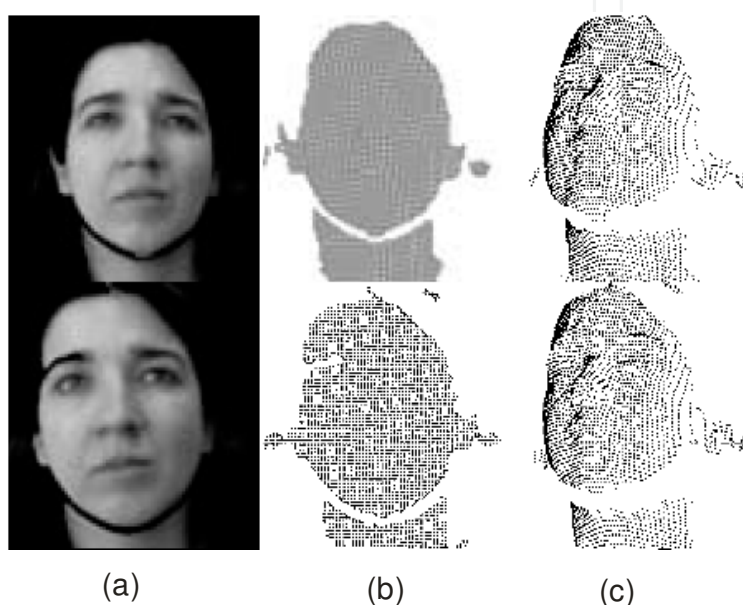


Fig. 11. Subject example of GavabDB (a) Textured 3D image (b) Same image without texture (c) Same image rotated in 3D.

## 6.2 Decision level fused classifiers

In the recognition stage of Fig.1, a query face model is aligned with all faces in the database and then classified for recognition based on Euclidian distance and voting classifiers. We compute the identification rate using the fusion of both Euclidean distance-based and voting-based classifiers at the decision level of the recognition system. Fig.12 shows a block diagram of the decision level classifier.

The Euclidean distance classifier, even though widely used, its performance can be greatly degraded in the presence of noise. The degradation is due to the equal summation of squared distances over all the features. Any noisy feature with a large distance can mask all other features and as a result the classification considers only the noisy feature, neglecting the information provided by the other features. To overcome this drawback, we use a voting classifier to decide on the final score of the recognition system. The voting classifier counts the maximum number of minimum distances of the features between corresponding features points. In this case the feature points are the 401 deformed vertices of the mesh model. In the voting classifier a face is recognized when it has the maximum number of feature (votes) when compared with the corresponding features of the other subjects in the

database. In the presented algorithm, when a query face model is given to the recognition system of Fig.12, it runs through both classifiers; a direct decision for a recognized face is made only when both classifiers' outputs agree on the same recognized face in the database ( $E=V_1$  in Fig.12). If the two classifiers are in disagreement, then a different procedure is taken before a final decision is made. In this case, the probe face is directly compared with the recognized face by the Euclidean and the voting classifier, using the voting approach. As a result, the second voting classifier is comparing only two faces. This approach reduces wrong decisions that might be taken by the Euclidean distance classifier, because of possible masking of noisy feature(s), and reroutes the final decision to another voting classifier for final recognition decision. In a scenario when both classifiers actually have the wrong decision, then there is no other clue and a wrong face is falsely recognized.

