Three-dimensional Facial Surface Modeling applied to Recognition

A. B. Moreno[1], A. Sánchez[1]*, E. Frías-Martínez[2], J. F. Vélez[1]

[1]*Departamento de Ciencias de la Computación, Universidad Rey Juan Carlos, C/ Tulipán s/n 28933 Móstoles (Madrid), Spain*
[2]*Biomedical Engineering IDP, 7523 Boelter Hall, HSEAS, University of California, Los Angeles, CA 90095, USA*

Abstract

Applications related to game technology, law-enforcement, security, medicine or biometrics are becoming increasingly important, which, combined with the proliferation of three-dimensional (3D) scanning hardware, have made that 3D face recognition is now becoming a promising and feasible alternative to 2D face methods. The main advantage of 3D data, when compared with traditional 2D approaches, is that it provides information that is invariant to rigid geometric transformations and to pose and illumination conditions. One key element for any 3D face recognition system is the modeling of the available scanned data. This paper presents new 3D models for facial surface representation and evaluates them using two matching approaches: one based on Support Vector Machines and another one on Principal Component Analysis (with a Euclidean classifier). Also, two types of environments were tested in order to check the robustness of the proposed models: a controlled environment with respect to facial conditions (i.e. expressions, face rotations, etc) and a non-controlled one (presenting face rotations and pronounced facial expressions). The recognition rates obtained using reduced spatial resolution representations (a 77.86 % for non-controlled environments and a 90.16% for controlled environments, respectively) show that the proposed models can be effectively used for practical face recognition applications.

Keywords: 3D Face Modeling; Voxel-based Face Representation; 3D Face Recognition; SVM; PCA

1. Introduction

Most research in face recognition has focused on two-dimensional (2D) images or sequences of images due to their standard acquisition and the computational advantages offered by their regular grid structure (Li, 2007). The main problem of using 2D patterns is that recognition accuracy is sensitive to: rigid geometric transformations, lighting conditions, facial expressions and/or a variety of occluding elements (such as hair, moustache, etc). Nevertheless, Automatic Face Recognition Systems (AFRS) using 2D images have provided excellent results when the image acquisition conditions are controlled (Zhao, 1999). This has motivated the development of techniques to reduce image acquisition restrictions (Zhao et al., 2003; Xie and Lam, 2005; Li and Zhang, 2007). Some methods have been proposed to tackle determined non-controlled variations, but they do not work well in arbitrary conditions (Bowyer et al., 2006).

Working with three-dimensional (3D) face images has some interesting advantages over 2D images (Hsu and Jain, 2001): (1) more geometric information can be obtained from 3D data than from 2D images because 2D images loose depth information as they are formed through projections of 3D objects; (2) the measured features from real 3D data are not affected by the scale and rotation of the face; and (3) if the 3D face recognition system does not consider texture information, the recognition is immune to the effect of illumination variations. One important problem of 3D face recognition was related with the hardware needed for face acquisition: (1) the cost of 3D range scanners; (2) the scanning speed; as well as (3) the size of this kind of digitizers. Nevertheless, advances in computational processing capacities as well as the reduction of both cost and size of 3D digitizers, have made possible the development of 3D face recognition systems. In recent years, the interest in 3D face recognition has increased (Huang et al., 2003; Lee and Ranganath, 2003; Rajagolopan et al., 2006; Wechsler, 2007; Mahoor et al., 2007).

Two main components of any face recognition architecture are: (1) face modeling, and (2) recognition or matching scheme. Face modeling transforms the original 3D surface meshes into a set of features to better represent the face in the AFRS. Matching scheme involves the selection (and training) of a classifier that using the extracted features from a test face compares them with the facial features of different people in a database to determine the test subject identity. While the matching scheme can be efficiently implemented using standard machine learning techniques like Neural Networks or Support Vector Machines, finding good 3D face features (models) for recognition is, in general, a complex problem.

3D Modeling approaches for face recognition can be classified into: a) those which combine 3D face models with 2D view-based techniques (Campbell and Flynn, 2001; Ma and Wong, 2006) in order to extract features and to determine the face pose and/or location, b) the use of 2D+3D multimodal data (i.e. 2D texture image in

---

* Corresponding author. Tel: +34 91 6647452; Fax: +34 91 4888578
 E-mail: angel.sanchez@urjc.es (A. Sánchez)

addition to the corresponding depth map) as input to the classifier in order to demonstrate how multimodal 2D+3D improves the 2D face recognition rates (Chang et al., 2005; Boyer et al., 2006) and c) the use of the captured 3D data exclusively to extract 3D shape descriptors as the local curvature, ridge lines, etc. (Hallinan et al., 1999; Moreno et al., 2006; Mahoor et al., 2007). Most of the existing 3D facial recognition methods focus on recognizing 3D faces with neutral expressions (Haar et al., 2008). Our work belongs to the third group because we use the captured 3D data (without the texture image) to extract 3D shape descriptors for face recognition, and it overcomes the recognition in presence of some head rotations and pronounced facial expressions.

This paper proposes and evaluates several 3D face modeling approaches. To validate the models, two matching schemes were used: (1) Support Vector Machines (SVM); and (2) Principal Components Analysis (PCA) in combination with a Euclidean distance classifier. We used our own 3D face database called GavabDB (which can be found in http://gavab.es) for the experiments because it presented a higher degree of variability among the images of each individual, especially related to facial expressions.

The rest of the paper is organized as follows. First, Section 2 describes the State of the Art of the 3D Automatic Face Recognition (AFR) Systems focused on the existing profile and contour curves and depth-map based approaches for modeling 3D faces. Section 3 includes the main features of our 3D face database and compares some works which have used it for their experiments. Section 4 presents the recognition results achieved with the profile and depth-maps models (both on a controlled and on a non-controlled environments) using SVM and PCA as matching schemes. Finally, conclusive remarks and future works are presented in Section 5.

2. 3D Automatic Face Recognition: State of the Art

The following subsections present the state of the art of 3D face recognition systems and the existing techniques related to our approach.
One of the main motivations of using 3D Automatic Face Recognition is to reduce some of the existing problems that 2D face images have (in particular, the illumination, pose and face expression dependences). Although intuitively 3D information (i.e. depth, curvature, etc) should provide better results than 2D intensity-based features, the evaluation results are still limited (Bowyer et al., 2006; Chang et al., 2005). There are some research works that compared 3D and 2D facial patterns (Chua et al., 2000; Lee and Ranganath, 2003; Bronstein et al., 2003; Huang et al., 2003a). An important part on research in 3D face recognition systems has focused on 3D face modeling and on image acquisition techniques.
This section reviews the 3D AFRS and classifies them according to the type of modeling technique into: a) those which use a 3D face model to extract features from

2D data, b) those which use 2D+3D multimodal data as inputs to the classifier and c) those which use the captured 3D data exclusively to extract 3D features from them.

2.1 3D Face models to extract features from 2D data

One common holistic technique is based on the correspondence between a set of 2D scene points and set of 3D model points in order to determine the object pose and its location, and also for extracting some facial features (Huang and Tang, 1996). In this context, a recent work is Ma and Wong (2006) which combines 3D face models with 2D view-based techniques for recognition.
The use of a generic 3D head model enables to obtain 3D information from 2D images. The generic model contains control points that can be displaced to adjust themselves to their corresponding points in the 2D images of the individual. These displacements deform the generic model generating a particularized model of the subject from the 2D images. The deformations are expressed in terms of modifications of certain parameters which characterize the 3D global face shape of the individual as well as local information. This technique avoids the 3D image acquisition stage of the test images. The need for searching 2D points in the images is a serious disadvantage because of the dependence on the acquisition conditions. Some applications of the use of a generic 3D head model are: analysis and synthesis of facial expressions (Lei et al., 1996), analysis of robustness to illumination or pose correction in an AFR systems (Lee and Ranganath, 2003), generation of synthetic faces to augment the number of variations of each individual of the training set (Vetter, 1999), among others.

