ROBUST REPRESENTATION OF 3D FACES FOR RECOGNITION

A. BELÉN MORENO, ÁNGEL SÁNCHEZ, ENRIQUE FRÍAS-MARTÍNEZ

Artificial Vision and Biometry Group - GAVAB
Escuela Superior de Ciencias Experimentales y Tecnología
Universidad Rey Juan Carlos
C/ Tulipán s/n 28933, Móstoles, Madrid, Spain
{belen.moreno,angel.sanchez,efrias}@urjc.es

Automatic face recognition is becoming increasingly important due to the security applications derived from it. Although the face recognition problem has focused on 2D images, recently, due to the proliferation of 3D scanning hardware, 3D face recognition has become a feasible application. This 3D approach do not use the colour of the faces, so, when it is compared to the more traditional 2D approach, has the main advantages of being robust under lighting variations and also of providing more relevant information. In this paper we present a new 3D face model based on curvature properties of the face surface. As will be presented, our system is able to detect the subset of characteristics of the face with higher discrimination power among a large set of them. The robustness of the model is tested by comparing recognition rates using both controlled and non-controlled environments regarding facial expressions and face rotations. The difference between the recognition rate of the two environments of only 5% proves that the model has a high degree of robustness against pose and facial expressions and that is robust enough to implement facial recognition applications, which can achieve up to 91% correct recognition rate.

Keywords: 3D Face Modelling; 3D Face Recognition; SVM; PCA.

1. Introduction

A biometric system is a pattern recognition system that determines the authenticity of an individual using some physical or behavioural features (fingerprints, face images, signature, voice, gait, etc.)[64]. The social demand for security and fraud control applications, in which it is necessary to establish personal authentication, has caused the need for increased research in such systems[9,42]. Within this trend, automatic face recognition has become one of the biometric systems where more attention has been focussed in recent years. It has many applications in areas like: personal identification, security applications, law enforcements and smart environments, among others. Of all the biometric modalities used for human recognition and verification, face recognition is one of the less intrusive[42], which has increased the interest and applications of these systems.

Traditionally, research in face recognition has focused on 2D intensity images. Under this context, the recognition accuracy is sensitive to the lighting conditions (position and kind of light), expressions, head orientation (rotation), and/or a variety of elements such as hair, moustache, glasses, makeup, etc. Among the proposed methods on 2D face recognition are those based on some local geometric features, such as shape of the eyes, mouth, as well as the relationship among them[11,32]. Another classic

method using geometric information is the template matching method[11,20]. Some of the most successful algorithms for 2D face recognition use texture information provided from 2D texture images, such as *eigenfaces*[38,60], *fisherfaces*[2], and *elastic graph matching*[65]. Automatic Face Recognition using 2D images provides excellent results when the image acquisition conditions (illumination, pose and face variations) are controlled[42]. Recent efforts are oriented to reduce image acquisition restrictions[43,66]. Some methods have been proposed to tackle non-controlled variations of pose and illumination, but they do not work well in arbitrary conditions[42].

Working with 3D face images has some advantages over 2D face images: (1) more geometrical information can be obtained from the 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 (as distances, angles, depth, areas,…) 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 use texture (colour) information, as in the case of this work, the recognition is immune to the effect of illumination variations. The advances in computational processing capacities as well as the reduction of size of the 3D digitizers, have contributed to their proliferation and to the development 3D face recognition systems.

In the last few years, the interest in 3D object recognition systems in general, and in 3D face recognition systems in particular, has increased[33,36,39]. One common technique using 2D images is based on the correspondence among 2D image points and their corresponding of a generic 3D object model in order to obtain a 3D particularized object model for the object location, pose determination and/or feature extraction[15]. Among the 3D free-form object descriptors to represent objects is the local curvature. Hallinan et al.[31] experimented a set of twelve 3D features extracted from segmented regions using curvature properties of the surface which were also used for face recognition using a small set of 8 individuals.

Two aspects that characterize a face recognition system are: (1) representation or face modelling and (2) recognition (or matching) technique. Face modelling transforms the information of the facial image into a compact and discriminating representation and matching scheme, involves the method that is used to select the best match from the set of identities of the face database. While the matching scheme can be efficiently implemented using standard machine learning techniques like Neural Networks[30] or Support Vector Machines[29], 3D Face Modelling is an open problem, specially in the areas of 3D face modelling[9,25,68]. In this paper, we develop a novel, robust and practical representation of 3D face images obtained from next stages: segmentation (using curvature information), feature extraction and analysis of the discriminating power. In order to check the robustness of the model designed, face recognition experiments have been run using (1) Support Vector Machines (SVM) and (2) Principal Components Analysis (PCA) in combination with a Euclidean distance classifier. Also, due to the lack of representative 3D face databases that present a high degree of variability among the images of each individual, specially related to facial expressions, and in order to test our face model using a real environment, we have proposed and developed a new 3D face database named *GavabDB*. This database, described in Moreno et al[47], can be found in *http://gavab.escet.urjc.es/*.

The paper is organized as follows: first, the state of the art of 3D face databases and 3D face modelling are presented. The next section introduces our face model using HK segmentation (based on local curvature properties). The fourth section studies the minimum set of features that better characterize each face. Section five, presents the design of a 3D face database that captures the variability needed to test face recognition applications in real environments, and checks the robustness of our face model by implementing face recognition system using SVM and PCA (in combination with a Euclidean classifier) as matching schemes. Finally, conclusive remarks and future works are given in section six.

2. State of the Art: 3D Face Databases and 3D Face Modelling

The following two subsections present the state of the art of 3D face databases and 3D face modelling focusing on their characteristics and main limitations.

2.1 3D Face Databases

The development of automatic face recognition systems needs from face image databases in order to train and evaluate the systems. There are some published face image databases orientated to different experimentation purposes: automatic face recognition, facial expression analysis, pose estimation, face detection, etc. Most of the databases only offer 2D face images. Some of the 2D facial image databases described in the literature are FERET database,[59,60] Yale DB database,[2] AR database,[44] MIT database,[60] ORL database (Olivetti),[54] CVL database,[22] PF01 database (Postech Faces '01),[37] and CMU database[55]. The number of 3D face databases available is very limited and, in general, they tend to be designed for specific aspects of recognition. To our knowledge, some of the most representative 3D face databases are:

• XM2VTS database:[45] It is a large multimodal database, specifically designed for testing multimodal verification systems. It includes a 3D face model for each individual.

• 3D_RMA database:[4] It contains 120 individuals captured in two different sessions, separated by 2 months. Digitisations consisted in 3 shots grabbing different and limited orientations of the head (frontal, at left or right and at up or down).

• York University 3D face database:[1] This database has images corresponding to 97 individuals. It contains 10 captures per individual including different poses. However, only 2 of the ten images of each individual present light facial expressions (happiness and frown) and one presents face occlusion.

• Notre Dame 3D database:[16] it contains 951 3D range images and its corresponding 2D images from 275 individuals without any kind of facial expression.

• CAESAR anthropometric database[13,25]: It contains 3D surface grid and colour map information of the full body (without facial expressions).

The study of the state of the art of 3D databases reveals that in general, there are not many varied changes among the different images of each individual. Some of databases offer variations related to some aspects but not to others, but in general, the

range of variation is very limited. Among the databases presented, just one captures images with some light facial expression. A common characteristic of the databases is that the scanning process of the images is usually done under strict rules in order to capture the image of the face as needed. Having databases with a high variability in the pose and in the facial expression is very important in order to test the robustness of a 3D face model and the efficiency of a recognition system in a real environment situation. Also, the strict scanning process used for creating the databases is not desirable to test 3D models under real environments where users will usually have some degree of freedom. For all these reasons, we felt that there was a need for a 3D facial database with a representative number of individuals and with a variety of pronounced facial expressions and different poses.

2.2 State of the Art: 3D Face Modelling

A current trend in face recognition is to use 3D face models to analyze if the 3D information improves the 2D face recognition rate and in this case, to quantify the increment they provide[68],[25]. 3D face recognition systems solve some of the existing problems of the 2D approaches, mainly illumination, pose and face variations dependence[10]. Although intuitively 3D information should provide better results that 2D, it has not still been proven with any rigorous study[18,68]. So far, research in 3D face recognition system has focussed on looking for descriptors that model the face (3D face modelling) and on image acquisition techniques[7], rather than in evaluation of these systems. The main reasons are the high cost of 3D digitizers and consequently the lack of 3D face databases which permit the evaluation of this kind of techniques. There are relatively few works about face recognition using 3D models,[10,21,33,39] and there are few works which present recognition rates evaluated using a great number of images.[8,68] In this section we present a review of the main approaches for 3D face modelling: local feature face modelling and holistic face modelling.

2.2.1 Local Feature-Based 3D Face Modelling

Differential geometry has been used for feature extraction in the context of free-form three-dimensional object recognition[27] and also used by some authors for facial feature extraction.[56] The local curvature of a surface 3D point, as well as the angle among surface normal vectors evaluated in a 3D point of the surface have been proposed as 3D free-form object descriptors[15,56,57] in the scope of 3D face recognition. Another technique of representation of 3D range face images using local information is the one based in computing the point signature over determinate 3D points in order to obtain descriptors. The point signature has been used to represent 3D free-form object, in particular 3D face images.[21] The rest of this section presents some representative techniques of local feature-based 3D face modelling.

Segmentation based in a point classification according to planar, spherical or revolution surface was proposed by Gordon in Ref 27. In this approach, a region classification according to concavity, convexity and saddle points was achieved to localize

nose, eyes and mouth. Neck, front and cheeks were also localized according to their smooth property. These features were used to normalize the images, scaling and posing them into a cylindrical mesh such that the volume between face and cylinder characterized the face for recognition. The proposed 3D face recognition system provided excellent recognition rates using local features and also using the whole face, but using a small set of images (24) corresponding to a small number of individuals (8) for the test experiments.

The local surface curvature evaluated in a point is characterized by the directions in which the normal of the surface changes more and less quickly. In Hallinan et al,[31] a set of twelve 3D descriptors extracted from segmented regions using curvature properties of the surface were experimented for face recognition over a database of 8 individual and 3 images per individual obtaining a high recognition rate (95,5%).

Some other techniques use 3D geometry data gather to the 2D intensity images captured simultaneously by the same digitizer.[5,63] A face verification system is proposed in Ref. 4 (using the 3D images captured by the active triangulation technique described in Ref. 7) which uses grey level information and two facial profile curves as face representation.

A representation based on local features, which gathers both the 3D shape information and the 2D texture information, is analyzed and described by Wang et al.[63] Features points are selected by Gabor filter responses in the 2D domain, and the signatures of these points over the same points in the 3D domain. Both of them are projected in their corresponding subspace using PCA. The obtained weight vectors are integrated in an augmented vector used to represent the face image in the recognition system. The identification system based in Support Vector Machines provides a 90% recognition rate using a database of 50 individuals presenting different facial expressions and poses.

2.2.2 Holistic 3D Face Modelling

This section describes some 3D face recognition techniques which employ a complete description of the 3D face surface instead of characteristics extracted from local features to represent the face in the recognition system. One common holistic technique on 3D object recognition is based on the correspondence among scene points and model points in order to determinate the object pose and location and to extract features to perform the recognition[35]. The use of a generic 3D head model[36] has some applications like analysis and synthesis of facial expressions[41], illumination or pose correction in a face recognition system[39], descriptor extraction[40] and generation of synthetic face variations (i.e. illumination changes, rotations...) to augment the number of training views of each individual[62,34] (to increase robustness when there exist such variations).

The use of a generic 3D head model lets to obtain 3D information from 2D images. This model contains control points that can be displaced to adjust themselves to their corresponding points of the 2D images of the individual. These displacements deform the generic model generating a particularized model of the individual of the 2D images. The deformations are expressed in terms of modifications of certain parameters

which characterize the three dimensional global shape of the individual as well as local information. This technique avoids the 3D image acquisition stage of the test images. The need of searching 2D points in the images is the great disadvantage of this technique, be cause of the dependence of the acquisition conditions.

Some recent approaches[18,59] use principal component analysis (PCA) to obtain a low dimensional representation of 3D images given by depth maps of the complete face. Their objective is to evaluate the influence of colour, depth and the combination of both in face recognition. They compared the use of 2D images, and their corresponding 3D depth maps separately and combined concluding that the combination of 2D and 3D information provided better result. These systems have not considered images presenting facial expressions. Pose variations experimented have been very limited[59] or they have not been considered[18]. Ref. 3 demonstrates that depth and texture play complementary roles in the coding of faces. In Godil et al.[25], it is demonstrated that for their data set, the recognition performance is not very different between 3D surface and colour map information using a PCA algorithm. It is also showed in this paper that using fusion approaches of both patterns there is significant improvement in the results.

Another work related with PCA using 3D information is[10], which proposes a representation based in geometrical invariants. This work obtains a canonical invariant under isometric deformations (bending-invariant canonical form), adding image texture to the canonical surface in the canonical form space and to apply PCA to the obtained representations (eigenforms). No recognition results were given in this work.

3. 3D Face Modelling Based on HK Segmentation

Our work is focussed on searching new 3D face modelling methodologies which improve the robustness of automatic face recognition systems. The objectives are to free the systems of some image restrictions of the image acquisition (illumination, pose, make-up, shades, facial expressions, etc.), as well as to look for 3D information extractable from the face images which provides more discriminating power among individuals. Our 3D face modelling system has three basic stages: (1) HK segmentation of regions and lines of interest, (2) feature extraction from the segmented regions and lines and (3) feature selection.

3.1 *HK Segmentation*

The HK algorithm[58] can be used to classify 3D points of facial meshes assigning to each point of the mesh a label describing the local shape of the surface using dictionary of shape classes. The algorithm, based on differential geometry properties, labels each point of 3D facial meshes as concave elliptical, convex elliptical, hyperbolic, concave cylindrical, convex cylindrical or planar. The classification is obtained using the signs of the median (H) and Gaussian (K) curvatures as criteria.

3.1.1 *Shape Classification*

Two curvature values are defined over each point of the 3D face surface. The mean curvature (*H*) of a 3D point is defined as the average of the principal curvatures of the 3D surface in that point and the Gaussian curvature (*K*) can be obtained as the product of the principal curvatures. These values, *H* and *K*, can then be used to identify the form of the curvature in a given point, considering that concavity and convexity are defined from the 3D scanner point of view. Table 1 presents the classification dictionary used to label each point of the 3D facial mesh. For example, in a cylindrical point, one of the two main curvatures is zero; in an elliptic point, the main curvatures have the same sign; in hyperbolic points, the main curvatures are non-zero and have different signs, being the local surface like a saddle. The previous classification is qualitative in the sense that the label assigned to each point depends only on the sign of the main curvatures and not on its absolute value. This increases the robustness of the classification, because the sign can be correctly classified even though some noise may have affected the original data.

Table 1. Dictionary of shapes classified according to the signs of the *H* and *K* curvatures.

	Hyperbolic	**Convex cylindrical**	**Convex elliptical**	**Planar**	**Concave cylindrical**	**Concave elliptical**
Shape						
Sign(*H*)	+/0/-	-	-	0	+	+
Sign(*K*)	-	0	+		0	+

3.1.2 *Obtaining H and K curvature values*

Given $P=f(u_0,v_0)$ a point of surface *f*, and $\boldsymbol{x}_i(P)$ the partial derivative of surface *f* with respect to *i* in *P*, the partial derivatives with respect to the main dimensions *x*, *y* and *z* can be obtained as:

$$\boldsymbol{x}_1(\boldsymbol{P}) = \left.\frac{\partial \boldsymbol{f}}{\partial \mathrm{x}}\right|_{(\mathrm{x}_0,\mathrm{y}_0)} = \frac{\boldsymbol{f}(\mathrm{x}_1,\mathrm{y}_1) - \boldsymbol{f}(\mathrm{x}_0,\mathrm{y}_0)}{\mathrm{x}_1 - \mathrm{x}_0} \tag{1}$$

$$\boldsymbol{x}_2(\boldsymbol{P}) = \left.\frac{\partial \boldsymbol{f}}{\partial \mathrm{y}}\right|_{(\mathrm{x}_0,\mathrm{y}_0)} = \frac{\boldsymbol{f}(\mathrm{x}_1,\mathrm{y}_1) - \boldsymbol{f}(\mathrm{x}_0,\mathrm{y}_0)}{\mathrm{y}_1 - \mathrm{y}_0} \tag{2}$$

$$\boldsymbol{x}_3(\boldsymbol{P}) = \frac{\boldsymbol{x}_1(\boldsymbol{P}) \times \boldsymbol{x}_2(\boldsymbol{P})}{\left\| \boldsymbol{x}_1(\boldsymbol{P}) \times \boldsymbol{x}_2(\boldsymbol{P}) \right\|} \tag{3}$$

Defining $\boldsymbol{x_{ij}(P)}$ as the second partial derivative of surface $\boldsymbol{f}$ with respect of the variable $\boldsymbol{j}$ of the first partial derivative $\boldsymbol{x_i}$ in $\boldsymbol{P}$, they can be obtained as:

$$\boldsymbol{x}_{11}(\boldsymbol{P}) = \left.\frac{\partial \boldsymbol{x}_1}{\partial \mathrm{x}}\right|_{(\mathrm{x}_0,\mathrm{y}_0)} \tag{4}$$

$$\boldsymbol{x}_{12}(\boldsymbol{P}) = \left.\frac{\partial \boldsymbol{x}_1}{\partial \mathrm{y}}\right|_{(\mathrm{x}_0,\mathrm{y}_0)} = \left.\frac{\partial \boldsymbol{x}_2}{\partial \mathrm{x}}\right|_{(\mathrm{x}_0,\mathrm{y}_0)} = \boldsymbol{x}_{21}(\boldsymbol{P}) \tag{5}$$

$$\boldsymbol{x}_{22}(\boldsymbol{P}) = \left.\frac{\partial \boldsymbol{x}_2}{\partial \mathrm{y}}\right|_{(\mathrm{x}_0,\mathrm{y}_0)} \tag{6}$$

where $\boldsymbol{x_1}$, $\boldsymbol{x_2}$ are the first derivative of the surface along x and y directions respectively, and $\boldsymbol{x_{11}}$, $\boldsymbol{x_{12}}$, $\boldsymbol{x_{22}}$ represent the corresponding second derivatives. The first and second fundamental forms of surface $\boldsymbol{f}$ in $\boldsymbol{P}$ are defined as a two *2x2* matrix. The first fundamental form $\boldsymbol{G}$ is defined as:

$$\boldsymbol{G} = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}, \tag{7}$$

where each element of the matrix is obtained using the scalar product:

$$\mathrm{G}_{\mathrm{ij}}(\boldsymbol{P}) = \left\langle \boldsymbol{x_i}(\boldsymbol{P}), \boldsymbol{x_j}(\boldsymbol{P}) \right\rangle. \tag{8}$$

The second fundamental form $\boldsymbol{L}$ is defined as:

$$\boldsymbol{L} = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}, \tag{9}$$

where each element of the matrix is obtained using the scalar product:

$$L_{ij}(\boldsymbol{P}) = \left\langle \boldsymbol{x_3}(\boldsymbol{P}), \boldsymbol{x_{ij}}(\boldsymbol{P}) \right\rangle. \tag{10}$$

The first and second fundamental forms of surface $\boldsymbol{f}$ can be used to obtain the following *2x2* matrix:

$$(dx_{3ij})_{\boldsymbol{P}} = \left(L_{ij}(\boldsymbol{P})\right)\left(G_{ij}(\boldsymbol{P})\right)^{-1}, \tag{11}$$

which can be used to obtain the Mean ($H_f(\boldsymbol{P})$) and Gaussian ($K_f(\boldsymbol{P})$) curvatures of surface $\boldsymbol{f}$ in $\boldsymbol{P}$ as:

$$H_f(\boldsymbol{P}) = -\frac{1}{2} tr(\boldsymbol{dx_3})_P, \tag{12}$$

$$K_f(\boldsymbol{P}) = \det(\boldsymbol{dx_3})_p = K_f(\boldsymbol{P}) = \frac{L_{12}^2 - L_{11}L_{22}}{G_{12}^2 - G_{11}G_{22}}, \tag{13}$$

where *tr* is the trace of a matrix and *det* is the determinant.

A problem of the 3D digitizer is the noise introduced in the images, as we can see in Figure 7(b). Mean and Gaussian curvatures can be locally affected by the noise of the 3D image, so it is very important to have a pre-processing stage that reduces the effect of noise. These problems have been satisfactorily solved trough a pre-processing stage applied over each image in which a median filter is applied to reduce impulse noise and a smoothing process implemented with a mean filtering, is applied

to attenuate possible local changes of the curvature signs. Also, the structure of 3D face meshes, which are formed by cells of four nodes, provides an easy mechanism to move in a horizontal and vertical direction within the mesh, which allows to compute approximate derivatives more efficiently in terms of time.

3.2 *Identification of Relevant Face Regions*

An example of H and K curvature signs obtained from a 3D face are presented in Figure 1 (a) and (b) respectively, where black points have a negative sign and grey points have a positive sign. The values of H and K can not be obtained for the border points of the mesh because it is not possible to obtain the second derivative on these points. Figure 1 (c) shows the results of applying the classification dictionary of Table 1 to each point of the surface that has both H and K values, where black represents hyperbolic points, dark grey convex elliptic and light grey concave elliptic points. Only three types of points are obtained because it is virtually impossible to obtain a value of *0* for H and K when working with real numbers. This fact causes that no concave cylindrical, convex cylindrical or planar points are obtained. Nevertheless these classes of points can be obtained when curvature thresholds are applied.

In order to isolate regions of high curvature and to obtain a greater variety of classification points, different curvature thresholds were experimentally tested. The goal of these thresholds is to isolate regions of interest that can be used to identify a face. The regions of interest obtained should comply a set of criteria: (1) should not change with different facial expressions, (2) these regions should be present in most of the images, (3) the regions should be easily detected and (4) the regions should be outside of areas susceptible of containing hear (moustache, beard, etc.). The first three requisites basically define facial regions with high curvature values, in which the skull is near the skin, so they are not affected by facial expression and can be easily identified. The fourth one restricts the face areas to those ones outside the mandible and the mouth.

After experimentally analyzing different thresholds values, the values selected were: H_t=±0.05 and K_t=±0.0005 for median and Gaussian curvatures, respectively.

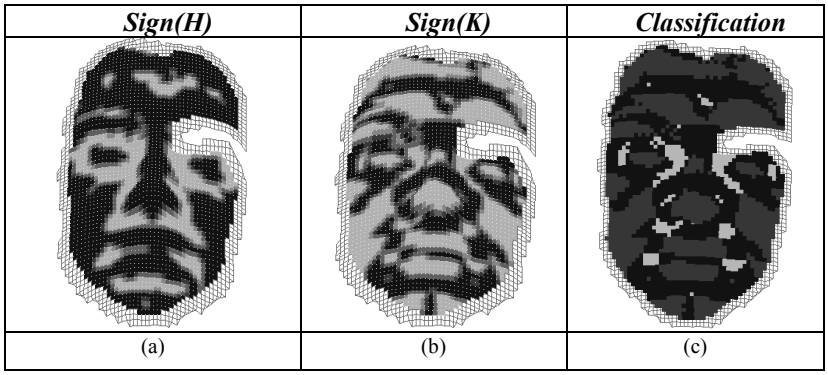


Fig 1. (a) Sign of H and (b) K curvatures of the points of the 3D faces corresponding to one subject, with the black points having negative sign and the grey points a positive sign, and (c) classification of the points using the classification dictionary, with black identifying hyperbolic points, dark grey convex elliptic and light grey concave elliptic.

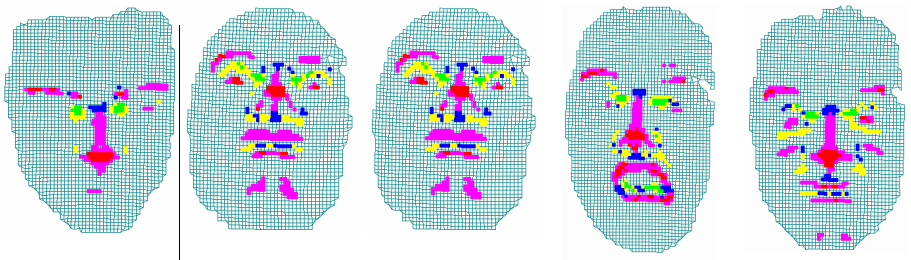


Fig. 2. Example of detected regions for some of the 3D images (looking down, looking up, neutral, random and smile gestures from left to right, respectively) for one subject after thresholding and classification stages.

These thresholds were used to assign a value of 0 to any value included within the corresponding intervals, in our case $-0.05<H_t<0.05$ and $-0.0005<K_t<0.0005$. After this process the same regions appear in the face surfaces. Figure 2 presents an example of those areas for some of the seven images taken from two individuals. As can be seen the same areas are highlighted by the process. Also planar, concave cylindrical and convex cylindrical points appear, increasing the richness of information.

From all the regions that appear after the application of the thresholds, taking into account the previous curvature criteria, seven regions have been selected:

- Region 1: Region contains the points of the tip of the nose. These points are classified as convex elliptical after the threshold stage.

- **Region 2**: Region composed by the frontal points of the nose. These points are classified as convex cylindrical after the threshold stage.

- Region 3: Region composed by the points situated in the upper bridge of the nose classified as hyperbolic after the threshold stage.

- Region 4: Region composed by concave elliptical points of the left eye cavity after applying the threshold stage.

- Region 5: Region composed by the concave elliptical points of the right eye cavity after applying the threshold stage.

- Region 6: Region composed by the convex elliptical points of the left eye.

- Region 7: Region composed by the convex elliptical points of the right eye.

Also two types of facial lines have been extracted:

- Line 1: Composed by points in which the sign of *H* curvature changes in the right side of nose.

- Line 2: Composed by the points in which the sign of *H* changes in the left side of the nose.

Figure 3 presents an example of the regions and lines selected. Five of those regions are obtained after applying thresholds, two before applying thresholds and the two lines are obtained from the sign of *H*. Most of these regions are from parts of the face where the skull and the face surface are separated by a thin layer of skin. This basically causes that this areas are almost not affected by facial expressions and aging, which make them ideal to characterize an individual. Notice that the selected regions are not areas with possibility of be occluded by facial hair.

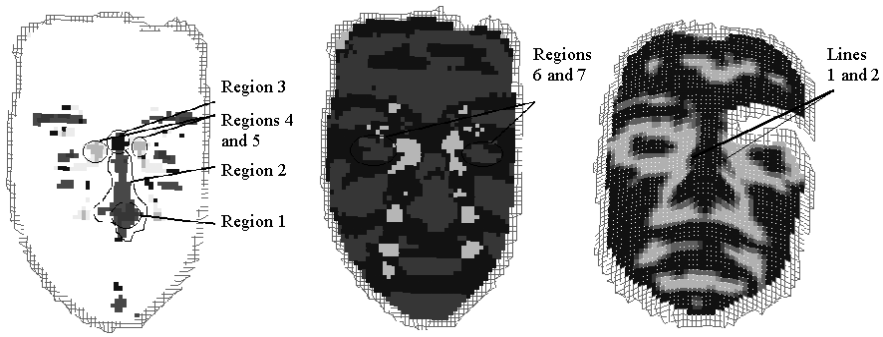


Fig. 3. Regions and lines selected after the classification and the threshold application stage: (a) regions obtained after the threshold stage. (b) Regions obtained before applying the threshold process, (c) the lines obtained from the sign of *H* curvature.

3.3 *Automatic Face Segmentation*

In order to automatically identify each one of the relevant regions and lines identified on the previous section, we propose an algorithm that first starts by identifying Region 1, which is the easiest to obtain, and, once this region has been located within the mesh, it is used to sequentially locate the other regions. The method is based in the following steps:
1) Region 1 is easily distinguishable because is the region which has the highest number of convex elliptic nodes in the threshold image.
2) Region 2 is formed by the convex cylindrical points. This region is neighbour to Region 1.
3) The points that constitute Region 3 are hyperbolic and they are located in a neighbour region of Region 2. Further relative position restrictions with respect to the eye regions can be used if more than one candidate region were obtained.
4) Regions 4 and 5 are formed by concave elliptic points. There are few candidate regions to them, and the selection is performed according to their relative positions with respect to the nose regions. 5) Regions 6 and 7 are formed by convex elliptic points in the non-threshold image, by eliminating candidate regions that not hold certain restrictions. Such restrictions are: their relative positions with respect to regions 3, 4 and 5 (for example, their points had to be found in a band containing these three regions) and height (major axis 1.7 mm. maximum).
6) Lines 1 and 2 are looked for from the external border of Region 2 at left and right directions respectively until a sign change of the median curvature is found.

3.4 *Feature Extraction*

The goal of this section is to extract a relevant set of features from the automatically segmented regions in order to characterize and model each 3D mesh. Using the automatically segmented areas, a total of eighty six non-independent features were defined. Table 2 describes and classifies these features in thirteen categories: (1) areas of the regions, (2) area relations, (3) mean of areas, (4) distance between mass centres of regions, (5) relations of distances between mass centres of regions, (6) mean of distances, (7) angles among mass centres of regions, (8) mean of angles, (9) mean of H of the points belonging to a region, (10) mean of K of the points belonging to a region, (11) variance of H of the points belonging to a region, (12) variance of K evaluated in the points belonging to a region, (13) lines 1 and 2 based features. Table 2 also includes the mathematical definition of each one of these features using the data provided by the segmented regions. Each feature has an associated identification number and a ranking position. This position corresponds to the ordered discriminating power of the feature (a smaller value in the ranking corresponds to a higher discrimination rate).

The discrimination power of each feature Φ has been computed using the Fisher coefficient[31], which represents the ratio of between-class variance to within-class variance, according to the formula:

$$CF_{\Phi} = \frac{\sum_{i=1}^{c} (m_i - m)^2}{\sum_{i=1}^{c} \frac{1}{n_i} \sum_{x \in \Phi_i} (x - m_i)^2} \tag{14}$$

where CF_{Φ} is the Fisher coefficient of the Φ descriptor, c is the number of classes (individuals in our case), Φ_i is the set of feature values for class i, n_i is the size of Φ_i, m_i is the mean of Φ_i, and m is the global mean of the feature over all classes, and x is each value of descriptor Φ in the class i.

There are 60 classes corresponding to the number of distinct individuals in the database. Although there are seven images per person, when computing the Fisher coefficients, only the 3D facial images having the whole regions and consequently the whole set of features correctly extracted have been employed (a total of 310 images).

When a feature could not be computed because of the non-existence of the region/s from which it is derived, it was zero valued, except when a symmetric feature exists. In this case, the non-existent feature is valued like its symmetric feature. Extracted features have been normalized to values in the interval [0, 1]. Figure 4 presents the Fisher coefficients corresponding to each feature rank position in the Fisher coefficient ordered list. Figure 5 shows the six most discriminating features.

4. Feature Selection

The goal of this section is to identify the minimum subset of features that maximize the recognition rate of a generic face recognition system. In order to select an optimal subset of features, eighty six feature subsets were tested. Each subset was composed by the first n features of the ordered features list according to Fisher coefficient, where $1 \leq n \leq 86$.

The training set consisted of 60 frontal images (one per individual) which, after segmentation, contained all the regions and lines necessary to obtain the feature vector. Six test image sets were used : (1) second set of frontal view images, which was composed of 47 images, (2) smiling images test, which contained 48 images, (3) laughing images test, which contained 45 images, (4) random gesture test, containing 47 images, (5) looking down images, with 19 images and (6) looking up images, with 41 images. Because the goal is to obtain the minimum set of features that produces the highest recognition rate, these test sets were constructed exclusively with the images of the individuals that contained all the regions and lines necessary to calculate all features. Figure 6 represents the recognition rates for some of the previous six test sets using as matching scheme the minimum Euclidean distance d_i:

$$d_i = (v_1 - a_{i1})^2 + (v_2 - a_{i2})^2 + \ldots + (v_n - a_{in})^2 \tag{15}$$

where $(v_1, v_2, \ldots, v_n)$ is the feature vector of the face being identified and $(a_{i1}, a_{i2}, \ldots, a_{in})$ is the feature vector of each face of the database.

Table 2. Extracted features from the segmented regions and lines, and their position in an ordered list according to their discriminating power. The used acronyms are: *Ri* = region *i*; *Li* = line *i*; *A_Ri* = area of region *i*; *C_Ri* = centroid of region *i*; *d*(*P1*, *P2*) = Euclidean distance between 3D points *P1* and *P2*; ang(*P1*,*P2*,*P3*) = angle defined by the 3D points *P1*,*P2* and *P3*, being *P2* the intermediate vertex; *H_Ri* and *K_Ri* are the respective averages of the Mean and Gaussian curvatures, evaluated in points belonging to the region *i*; *VH_Ri* and *VK_Ri* are the respective variances of the Mean and Gaussian curvatures evaluated in points belonging to the region *i*.

Id	FEATURE DESCRIPTION	rank
	Area of the regions	
0	A_R1	52
1	A_R4	41
2	A_R5	47
3	A_R2	30
4	A_R3	35
5	A_R6	45
6	A_R7	50
	Area relations	
7	A_R2/A_R1	59
8	A_R1/A_R3	71
9	A_R1/(A_R4+A_R5)	70
10	A_R1/(A_R6+A_R7)	58
	Mean of areas	
11	(A_R4 + A_R5)/2	38
12	(A_R6+A_R7)/2	42
	Distance between mass centres of regions	
13	d(C_R4, C_R5)	18
14	d(C_R6, C_R7)	81
15	d(C_R1, C_R3)	6
16	d(C_R4, C_R3)	8
17	d(C_R5, C_R3)	9
18	d(C_R6, C_R3)	36
19	d(C_R7, C_R3)	34
20	d(C_R6, C_R1)	61
21	d(C_R7, C_R1)	60
22	d(C_R4, C_R1)	15
23	d(C_R5, C_R1)	19
	Relations of distances between mass centres of regions	
24	d(C_R4, C_R5)/ d(C_R1, C_R3)	10
25	1/2 [d(C_R4,C_R3)+d(C_R5,C_R3)]/d(C_R3, C_R1)	32
26	d(C_R6, C_R7)/d(C_R3, C_R1)	49
27	1/2 [d(C_R6,C_R3)+d(C_R7,C_R3)]/d(C_R3, C_R1)	40
	Mean of distances	
28	1/2[d(C_R4, C_R3)+d(C_R5,C_R3)]	5
29	1/2[d(C_R6, C_R3)+d(C_R7,C_R3)]	29
76	1/2[d(C_R6, C_R1)+d(C_R7,C_R1)]	57
77	1/2[d(C_R4, C_R1)+d(C_R5,C_R1)]	14
	Angles among mass centres of regions	
30	ang (C_R4, C_R3, C_R5)	1
31	ang (C_R4, C_R3, C_R1)	2
32	ang (C_R5, C_R3, C_R1)	4
33	ang (C_R4, C_R1, C_R5)	31
34	ang (C_R6, C_R1, C_R7)	76
35	ang (C_R6, C_R3, C_R7)	27
36	ang (C_R6, C_R3, C_R1)	23
37	ang (C_R7, C_R3, C_R1)	26
	Mean of symmetric angles	
38	1/2[ang(C_R4,C_R3,C_R1)+ ang(C_R5, C_R3,C_R1)]	3
39	1/2[ang(C_R6,C_R3,C_R1) +ang(C_R7, C_R3,C_R1)]	25

Id	FEATURE DESCRIPTION	rank
	Mean of H of the points belonging to a region	
40	H_R1	28
41	H_R2	24
42	H_R3	16
43	H_R4	17
44	H_R5	21
45	H_(R4 U R5)	12
46	H_R6	53
47	H_R7	63
48	H_(R6 U R7)	48
	Mean of K region of the of a region	
49	K_R1	33
50	K_R2	37
51	K_R3	11
52	K_R4	20
53	K_R5	22
54	K_(R4 U R5)	13
55	K_R6	55
56	K_R7	64
57	K_(R6 U R7)	51
	Variance of H of the points of a region	
58	VH_R1	74
59	VH_R2	85
60	VH_R3	62
61	VH_R4	83
62	VH_R5	72
63	VH_(R4 U R5)	68
64	VH_R6	79
65	VH_R7	86
66	VH_(R6 U R7)	78
	Variance of K of the points of a region	
67	VK_R1	73
68	VK_R2	69
69	VK_R3	82
70	VK_R4	77
71	VK_R5	67
72	VK_(R4 U R5)	66
73	VK_R6	75
74	VK_R7	84
75	VK_(R6 U R7)	80
	Lines 1 and 2 based features	
78	d(upper end point L1, upper end point L2)	43
79	d(lower end point L1, lower end point L2)	65
80	Lenght of L1	44
81	Lenght of L2	46
82	1/2 [Lenght of L1+Lenght of L2]	39
83	Area of the polygon formed by the 4 ends of L1,L2	54
84	ang(lower point of L1,C_R1, lower end of L2)	56
85	ang(upper point of L1,C_R3, upper end of L2)	7

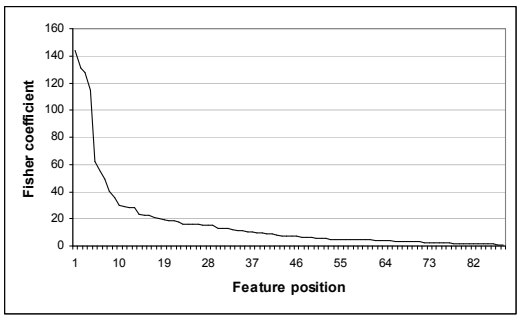


Fig. 4. Ordered list of the Fisher coefficient values corresponding to each extracted feature.

The matching scheme sequentially compares each testing image with each image of the database, by identifying the more similar ones. Each graph also presents the correct recognition rate in five different situations: (1) considering just the most similar face (the minimum distance), (2) considering the two more similar faces, (3) the three more similar faces, (4) the five more similar faces and (5) the eight more similar faces.

The reason to present these results in order to decide the length of the feature vector is to have more data for the decision, and because typically in face recognition applications, the solutions are semi-automated, where the output of the system is not just the more similar element, but a set of the most similar elements found, and after that the identification is refined manually.

The best recognition rates are obtained when the test set used is the frontal view of images, and the second one when the test set used is the smiling faces. After that recognition rates rapidly deteriorate, especially when using as test set any of the two

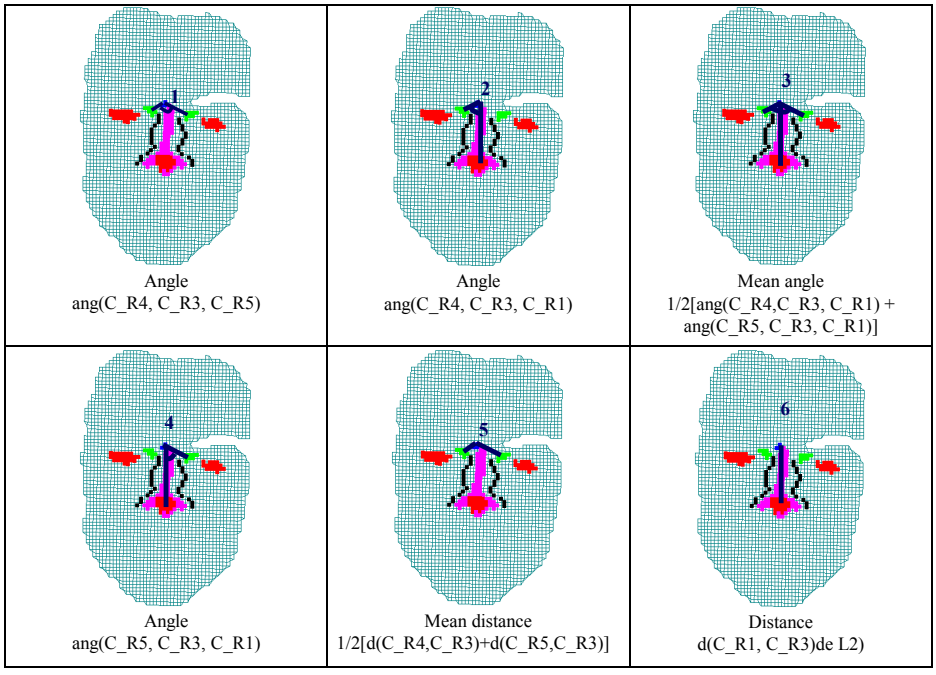


Fig. 5. Representation of the six most discriminative features according to their Fisher coefficients.

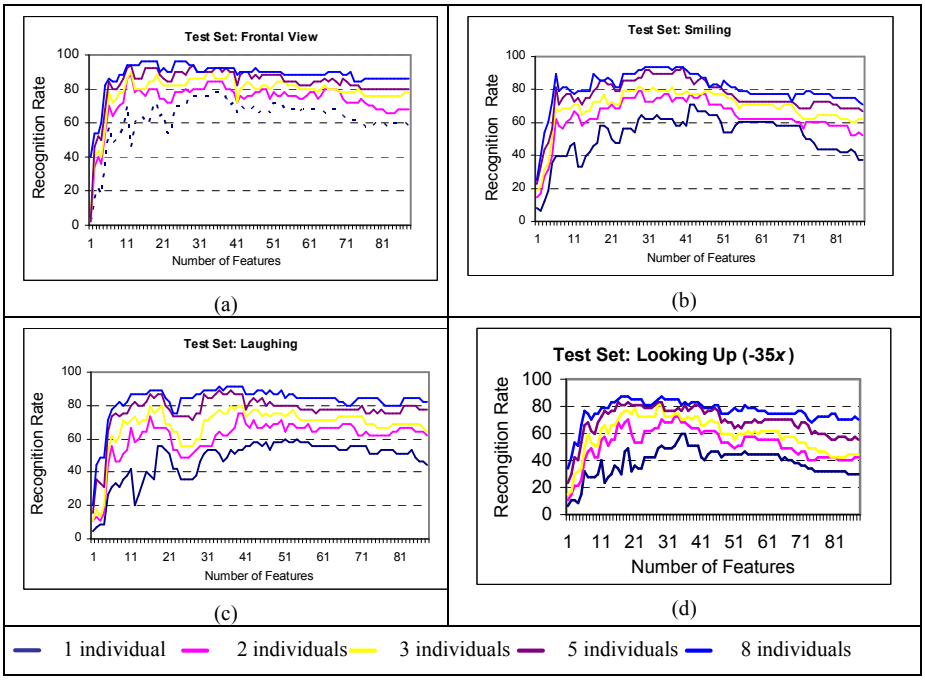


Figure 6. Recognition rates when using a feature vector of length n with $1 \leq n \leq 86$ for each test set: (a) frontal view, (b) smiling, (c) laughing and (d) looking up.

rotated images. By observation of the graphs, the conclusion is that, in these six test sets, the best recognition results were obtained when the number of features used to represent the individuals was between the first 30 and 37 features (those of lower positions in the ordered list according to Fisher coefficients).

5. Robustness of HK Face Modelling: Experiments and Result Analysis

The quality of any face model is measured by its ability to cope with a non-controlled environment in which users have certain degrees of freedom regarding facial expressions, rotations of the face, etc. In order to efficiently implement recognition systems it is necessary to have a robust face model able to cope with the variability of the problem. This section first presents the characteristics of our 3D face database designed to capture face variability. This database is necessary to test our model under some acquisition variations related to pose and pronounced facial expressions. In the other hand, our model is not affected by illumination variations (shades, kind of light) or makeup, cause it does not contain colour information (only geometrical data of the 3D meshes of face surfaces were used to obtain the model). After that, the results of applying the automatic segmentation algorithm to our database are given. The robustness of the model is checked by comparing the correct recognition rates using a con-

trolled and a non-controlled environment for two pattern classification methods: (1) PCA in combination with a Euclidean distance classifier and (2) SVM.

5.1. *3D Face Database*

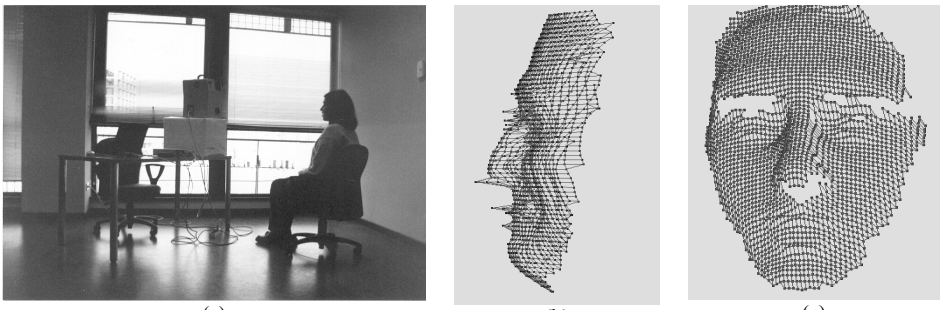
In order to evaluate the robustness of a 3D face model it is needed a database that captures the possible variations in pose and facial expression that an individual can have when interacting with a face recognition system, and to create that database using a flexible scanning environment. As the state of the art revealed there is no database that captures this variability. This has motivated the design and creation of a 3D database, called *GavabDB,* that can be used to test our 3D face model in a real environment. Database images can be found (without colour or texture) in *http://gavab.escet.urjc.es*. *GavabDB* contains 549 three-dimensional facial surface images corresponding to 61 Caucasian individuals with 9 different images per each person. Detailed database creation and its main features are described in[47].

In order to contain the maximum variability our database has been designed to capture both sources of variations. Each set of samples of an individual contains 9 images, being the 7 first ones (corresponding to 61 individuals, 427 in total) the images used in the experiments of this work: three frontal images with some kind of expression (one is a random expression chosen by the individual, one smiling and one laughing), two frontal images with neutral expression, one in which the individual is looking down (-35º *x* rotation approximately) with a neutral expression, one in which the individual is looking up (+35º *x* rotation approximately) with a neutral expression, one in which the individual is looking to the left (-90º *y* rotation approximately) with neutral expression and one in which the individual is looking to the right (+90º *y* approximately rotation) with neutral expression.

Facial surfaces are represented by meshes as provided by the range laser sensor. Thousand of points which approximate the surface of each sampled face, and the connections of these points, form a mesh representing the scanned face. Cells of the mesh have four non-coplanar nodes, and occasionally three in the contour. The grid provides an easy way of establishing two "orthogonal" directions (horizontal and vertical) along which it is possible going from one node to its neighbour in the mesh. The average points per face mesh in our database is 2,186 (at 1/4 of the scanning resolution). The background of the images is empty because the scanner does not sample the objects (i.e. the wall) placed out of focus. The scanner used for creating the database is a VI-700 Minolta.

When scanning each individual was positioned at about $1.5 \pm |0.5|$ meters from the scanner, the head did not have a fixed position. In order to capture each one of the images, each individual was asked to look at a specific point. Figure 7 shows the scanning environment. The captured images have three main problems: noise, holes and uncompleted contour. The noise present in the images can be solved using filters in the pre-processing stage. The problem of images with holes is caused because dark parts of the face (like eyelids) do not reflect light and consequently sometimes they are not scanned. The third problem is caused by the auto-occlusion of the face when it is

(unconsciously or not) rotated. Figure 7(b) presents an example of noise and Figure 7(c) presents one example of holes and uncompleted contour.



(a) (b) (c)

Fig. 7. (a) The capture environment and examples of the problems of the scanning: (b) noise and (c) holes.

5.2 *Automatic Face Segmentation: Result Analysis*

The Automatic Face Segmentation algorithm proposed in subsection 3.3 has been tested using all the faces (427 corresponding to 61 individuals) of our database. A region or line location result can be labelled as: (1) failure (F), (2) non-existent (NE) or (3) correct location (C). Failure means that a searched region or line was found in an incorrect location. Non-existence means that no result was obtained from searching for that region according to the automatic face segmentation algorithm. Correct location occurs when the segmented region was correctly found. Failure, non existence and correct percentage results for the different regions and for the different views of the images of the database after the segmentation are shown in Table 3.

The poorest segmentation results were obtained for the eye region, in particular, when the view of the individuals presented a light rotation when the individual was looking down. The main reason for that is caused by the fact that occluded eye regions were automatically reconstructed by the software of the 3D digitizer using an interpolation algorithm. This produced a pernicious effect in the surface, smoothing it and losing information of the real curvature value in that region. This phenomenon also occurred in Region 3 when individuals were looking up. Nose point region (Region 1) offers better results than the other regions, failing only when the image has a hole of the surface in this place. A failure of this region location produces failures in the rest of region locations, because the rest of the regions are searched according to its position. While a failure (F) on a region searched produces a recognition error, a non existence (NE) of some region makes possible the face recognition stage.

The last row of Table 3 presents the total correct location rate for each Region. We conclude that the proposed automatic face segmentation algorithm works satisfactorily, producing a minimum location rate of 83.6% and 87.65% for Region 5 and 6 (both eyes), and up to a 98% correct recognition rate for Regions 1 and 2 and Lines 1 and 2. In conclusion, the algorithm provides enough information to construct a robust model of the face.

Table 3. Failure (F), non-existence (NE) and correct location (C) percentages obtained in the region and line location stage. Results are grouped per searched regions and lines over each set of 60 images corresponding to each kind of view. Last rows are the percentages per searched regions and lines over the 420 images of the database.

		Region 1	Region 2	Region 3	Region 4	Region 5	Region 6	Region 7	Line 1	Line 2
looking down	F	0%	0%	0%	0%	8.3%	6.6%	0%	0.6%	0%
	NE	0%	0%	1.2%	15%	28.3%	51.6%	0%	0%	15%
	C	100%	100%	98.8%	90%	63.4%	41.8%	100%	99.4%	90%
looking up	F	0.6%	0.6%	1.2%	0.6%	1.2%	0.6%	0.6%	0.6%	0.6%
	NE	0%	0%	18.3%	6.6%	8.3%	15%	0%	1.2%	8.3%
	C	99.4%	99.4%	80.5%	92.8%	90.5%	84.4%	99.4%	98.2%	91.1%
frontal view 1	F	0%	0%	0.6%	0%	0.6%	0%	0%	0%	0%
	NE	0%	0%	6.6%	5%	6.6%	1.2%	0%	0%	1.2%
	C	100%	100%	92.8%	95%	92.8%	98.8%	100%	100%	98.8%
frontal view 2	F	0%	0%	0.6%	0%	0.6%	0.6%	0%	0%	0%
	NE	0%	0%	5%	8.3%	6.6%	5%	0%	0%	5%
	C	100%	100%	94.4%	91.7%	92.8%	94.4%	100%	100%	95%
Random gesture	F	0%	0%	0%	0%	0.6%	0.6%	0%	0%	0%
	NE	0%	0%	8.3%	0.6%	5%	10%	0%	0%	1.2%
	C	100%	100%	91.7%	99.4%	94.4%	89.4%	100%	100%	98.8%
laugh	F	0%	0%	0%	0%	1.2%	0.6%	5%	1.2%	0%
	NE	0%	0%	6.6%	0.6%	5%	10%	0%	0%	1.2%
	C	100%	100%	93.4%	99.4%	93.8%	89.4%	95%	98.8%	98.8%
smile	F	0%	0%	0%	0%	0%	0%	0.6%	0%	0%
	NE	0%	0%	6.6%	0.6%	6.6%	6.6%	0%	0%	0.6%
	C	100%	100%	93.4%	99.4%	93.4%	93.4%	99.4%	100%	99.4%
Total views	F	0.2%	0.2%	0.95%	0.2%	2.85%	1.9%	1.19%	0.95%	0.2%
	NE	0%	0%	7.85%	5.71%	9.5%	14.5%	0%	0.2%	5.71%
	C	99.8%	99.8%	91.2%	94.0%	87.6%	83.6%	98.8%	98.8%	94.0%

5.3 *Robustness of HK Modelling: Applications to Face Recognition*

In this section the robustness of our 3D face modelling system for face recognition is tested. The robustness of the model will be measured by how the changes in pose and gesture of the faces affect the recognition rate of a face recognition system. A robust model will be able to capture those changes without a relevant variation of the recognition rate. The robustness of the model has been tested using a face recognition system that uses as matching schemes PCA with a Euclidean classifier and SVM. Both PCA and SVM have produced very good results in 2D recognition problems. In order to test the robustness, for each matching scheme two experiments have been run: (1) using a non-controlled environment and (2) using a controlled environment. The comparison of the recognition rates obtained by the recognition system in the controlled and non-controlled environment will gave us an indication of the robustness of the designed model.

5.3.1 *Face Recognition Architecture*

Our face recognition architecture has four steps: (1) Pre-processing, (2) Segmentation, (3) Feature Extraction and (4) Classification or Matching. This architecture is valid both for training and testing, but while in the training phase (off line) the information provided to the matching scheme is used to create a classification system, in the testing phase (on line), the information is used to identify the class, in our case, the indi-

vidual. Figure 8 represents this generic architecture. The main steps of the system are:

- **Pre-processing:** This stage removes the information regarding the neck, ears and hair if they appear (which are not relevant for the 3D model) and successively applies both a median filter to reduce noise and a mean filter to smooth the surface. The median filter transforms the original 3D mesh into another one which substitutes the original *z* coordinate of each vertex by the median value of the surrounding. This non-linear filter cancels impulsive noise from the original mesh face. This type of noise is characterized by extreme values of *z* randomly distributed through the mesh, as shown in Figure 7(b). The mean filter is implemented as a low-pass filter which smoothes the image by attenuating high spatial frequencies. This filter has been implementing with a 3×3 convolution mask with all its weights with a value of 1/9.
- **Segmentation:** Automatic detection of facial regions. In our case it has been implemented using an HK algorithm for the classification of the points, using the procedure explained in section 3.3.
- **Feature Extraction:** Automatic calculation of features using the regions obtained by the segmentation algorithm in order to construct a feature vector of each face, as it was explained in section 3.4.
- **Matching:** Using the feature vectors obtained from the training set of images we construct a classification system to identify each subject. This system is applied to the feature vectors of the image test set. The used matching schemes were: (1) PCA in combination with a Euclidean classifier and (2) SVM, respectively.

5.3.2 *SVM vs. PCA for 3D Face Recognition*

Principal Component Analysis (PCA) and Support Vector Machines have been used with very good results in high-dimensional pattern recognition problems[12,51,61]. PCA is used in pattern recognition problems for reducing the dimensionality of the feature vectors in order to obtain a compact representation. This compressed representation is the used to implement recognition by a pattern classifier, typically a distance based classifier, neural networks or Bayesian classifiers.

SVM is a classifier derived from Statistical Learning Theory[61]. The problem that SVM try to solve is to find an optimal hyperplane that correctly classifies data points

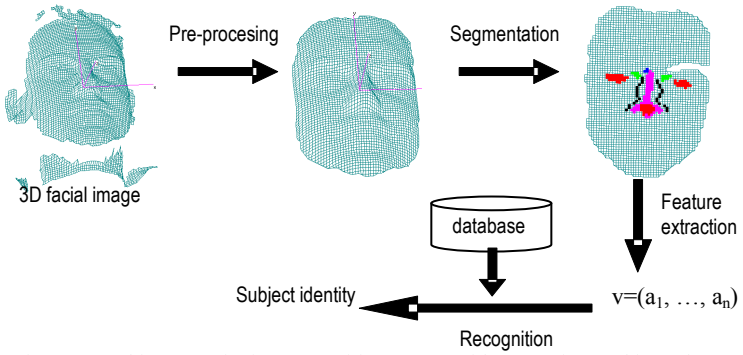


Figure 8. Architecture of a face recognition system with a generic matching scheme.

and separates the points of two classes as much as possible. The main parameter that defines a SVM classification system is the kernel used. Intuitively a kernel should represent the notion of similarity between any two individuals of the data base. Among the more common used kernels are the linear, polynomial and Gaussian kernels. A kernel can be also specifically designed for a 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) by construction, SVM have a relationship between the structure (the support vectors) and the classification tasks and (4) SVM optimize the separation surfaces between two classes. SVM have been used very effectively for recognition applications[28,52] in general, and for 2D face recognition[29] and 3D face recognition[52] in particular. Ref. 12 presents an extensive review of SVM pattern recognition applications.

5.3.3 *Experimental results*

The matching schemes have been implemented using a free PCA implementation[53] and SVMTorch[23]. The PCA implementation used Euclidean distance to identify a given face. The kernel used 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[12]. The only parameter a Gaussian kernel has is the standard deviation (*std*) and ideally should represent the minimum distance between any two elements of two different classes. In each case the classification system has been constructed using a vector with the 30 more discriminating features presented in section 4.

Each matching scheme has been tested with a controlled and a non-controlled environment. In this context, a controlled environment means that the images used for training and testing, are selected in such way that it has been ensured that there is one image in the train set having the same pose and expression than the image used for testing (in particular, frontal and neutral expression). As it was commented before, the lighting conditions are not controlled. In the non-controlled experiments, from the seven images of each individual, five have been randomly chosen to be in the training set, an the other two in the testing set, this produces a training set with 305 images (5 per each one of the 61 individuals) and 122 test images. In the controlled environment, six images are used for training, which include one of the frontal images, and the other frontal image is used for testing. This produces a training set with 366 images and a testing set with 61 images. The reason to run these two experiments is to test the robustness of the 3D face modelling proposed by testing how a non-controlled environment affects the recognition rate. This is an indication of the robustness of the model against variations of the pose and gesture. It have been showed in preceeding works that the recognition success goes down under non-controlled environments in such way that some systems become inoperative[68]. Table 4 presents the results of the experiments, including recognition rate, training time and testing time. The experiments were run in a AMD K-6 500MHz running Windows for the PCA environment and Linux for SVMTorch. Both training and testing time are given only for the non-controlled experiment, because this case represents a real application simulation. The

correct recognition rate indicates when the classification provided by the matching scheme identified the individual correctly (in this case just one image, the most similar one, was considered).

In general SVM produces a better recognition rate that PCA (the last one as a previous stage to a distance classifier), but, while in the non-controlled environment this increment is of about 1.5%, in the controlled environment is almost 9%. Also, both training and testing time are smaller for SVM, which has an average response time for each image of 0.003 sec., compared to 0.14 sec. of PCA. For these processing times it has to be taken into account the different operative systems and the different optimization of the implementations used. Going from a non-controlled environment to a controlled environment increases SVM correct recognition rate by 12.26%, while PCA only improves 5.7%. In any case, the small difference in correct recognition rate between a controlled and a non-controlled environment implies that the 3D face model created is robust enough to capture to a large extent of variations that real face recognition systems have. These experiments also show that the correct recognition rates obtained when using the model are enough to implement real face recognition system applications.

Our 3D model has advantages over existing methods in different aspects. First of all, results from extent studies[49] clearly show that the performance of traditional 2D face recognition approaches are adversely affected by varying lighting conditions and with respect to varying pose. Our proposed method uses the surface mesh without texture (colour) in order to avoid the influence of the variations on the lighting conditions. The features have been extracted from the curvature properties of the surface in order to obtain pose-invariance features (avoiding a pose normalization stage). With respect to other 3D techniques, most presented works either use a small database[3,31] for the experiments either experiment over a database in which the acquisition conditions of the images are more controlled than ours[18,21,25,31,59], or need a normalization stage[18,31], or in the case of systems that uses geometrical descriptors a number more reduced of them have been analyzed[31]. Some existing techniques were described in Section 2.2. Now, we are going point out some benefits of our system in comparison with some representative recent works. In Godil et al.[25] it was obtained a 70% of recognition success using 3D shape and 82% using a multimodal biometric recognition system with achieve the fusion at the scoring level of 3D surface and colour map information. Only images without facial expressions were considered in this work, and the resolution of the meshes was 4000 3D points per face (the double than ours). Ref. 3 present a 3D face recognition system based in depth map and texture map demonstrating that there is a significant improvement when using both together than when using either one alone. Although they reached 100% of recognition success, their

Table 4. Recognition rate of PCA and SVM under a controlled and a non-controlled environment.

	PCA + Euclidean Distance Classifier	**SVM**
Recognition Rate – Controlled Environment	81.96%	90.16%
Recognition Rate – Non-Controlled Environment	76.2%	77.9%
Training Time	6 min 31 sec.	50s
Testing Time	18 sec.	.4s

experiments were based on a very high-resolution data obtained with completely controlled environment and on manually marked facial features. Ref. 34 uses two 2D input images from each person and a 3D morphable model to generate a vast number of synthetic face images under varying poses and illumination conditions for the training database. Their system achieved recognition rates around 98%, but over a database which contains only 6 subjects. Other existing works which use geometrical descriptors to represent the face[21,31], have analyzed their discriminating power over reduced set of images, without facial expressions and/or without pose rotations. Here we present a novel evaluation of a high number of descriptors (86) over a great database of 420 images (60 individuals).

6. Conclusion and Discussion

3D face modelling has attracted a lot of attention in recent years due to the security applications derived from it and to the proliferation and improvement of the 3D scanners. A 3D face recognition system depends on a 3D face model robust against the variations in position and gesture that an individual may have when interacting with it. The main problems and limitations of 3D face modelling are: (1) the lack of models able to capture the variability in position and facial expression that real applications should have; and (2) the lack of standard 3D face databases that capture the variability in gesture and position.

This paper has presented a robust 3D face modelling system based on HK segmentation. The proposed system is independent of the lighting conditions, robust under facial expressions and it does not require normalizing the pose of the images. The model starts by automatically identifying eight regions and lines of each 3D facial mesh, which are then used to obtain the most relevant features according to Fisher's coefficient. The obtained feature vector produces a robust representation of an individual. The robustness of our model has been tested implementing two face recognition system based on PCA and SVM as matching scheme under controlled and non-controlled environments. Experimental results showed that the correct recognition rates are enough to implement real face recognition applications and that the proposed model is robust. Using SVM in a non-controlled environment produces 77.9% correct recognition rate, while a 90.16% is obtained in a controlled environment. We can conclude that a system of this kind has a recognition rate inside the interval [78%, 90%] for the considered variations. The proposed system works with all the images including those for which the segmentation stage does not provide all the regions and then, all descriptors can not be extracted. This fact reduces the recognition rate around 5% but permits to work with all images of the database.

In order to solve the lack of standard databases to test the robustness of 3D face models, this paper has also presented a 3D face database that captures the variability in gesture and position of users. The database contains 420 3D facial images of 60 different individuals, with seven images per individual: two frontal with no gesture, one rotated one looking up, one rotated looking down, one smiling, one laughing and one with a random gesture.

Future work includes choosing other features different from those belonging to the eye regions when faces are rotated looking up and looking down because their corresponding regions appear occluded in many patterns. Complete images have been used independently of the segmentation results in the recognition experiments. To develop a validation procedure to reject false positives in the region location stage is a critical task to improve the obtained results. A failure in a region or line location is easily detectable because usually the chosen candidate region is far from the correct location, and the size of the wrong region candidate differs a lot from the correct size of the searched region. Another future work is to apply a smoothing filter in order to loose the least curvature information possible to improve the recognition result.

References

1. http://www-users.cs.york.ac.uk/~tomh/3DFaceDatabase.htm.
2. P. Belhumeur, J. Hespanha, D. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection", *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**(7), 1997, 711-720.
3. C. Ben Abdelkader, P.A. Griffin, "Comparing and combining depth and texture cues for face recognition", *Image and Vision Computing* **23**, 2005, 339-352.
4. http://www.sic.rma.ac.be/~beumier/DB/3d_rma.htm
5. C. Beumier and M. Acheroy, "Automatic 3D face Authentication", *Image and Vision Computing* **18**, 2000, 315-321.
6. C. Beumier and M. Acheroy, "Automatic Face Verification from 3D and Grey Level Clues", *Portuguese Conference on Pattern Recognition* (RECPAD), 2000, 95-101.
7. C. Beumier and M. P. Acheroy, "Automatic face Authentication from 3D surface", *Proc. British Machine Vision Conferenc*e (BMVC), 1998, 449-458.
8. The Biometric Consortium, http://www.biometrics.org.
9. Biometric Technology Today, "Face Recognition: a new dimension", Nov-Dec 2002.
10. A. M. Bronstein, Michael M. Bronstein and Ron Kimmel, *Expression-Invariant 3D Face Recognition*, AVBPA, LNCS 2688, Spinger-Verlag, Berlin Heidelberg, 2003, 62-70.
11. R. Brunelli and T. Poggio, "Face Recognition: Features versus Templates", *IEEE Trans. Pattern Anal. and Mach. Intell.* **15**, 1993, 1042-1052.
12. H. Bynm, "A survey on pattern recognition applications of Support Vector Machines", *Intl. Journal of Pattern Recognition and Artificial Intelligence* **17**(3), 2003, 459-486.
13. http://www.sae.org/technicalcommittees/caesumm.htm
14. P. Campadelli and R. Lanzarotti, "Fidual point localization in color images of face foregrounds", *Image and Vision Computing* **22**, 2004, 863-872.
15. R. J. Campbell and P. J. Flynn, "A Survey of Free-Form Object Representation and Recognition Techniques", *Computer Vision and Image Understanding* **81**, 2001, 166-210.
16. http://www.nd.edu/~cvrl/
17. C. C. Chang and Y. H. Yu, "An Efficient Approach for Face Detection and Facial Feature Location using Prune-and-Search Technique", *International Journal of Pattern Recognition and Artificial Intelligence* **18**(6), 2004, 987-1005.
18. K. I. Chang, K. W. Bowyer y P. J. Flynn, "Face Recognition Using 2D and 3D Facial Data", *Workshop on Multimodal User Authentication* (MMUA), 2003.
19. K. I. Chang, K. W. Bowyer and P. J. Flynn, "An Evaluation of Multimodal 2D+3D Face Biometrics", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **27(**4), 2005.

20. R. Chellapa, C. L. Wilson and S. Sirohey, "Human and Machine Recognition of Faces: A Survey", *Proc. of the IEEE* **83** (5), May 1995, 704-740.
21. C.-S. Chua, F. Han and Y.-K. Ho, "3D Human Face Recognition Using Point Signature", *Fourth IEEE Intl. Conf. on Automatic Face and Gesture Recognition*, 2000.
22. http://lrv.fri.uni-lj.si/facedb.htm.
23. R. Collobert and S. Bengio, "SVMTorch: Support Vector Machines for Large-Scale Regression Problems", *Journal of Machine Learning Research*, vol. 1, pp. 143-160, 2002.
24. M. N. Dailey and G. W. Cottrell, "Organization of face and object recognition in modular neural network models", *Neural Networks* **12**, 1999.
25. A. Godil, S. Ressler and P. Grother, "Face Recognition using 3D Facial Shape and Color Map Information: Comparison and Combination", *Proc. SPIE* **5404**, 2004, 351-361.
26. S. Gong, S. J. McKenna, A. Psarrou, *Dynamic Vision: From Images to Face Recognition*, ImperCooege Press, 2000.
27. G. Gordon, "Face recognition based on depth maps and surface curvature", *SPIE Proc.: Geometric Methods in Computer Vision* **1570**, 1991, 234-247.
28. D. Gorgevik, D. Cakmakov, and V. Radevski, "Handwritten digit recognition by combining support vector machines using rule-based reasoning", *23rd Int. Conf. Information Technology Interfaces*, 2001, 139-144.
29. G. Guo, S. Z. Li and K. Chan, "Face Recognition by Support Vector Machines*", Proc. 4th IEEE Intl. Conf. on Face and Gesture Recognitio*n, Grenoble, March 2000.
30. J. Haddadnia, M. Ahmadi, "N-feature neural network human face recognition", *Image and Vision Computing*, **22**, 2004, 1071-1082.
31. P. Hallinan, G. Gordon, A. L. Yuille, P. Giblin and D. Mumford, *Two-and Three-dimensional patterns of the face*, Ed. A. K. Peters, 1999.
32. C. L. Huang and C. W. Chen, "Human Facial Feature Extraction for Face Interpretation and Recognition", *Pattern Recognition* **25**(12), 1992, 1435-1444.
33. J. Huang, B.Heisele and V.Blanz, "Component-based face recognition with 3D morphable models", *4th Conf. Audio and Video based Biometric Person Autenthication*, 2003, 62-69.
34. J. Huang, V. Blanz and B. Heisele, "Face Recognition with Support Vector Machines and 3D Head Models", *Intl. Conf on Computer Vision*, Vol. **2**, 2001, pp. 688-694.
35. T. Huang and L. Tang, "3-D Face Modeling and Its Applications", *International Journal Pattern Recognition and Artificial Intelligence* 10 (5), 1996, 491-520.
36. R. L. Hsu and A. K. Jain, "Face Modelling for Recognition", *Proc. IEEE Intl. Conf. on Image Processing*, Vol. 2, 2001, pp. 693-696.
37. Intelligent Multimedia Lab., "Asian Face Image Database PF01", Technical Report (Department of Computer Science and Engineering, Pohang University of Science and Technology, Korea), 2001.
38. M. Kirby and L. Sirovich, "Application of the Karhunen-Loève Procedure for the Characterization of Human Faces",*IEEE Trans. Patt. Anal. and Machine Intell.***12**,1990,103-108.
39. M. W. Lee and S. Ranganath, "Pose-invariant face recognition using a 3D deformable model", *Pattern Recognition* **36**, 2003, 1835-1846.
40. R. S. T. Lee, "Elastic Face Recognizer: Invariant Face Recognition based on Elastic Graph Matching Model", *International Journal Pattern Recognition and Artificial Intelligence* **16** (4), 2002, 463-479.
41. Y.W. Lei, J. L. Wu and M. Ouhyoung, "A three-dimensional muscle-based facial expression synthesizer for model-based image coding", *Signal Processing: Image Communication* **8**, 1996, 353-363.
42. S. Z. Li and A. K. Jain (editors), *Handbook of Face Recognition*, Springer, 2005.
43. A. M. Martínez, M.-H. Yang and D. J. Kriegman, "Introduction. Special issue on face recognition", *Computer Vision and Image Understanding* **91**, 2003, 1-5.
44. A. M. Martínez and R. Benavente, "The AR Face Database",*CVC Tech. Report* **24**, 1998.

45. J. Matas, M. Hamouz, K.Jonsson, J. Kittler, Y. Li, C. Kotropoulos, A. Tefas, I. Pitas, T. Tan, H. Yan, F. Smeraldi, J. Bigun, N. Capdevielle, W. Gerstner, S. Ben-Yacoub, Y. Abdeljaoued and E. Mayoraz, "Comparison of face verification results on the XM2VTS database", *Proc. 15th International Conference on Pattern Recognition*, Barcelona (Spain), vol. 4, September 2000, 858-863.
46. K. Messer, J. Matas, J. Kittler, J. Luettin and G. Maitre. "Xm2vtsdb: The extended m2vts database", *2nd Intl. Conf. on Audio and Video-based Biomet. Person Authentication*, 1999.
47. A. B. Moreno and A. Sánchez, "GavabDB: a 3D Face Database", *Workshop on Biometrics on the Internet* COST275, Vigo, March 25-26, 2004, 77-85.
48. S. Pankanti, R. M. Bolle and A. K. Jain, "Biometrics: The Future of Identification*", IEEE Computer*, February 2000, 46-49.
49. P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, J. M. Bone, "Face Recognition Vendor Test 2002: Overview and Summary", March 2003.
50. P. J. Phillips, H. Wechsler, J. S. Huang and P. J. Rauss, "The FERET Database and Evaluation Procedure for Face-Recognition Algorithms", *Image and Vision Computing* **16**(5), 1998, 295-306.
51. P. J. Phillips, "Support Vector Machines applied to face recognition", *Advanced Neural Information. Processing Systems* **11**, 1998, 803-809.
52. M. Pontil and A. Verri, "Support vector machines for 3-D object recognition", *IEEE trans. On Pattern Analysis and Machine Intelligence* **20**, 637-648,, 1998.
53. S. Romdhani, http://www.vision.im.usp.br/~teo/pca/, 1996.
54. F. Samaria, "Face Recognition Using Hidden Markov Models", Ph. D. Thesis, University of Cambridge, 1994.
55. T. Sim and S. Baker, "The CMU Pose, Illumination, and Expression Database", *IEEE Trans. Patt. Anal.. Mach. Intell.,* **25** (2003) 1815-1818.
56. H. T. Tanaka, M. Ikeda and H. Chiaki, "Curvature-based face surface recognition using spherical correlation", *Proc. 3rd IEEE Intl. Conf. on Automatic Face and Gesture Recognition,* 1998, 372-377.
57. J. P. Thirion, "The extremal mesh and the understanding of 3D surfaces*", Intl. J. Comput. Vision* **19**, 1996, 115-128.
58. E. Trucco, A.Verri,*Introductory Techniques for 3-D Computer Vision*, Prentice-Hall,1998.
59. F. Tsalakanidou, D. Tzovaras and M. G. Strintzis, "Use of depth and colour eigenfaces for face recognition", *Pattern Recognition Letters* **24**, 2003, 1427-1435.
60. M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces", Proceedings Int. Conf. on Pattern Recognition, 1991, 586-591.
61. V. Vapnik, *Statistical Learning Theory*, John Wiley and Sons, New York, 1998.
62. T. Vetter, "Flexible Models of Human Faces for the Analysis and Synthesis of Images", *Handbook of Comp. Vision and Appl, 3, Systems and Applications*, Academic Press, 1999.
63. Y. Wang, C.-S. Chua and Y.-K. Ho, "Facial feature detection and face recognition from 2D and 3D images", *Pattern Recognition Letters* **23**, 2002, 1191-1202.
64. J. L. Wayman, "Biometrics – Now and then: The development of biometrics over the Last 40 Years", *Proc. Second BSI Symposium on Biometrics*, 2004, 9-18.
65. L. Wiskott et al., "Face Recognition by elastic Bunch Graph Matching", *Trans. IEEE Pattern Analysis and Machine Intelligence* **12**, 1997, 775-779.
66. X. Xie and K.-M. Lam, "Face recognition under varying illumination based on a 2D face shape model", *Pattern Recognition* **38**, 2005, 221-230.
67. http://www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb/
68. W. Zhao, "Face Recognition: A Literature Survey", UMD-CFAR, *Tech. Rep. CAR-TR-948*, 2000.