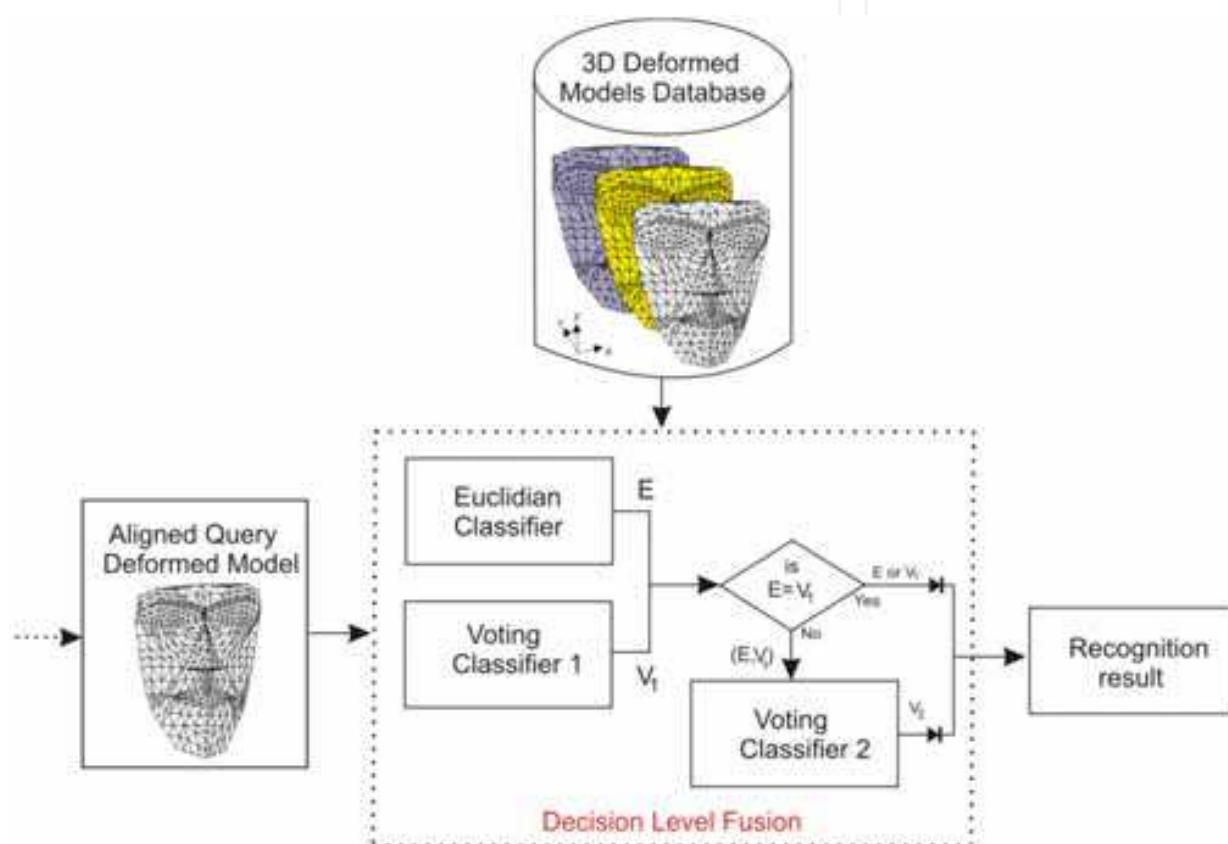


Fig. 12. Decision level fused 3D face recognition system.

### 6.3 Face recognition experiments

Following the procedure illustrated by the recognition stage in Fig.1 and the classifier of Fig.12, we test the recognition algorithm separately using the Euclidian distance-based classifier, the voting-based classifier, and the fused classifier of the recognition system. Fig.13 shows the overall Cumulative Match Curve (CMC) identification rate for the 61 subjects of the GAVAB database. From the performance figure, the fused rank one identification rate achieves a 90.2% compared to a lower single classifier rate of 85.2% or 65.6% by the Euclidean or the voting classifier, respectively. The fusion obviously gives superior performance at all ranks. It has been reported that the same database was used in (Moreno et al., 2003) achieving 78 % rank one identification rate for 60 out of 61 subjects using 68 curvature-based extracted features.