2.2 2D+3D multimodal data

These techniques use 3D geometry data together with the 2D intensity images captured simultaneously by the same digitizer (Beumier and Acheroy, 2000b; Wang et al., 2002; Godil et al. 2004). (Chang et al., 2005) and (Bowyer et al., 2006) present studies of how multimodal 2D+3D improves face recognition rates.
A face verification system was proposed in Beumier et al. (2000a) for 3D images captured using active triangulation. A representation technique based on local features, which gathers both the 3D shape information and the 2D texture information, was analyzed and described in Wang et al. (2002). Feature points were located using Gabor filter responses in the 2D transformed domain, and the signatures of these points were combined with the corresponding points in the 3D domain. Both representations were projected on their corresponding subspace using PCA. The obtained weight vectors were integrated in an augmented vector used to represent the face image in the recognition system. The identification system was constructed using Support

Vector Machines and achieved a 90% correct recognition rate using a database of 50 individuals presenting different facial expressions and poses.

Other works (Tsalakanidou et al., 2003; Chang et al., 2003) used PCA to obtain a low dimensional representation of 3D images (consisting in depth maps of the complete face) for recognition. The goal was to evaluate the influence of color, depth (defined as the distance from a scene point to the sensor) and a combination of both types of data in a 3D face recognition system. These authors compared the use of 2D complete images and their corresponding 3D depth maps separately and also both representations together. The fusion of 2D and 3D information provided better results (Chang et al., 2005). These systems have not considered images presenting facial expressions. Pose variations were very limited (for example, rotations of 2º, 5º and 10º (Tsalakanidou et al., 2003) or they were not considered (Chang et al., 2003). Ben Abdelkader and Griffin (2005) demonstrated that depth and texture information played complementary roles in coding the 3D face patterns.

2.3 3D captured data to extract 3D features

Differential geometry has been used for feature extraction in the context of free-form 3D object recognition and also used by some authors for facial feature extraction (Tanaka et al., 1998). The local curvatures of 3D surface points and the angle among surface normal vectors have been proposed as 3D free-form object descriptors for recognition (Stein and Medioni, 1992; Thirion, 1996; Tanaka et al., 1998; Hallinan et al. 1999; Campbell and Flynn, 2001; Moreno et al. 2006).

A classification of segmented points according to planar, spherical or revolution surfaces was proposed in (Gordon et al. 1991). A region classification, according to concavity, convexity and saddle points, was used to localize nose, eyes and mouth. Neck, front and cheeks were also localized according to their smoothness properties. These features were used to normalize the images by scaling and posing them into a cylindrical mesh such that the volume between face and cylinder represented a face for recognition. The system was tested using 24 images corresponding to 8 individuals having 3 images of each one of them. Feature detection was 100% correct, and the recognition rate was 97% using individual features and 100% using the whole face as pattern.

The local surface curvature evaluated at a point is described by the directions in which the normal of the surface changes more and less quickly (Gray et a., 1999). In (Hallinan et al. 1999), a set of twelve 3D descriptors extracted from segmented regions using curvature properties of the surface were tested for face recognition using a database of 8 subjects and 3 images per subject obtaining 95.5% recognition rate assuming a previous 100% correct feature extraction.

Another representation of 3D range face images using local information is based on computing the point signature over specific 3D points in order to obtain descriptors. The point signature has been used to represent 3D free-form objects; in particular, 3D face images (Chua et al., 2000).

Other techniques belonging to this category are those which use 3D profiles and surface contours (curves extracted from the 3D face surface) to represent the faces in the AFRS. Profile and contour-based face recognition is a very intuitive idea used by different authors (J. Y. Cartoux et a., 1989; T. Nagamine et al. 1992; Beumier et al., 2001; Pan et al. 2003). These techniques reduce the amount of face data to one or a few 3D curves. The selection of the optimal set of 3D curves and the best way to match them is still a research area (Haar et al., 2008). More recent works also propose the use of these types of features for face recognition. In (Li, C. and Barreto, A., 2004) different combinations of profiles and contours were extracted from 3D range faces and tried for face recognition using a set of 27 subjects (with 27 neutral images in the gallery and 27 neutral test images both captured in different times). Each face had an average of 18,000 vertices. Their results showed that the central vertical profile is the most powerful profile to characterize individual faces among the different profiles they tested. In (Li, C. and Barreto, A. et al., 2005) the central vertical profile and the face profile (curve where the face cut a plane placed at 30 mm. below the tip of the nose) are both used as very useful features for face recognition. When both are combined, better recognition rates (around 81%) were obtained than just using any of them alone. They used a database of 80 subjects for the experiments containing two frontal neutral images per subject (one gallery image and one test image) captured with a lapse between them between one to thirteen weeks. These images consisted in 640×480 array of range data.

Haar et al. (2008) proved that a combination of contours with profiles and a second combination of different contours achieved a performance of 0.76 and 0.79, respectively on the 2007 Shape Retrieval Contest of 3D Face Scans (SHREC 2007).

Our research aims the analysis of new 3D face models that combine efficiency and robustness on presence of pronounced face gestures and head rotations.

3. *GavabDB*: description and reported works

This Section presents our 3D face database and describes a relation of works in the literature that used GavabDB for their experiments.

3.1 *GavabDB:* justification and description

AFR systems need from face databases in order to train and evaluate such systems. These databases should have a degree of variability among the different patterns of the same subject with respect to pose, illumination, facial expressions, etc.. There are different facial

Fig. 1. Seven 2D images of the same person that corresponds to the non-textured 3D meshes.

databases designed for different experimental purposes: face recognition, facial expression analysis, pose estimation, face detection, etc. However, most of the databases only provide 2D face images. The number of 3D face databases available is limited and, in general, they are designed for specific aspects of recognition. To our knowledge, the most representative 3D face databases are: XM2VTS multimodal database (xm2vtsdb, 1999; Messer et al., 1999; Matas et al., 2000), 3D_RMA database (Beumier Database, 2000), York University 3D face database (3DFaceDBYork, 1999), Notre Dame 3D database (Notre Dame, 2003), FRGC dataset (Phillips et al., 2005) and *GavabDB* (Moreno et al., 2004).

Most of the existing 3D databases contain few 3D meshes per subject in some cases and there are not many variations among the different 3D images of a person. Some of the databases offer variations related to some particular aspects but not with respect to others, and, in general, the variation rank is reduced. Among the presented databases, very few capture images with some kind of 3D facial expression. The database which offers most pronounced facial expressions and that is the nearest to the variability that an individual can have when interacting with a real face recognition system is *GavabDB*. (Li et al. 2007) expressed this as: "we focus on the GavabDB 3D face database in our studies as it offers special challenges in terms of data artefacts and severe facial expressions". This high variability in pose and in facial expression is very important in order to test the robustness and the efficiency of a particular face representation in practical conditions. This database has been created to test the performance of face recognition systems. Moreno and Sánchez presented a preliminary description of the GavabDB database in (Moreno et al., 2004). This database can be found in http://gavab.es.

In our experiments, we have used a subset of 427 3D facial surface meshes from the 549 ones in *GavabDB*. These patterns correspond to 61 subject (having 7 meshes per subject) without using the color (texture) information. Fig. 1 shows seven 2D images of the same

person corresponding to the non-textured 3D meshes for each individual used in this work. The set of samples of each individual contains: two frontal images with neutral expression, one image in which the subject is looking down (+35º *x*-rotation approximately) with a neutral expression, one image in which the individual is looking up (-35º *x*-rotation approximately) with a neutral expression and three frontal images with some kind of expression (one random expression chosen by the individual, one smiling and other laughing). The remaining left and right subject's profile images in the *GavabDB* face database (±90º *y*-rotation) have not been used in this work.

Occlusions are present in 12 subjects (19% of the individuals in the database): by fringe (3), braids (1), beard (4), moustache (1) and both goatee and moustache (3). They appear in Fig. 2. On the other hand, there are occlusions in some gesture image which appear in the last column of Fig. 2. There are no individuals wearing glasses in the database.