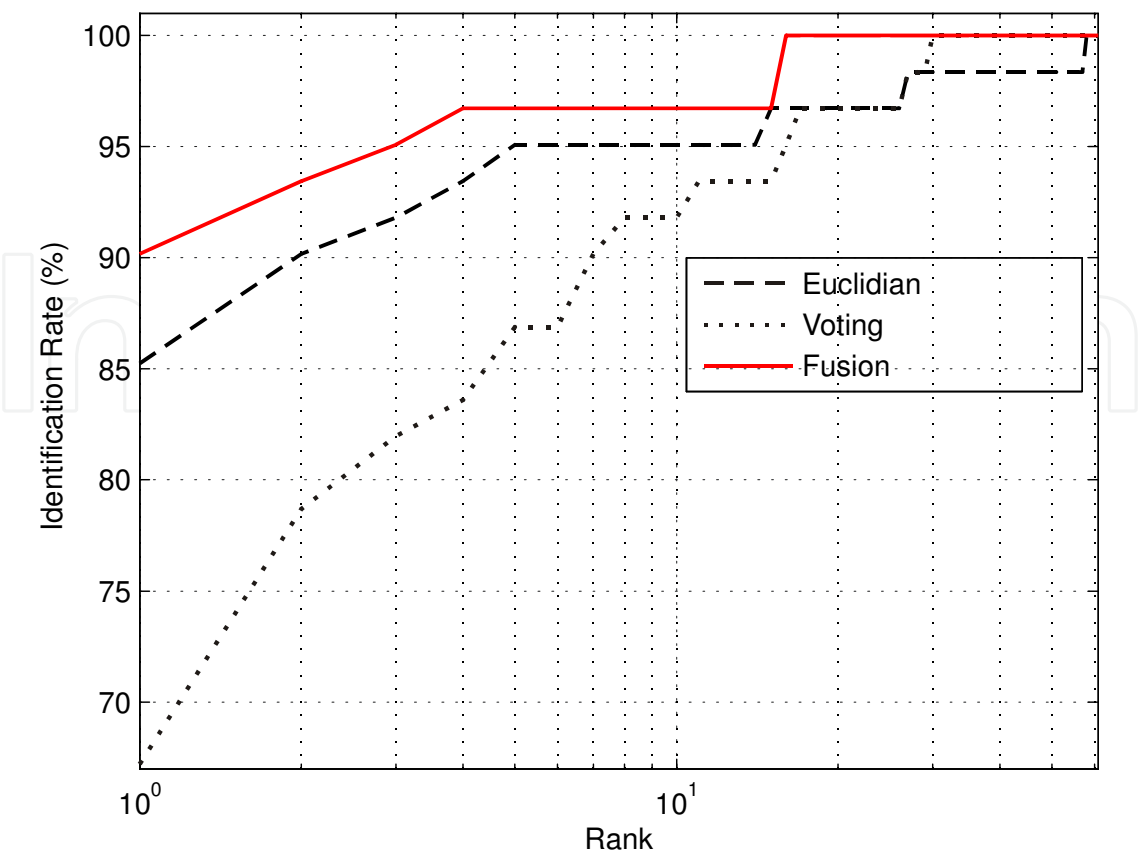


Fig. 13. CMC curves using single and fused classifiers for 401 model’s vertices.

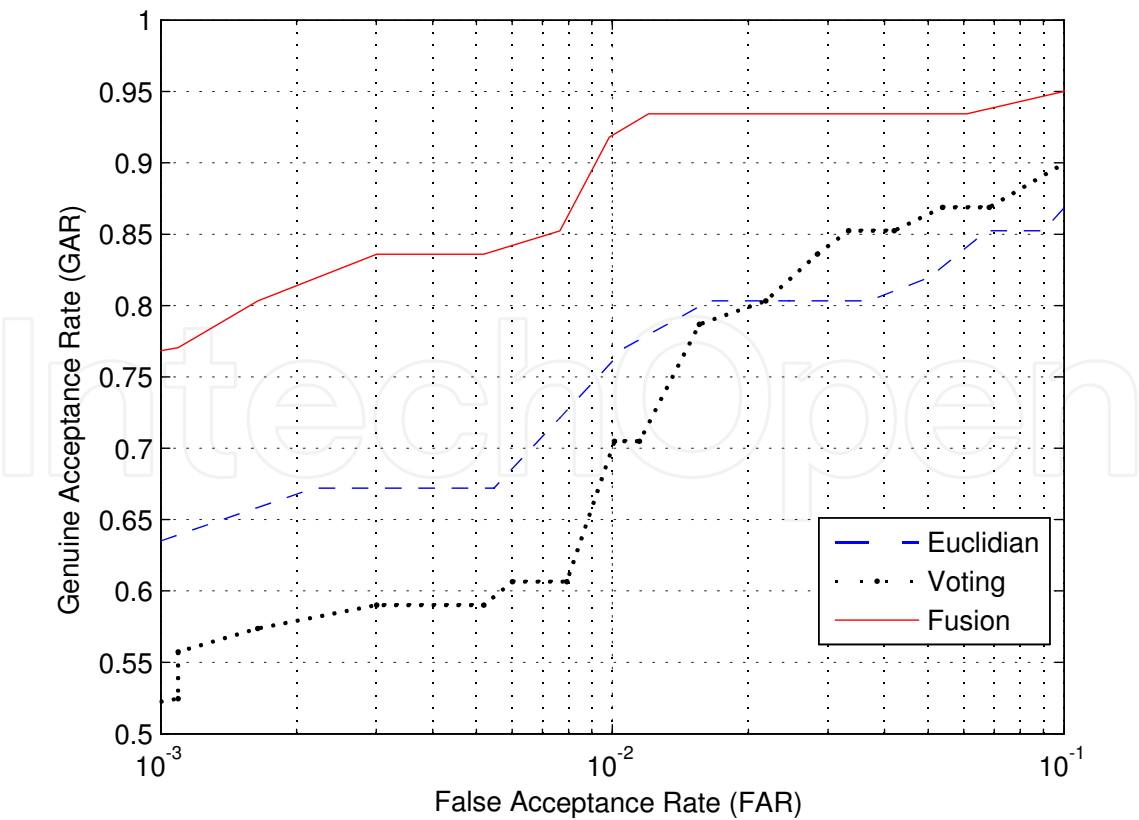


Fig. 14. ROC curves using single and fused classifiers for 401 model’s vertices.

Similarly, testing the system in the verification mode, Fig.14 shows the Receiver Operating Characteristic ROC performance curves of the recognition system. At false acceptance rates of 0.1% and 1%, the fused result of the recognition system achieves genuine acceptance rates of 76% and 92%, respectively.

7. Conclusion and discussion

A model-based algorithm for 3D face recognition from range images is presented. The algorithm relies on deforming the triangular meshes of the model to the range data establishing direct model vertices correspondences with other deformed models in the database. These features correspondences greatly facilitate faster computational time, accuracy, and recognition comparisons. By only detecting three facial features and a generic model, we achieved a 90.2% rank one identification rate using a noisy database. The presented method is proved to be useful for face recognition. However, the method can also be sensitive to noisy or missing data under the mesh model. In the conducted experiments, six subjects out of the 61 were not correctly recognized. The wrong recognition was mainly due to the dataset being either very noisy, incomplete, or the query range image set looks very different from the database set. Unfortunately, the range data pre-processing and