Sometimes dark parts of the face surface do not reflect the projected light from the 3D digitizer and then, they can not be sampled. Also there are occluded parts from the scanner viewpoint that can not be sampled. The scanner software offers the possibility of filling the holes in the meshes, which has been applied to the *GavabDB* database images. Holes are patched by linearly interpolating adjacent pixel values. This is automatically performed by the 3D digitizer software and it causes a surface reconstruction error but in compensation the resulting mesh offers us the possibility of moving from all new mesh nodes to the adjacent 4-connected or 8-connected nodes (Moreno et al., 2004). Fig. 3 shows an example of the introduced reconstruction error by the filling holes stage. Also, the captured 3D faces are incomplete because of their auto-occlusion at certain face orientations.

Our database was created using the scanner Konica Minolta VI-700 (Konica Minolta). Thousand of points approximate the surface of each sampled face and the local connections between these points form a mesh
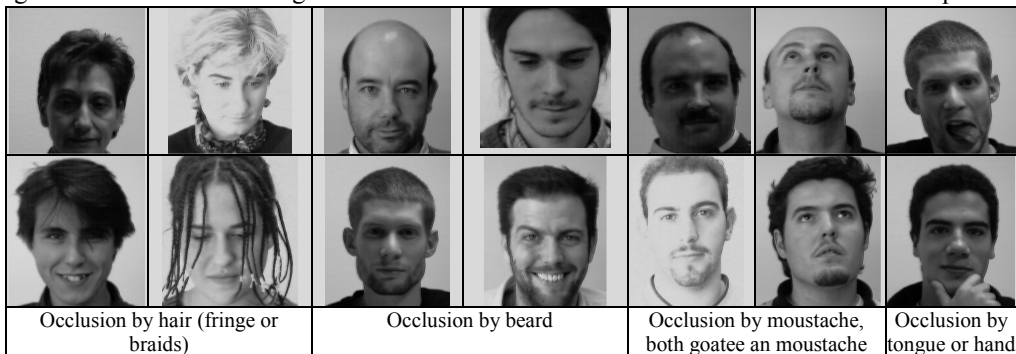

Occlusion by hair (fringe or braids) | Occlusion by beard | Occlusion by moustache, both goatee an moustache | Occlusion by tongue or hand

Fig. 2. Examples of 2D images captured by the 3D digitizer whose corresponding meshes belong to the G*avabDB* database, which present some kind of occlusion.

representing the scanned face. Cells of the mesh have four nodes, and occasionally three (in the border of the mesh).

The 3D facial meshes have been stored forming two datasets containing each one images with one of the two following resolutions: (1) the original resolution (abbreviated as 1-1 resolution) which maintains all the points provided by the scanner, and (2) ¼ of the original resolution, which maintains only one of each 4 original points (abbreviated as 1-4 resolution). The number of points varies from one mesh to another one. The average number of vertices in the meshes is 2,350 at the 1-4 resolution (after manually removing neck, ears and/or others possible patches not belonging to the face), being 3,420 the maximum and 1,538 the minimum number of vertices, respectively. The goal of producing both datasets for this work is studying how different resolutions could affect the robustness of each face model and recognition method, regarding the corresponding correct recognition rate achieved and its processing time. Fig. 4 shows an example of a 3D face mesh at the original 1-1 resolution (left) and at the 1-4 resolution (right).
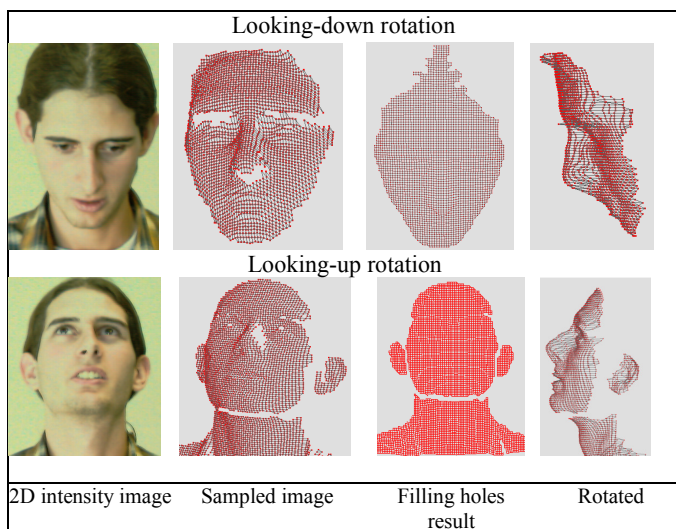

Fig. 3. Error after the filling holes stage. Last column shows the reconstructed base of the nose of an individual looking down and the nose of the same individual captured when it was visible.

3.2 *GavabDB*: related works

GavabDB database has been used in a variety of works focusing on 3D face modelling and recognition, among them: Joonsoo Lee (2005); Berretti et al. (2006); Berretti et al. (2007); Ansari et al. (2007); Mahoor et al. (2007); and Li et al. (2007). Different works have considered different subsets of the *GavabDB* patterns for their experiments. Some of them do not include pose variations; others do not include the "random gesture" and/or the "laugh" images. Besides, different numbers of images in the gallery and the test datasets have been used. Berretti et al (Berretti et al. 2006; Berretti et al. 2007) propose a model based on Iso-Geodesic Stripes

(IGS) extracted from the 3D face surface. An IGS is formed by the surface points in which the value of a real valued function is the same. The function they used was the geodesic distance from the point to a reference point (nose tip), divided by the Euclidean eye-to-nose distance. The geodesic distance (length of the shortest piecewise linear path on the mesh vertices) was accomplished through the Dijkstra algorithm. Once values of the function are computed for every surface point, iso-geodesic stripes were identified by quantizing the range of the function values into *n* intervals. The face was characterized by the relationships among every pair of stripes, which were computed by the relative position of their points in the 3D space. They have been computed using 3D weighted walkthroughs (3DWW, formed by 27 weights organized in a 3×3×3 matrix) between every voxel pair. Every face was represented by a graph where graph nodes represent iso-geodesics and graph edges represent their spatial relationships. The face recognition is achieved by graph matching. They reported 94.5% and 82% of recognition rate for neutral and non-neutral expression, respectively, considering 1-1 mesh resolution. Our proposed approach presents some positive aspects comparing to this approach, for example, the possibility of obtaining efficient face representations at much lower voxel resolutions. For example, we reach more than 90% of recognition rate for 45×45×45 voxels for controlled environments. This fact is very important for this kind of systems.

Other work which uses the *GavabDB* database is (Ansari et al. 2007) which uses three facial feature points extracted from the range images to align them to a 3D generic face model. Each plane defined by each triangle vertices of the model is fitted to the corresponding interior 3D range data using least squares plane fitting. Next, the model is subdivided into a higher resolution mesh which is deformed again to fit the range data. The final individualized model captures the surface characteristics of the face surface. This work does not experiment with images presenting facial expressions nor pose variations, reporting a 90.3% recognition rate for frontal and neutral images whose resolution is also 1-1.

(Mahoor et al. 07) presents an approach for identification of both neutral faces and faces with smile expression using the corresponding images of the G*avabDB* database, reporting a 93.5% recognition rate for neutral expression and a 82% for the faces with smile expression (being the gallery formed by frontal images). For pose alignment, they extracted the locations of three feature
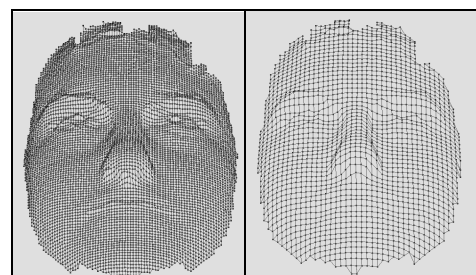

Fig. 4. (Left) Original mesh produced by the 3D scanner and (right) the same mesh at 1-4 the original resolution.

points, the inner corners of the two eyes and the tip of the nose. For face identification, they found the location of the ridge lines in the range images. These are the lines where the principal curvature has positive value, upper than a threshold. They obtained a 3D binary image that contained only ridges. For recognition they used Hausdorff distance to find the best match for a given probe image from the facial range images in the gallery. Their process of locate three feature points failed for a 15% of the images, and their first approximation of the nose as the nearest point from the scanner failed for 10% of the images, where a manual selection of the points was achieved. Their proposed system do not has been proved using pronounced expressions which have higher variations from the neutral expression of the gallery images.