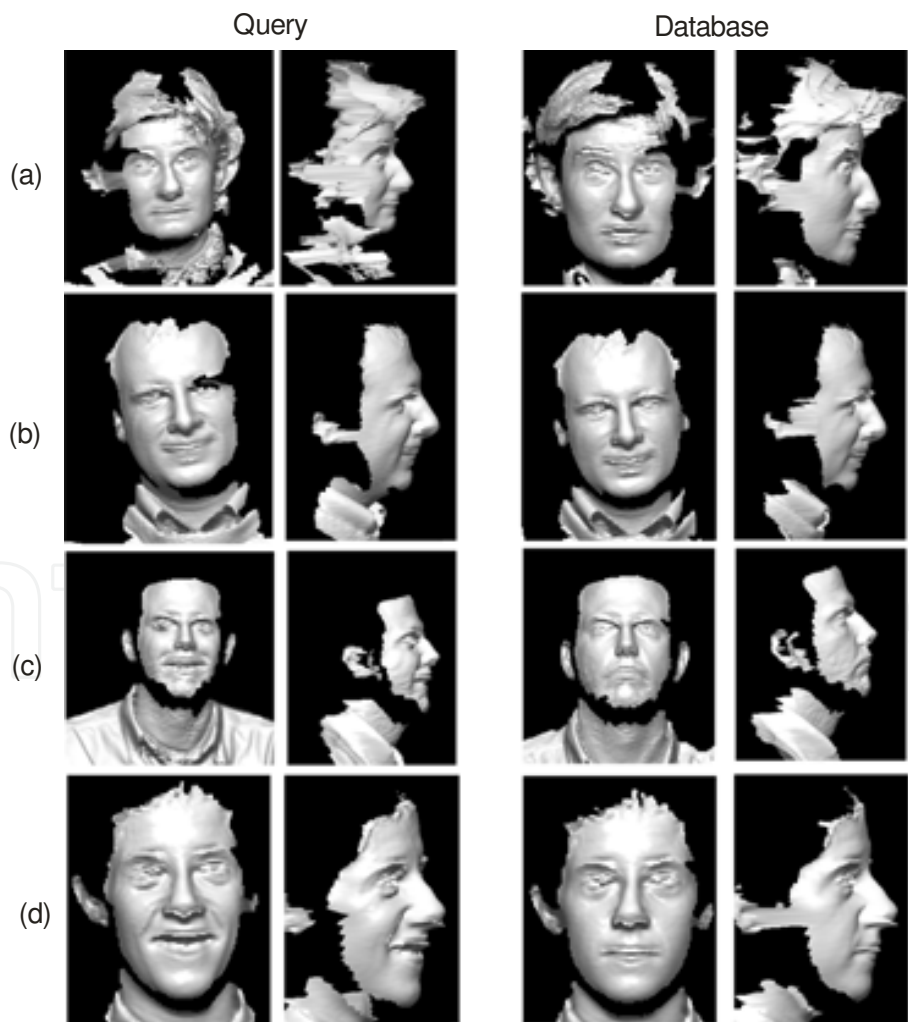


Fig. 15. Probe and database range images of four out of six subjects misrecognized.



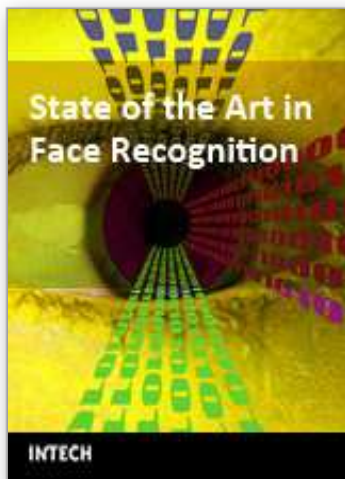
filtering, presented in section 4, cannot always cope with large areas of holes or spikes. Fig.15 shows four of the six subjects that were not correctly recognized. Both the query and database sets of Fig.15.a show noisy and incomplete facial scan at the left and right side of the face. Fig.15.b shows similar incomplete data at the eye location. Fig.15.c and Fig.15.d show not only noisy data but also facial expression between the compared query and database images. These factors make the query set of images very different from the database set. In order to demonstrate the robustness of the presented algorithm, better data or another database must be attempted on a large scale datasets captured by high quality 3D scanners. The noise introduced around the subjects' eyes of all subjects in Fig.15 is typical of a lower quality and an older type of 3D scanners. At the time of publishing this chapter, the authors are in process of obtaining a license for the Face Recognition Grand Challenge (FRGC) database (Phillips et al., 2005) in order to apply the algorithm to a much better and cleaner database.

## 8. References

- Achermann, B. & Bunke, H. (2000). Classifying range images of human faces with Hausdorff distance. *The 15-th International Conference on Pattern Recognition*, pp. 809–813, September 2000.
- Ansari, A. & Abdel-Mottaleb, M. (2005). Automatic facial feature extraction and 3D face modeling using two orthogonal views with application to 3D face recognition. *Pattern Recognition*, Vol. 38, no. 12, pp. 2549–2563, December 2005.
- Ansari, A. (2007). 3D face mesh modeling with applications to 3D and 2D face recognition. Ph.D. dissertation, *University of Miami*, Coral Gables, Florida, USA.
- Ansari, A.; Abdel-Mottaleb M., & Mahoor M. (2006). Disparity-based 3D face modeling using 3D deformable facial mask for 3D face recognition. *Proceedings of IEEE ICME*, Toronto, Ontario, Canada, July 2006.
- Besl, P.; & McKay, N. (1992). A method for registration of 3-D shapes, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.14, pp.239–256, 1992.
- Blackburn, D. ; Bone, J & Phillips, P. (2000). Face recognition vendor test 2000. Technical report, [http:// www.frvt.org/ FRVT2000/ documents.htm](http://www.frvt.org/FRVT2000/documents.htm).
- Blanz, V. & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *Proceedings of SIGGRAPH*, pp. 187–194. 1999.
- Bolme, D. (2003). Elastic bunch graph matching, Master thesis, *Colorado State University*, Summer 2003.
- Bowyer, K.; Chang, K. & Flynn, P. (2004). A Survey of Approaches to Three-Dimensional Face Recognition. *Proceedings of the International Conference on Pattern Recognition*, Cambridge, England, August 2004.
- Bowyer, K.; Chang, K. & Flynn, P. (2006). A Survey of Approaches and challenges in 3D and multi-modal 3D + 2D face recognition. *Computer Vision and Image Understanding*, vol. 101, pp 1–15, 2006.
- Chan, H. & Bledos, W. (1965). A man-machine facial recognition system: some preliminary results. Technical report, *Panaromic Research Inc*, California 1965.
- Chang, K.; Bowyer, K. & Flynn, P. (2005). Adaptive Rigid Multi-Region Selection for Handling Expression Variation in 3D Face Recognition,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 157–164, 2005.