In Li et al. (2007), each face is represented by a triangle mesh, pre-processed to have a uniform connectivity. This needs to pick manually 43 vertices corresponding to a face mask. The face symmetry is used to complete missing data. They investigate the use of geometric attributes including angles, geodesic distances, triangle areas, etc. They adapt the attributes to instabilities introduced by facial expressions and test combinations of them during the training stage, to obtain an optimized face recognition scheme. A one-parameter weighting function is proposed to stabilize a descriptor element, being this parameter computed in the training stage as the one which gives the best recognition rate among all tests. Five frontal images per individual were used (two neutral faces and three with expressions) but they do not considered rotated images. A set of 60 subjects of the *GavabDB* was partitioned into two groups of 30 subjects with 5 images per person. One of the groups was used to adjust the system parameters during the training, while the other group was used for testing. The experiment was repeated changing the roles of both groups. The average recognition rate achieved was a 93%.

4. Voxel-based 3D Face Modeling

In profile and contour-based face modeling, a 3D facial mesh can be represented by curves describing its surface. As an example, we can use a contour curve at the nose level, a contour curve at the eyes level, or a combination of both. Our proposed face descriptors are profiles (alone or in combination with others) given by curves corresponding to the intersection of a given cut-plane with the face surface.

On the other hand, the number of points and the sampled area are different for each mesh, depending on the face shape and the head pose. Because we use classifiers that need from patterns with a fixed dimension, we have developed a framework that (1) has a fixed length for all the patterns, and (2) intuitively describes the selected patterns. This framework represents 3D faces using voxels as the basic mechanism to define the proposed profiles.

4.1 Voxel-based Face Representation

The voxel representation of a face is obtained when the original 3D mesh is transformed into a discrete and regular representation with a fixed number of volume elements (voxels). The process starts creating a 15x15x15 cm^3 cube to contain the relevant facial parts of all meshes. The cube size has been selected after analyzing the sizes of the faces in the database.

In order to place a facial mesh into the cube, it has to be normalized with respect to the pose. Then, it is translated until the pronasale point (nose tip) of the facial mesh is placed in the coordinates x=7.5 cm, y=7.5 cm and z=15 cm of the cube. Fig. 5 shows an example of cube translation to capture the face. In general, the upper and lower parts of the facial mesh are located outside the cube, which will imply that information of these parts will not be captured by the voxel representation. This is not relevant since the upper part of the mesh could contain hair and the lower part of the facial mesh has high variability under facial expressions.

Once this facial mesh has been placed in the cube, it is subdivided into $n \times n \times n$ equally spaced voxels, where n represents the voxel resolution. Fig. 6 presents an example in which the cube containing the face is subdivided into: (a) $2 \times 2 \times 2$ voxels, and (b) and (c) into $3 \times 3 \times 3$ voxels. This voxel-based representation enables us to convert a 3D facial mesh into a binary representation. If a voxel contains vertices of the facial mesh it is represented by a '1'; otherwise, the corresponding voxel is represented by a '0'. Fig. 6 (d) shows a section of the cube occupied by the mesh and divided into voxels, and (e) the binary values produced by the voxels using a cutting plane defined at y=7.5 cm.

In this work, we used three voxel resolutions to produce the representations of facial meshes: $30 \times 30 \times 30$, $45 \times 45 \times 45$ and $55 \times 55 \times 55$, respectively. Each one is applied to both considered facial mesh resolutions: 1-1 (original) and 1-4 (one quarter), respectively.

4.2 Profile Modeling

Using the previous mechanism to obtain profiles, the following five potential types of single profiles (shown in Table 1) were identified:

(1) Horizontal cut at mouth level. This profile is obtained by cutting the cube with a horizontal plane at y=$n/3$, being n the length of the cube in voxels. This cut captures a very relevant and distinguishable
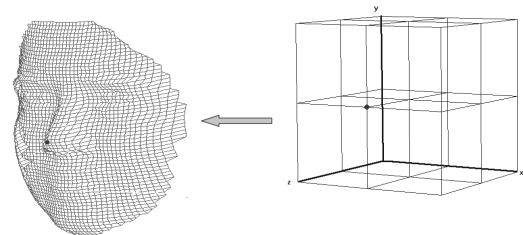

Fig. 5. Cube translation to contain the facial mesh using the nose tip as reference.
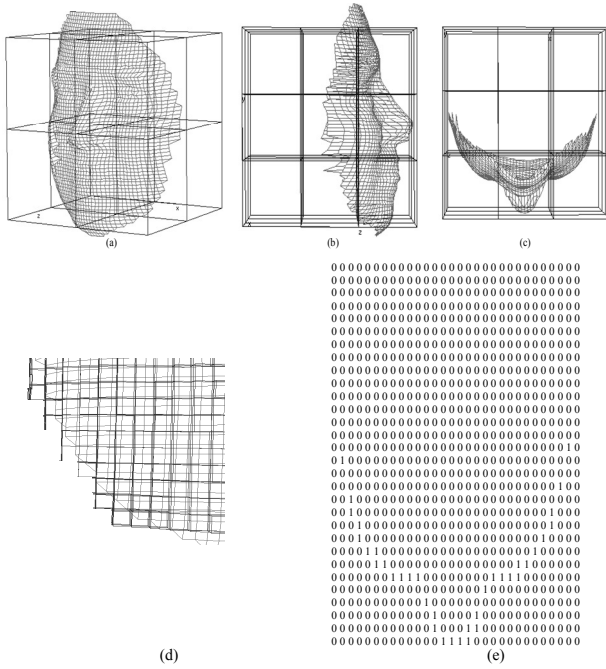
Fig. 6 (a) 2×2×2 voxel division of the reference cube (containing a facial mesh); (b) and (c) 3×3×3 voxel division of the rotated cube; (d) voxels (dark grey) containing part of a facial mesh, and (e) example of the binary voxel representation corresponding to a horizontal section at nose level (the contour is represented by the 1-values).

part of the face: the mouth. Nevertheless its main drawback is its sensitivity with facial expressions.

(2) Horizontal cut at the nose level. This profile is defined as the plane obtained by a horizontal cut at $y=n/2$. It presents the horizontal profile of the nose.

(3) Horizontal cut at the eyes level. This cut is obtained by the horizontal plane with $y=2n/3$. This profile contains the curve of the eyes and part of the nose.

(4) Vertical cut at the nose level. This cut is obtained by defining a plane with $x=n/2$.

(5) Vertical cut at one eye level. This cut is obtained by the vertical cut at $x=2n/3$. It shows the profile of the face passing by one eye.

Horizontal cut at mouth level (y=n/3)	
Horizontal cut at nose level (y=n/2)	
Horizontal cut at eye level (y=2n/3)	
Vertical cut at nose level(x=n/2)	
Vertical cut at eye level (x=2n/3)	

Table 1. Examples of the defined single profiles.

These five profiles have been obtained for each mesh

using three different voxel resolutions and the two considered mesh resolutions(1-1 and 1-4, respectively). An example of the horizontal cut at nose level for an individual at the considered resolutions is presented in Table 2.

	30×30×30	45×45×45	55×55×55
1-1 Resolution			
1-4 Resolution			

Table 2. Examples of horizontal cut at nose level for the considered voxel and scanning resolutions.

We have also defined two profile combinations in order to experiment multiple-profile modeling approaches. It aims at representing each facial mesh by a combination of more than one single profile in order to capture more relevant information. The combinations proposed (shown in Table 3) are:

(1) Combination of "eyes and nose" profiles. This approach models each face with the concatenation of two profiles: horizontal cut at the eye level and vertical cut at the nose level.

(2) Set of horizontal cuts between nose and eyes. This representation models each face as the concatenation of all the horizontal cuts present between $y=n/2$ and $y=2n/3$, which approximately represents all cuts between the forehead and the nose for each considered resolution.

Table 3 shows an example of both multiple profile models. For each combination, the considered resolutions have been stated in the experiments.

Combination "eyes and nose" profiles	Horizontal cuts between eyes and nose

Table 3. Examples of combination profiles.