- Cook, J.; Chandran, V. & Fookes, C. (2006). 3D face recognition using log-gabor templates, *The 17th British Machine Vision Conference*, September 2006.
- Coxeter, M. (1969). *Introduction to geometry*, New York, NY, Wiley, 2nd edition, pp. 216-221, 1969.
- Chua, C.; Han, F. & Ho Y. (2000). 3D human face recognition using point signature, *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, p.233-238, 2000.
- Dorai, C. & Jain, A. (1997). COSMOS- A representation scheme for 3D free-form objects, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 10, pp. 1115-1130, October, 1997.
- Hesher, C.; Srivastava, A. & Erlebacher, G. (2003). A novel technique for face recognition using range images, *The Seventh International Symposium on Signal Processing and Its Application*, 2003.
- Hsu, R. & Jain, A. (2001). Face modeling for recognition, *Proceedings of the IEEE International Conference on Image Processing*, Greece, October, 2001.
- Huttenlocher, D.; Klanderman, G. & Rucklidge, W. (1993). Comparing images using the Hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 9, pp. 850–863, 1993.
- Jiang, D.; Hu, Y., Yan, S., Zhang, L., Zhang, H. & Gao, W. (2004) Efficient 3D reconstruction for face recognition. *Special Issue of Pattern Recognition on Image Understanding for Digital Photos*, 2004.
- Johnson, A. & Hebert, M. (1999). Using spin images for efficient object recognition in cluttered 3D scenes, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 21, no.5, pp.433-449, 1999.
- Kendrick, K.; Da Costa, A., Leigh, A., Hinton M., & Peirce J (2000). Sheep don't forget a face. *Nature*, pp.165–166, 2000.
- Lee, Y. & Shim, J (2004). Curvature-based human face recognition using depth-weighted Hausdorff distance. *Proceedings of the International Conference on Image Processing*, pp. 1429–1432, 2004.
- Lu, X.; Jain, A. & Colbry, D. (2006). Matching 2.5D face scans to 3D models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no.1, pp. 31-43, 2006.
- Lu, X.; Colbry, D., Jain, A. (2004). Matching 2.5D scans for face recognition. *Proceedings of the International Conference on Pattern Recognition*, pp. 362–366, 2004.
- Lu, X.; Jain, A. (2005). Deformation analysis for 3D face matching. *The 7th IEEE Workshop on Applications of Computer Vision*, pp. 99–104, 2005.
- Pan, G.; Han, S., Wu, Z. & Wang, Y. (2005). 3D face recognition using mapped depth images. *IEEE Workshop on Face Recognition Grand Challenge Experiments*, June 2005.
- Russ, T.; Koch, K. & Little, C. (2005). A 2D range Hausdorff approach for 3D face recognition, *IEEE Workshop on Face Recognition Grand Challenge Experiments*, 2005.
- Russ, T.; Koch, K. & Little, C. (2004). 3D facial recognition: a quantitative analysis. *The 45-th Annual Meeting of the Institute of Nuclear Materials Management*, July 2004.
- Mahoor, M.; Ansari, A. & Abdel-Mottaleb, M. (2008). Multi-modal (2D and 3D) face modeling and recognition using attributed relational graph, *Proceedings of the International Conference of Image Processing*, October 2008.