4.3 Depth Map Modeling

A depth map is a two-dimensional matrix in which the value of each component represents the distance from a plane to the face. In our case, we have placed the origin

plane at the pronasale point to obtain depth maps in vertical position, which means that for this point the depth is 0. Around the nose region the distance is small, while in other face areas the distance increases. We have measured this distance using the number of empty voxels from the origin plane to the facial mesh. In order to study the modeling capabilities of depth maps we have defined three types of maps (Table 4):

Full face	Upper-half	Left-side

Table 4. Examples of the considered depth maps.

(1) Full facial depth map. This modeling approach presents a matrix containing the distance measured in voxels from the vertical origin plane to the corresponding point of the face.

(2) Upper-half facial depth map. Because the mouth area is very sensible to facial expressions, modeling a face considering only the upper-half of the face can be appropriate to obtain more robust models. The upper-half of the face is defined by the horizontal cut at the pronasale point.

(3) Left-side facial depth map. The goal of this model approach is to eliminate redundant facial information due to face symmetry. In theory, both halves of the face contain the same information, so only one of them should be enough to represent the face.

Three kinds of depth-maps have been obtained at the previously described voxel and mesh resolutions. Table 5 presents one example of the "upper-half" facial depth-map obtained for the face of one individual at the considered cube resolutions and at 1-1 mesh resolution. Shorter distances are visualized by darker greylevels (the black colour means a distance of 0 voxels).

	30×30×30	45×45×45	55×55×55
1-1 Resolution			

Table 5. Example of upper-half depth map for the considered voxel resolutions at 1-1 mesh resolution.

5. 3D Face Recognition Experiments

In this Section, we present the AFR system architecture implemented and the results obtained using the proposed 3D profile and depth-maps. Two matching schemes have been tested: Support Vector Machines (SVM) and Principal Component Analysis in combination with a Euclidean classifier (referred as PCA in Fig. 7). In order to test the robustness of our face recognition system, experiments have been run in a controlled and in a non-controlled environment to measure how the non-controlled conditions affect the recognition rate.

5.1 Face Recognition Architecture

The proposed face recognition architecture has four stages pre-processing, (2) pose normalization, (3) representation of the facial mesh using voxels, producing the set of profile models previously presented, and (4) classification or matching. In the face modeling (off-line) the three first steps are applied. In the matching, the system is trained and tested for the two considered types of classifiers. Fig. 7 shows this generic architecture.

5.1.1 Face Pre-processing

A pre-processing stage to remove the regions corresponding to the neck, the ears and the hair (which
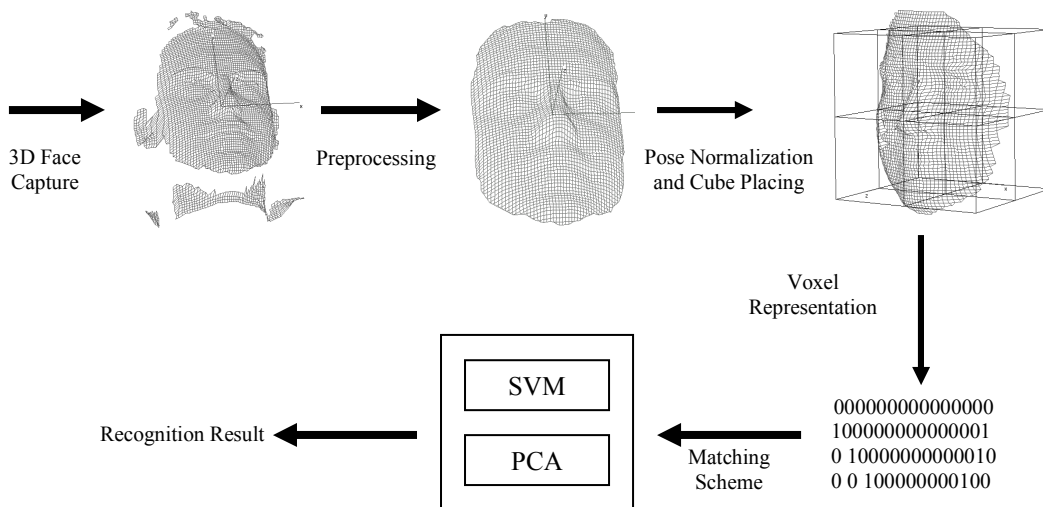

Fig. 7. Architecture of our 3D facial recognition system with two generic matching schemes.

are not relevant for 3D recognition) has been manually applied. After that, the impulsive noise present in the meshes (characterized by extreme values of z randomly distributed through the mesh) has been attenuated automatically by the application of a 3×3 median filter. Also, a 3×3 mean filter has been applied to smooth the z coordinates of the image. No interaction to remove another kind of noise has been done. Fig. 8 shows an example of noise filtering and smoothing. Although the noise filtering is not perfect, the noise attenuation obtained is acceptable for processing our mesh images.
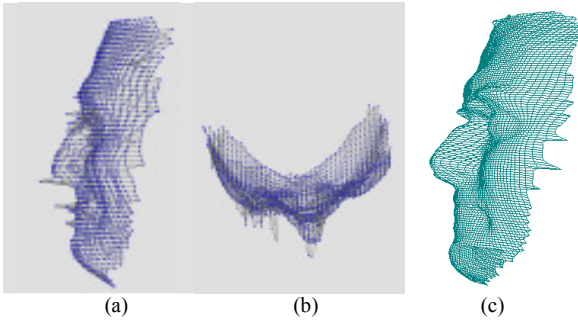

<div align="center">(a)   (b)   (c)</div>
Fig. 8. (a)(b) Noisy mesh from two different viewpoints; (c) Mesh after applying noise filtering and smoothing.

5.1.2 Pose Normalization.

The goal of pose normalization is to have all the facial meshes with the same position and orientation in the 3D space. Scale normalization is not computed since the scale of the 3D images is real, and it characterizes the differences among individuals. In order to normalize each 3D mesh, it is multiplied by a roto-translational matrix that rotates and translates the original mesh to adjust it to a reference face selected from the database (a frontal neutral image). This matrix is computed for each image using a determined set of 3D feature points that have been extracted from the facial mesh to be normalized. The same set of points extracted from the reference face. Both sets of corresponding 3D feature points define an overdetermined system of linear equations of the geometrical transforms that has been approximated by the iterative least-squares method.

The set of points first used were the mass centres of six homogeneous regions according to the local shape computed using the signs of the mean and Gaussian curvatures, using the HK algorithm (Trucco and Verri, 1998). The selected regions are shown in Fig. 9. Region 1 (elliptical convex points of the nose tip), Region 3 (hyperbolic points of the upper-nose bridge) and Regions 4-5 (elliptical concave points of the inner corner of the eyes cavities) were located after the point classification using the threshold curvature values (assigning 0 curvature value when this is lower than a threshold). Regions 6-7 (elliptical convex points of the eyes) were obtained from the point classification of the point shape without applying the curvature threshold process. These normalization regions and other segmented ones are described with more detail in (Moreno et al., 2006).
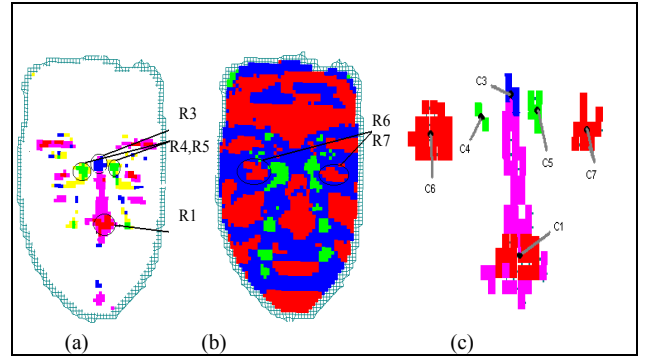

<div align="center">(a)   (b)   (c)</div>
Fig. 9. Regions of an image whose mass centres have been used for pose normalization from (a) the curvature threshold image and from (b) the non threshold image; c) mass centres of these regions.

The selected points are invariant to geometrical transforms. The regions are present in most of the face images and they are scarcely affected by facial expressions. A first normalization process using these six points was applied for the images in which at least three of them were located. It let us to normalize the 95.5% of the images (408 images from the set of 427 ones). The remaining 4.5% of images (a total of 19) were automatically normalized by using another set of nine points located by using a second normalization method. These nine points correspond to landmarks (pronasale, left and right entocanthions, left and right ectocanthions, nasion, subnasale, and finally left and right alares). These points have been located by using another method based on differential properties of the surface (Moreno et al., 2002). The use of both point location methods made possible to automatically normalize the whole 427 images. Table 6 shows the average distance in mm. between corresponding feature points of each normalized view of each individual with these points of the reference face, using the second normalization method for the complete database. Fig. 10 shows the average percentage of the coincident 1-labeled voxels between frontal and another different view of the same individual averaged over all individuals of the database per each view, using the second normalization method for the complete database. First normalization method has to provide better results because the curvature used for the region extraction is invariant under geometrical transforms.

Looking down	Looking up	Frontal 1st	Frontal 2nd	Random gesture	Laugh	Smile
7.0	5.1	4.7	4.7	4.9	4.9	4.8

Table 6. Average distances in mm. between corresponding feature points of each normalized view of each individual to these points of the reference face.
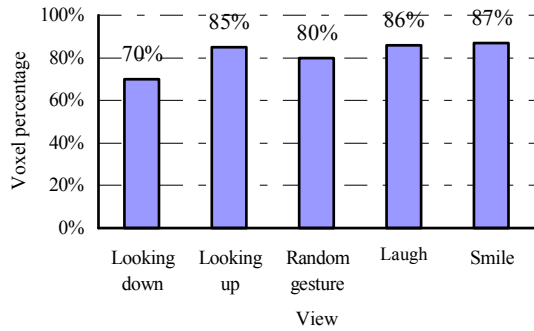
Fig. 10. Average percentage of the coincident 1-labeled voxels between frontal and other different view of the same subject, for each view and averaged by all individuals of the database.

5.1.3 Voxel Representation and Profile Modeling

Once all facial meshes have the same pose they can be transformed into a voxel representation using the procedure explained in Section 4.1. Therefore, all the facial meshes are represented by constant-length vectors. The length of the pattern vectors depends on the voxel resolution used to divide the cube. Each face produces a set of vectors defined by the profiles and depth maps (presented in subsections 4.2 and 4.3). Different voxel and scanning resolutions have been considered to test the recognition capabilities of the presented 3D face surface models.

5.1.4 Matching Schemes

Once the feature vectors are obtained from the facial meshes, a recognition system is trained to classify them. We have implemented recognition systems using two matching schemes that have produced very good results for recognition applications: Principal Component Analysis with a Euclidean Classifier (PCA) and Support Vector Machines (SVM).

PCA is applied for reducing the dimensionality of the feature vectors in order to obtain a more compact representation (Duda, Hart and Stork, 2000). This compressed representation is the input for a pattern classifier, typically a distance based classifier.

SVM is a classification technique derived from Statistical Learning Theory that was originally presented in Boser, Guyon and Vapnik (1992). The problem that SVM tries to solve is to find an optimal hyperplane that correctly classifies data points and separates the points of two classes as much as possible. One important component that defines a SVM classification system is the type of kernel. Intuitively, a kernel should represent the notion of similarity between any two individuals of the database. The most common kernels are linear, polynomial and Gaussian. A kernel can also be specifically designed for a particular problem. The main advantages of SVM, when used for image classification problems, are: (1) ability to work with high dimensional data; (2) high generalization performance without the need to add a-priori knowledge, even when the dimension of the input space is very high; (3) a relationship between the structure (the support vectors) and the classification tasks; and (4) optimal separation surfaces between classes. Bynm (2003) presents an extensive review of SVM for pattern recognition. Excellent introductions to SVM can be found in Vapnik (1998), and Cristianini and Shawe-Taylor (2000).

The proposed matching schemes have been implemented using a PCA distribution (Romdhani, 1996) and the SVMTorch software (Collobert and Bengio, 2002), respectively. The PCA output vectors were classified using a Euclidean distance measure to identify each test face. The kernel included to run all the SVM experiments was a Gaussian kernel. The reason for choosing this kernel is that it has been widely used with very good results for pattern recognition applications (Bynm, 2003).

5.2 Face Recognition Experiments in a Non-Controlled Environment

The goal of this approach is to test how the proposed profile models work in real 3D recognition environments, where users have a certain degree of freedom regarding facial gestures and head rotations. The non-controlled environment has been built by taking five randomly selected facial meshes (of the total of seven per each individual) to train the classifier and the other two remaining images for testing purposes. Therefore, a total of 305 profile vectors were used for training and 122 for testing. The following subsections the most significant recognition rates obtained by the different models at the considered voxel and scanning resolutions.

5.2.1 Single Profile Modeling

Single profile modeling groups five types of face representations presented at Table 1: (1) horizontal cut at the mouth level, (2) horizontal cut at the nose base level, (3) horizontal cut at the eyes level, (4) vertical cut at the nose level and (5) vertical cut at one eye level.

The horizontal cut at mouth level did not produce good results in non-controlled environments due to the variations within the same individual caused by gestures. The horizontal cut at the nose base level did not produced relevant results. We think this is caused by the produced occlusions in the base of the nose when the head is lightly rotated downwards and the reconstructed error produced by the scanner in the filling-holes stage. The horizontal cut at eye level improves the two previous profiles achieving up to 60% correct recognition rate for the 1-1 mesh resolution and for the 45×45×45 cube resolution. We think that this model could provide better results at higher cube resolutions if a Bresenham algorithm were applied to connect the 1-voxels in this profile model, but it also would increase the cost in time and memory.

The recognition results obtained with vertical cut, at the eye level did not produce good results, and the vertical cut at the nose level also achieved up to 60% correct recognition rate.

5.2.2 Multiple Profile Models

Multiple profile models were designed to capture the most discriminating information from each facial mesh in order to improve recognition rate. Two types of multiple profiles are proposed: (1) combination of eye and nose profiles, and (2) combination of all horizontal cuts between nose and eyes. These two approaches include nose and eyes, which are the facial features less affected by facial expressions and gestures. Figs. 11 and 12 present the achieved recognition rates for both multiple priles using each considered matching scheme, voxel and scanning resolutions.

The combination of "eyes and nose" correctly identified 77.86% of the considered 3D test faces, which means an important increase when compared to the 50% and 60% recognition rates achieved by each one of these two profiles independently. Also, in this case, as observed in Fig. 11, the 1-1 resolution provided much better results than the 1-4 resolution. The combination of all horizontal cuts between nose and eyes (Fig. 12) did not really increased recognition rates when compared with the previous two-profile combination, achieving a 77.05% of recogntition success. Nevertheless, in this case, we faced the difficulty of training a SVM classifier using voxel resolution higher than 30×30×30 due to size limitations imposed by SVMTorch software files (for example, a 45×45×45 voxel resolution produced files of about 16.4 Mbytes per face pattern).
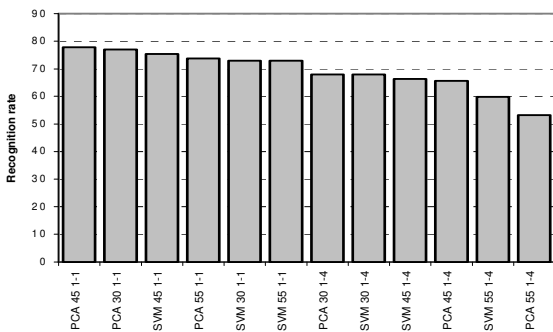

Fig. 11. Best recognition rates for the combination of "eyes and nose" profile.

5.2.3 Depth Map Modeling

Depth map modeling encompasses three surface models: (1) full facial depth map, (2) upper-half facial depth map and (3) left-side facial depth map. Table 7 presents the best recognition rates obtained for each model. Some references of 3D face recognition used PCA and depth maps (Tsalakanidou et al., 2003; Chang et al., 2003) but in these approaches no facial expressions were considered and the capture environment was not flexible.
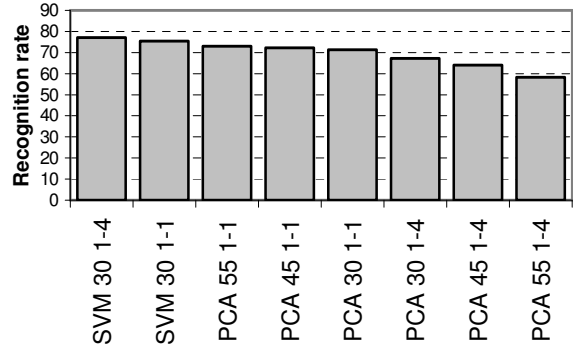

Fig. 12. Correct recognition rate for all the horizontal cuts between nose and eyes profile.

The use of full facial depth maps did not produce good recognition results, achieving a maximum of 58,2% correct recognition rate. In this approach SVM worked much better that PCA as matching scheme. Upper-half facial depth maps produced better results than left-side facial depth maps, because the mouth area is very sensible to facial gestures. In general, SVM worked much better than PCA, and the resolution 1-1 also produced better results than the 1-4 resolution. This lower resolution negatively affected to the recognition results by using the best profiles. These experiments confirm that the proposed multiple profiles provide much better results than depth maps for the same resolution. Other works have proved that depth maps improve these recognition results for higher resolutions. Lee (2005) obtained a 73.77% of recognition rate using depth maps extracted from the 3D meshes of *GavabDB* and the PCA scheme.

Full-face	58,2%
Upper-half	54,1%
Left-side	44,2%

Table 7. Recognition rates using facial depth maps in a non-controlled environment at 55×55×55 and 1-1 voxel and scanning resolutions respectively using SVM.

5.3 Face Recognition Experiments in a Controlled Environment

The goal of this experiment was to test how the variability of faces affects the recognition rate in non-controlled environments compared to controlled environments. In other words, how the proposed 3D facial profile models are robust regarding facial gestures and rotations. For this aim, we assumed a controlled environment in which one of the frontal and neutral images of each subject were part of the training set and the other frontal and neutral images were used for the tests. Because we worked with 3D meshes without texture, a controlled environment did not mean any restriction on the illumination conditions. Tests were run for the profile models that provided better results in the non-controlled environment (i.e. combination of eyes and nose profiles). Table 8 shows in

decreasing order the recognition rates achieved with the proposed profile models for the different combinations of matching schemes, voxel and scanning resolutions.

It can be observed that recognition rates under controlled conditions improved on average a 10% over the non-controlled environment, thus obtaining a best recognition rate of about 90% with the PCA matching schema and an 88% with SVM. This 10% difference between the controlled and the non-controlled environment shows that the combination of "eyes and nose" profile model is robust since it copes with the variability regarding face position and facial gestures.

Matching Scheme	Recognition Rate (%)
PCA 45 1-1	90.16
SVM 45 1-1	88.52
SVM 55 1-1	88.52
SVM 30 1-1	86.89
PCA 55 1-1	86.88
PCA 30 1-1	80.32
SVM 30 1-4	78.69
SVM 45 1-4	75.41
SVM 55 1-4	73.77
PCA 45 1-4	72.13
PCA 30 1-4	65.57

Table 8. Best recognition rates for the combination of "eyes and nose" profile in a controlled environment.

6. Conclusion and Discussion

3D Automatic Face Recognition (AFR) systems have attracted a lot of attention in recent years due to the interesting applications derived from it and to the proliferation and increasing performance of 3D scanners. Such 3D face recognition systems are applicable if they are robust against the variations in position and gestures that a subject can produce while he/she is interacting with the system. Some of the problems and limitations of 3D face recognition systems are: (1) there are not many systems that consider the face variability in position and facial expressions that real applications demand, and (2) the lack of 3D face databases with significant variability in pose and gestures.

This paper has proposed and tested a set of 3D face profile models to represent individuals for recognition purposes. In general, for both matching techniques used, the best results obtained in non-controlled environments were using a combination of profiles (eye and nose), achieving a 77.86% recognition rate. Taking into account the variability in pose and gesture of the images used and the fact that there is just a 10% difference of correct recognition rate when compared to a controlled environment, these results show that the proposed 3D face surface modeling approaches can be suitable for practical non-controlled recognition applications. The worst results were obtained when using the mouth region due to its high variability regarding facial expressions and/or gestures. Other conclusion obtained from this work is that the proposed profile combinations provide better recognition results than the corresponding depth maps at the same resolutions.

We have also shown that this model is, to a large extent, robust against variations in pose and facial gestures. The system was tested with two matching schemes: SVM, and PCA combined with a Euclidean distance classifier. Experimental results demonstrated that while there is not a significant difference in the recognition rates achieved, and that these differences were important regarding training time in favour of SVM. On average, when using the same type of model, training the recognition system on the same computer took about four times less when using SVM.

The results pointed out that some of the proposed multiple-profile models are also robust against variations in illumination, gestures and facial expressions.

In order to solve the lack of databases to test 3D face recognition systems under real-world environments, this paper presents the *GavabDB* as a facial database that captures a high degree of variability in gestures and pose.

In our approach, 3D facial meshes from the same person can have a different facial area surface. This is caused by auto-occlusions produced in the borders of the mesh. We consider that these variations in the images of the same individual deeply affect recognition rates. As future work we plan to implement a 3D face recognition system focused only on a predetermined central area of the face in order to avoid occluded areas and also to ensure that the facial meshes of the same individual have the same size. This will probably improve the recognition rates of the proposed 3D face recognition models.

We are compiling another database of new individuals for testing the system with unknown individuals. In this work we have put the emphasis in the comparison of the discrimination power of different profile-based and depth-map features to be used as descriptors in a 3D face recognition system.

Another future work is the experimentation with distances between profiles and cuts in the depth map as features, instead of using binary data. We think that some proposed models could provide better results for higher cube resolutions if a Bresenham algorithm were applied to connect the 1-voxels in our profile models.

Acknowledgments

References

3DFaceDBYork, 1999
http://www-users.cs.york.ac.uk/~tomh/3DFaceDatabase.htm
Ansari A.N., Abdel-Mottaleb M. and Mahoor M.H., 2007, 3D Face Mesh Modeling from Range Images for 3D Face Recognition, 14th IEEE Intl. Conf. on Image Processing (ICIP'07).

Ben Abdelkader, C., Griffin, P.A., 2005. Comparing and combining depth and texture cues for face recognition. Image and Vision Computing 23, 339-352.

Berretti S., Del Bimbo A., Pala P., 2006, Description and retrieval of 3D face models using iso-geodesic stripes, Multimedia Information Retrieval (MIR'06), pp.13-22.

Berretti S., Del Bimbo A., Pala P., 2007, 3D Face Recognition using iso-Geodesic Surfaces, IRCDL 2007, pp. 111-116.

Beumier Database, 2000.
http://www.sic.rma.ac.be/~beumier/DB/3d_rma.htm

Beumier, C., Acheroy, M., 2000a. Automatic 3D face authentication. Image and Vision Computing 18, 315-321.

Beumier, C., Acheroy, M., 2000b. Automatic Face Verification from 3D and Grey Level Clues. Portuguese Conference on Pattern Recognition RECPAD, 95-101.

Beumier, C., Acheroy, C., 2001, Face verification from 3D and grey level clues, Pattern Recogntion Letters, 22 (12), pp. 1321-1329.

Boser, B.E., Guyon, I.M., Vapnik, V.N., 1992. A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory, 144-152.

Bowyer, K. W., 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 101, 1-15.

Bronstein, A.M., Bronstein, M.M., Kimmel, R., 2003. Expression-Invariant 3D Face Recognition. AVBPA, LNCS 2688, Spinger-Verlag, 62-70.

Bynm, H., 2003. A survey on pattern recognition applications of Support Vector Machines. Intl. Journal of Pattern Recognition and Artificial Intelligence (17) 3, 459-486.

Campbell, R.J., Flynn, P.J., 2001. A Survey of Free-Form Object Representation and Recognition Techniques. Computer Vision and Image Understanding 81, 166-210.

Cartoux, J. Y., LaPreste, J. T. and Richetin, M., 1989, Face authentication or recognition by profile extraction from range images, Proceedings of the Workshop on Interpretation of 3D Scenes, pp. 194-199.

Chang, K.I., Bowyer, K.W., Flynn, P.J., 2003. Face Recognition Using 2D and 3D Facial Data. Workshop on Multimodal User Authentication (MMUA).

Chang, K.I., Bowyer, K.W., Flynn, P.J., 2005. An evaluation of multimodal 2D+3D face biometrics. IEEE Trans. on Pattern Analysis and Machine Intelligence 27, 619-624.

Chua, C.S., Han, F., Ho, Y.K., 2000. 3D Human Face Recognition Using Point Signature. Fourth IEEE Intl. Conf. on Automatic Face and Gesture Recognition.

Collobert, R., Bengio, S., 2002. SVMTorch: Support Vector Machines for Large-Scale Regression Problems. Journal of Machine Learning Research (1), pp. 143-160.

Cristianini, N., Shawe-Taylor, J., 2000. An Introduction to Support Vector Machines. Cambridge University Press.

Duda, R.O., Hart, P.E., Stork, D.G., 2000. Pattern Classification. 2nd Edition. Wiley-Interscience.

Godil, A., Ressler, S., Grother, P., 2004. Face recognition using 3D facial shape and color map information: Comparison and combination. SPIE Biometric Technology for Human Identification Conference 5404, pp. 351-361.

Gordon, G., 1991. Face recognition based on depth maps and surface curvature. SPIE Proc.: Geometric Methods in Computer Vision 1570, 234-247.

Gray, A., 1999. Modern Differential Geometry of Curves and Surfaces with Mathematica, Second edition, CRC Press.

Haar, Frank B. and Veltkamp, Remco C., 2008, A 3D Face Matching Framework, Institute of Information and Computing Sciences, Utrecht University, Technical Report UU-CS-2007-047, www.cs.uu.nl.

Hallinan, P., Gordon, G., Yuille, A.L., Giblin, P., Mumford, D., 1999. Two-and Three-dimensional patterns of the face. Ed. A. K. Peters.

Hsu, R.-L., Jain, A.K., 2001. Face modeling for recognition. Proc. Intl. Conf. on Image Processing, 2, 693-696.

Yuxiao Hu, Zhenqiu Zhang, Xun Xu, Yun Fu and Thomas S. Huang, 2007. Building Large Scale 3D Face Database for Face Analysis, Intl. Workshop on Multimedia Content Analysis and Mining, LNCS 4577, Springer, 343-350.

Huang, J., Heisele, B., Blanz, V., 2003a. Component-based face recognition with 3D morphable models. 4th Conf. on Audio and Video based Biometric Person Autenthication, 62-69.

Huang, J., Blanz, V., Heisele, T., 2003b. Face Recognition with Support Vector Machines and 3D Head Models. International Workshop on Pattern Recognition with Support Vector Machines (SVM2002), 334-341.

Huang, T., Tang, L., 1996. 3-D Face Modeling and Its Applications. International Journal Pattern Recognition and Artificial Intelligence 10 (5), 491-520.

Konica Minolta. http://konicaminolta.com

Lee, M. W., Ranganath, S., 2003. Pose-invariant face recognition using a 3D deformable model. Pattern Recognition 36, 1835-1846.

Lee, Joonsoo, 2005. 3D Face Recognition Using Range Images, Final Report, Multidimensional DSP, Spring 05.

Lei, Y.W., Wu, J.L., Ouhyoung, M., 1996. A three-dimensional muscle-based facial expression synthesizer for model-based image coding. Signal Processing: Image Communication 8, 353-363.

Li, Chao and Barreto, Armando, 2004, Profile-based 3D Face Registration and Recognition, Proceedings of 7th International Conference on Information, Security and Cryptology (ICISC), pp. 456-467.

Li C., Barreto, A., Jing, Z. and Chin, C., 2005, Exploring Face Recognition by Combining 3D Profiles and Contours, IEEE, pp. 576-579.

Li X. and Zhang H., 2007. Adapting Geometric Attributes for Expression-Invariant 3D Face Recognition, IEEE Intl. Conf on Shape Modeling and Applications (SMI'07).

Ma, B., Wong, H.-S., 2006. 2D clustering based discriminant analysis for 3D head model classification. Pattern Recogniton 39, 491-494.

Mahoor, M. H. and Abdel-Mottaleb M., 2007, 3D Face Recognition based on 3D ridge lines in range data, 14th IEEE Intl. Conf. on Image Processing (ICIP'07).

Matas, J., Hamouz, M., .Jonsson, K., Kittler, J., Li, Y., Kotropoulos, C., Tefas, A., 2000. Comparison of face verification results on the XM2VTS database. Proceedings of the 15th International Conference on Pattern Recognition (4), 858-863.

Messer, K., Matas, J., Kittler, J., Luettin, J., Maitre, G., 1999. Xm2vtsdb: The extended m2vts database. Second International Conference on Audio and Video-based Biometric Person Authentication.

Moreno A. B, Ángel Sánchez and José F. Vélez, 2002, Robust Normalization of 3D Facial Meshes using Automatically Located Feature Points, 12th Portuguese Conference on Pattern Recognition, Aveiro (Portugal), June 27-28.

Moreno, A.B., Sánchez, A., 2004. GavabDB: a 3D Face Database. Workshop on Biometrics on the Internet COST275, 77-85.

Moreno, A.B., Ángel Sánchez and Enrique Frías Martínez, 2006, Robust Representation of 3D Faces for Recognition,

International Journal of Pattern Recognition and Artificial Intelligence (20) 8, 1159-1186.

Nagamine, T., Uemura, T. and Masuda, I., 1992, 3D facial image analysis for human identification, Proceedings of Intl. Conf. pm Pattern Recognition (ICPR), pp. 324-327.

Notre Dame, 2003. http://www.nd.edu/~cvrl/

Pan, G., Wu, Y. and Wu, Z., 2003, Investigating Profile Extraction from Range Data for 3D Face Recognition, Proc. of IEEE Intl. Conf. on Systems Man and Cybernetics, pp. 1396-1399.

Phillips, P. J., et al., 2005. Overview of the face recognition grand challenge. Computer Vision and Pattern Recognition (CVPR), 1, 947-954.

Romdhani, S., 1996, Face Recognition Using Principal Components Analysis http://www.vision.im.usp.br/~teo/pca.

Stein, F., Medioni, G., 1992. Structural indexing: Efficient 3-D object recognition. IEEE Trans. Pattern Anal. Mach. Intell. 14, 125-145.

Tanaka, H.T., Ikeda, M., Chiaki, H., 1998. Curvature-based face surface recognition using spherical correlation. Proc. 3rd IEEE Intl. Conf. on Automatic Face and Gesture Recognition, 372-377.

Thirion, J.P., 1996. The extremal mesh and the understanding of 3D surfaces. Intl. J. Comput. Vision 19, 115-128.

Trucco E. and Verri, A., 1998, "Introductory Techniques for 3-D Computer Vision", Prentice-Hall.

Tsalakanidou, F., Tzovaras, D., Strintzis, M.G., 2003. Use of depth and colour eigenfaces for face recognition. Pattern Recognition Letters 24, 1427-1435.

Vapnik,V.,1998. Stratistical Learning Theory, John Wiley and Sons, New York.

Veltkamp, R. C. and Haar, 2007, F. B. SHREC2007: 3D Shape Retrieval Contest, Technical Report UU-CS-2007-015, Utrecht University.

Vetter, T., 1999. Flexible Models of Human Faces for the Analysis and Synthesis of Images. Handbook of Computer Vision and Applications, 3, Systems and Applications, Academic Press.

Wang, Y., Chua, C.S., Ho, Y.K., 2002. Facial feature detection and face recognition from 2D and 3D images. Pattern Recognition Letters 23, 1191-1202.

Wechsler, H., 2007. Reliable Face Recognition Methods: System Design, Implementation and Evaluation, Springer.

Xie, X., Lam, K., 2005. Face recognition under varying illumination based on a 2D face shape model. Pattern Recognition 38, 221-230.

Xm2vtsdb,1999, http://www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb/

Zhao, W., 1999. Improving the Robustness of Face Recognition. Proc. Intl. Conf. on Audio and Video Based Person Authentication, 78-83.

Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A., 2003. Face Recognition: A Literature Survey. ACM Computing Surveys 35 (4), December 2003, 399-458.