- Mahoor, M. & Abdel-Mottaleb, M. (2007). 3D face recognition based on 3D ridge lines in range data, *Proceedings of the IEEE International Conference on Image Processing*, San Antonio, Texas, September 16-19, 2007.
- Medioni, G. & Waupotitsch, R. (2003). Face recognition and modeling in 3D. *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, pp. 232-233, October 2003.
- Messer, k.; Matas, J., Kittler, J., Luettin, J. & Maitre, G. (1999). XM2VTSDB: The extended M2VTS database. *Proceedings of Audio and Video-based Biometric Person Authentication*, pp. 72-77, March 1999, Washington, D.C.
- Moreno, A.; Sánchez, Á., Vélez, J. & Díaz, F. (2003). Face recognition using 3D surface-extracted descriptors, *The Irish Machine Vision and Image Processing Conference*, September 2003.
- Moreno, A. & Sanchez, A. (2004). GavabDB: A 3D face database," *Proceeding of the 2nd COST275 Workshop on Biometrics on the Internet: Fundamentals, Advances and Applications*, Vigo, Spain, March, 2004.
- Phillips, P.; Moon, H., Rizvi, S. & Rauss, P. (2000). The FERET evaluation methodology for face recognition algorithms. *IEEE Transaction on Pattern Analysis Machine Intelligence*, Vol. 22, No. 10, October 2000.
- Phillips, P.; Grother, P. , Micheals, R., Blackburn, D., Tabassi, E. & Bone, J. (2003). Face recognition vendor test 2002. <http://www.frvt.org/FRVT2002/Default.htm>. Evaluation Report, 2003.
- Phillips, P.; Flynn, P., Scruggs, T., Bowyer, K., Chang, J., Hoffman, K., Marques, J., Min, J. & Worek, W. (2005). Overview of the face recognition grand challenge, *Proceedings of IEEE on CVPR*, San Diego, pp. 947-954, June 2005.
- Uchida, N.; Shibahara, T., Aoki, T., Nakajima, H. & Kobayashi, K. (2005). 3D face recognition using passive stereo vision. *Proceedings of the IEEE International Conference on Image Processing*, pp. II-950-II-953, Sep. 2005.
- Wiskott, L.; Fellous, J., Kruger, N. & Malsburg, C. (1997). Face recognition by elastic bunch graph matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 775-779, 1997.
- Wu, Z.; Wang, Y. & Pan G. (2004). 3D face recognition using local shape map, *Proceedings of the IEEE International Conference on Image Processing*, Vol. 3, pp. 2003- 2006, Oct. 2004.
- Xu, C.; Wang, Y., Tan, T. & Quan, L. (2004a). Automatic 3D face recognition combining global geometric features with local shape variation information. *Proceedings of the Sixth International Conference on Automated Face and Gesture Recognition*, pp. 308-313, May 2004.
- Xu, C. ; Wang, Y, Tan, T. (2004b). Three-dimensional face recognition using geometric model," *Proceedings of the SPIE*, Vol. 5404, pp. 304-315, August 2004.
- Zhaho, W.; Chellappa, R., Phillips, P. & Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM Computing Survey*, pp. 399-458, December 2003.



### **State of the Art in Face Recognition**

Edited by Julio Ponce and Adem Karahoca

ISBN 978-3-902613-42-4

Hard cover, 436 pages

**Publisher** I-Tech Education and Publishing

**Published online** 01, January, 2009

**Published in print edition** January, 2009

Notwithstanding the tremendous effort to solve the face recognition problem, it is not possible yet to design a face recognition system with a potential close to human performance. New computer vision and pattern recognition approaches need to be investigated. Even new knowledge and perspectives from different fields like, psychology and neuroscience must be incorporated into the current field of face recognition to design a robust face recognition system. Indeed, many more efforts are required to end up with a human like face recognition system. This book tries to make an effort to reduce the gap between the previous face recognition research state and the future state.

#### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Ansari A-Nasser, Mahoor Mohammad and Abdel-Mottaleb Mohamed (2009). 3D Face Mesh Modeling for 3D Face Recognition, State of the Art in Face Recognition, Julio Ponce and Adem Karahoca (Ed.), ISBN: 978-3-902613-42-4, InTech, Available from:

[http://www.intechopen.com/books/state\\_of\\_the\\_art\\_in\\_face\\_recognition/3d\\_face\\_mesh\\_modeling\\_for\\_3d\\_face\\_recognition](http://www.intechopen.com/books/state_of_the_art_in_face_recognition/3d_face_mesh_modeling_for_3d_face_recognition)

**INTECH**  
open science | open minds

#### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

#### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2009 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen