5

Processing and Recognising Faces in 3D Images

Eyad Elyan and Daniel C Doolan
School of Computing, Robert Gordon University
United Kingdom

1. Introduction

Face recognition is one of the most active research areas in computer vision, statistical analysis, pattern recognition and machine learning (Huq et al., 2007). Significant progress has been made in the last decade, in particular after the FRVT 2002 (Phillips et al., 2003). For example (O'Toole et al., 2007) showed that face recognition systems surpassed human performance in recognizing faces under different illumination conditions. In spite of recent progress the problem of detecting and recognizing faces in un-controlled biometric environments is still largely unsolved.

The use of other biometric techniques, such as fingerprinting and iris technology appear to be more accurate and popular from a commercial point of view than face recognition (Abate et al., 2007). This is due to the inherent problems with 2D-image based FR systems. These include the viewing point of the face, illumination and variations in facial expression. These problems exhibit a great challenge for such systems and significantly affect performance and accuracy of algorithms.

In an overview of the Face Recognition Grand Challenge (FRGC) Phillips et al. (2006), the authors pointed out some of the new techniques used in Face Recognition that essentially hold the potential to improve performance of automatic face recognition significantly over the results in FRVT 2002. Among these techniques the use of 3D information to improve the recognition rates and overcome the inherent problems of 2D image based face recognition has become a current research trend.

In this chapter we present a novel technique for 3D face recognition using a set of parameters representing the central region of the face. These parameters are essentially vertical and cross sectional profiles and are extracted automatically without any prior knowledge or assumption about the image pose or orientation. In addition, these profiles are stored in terms of their Fourier Coefficients in order to minimize the size of the input data. The algorithm accuracy is validated and verified against two different datasets of 3D images covers a sufficient variety of expression and pose variation. Our computational framework is based on concepts of computational geometry which yield fast and accurate results. Here, our first goal is to automatically allocate the symmetry profile along the face. This is undertaken by means of computing the intersection between the symmetry plane and the facial mesh, results in a planner curve that accurately represents the symmetry profile.

Once the symmetry profile and few features points are allocated, then it is used to align the scanned images within the Cartesian coordinates with the tip of the nose residing at the origin.

Aligning the 3D images within the same Cartesian coordinates makes it possible to compare images against each other. Finally, profile-based comparisons are carried out, where profiles (space curves) of different faces are compared. Here, only part of these profiles are considered for the comparisons. These parts are assumed to be less sensitive to facial expression variation. This chapter is organised as follows: section 2 examines the current state of face recognition research. In section 3 the algorithm for processing 3D facial data and extracting facial features will be presented. Following on from this in section 4 profile-based comparisons will be briefly introduced. Experiments and results will be shown and discussed in the second last section. Finally the conclusions and limitations of our method and direction for future work are presented.

2. Previous work

Formally, face recognition maybe defined as: "given a still or video images of a scene, identify or verify one or more persons in the scene using a stored database of faces" (Zhao et al., 2003). In other words, for any facial recognition system, it is usually initialized with a set or a database of images of known persons. This image repository is usually termed as the "gallery". In a recognition scenario an incoming image of a certain person termed as "probe" is matched against the gallery for recognition purposes. This matching scenario is a one-to-many relation where the probe is matched against the entire set of images in the gallery to find out the best match based on some criterion or threshold.

The very initial step in an automated face recognition system is to detect the face in an image. Although it is very likely that more than one face might exist per image, usually it is assumed that only one face exists per image. Detecting the face, and identifying the region of interest on that face (the region which contains main facial characterises such as eyes, nose, and mouth) is essentially a critical step in order to align the face image with a certain coordinate system (often called image registration), and thereby make comparisons between faces feasible, and more likely to produce accurate results. This step is also important to allow the extraction of some facial features that will be used for comparisons purposes.

2.1 Types of face recognition systems

The vast majority of work that has been done in the area of FR has been based on 2D intensity images. In other words face recognition techniques that are solely based on 2D intensity images, usually acquired by 2D digital cameras. Such systems have several advantages, including the availability of cheap over the counter equipment and wide range of algorithms and existing solutions. Although it is difficult to categorize face recognition systems, it is often found in the literature that they are sometimes categorized based on the type of images used, for example 2D or 3D image-based FR systems. Following (Zhao et al., 2003) 2D-image based FR systems can be broadly categorized as holistic, feature-based or hybrid approach approaches. This categorization of 2D-image based FR system could also be generalized on systems that utilize 3D images.

Holistic approaches use the whole face region as an input to recognition. The work proposed early by Turk & Pentland (1991) serve as a corner stone for holistic-based face recognition approaches which is based on Principal Component Analysis (PCA). This in turn, is a dimensionality reduction technique, which treats the image as a point or a vector in a high dimensional space.

Other methods include the use of spatial-frequency techniques, such as Fourier transformations. In these methods, face images are transformed to the frequency domain, and only the coefficient in the low frequency band are preserved for recognition purposes. These approaches have been successfully applied to face recognition, however, the accuracy of such algorithms drops significantly under pose or light variation. In addition, for real-life applications with large databases, holistic methods may not provide sufficient discriminant information.

Unlike holistic-based methods, feature-based approaches utilize facial features such as eyes, nose, and mouth as an input for recognizing faces. In feature-based approach, the recognition rate is highly dependent on the accuracy of the face and facial feature localizations techniques. An example of such work is demonstrated by (Asteriadis et al., 2009) who used geometrical information to localize faces in images and several facial features, such as eyes, nose and mouth. First the face in an image was detected using the Boosted Cascade Method, then a Distance Vector Field was used to detect facial features such as eyes and mouth were geometric information about each pixel are encoded in the feature space.

Hybrid approaches utilize both holistic and local feature representation of the face images (Su et al., 2009). This technique is inspired by psychophysics and neuroscience literature which shows that human beings perceive faces based on both global and local features (Sinha et al., 2006). An example of such an approach is presented by (Su et al., 2009). In this work, the face image has been globally represented by means of Fourier transform defined by the authors as Global Fourier Feature Vector (GFFV).Gabor wavelets were used to extract local features from faces to form a vector space called Local Gabor Feature Vector (LGFV). Thus, representing each face image by one GFFV and multiple LGFVs. With such form of hybrid representation of the facial data, more diverse discriminatory information was encoded in the feature space.

2.2 Challenges

The accuracy of face recognition systems are significantly affected by various challenges. Although the revision of each of these challenges is beyond the scope of this chapter, it is worth highlighting some of these:

- Illumination is considered as one of the challenges that may hinder the robustness of 2D FR in an unconstrained environment.
- Pose variation is another important issue that has been the subject of extensive research. In spite of the major advances that took place in the past decade, handling varying head poses is still considered as one of the major challenges encountered by face recognition techniques. (See for a good review of various techniques that address this issue (Zhang & Gao, 2009)). Various solutions have been proposed to address this problem. One of the simplest solutions is to enrol the gallery with different images per individual that correspond to various poses. Several experiments show that enrolling more than one image, increases recognition rates. Eigenfaces, self organizing map and convolution network approaches both performed better when five gallery images per person were available rather than just one (Zhang & Gao, 2009). However, such an approach is not always possible due to the difficulties in obtaining suitable image data. In addition, it is quite impossible to represent all face poses in a database. Moreover, enrolling the gallery with multiple images per individual would add computational and storage costs, thereby impacting the overall performance of the system.

- Facial expression variation is still considered as one of the challenging problems for FR it has been estimated that the face could generate up to 55,000 different actions (Zhiliang et al., 2008). Addressing this problem will not only improve recognition rates, but will also have a positive impact on other domains, such as facial modelling, animation, and speech synthesis.
- Occlusions due to other objects such as sunglasses and hats.
- Aging, is considered as one of the main challenges to Face Recognition, as it causes significant alteration to the appearance of faces (Lanitis et al., 2002).

Due to the above challenges, the trend has shifted toward using 3D images in FR systems. It is strongly believed in the research community that using 3D information will improve recognition rates and overcome some of these challenges.

2.3 3D face recognition

(Phillips et al., 2006) point out to some of the new techniques used in face recognition. These include recognition from 3D images, high resolution still images, multiple still images and multi-modal techniques. Such techniques essentially hold the potential to improve performance of automatic face recognition significantly over the results in FRVT 2002. Among these techniques, possibly due to the recent development in 3D capturing devices and in order to overcome inherent problems of 2D-based face recognition systems, the trend is shifting toward utilizing 3D information to improve recognition rates. Here, the recognition is performed by matching the 3D models representing the shape of the faces. It is believed that the 3D representations of facial data will essentially overcome problems such as pose and illumination variations.

3D face recognition is attracting more attention in the recent years due to two important factors. Firstly because of the inherent problems with 2D face recognitions systems that appear to be very sensitive to facial pose variations, variant facial expressions, lighting and illumination. Xu et al (Chenghua et al., 2004) compared 2D intensity images against depth images and from their experiments they concluded that depth maps give a more robust face representation, because intensity images are significantly affected by changes in illumination. Secondly, due to the recent development in 3D acquisition techniques such as 3D scanners, infrared, and other technologies that makes obtaining 3D data acquisition relatively easy and accurate (Bowyer et al., 2006).

Although the utilization of 3D data in the area of face recognition started early (Y. et al., 1989), in comparison with image-based face recognition, the use of 3D information is relatively new in terms of literature, algorithms, commercial applications, and datasets used for experimentations (Bowyer et al., 2004). The number of persons represented in datasets for 3D face recognition experiments didn't reach 100 until 2003 (Bowyer et al., 2006), with little experimentation explicitly incorporating pose and expression variations. In this review paper (Bowyer et al., 2006) the authors surveyed some techniques which reported 100% recognition rates. However, the authors pointed out that this is due the limited size of the databases used. In addition, it was clear from the review that in early work only few published results have dealt with datasets that explicitly incorporate pose and/ or expression variation. In the past few years, this has changed, data sets have become larger and algorithms become more sophisticated. Therefore, it is not surprising to see that recent reported recognition rates are not as high as early work.

Several approaches are used in the literature for 3D face recognition. Some of these are based on the segmentation of the face into meaningful points, lines and regions. Others are considered as model based approaches using information about texture, edges, and colours. Profile-based techniques where multiple profile comparisons are carried out, by which a set of profiles are compared against each other. Such profiles might be symmetry ones, transverse, vertical or even cross-sectional.

Among the existing approaches for addressing 3D face recognition systems is the use of Extended Gaussian Image (EGI). Early work by (Lee, 1990) segment convex regions in a range image based on the sign of the mean and Gaussian curvature, and creates an extended Gaussian image. The matching algorithm is done by correlating the EGIs between the probe and an image in the gallery. The EGI in turn, describes the shape of an object by distribution of surface normal over the object structure.

A 2005 article (Gökberk et al., 2005) compared five approaches to 3D face recognition. They compared methods based on EGI, ICP matching, Range Profile, PCA, and Linear Discriminate Analysis (LDA). They used a database of 160 people consisting of 571 images. They found out that ICP and LDA approaches offer the best performance, although performance is relatively similar among all approaches but PCA.

One of the earliest attempts to utilized depth information was possibly proposed by Gordon (Gordon, 1992) who segmented the face based on curvature description, he then extracted a set of features that describes both the curvature and metric size properties of the face. Thus, each face becomes a point in the feature space and the comparisons were carried out using the nearest neighbouring algorithm.

Nagamine (Nagamine et al., 1992) extracted five feature points and used it to standardize face pose, matching various curves or profiles though the face data. According to this experiment the best recognition rates were achieved using vertical profiles that pass through the central region of the face.

Achermann (Achermann & Bunke, 2000) approached 3D face recognition based on an extension of Hausdorff distance matching. A database of 240 images were used, and 100% recognition rate was reported.

Lee et al. (Lee et al., 2005) Approached the problem based on the curvature values at eight feature points on the face. Using support vector machine for classifications they reported a rank-one recognition rate 96% for a data set representing 100 persons. The feature points were manually allocated.

Mahoor et al. (Mahoor & Abdel-Mottaleb, 2009) presented an approach for 3D face recognition from frontal range data. In this work each image was represented by ridge lines. These are points around the eyes, the nose and the mouth. The lines were defined based on the mean and Gaussian curvature computation, and used for representing the face images. Then for matching ridge images, robust Hausdorff Distance (HD) and ICP were used. It was reported by the authors that ICP outperformed HD in experiments carried out using GavabDB and FRGC 2.0 databases (third experiment) where neutral 3D face images of the FRGC2.0 were used, 58.9% rank-one identification using HD was reported while 91.8% using the ICP. Note that the authors here only used frontal images from GavaDB and FRGC 2.0 databases. In other words, results presented here may change by incorporating more pose and expression variation in the experiments.

Facial profile (vertical, cross-sectional, etc.) were also explored in 3D face recognition. Zhang et al. (Zhang et al., 2006) approached face recognition by utilizing 3D triangulated polygonal

meshes. Their approach starts by first identifying the symmetry plane (assuming that the facial data is symmetric), and then by computing the symmetry profile. Based on the mean curvature plot of the facial surface, and symmetry profile, they recovered 3 feature points on the nose area to define what they called facial intrinsic system (namely, the nose tip, nose bridge, and nose point at the lower nose edge), which were used to standardize the faces. For detection purposes the symmetry profile with another two transverse profiles provide a compact representation of the face and were used for comparison purposes. A database of 382 different scans was used consisting of 166 individuals of which 32 individuals have multiple scans, and others have just a single. EER for face authentication with variant facial expression reported was 10.8%. For scans with normal expressions 0.8% EER was reported. The symmetry profiles of two models to be compared are fist registered by mean of ICP algorithm. Then translation is done to make the cheek, forehead, and symmetry profiles coincide in the two models. The comparison is done by a set of sampling points on the corresponding profiles. A semiautomatic pre-processing procedure is used to trim of the non-facial regions in the raw mesh.

3. Automatic feature extraction

One of the main challenges in processing and determining certain facial features for a given raw 3D facial mesh is due to the resulting scanned image, which usually contains unwanted geometry that need to be identified and discarded at a pre-processing stage as shown in Figure 1. In certain applications semi-automatic approaches have been introduced to overcome this problem. For example (BenAbdelkader & Griffin, 2005) used seven manually selected land-mark points. Similarly in (Nagamine et al., 1992) it is necessary to identify several landmark points for example nose tip and eye corners which can then be used to register the face.

Fig. 1. Typical Examples of 3D Face Images

A key component within facial data is the symmetry characteristic that is defined by a symmetry plane which divides the face into two similar halves. Wide range of methods are available in the literature that deals with symmetry detection, in particular for 3D face shapes (Zhang et al., 2006), (Colbry & Stockman, 2007), (Pan et al., 2006)-(Gökberk et al., 2006).
Sun et. al. (Sun & Sherrah, 1997), for example, assume that the symmetry plane passes through the center of mass of a given object and uses Extended Gaussian Image (EGI) based technique to detect reflection, and rotational symmetry of objects. For facial data such an assumption might not hold, especially that 3D facial data acquired by laser scanners might be highly asymmetric since it would contain noise, and undesired geometry such as neck and the shoulder.

Zhang et al. (Zhang et al., 2006) detected the pose of a raw mesh by means of PCA, and then detected the symmetry plane by determining certain facial features (e.g. nose ridge points). They reported that out of 120 images, 117 model were correctly characterized by its symmetry profiles and few feature points along the nose area, with an average processing time of 10 seconds.
Colbry and Stockman (Colbry & Stockman, 2007) identified the symmetry plane of a facial scan by matching that scan with a mirror image of itself using face surface alignment algorithm assuming that pose variation is up to 10 degree in roll and pitch and up to 30 degree in yaw.

3.1 Main Method

In this section we will describe our technique to automatically extract the main facial features. First, we will give some definitions and terminologies that are used from this point on.

3.1.1 Definitions

3D images are either produced as point clouds or polygonal meshes (usually triangular). A point cloud is simply a set of n vertices $V = \{p_i \mid p_i \in R^3, 1 \leq i \leq n\}$. A triangular mesh S on the other hand, includes the set of vertices and adjacency information and is defined as $S = \{V, E, F\}$, where E is a set of edges defined as $E = \{(p_i, p_j) | p_i, p_j \in V\}$ and F is a set of facets defined as $F = \{(p_i, p_j, p_k) | p_i, p_j, p_k \in V\}$. The Euclidean distance between two points v_1, v_2 denoted by $v_1 = (x_1, y_1, z_1)$, and $v_2 = (x_2, y_2, z_2)$ is defined as $d(v_1, v_2) = ||v_1 - v_2 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$. If we Let $f_i \in F$ be a facet on the surface mesh defined by the triplets $f_i = \{v_0, v_1, v_2\}$ then the circumference of the f_i is defined as $d(f_i) = d(v_0, v_1) + d(v_1, v_2) + d(v_2, v_3)$. Based on this arrangement we could approximate the tolerance value of the surface mesh as,

$$S_t = \frac{1}{C} \sum_{i=1}^{nf} d(f_i) \tag{1}$$

where nf represents the number of facets in the triangular mesh, and $d(f_i)$ is the circumference of the i^{th} facet and C is a constant computed based on an average estimation of the number of common edges between adjacent facets on the surface mesh. A normalized and registered raw mesh means that all values of the vertices are scaled to be in the range between 0.0 and 1.0. In addition, the facial data is aligned with the Cartesian coordinate system, such that the nose tip is located at the origin and the face is looking towards the positive z-axis.
A plane is defined by a point and its normal vector, hence a plane will be denoted in the form of $\prod(p_0, n)$ where p_0 is a point on the plane, and n is its unit normal vector. A reference depth plane is used as the reference for measuring the depth of a given surface point on the mesh. The depth of any point denoted by $p_0 = (x_0, y_0, z_0)$ on the surface mesh is measured as the distance between that point and its projection on the depth plane $\prod$ which is defined as

$$d(p_0, \prod) = \frac{n_x x_0 + n_y y_0 + n_z z_0}{\sqrt{n_x^2 + n_y^2 + n_z^2}} \tag{2}$$

where the normal vector of the plane is defined as $n = (n_x, n_y, n_z)$. A planner curve is defined as a set of points in the 3D space that belongs to the mesh, and intersects a certain plane. The

length of a planner curve is defined as $\sum_{i=1}^{m} d_i(v_i + v_{i+1})$, where $d_i(v_i + v_{i+1})$ is the Euclidean distance between the two position points $v_i + v_{i+1}$ where $v_i + v_{i+1}$, are point positions on the planner curve, and v_i, v_m are the first and last point respectively on the curve.
The 3D face is said to be symmetric, if there is a plane, such that the face is invariant under reflection about it. Essentially a symmetry plane will pass through the tip of the nose. Thus, if the tip of the nose and another two position points are identified on the face then one could define the symmetry plane.
The centroid position of a facial surface mesh with vertices is denoted by $c = (c_x, c_y, c_z)$ where $c_x = \frac{1}{n}\sum_{i=1}^{n} x_i, c_y = \frac{1}{n}\sum_{i=1}^{n} y_i, c_z = \frac{1}{n}\sum_{i=1}^{n} z_i$. For a well characterized facial data set, the centroid point of a mesh usually lies within the region of interest which includes the nose, eyes and mouth features. Thus, it is highly unlikely that such a point would lie outside this region, for example near the neck area or the hair. Figure 1 shows various 3D facial scans (Moreno & Sánchez, 2004) with irregular outliers where the above assumption about the centroid position is still true.

3.1.2 Method outline

The tip of the nose is considered as one of the easiest feature points to recover from a facial image. In addition, we assume that the symmetry plane of the face passes through the tip of the nose. For human faces this is a very reasonable assumption which is widely accepted in the research community (Zhang et al., 2006), (Colbry & Stockman, 2007), (Sun & Sherrah, 1997). Our methodology is focused on identifying the symmetry plane based on the determination of the tip of the nose. The basic structure of the proposed algorithm is as follows,

- The central region of a 3D scan is initially approximated based on the center of mass and some extreme points.
- The tip of the nose is determined as the point on the facial surface with maximum perpendicular distance from a certain depth plane.
- The symmetry plane that passes through the pre-determined nose tip is then determined.
- A planner curve that accurately represents the symmetry profile is then extracted.
- Some feature points are then automatically determined on the symmetry profile. These feature points include the nose bridge and lower part of the nose.
- The central region is then extracted, based on approximating the positions of the outer corners of the eyes.

3.1.3 Nose tip identification

The first step in this process is to identify the tip of the nose. This is considered as the easiest point to recover on a facial scan. In order to determine this point, we fit a bilinear blended Coon's surface patch. Coon's patch is simply a parametric surface defined by four boundary curves(Farin & Hansford, 1999). The four boundaries of the Coon's patch are determined based on the boundary curves that enclose an approximated central region of the face.
In order to approximate the region of interest we take the centroid and all points that lie within a pre-determined distance from that point. It is important to highlight that the central region identified here is not an accurate representation of central region of the face. Rather it is an approximation which can be used to identify a "minimum" region of the face which can provide a smooth boundary on which it includes certain facial features, in particular the

nose region. Once this region is approximated, its boundary is sorted and organized so that it represents the four boundary curves of a Coon's patch. Finally, a surface patch within the boundary curves is interpolated based on Coon's patch definition (see (Farin & Hansford, 1999) for more information).

Having the Coon's surface generated as a reference to the facial points on an approximated central region, it becomes straightforward to recover an initial estimation of the nose tip as the one with the maximum depth from the patch. If we let $'$ to denote the set of all vertices within the approximated region of interest of the facial data and let C denote the set of vertices of the Coon's surface patch, then the initial approximation of the nose tip could be formulated as follows,

$$NTIP_{init} = \max\{d(p_i, e_j) : \forall p_i \in \forall', e_j \in C\} \tag{3}$$

Provided that, $e_j = \min\{d(p_i, e_j) : \forall e_j \in C\}$. Since the Coon's surface is composed of relatively small number of vertices in order to keep computation to a minimum, the above formulation only gives an approximation of the nose tip position. To improve our approximation we fit a plane using the points e_j recovered in Equation 3 and its neighbors e_{j0}, e_{j1} and compute the nose tip position as the point with maximum depth from the constructed plane. Figure 2(b) illustrates this concept. Assuming that the nose tip is denoted by N_{TIP} , the constructed depth plane fitted is defined as $\prod_{depth}$ and n is the normal unit vector to the plane, then the tip of nose is formulated as follows,

$$N_{TIP} = \max\{d(v_i, \prod_{depth}) : \forall v_i \in \forall'\} \tag{4}$$

where $d(v_i, \prod_{depth})$ is the Euclidean distance between a point v_i on the surface of the face and the constructed depth plane $\prod_{depth}$.

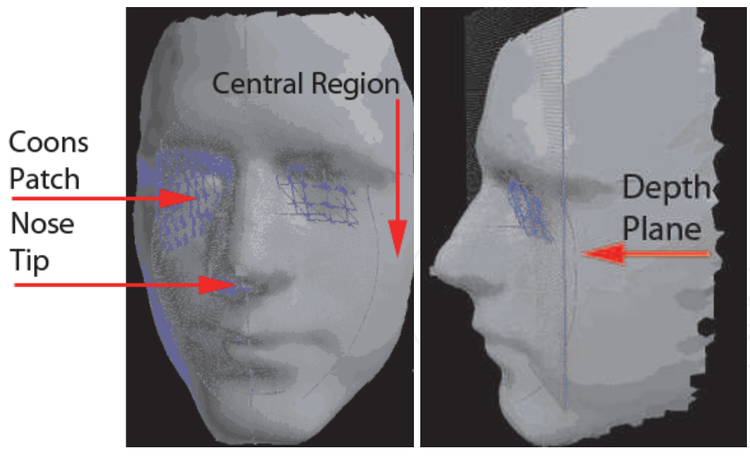


Fig. 2. Nose tip identification. (a) Initial estimation of nose tip based on depth measured relative to the Coon's patch. (b) Improving accuracy of nose tip positing based on fitting a plane

This procedure enables us to neutralize the facial data with the tip of the nose residing at the origin of a right hand coordinate system. In addition, the facial data can now be transformed in the Cartesian coordinate system with a rotation vector r defined by two points $N_{TIP}, Nproj_u$

that respectively represents the nose tip and its projection on the depth plane with normal unit vector u. Thus, once we identify the nose tip correctly, we then rotate the facial data such that r becomes aligned with the z-axis of the Cartesian coordinate system.

3.1.4 Symmetry plane detection

To identify the symmetry plane, we assume that N_{TIP} point lies on the symmetry plane. In addition, we let a point p_{s1} be any arbitrary point that lies on the depth plane such that $n_s = (N_{TIP} - p_{s1}) \times (N_{TIP} - Nproj_u))$ where $(N_{TIP} - p_{s1}), (N_{TIP} - Nproj_u))$ are two vectors such that $Nproj_u$ is the projection of N_{TIP} into the depth plane and n_s is the normal unit vector resulting from their cross-product. Figure 3 illustrates this arrangement. Clearly both depth plane and the initial symmetry plane with normal n_s are perpendicular to each other. Assuming that the initial symmetry plane defined by the point p_{s1} and its normal unit vector n_s denoted as $\prod(p_{s1}, n_s)$ and recalling that p_{s1} is one of the points lying on the depth plane then we make the following observations:

1. For a human face, the height dimension of the face is greater than its width.
2. It is clear that if the upper part of the face was considered, and the initial symmetry plane was rotated around the z-axis, then the planner curve that is identified as the intersection between the facial points and the initial symmetry plane with the minimum length will be the symmetry profile.

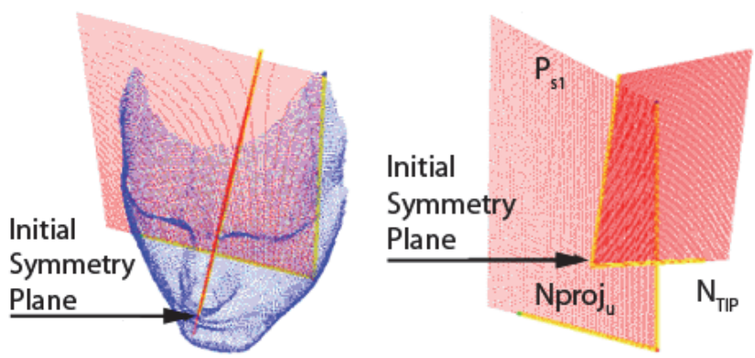


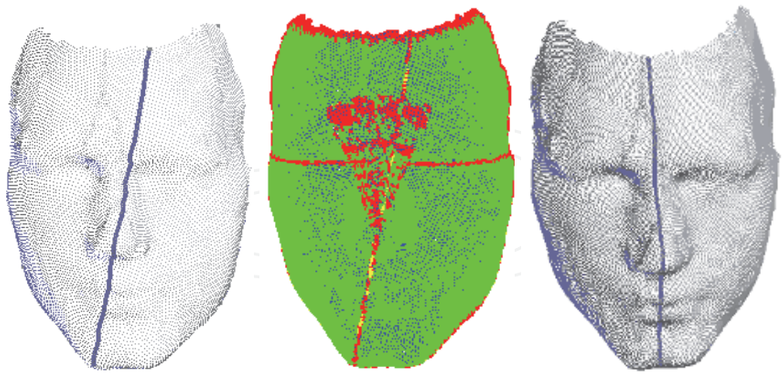
Fig. 3. Symmetry plane identification. (a) Facial surface with depth plane, and initial symmetry plane. (b) Initial symmetry plane results from the nose tip, its projection into depth plane, and an arbitrary point on the depth plane

Based on this arrangement, the initial symmetry plane is rotated by 2π around the z-axis and computation is performed to verify the correct allocation of the symmetry plane. In order to perform the rotation, we compute an angle θ where θ is the angle by which the symmetry plane should be rotated each time. Recall that, the facial data is already aligned within the Cartesian coordinate system with the nose tip residing at the origin. Therefore, if we assume that the initial symmetry plane is defined by the three points $N_{TIP}, p_{s1}, Nproj_u$ and recalling that p_{s1} is one point on the depth plane then θ could be defined as $\theta = \cos^{-1}(\frac{d^2}{S_t^2 + d^2})$, where $d = d(p_{s1}, NTIP)$, and S_t is the tolerance value of the mesh. Defining θ to be dependent on

the tolerance value of the mesh makes it more accurate regardless of how the meshes vary in terms of their density. In addition, rotation of the initial symmetry plane based on a very small value for θ minimizes the error value of the detected symmetry plane. Based on θ, the number of rotations that need to be performed is then approximated as $n = \frac{2\pi}{\theta}$, Working out the degree of rotations and the number of rations to validate the symmetry plane, the algorithm proceeds as shown in Algorithm 1. Figure 4, provides an illustration for symmetry plane detection algorithm.

Algorithm 1: Approximating Symmetry Plane

Let $height = -100.0, length = 100.0$ Let $V^{'}$ be a subset of the facial data that represents an approximated central region of the face. ;
while *Number of rotations* $\leq n$ **do**
  Find $v_0, v_1 \in V^{'}$ such that they both intersect initial symmetry plane at both ends of the central region.;
  Let v_{0p}, v_{1p} be the projected point of v_0, v_1 respectively into the depth plane, and construct an initial symmetry plane based on the three points v_{0p}, N_{TIP}, v_{1p}. ;
  Find the planner curve $p(l)$ that is resulting from the intersection of the Facial points of the central region with the initial symmetry plane, and let p_{lenght} be the length of its upper part. ;
  if $d(v_{0p}, v_{1p} > height\ and\ P_{length})$ **then**
    set $height = d(v_{0p}, v_{1p}), length = P_{length}$ and store v_0, v_1 , as possible candidate for symmetry plane points.;
  end
  rotate the initial symmetry plane by θ.;
end



(a) (b) (c)

Fig. 4. Improving symmetry plane identification. (a) initial symmetry plane, (b) rotating symmetry plane and computing length of the curve. (c) the final symmetry profile

In order to analyze the symmetry profile extracted from the facial data, we fit a spline of the form $P_i = \sum C_i B_i$ where $B - i$ is a cubic polynomial and C_i are the corresponding control points. This process of curve fitting to the extracted discrete symmetry profile data enables us

to have a smooth curve passing through the discrete data. Once we have a smooth symmetry profile, we analyze the profile by identifying local extreme points that corresponds to the nose bridge and the lower point of the nose Figure 5.
Based on the symmetry profile, a profile that passes through the nose bridge and through the eyes area can be extracted. Figure 5 shows the relation between the cross-sectional eyes profile which passes through the point NB (nose bridge), and the symmetry profile.

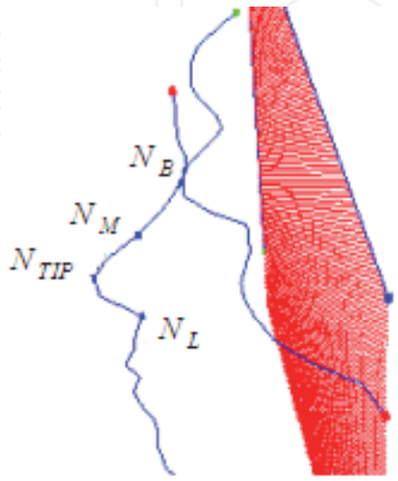


Fig. 5. Symmetry profile analysis

An initial testing of the accuracy of the algorithm for detecting symmetry profile was carried out based on reflective symmetry. Figure 6 illustrates this approach, which is similar to the one used in (Gökberk et al., 2005) to detect a symmetry plane. As shown in Figure 6, it is assumed that $n = (n_x, n_y, n_z)$ is the unit normal vector of the detected symmetry plane. If we define a set of points $P = \{p_i\}$ to be the set of vertices that exist at one side of the symmetry profile, and reflect these points around the symmetry plane, another set of Points Q, will be obtained. Assuming the facial data is perfectly symmetric and the identified symmetry plane is the correct one, then the average mid points of each point and its image $\{p_i, q_i\}$ which is denoted by m_i could be computed using the parametric equation of line $m = p_i + \alpha(q_i - p_i)$ where $\alpha = 0.5$ and the average error value is computed as $Err = \frac{1}{n}\sum_1^n d(\prod, m_i)$, where $d(\prod, m_i)$ is the distance between the i^{th} mid point and the detected symmetry plane $\prod$ in the Euclidean space.
Figure 7 shows some extreme examples of the results of our algorithm on a set of images taken from (Moreno & Sánchez, 2004). In the following section this technique will be further tested by using the resulting features in comparing face images.

4. Profile-based face recognition

Scanned images can be of different poses within the coordinate system. Thus, in order to carry out comparisons between these different scans, the scanned images have to be properly aligned within the Cartesian coordinate. This process is carried out automatically by relying in the proposed algorithm discussed in the previous section. Three feature points namely the nose tip, Nose Bridge, and the lower edge of the nose are used to align the scanned image within the Cartesian product. It is important to stress that the identification of these feature

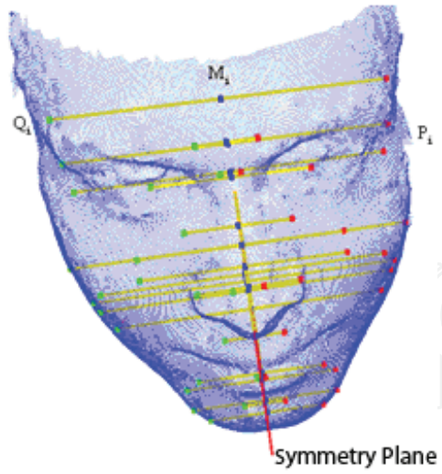


Fig. 6. Symmetry accuracy by calculating the midpoint between a surface point and its image around the symmetry plane

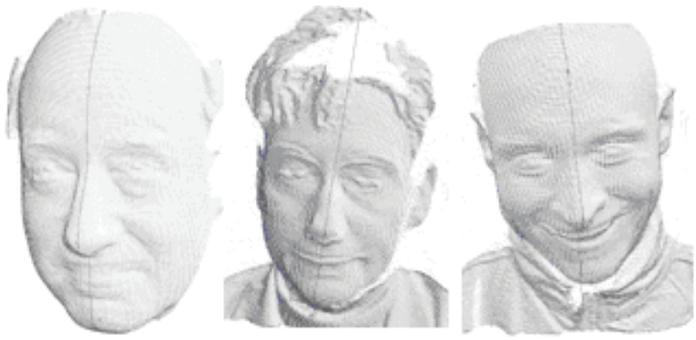
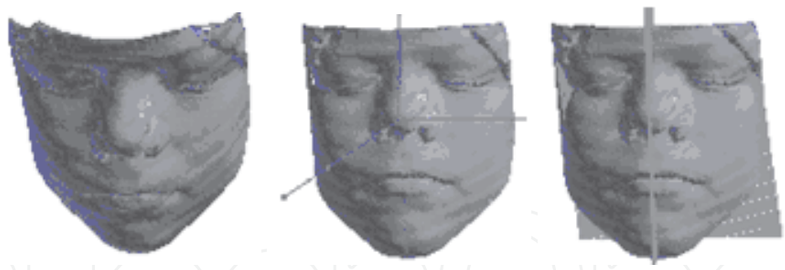


Fig. 7. Visualizing correct identification of symmetry profile in different sample images

points on the symmetry profile is an approximation, in other words the allocated points may not be very precise. However they are good enough for matching and registration purposes as will be discussed and validated in the following section. The alignment of the images is done by carrying out a rigid transformation of the dataset of the 3D points that make the image. The transformation is carried out based on the symmetry profile and the nose tip and is composed of a series of simple translations and rotations to end up with an image aligned within the Cartesian coordinate with the nose tip residing at the origin and facing the positive Z-direction as shown in Figure 8.

For comparisons purposes, it is important to point out that some facial regions are considered more rigid and less sensitive to facial expression variation than others. Nose region for instance, is considered relatively rigid compared with other regions such as the mouth. For profile-based face recognition, the sensitivity of facial regions is even increasing and hence seriously affecting the recognition accuracy, because regions are represented by space profiles. The lower part of the symmetry profile for example is highly sensitive to facial expression variations, while it is more rigid within the area bounded by NL and NB as shown in Figure 9.



(a) (b) (c)

Fig. 8. Processing and registering 3D images (a) loaded face in arbitrary pose and orientation (b) face is automatically processed and aligned with the nose tip residing at the origin (c) the symmetry plane of the face

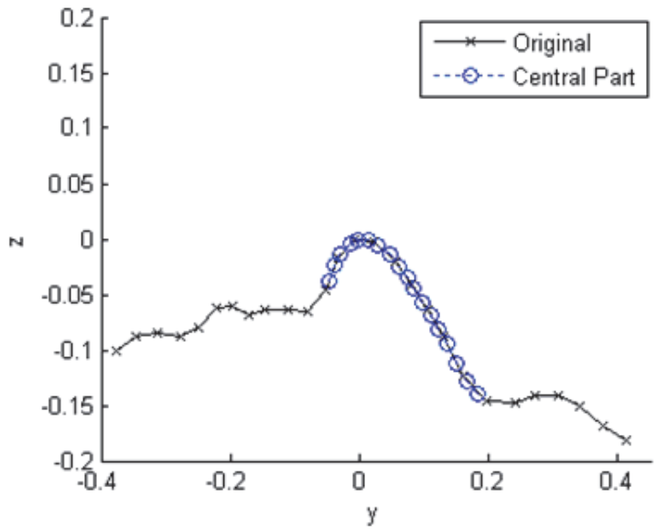


Fig. 9. Central part of the Symmetry profile

Similarly, cross-sectional profles that pass through the eyes area are highly sensitive to facial variations. Thus, it is not reliable to use it for recognition purposes. So for our recognition algorithm we use the central part of the symmetry profile which lies on the nose region (Figure 9) and central part of the cheeks profile. The cheeks profile is simply the profile that crosses the nose area at the mid distance between the points NB and NL. In order to minimize the input data, we compute the Fourier coefficients of the designated profiles and store it in a database, other than storing the actual points of the profile. Thus, having a database of images representing different individuals where each person is represented by two profiles stored by means of their Fourier's. In real time the database file would be loaded into memory and the profiles would be reconstructed according to the general form of Fourier series expansion Equation 5.

$$f(t) = \frac{1}{2}a_0 + \sum_{n=1}^{M} a_n \cos(nt) + b_n \sin(nt) \tag{5}$$

M is chosen to be relatively small, such that the number of coefficients required to reconstruct the curve is relatively much smaller than the number of 3D points that represent the profile. Matching faces against each other is carried out by a profile-by-profile comparisons with closest match selected. The comparison of the profiles is done point-by-point in the 3D space similar to (Zhang et al., 2006). If we let Lp and Lg be two profiles representing the central part of the symmetry profile of a probe and an image in the gallery respectively as shown in Figure 10.

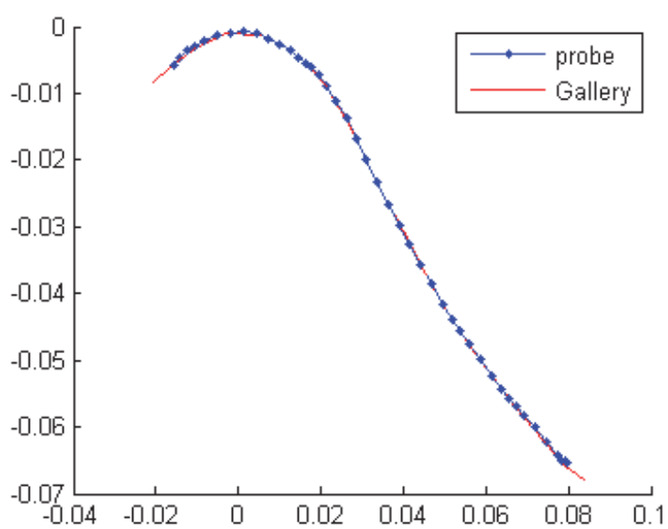


Fig. 10. Profile comparisons

Clearly the distance between the two polylines is directional, in other words the mean distance between the two profiles L_p to L_g is not necessarily the same as the distance from L_g to L_p. These distances are defined as

$$d_{pg} = \frac{1}{n} \sum_{p1 \in L_p} \min_{p2 \in L_g} d(p_1, p_2) \tag{6}$$

$$d_{pg} = \frac{1}{m} \sum_{p2 \in L_p} \min_{p1 \in L_p} d(p_2, p_1) \tag{7}$$

where n, and m represent the number of positions points on the profiles L_p and L_g , and $d(p_1, p_2)$ is the Euclidean distance between p_1, p_2. Thus, the similarity measure between the two profiles can be formulated as

$$E = \frac{1}{2}(d_{pg} + d_{gp}) \tag{8}$$

Based on Equation 8 the measure between two images is computed as

$$E_{total} = E_{cs} + E_{cc} \tag{9}$$

where E_{cs} represents the similarity measure between the central part of the symmetry profiles of the two images, and E_{cc} represents the similarity between the central part of the cheeks profiles of the two images, and both measures are computed as in Equation 8.

5. Experiments and results

For experimental purposes a 3D platform has been developed using Microsoft Foundation Classes (MFC), c++ and openGL. The platform is used to load 3D images, carry out the features extraction and conduct the matching algorithm to search for the best match in Database (Figure 11). In testing our processing and matching algorithm two experiments were carried out using two different databases.

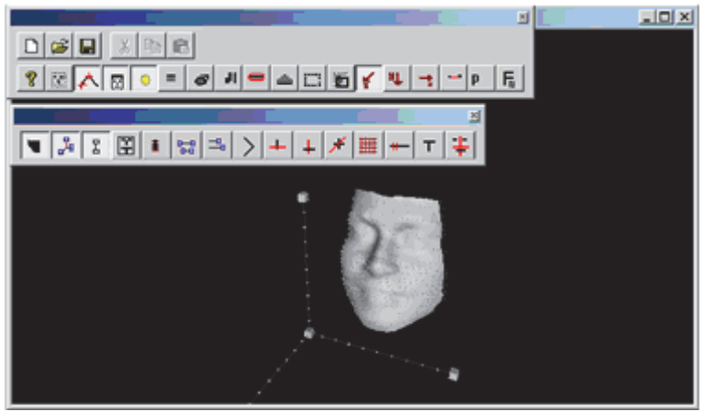


Fig. 11. Software for loading, processing and recognizing 3D Face images

5.1 Experiment 1

In the first experiment a database representing 22 different individuals was used. Each individual in the database is represented by 5 images, each represent a different pose (Figure 12). Only one profile was used to for comparing images, namely the central part of the symmetry profile.

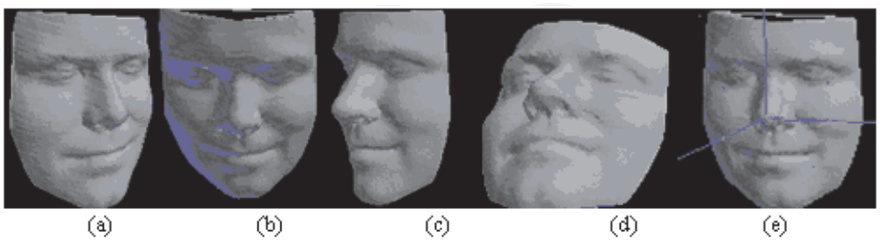


Fig. 12. Facial scans with different poses, (a) pose rotated by degree around the y-axis, (b) rotated by degree around the z-axis, (c) rotated by degree around the y-axis, (d) rotated by degree around the x-axis, and (e) aligned with the Cartesian coordinate.

Figure 13 shows a screen shot for the database file that we used in our experimentation. In this experiment the first line in the file represents the number of images in the database. Individual images are numbered consecutively and each number is followed by the Fourier coefficients representing the central profiles of that person. Hence, in a recognition transaction, an image

is loaded into our system, its features are extracted, the pose is aligned within the Cartesian coordinate and finally the database file is loaded into memory and profiles are constructed and compared against the image. In this particular experiment a 100% recognition rate was achieved. This result was expected as the main face variation is due to the pose not to the facial expression variation.

```
22
1.obj
-0.000822  0.022472 -0.043260
-0.001687  0.047064 -0.019793
-0.000342  0.000924 -0.000100
-0.000207  0.004315 -0.000827
-0.000152  0.000727 -0.000179
 0.000127  0.000470 -0.000558
-0.000080  0.000459 -0.000011
-0.000062  0.001190 -0.000574
-0.000033 -0.000401  0.000443
 0.000114  0.000532 -0.000341
```

Fig. 13. Typical database file for 3D images used in the experiments

5.2 Experiment 2

In the second experiment we used gavabDB (Moreno & Sánchez, 2004) which is a public 3D database of human faces. The database covers enough systematic variation in terms of facial poses and facial expressions. Total number of individuals represented in the database is 60, out of these, 45 are male with the remainder being female. Each individual in the database is represented by 9 different images. In our experiment we only consider 7 images per person and discarded two images/person from the database as only part of the face is available in the image such as the left side or right side or the face.

Both profiles (central parts of cheeks and symmetry) are used in this experiment as shown in Figure 14. In total 365 images were tested using our algorithm and were correctly identified, which corresponds to an accuracy recognition rate equal to 86.90%. Inaccurate results were due to the failure of the feature extractions algorithm to standardize the pose and hence extracting the required profiles for comparing the images. In other words 55 different images that were incorrectly identified were actually falsely rejected by the matching algorithm. This raises the False Rejection Rate (FRR) of this experiment to 13.0% which is due to the inaccurately identified features which result in relatively large error value between the profiles of the compared images and result in rejecting the image.

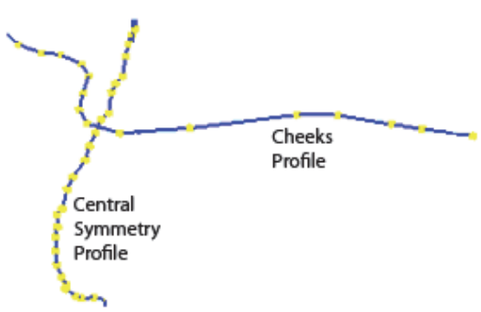


Fig. 14. Two profiles used to compute similarity measure between two face images

6. Conclusions and future work

In this chapter we introduced a new technique for processing 3D images of human faces and extract certain features to be used for recognition purposes. In addition, we have successfully demonstrated that utilizing rigid regions of a human face is very useful in terms of improving recognition rates and minimizing the search space. The average processing time for recognition was 10 seconds. This time includes, loading an image, processing it, extract facial features, standardize the pose, load the database file and conduct the profile comparisons.

Possible improvements to the current recognition system would include improving the features extraction algorithm so that more features points are extracted automatically. In addition, the algorithm should be improved to deal with low quality images. In our experiments the algorithm failed when the images contains holes or spikes, simply because this would lead to false identification of the tip of the nose and would essentially lead to false identification of the rest of the features required.

7. References

Abate, A. F., Nappi, M., Riccio, D. & Sabatino, G. (2007). 2d and 3d face recognition: A survey, *Pattern Recognition Letters* 28(14): 1885 – 1906. Image: Information and Control.

Achermann, B. & Bunke, H. (2000). Classifying range images of human faces with hausdorff distance, *Pattern Recognition, International Conference on* 2: 809–813.

Asteriadis, S., Nikolaidis, N. & Pitas, I. (2009). Facial feature detection using distance vector fields, *Pattern Recognition* 42(7): 1388 – 1398.

BenAbdelkader, C. & Griffin, P. A. (2005). Comparing and combining depth and texture cues for face recognition, *Image and Vision Computing* 23(3): 339 – 352.

Bowyer, K. W., Chang, K. & Flynn, P. (2004). A survey of approaches to three-dimensional face recognition, *17th International Conference on Pattern Recognition*, pp. 358–361.

Bowyer, K. W., Chang, K. & Flynn, P. (2006). A survey of approaches and challenges in 3d and multi-modal 3d + 2d face recognition, *Comput. Vis. Image Underst.* 101(1): 1–15.

Chenghua, X., Yunhong, W., Tieniu, T. & Q., L. Q. A. L. (2004). Depth vs. intensity: which is more important for face recognition?, *17th International Conference on Pattern Recognition*, pp. 342–345.

Colbry, D. & Stockman, G. (2007). Canonical face depth map: A robust 3d representation for face verification, *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pp. 1 –7.

Farin, G. & Hansford, D. (1999). Discrete coons patches, *Comput. Aided Geom. Des.* 16: 691–700.

Gökberk, B., Irfanoglu, M. O. & Akarun, L. (2006). 3d shape-based face representation and feature extraction for face recognition, *Image and Vision Computing* 24(8): 857 – 869.

Gökberk, B., Salah, A. A. & Akarun, L. (2005). Rank-based decision fusion for 3d shape-based face recognition, *LNCS 3546: International Conference on Audio and Video-based Biometric Person Authentication (AVBPA 2005*, pp. 1019–1028.

Gordon, G. (1992). Face recognition based on depth and curvature features, *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR '92., 1992 IEEE Computer Society Conference on*, pp. 808 –810.

Huq, S., Abidi, B., Kong, S. G. & Abidi, M. (2007). A survey on 3d modeling of human faces for face recognition in computational imaging and vision series, *3D Imaging for Safety and Security* 35: 25–67.

Lanitis, A., Taylor, C. & Cootes, T. (2002). Toward automatic simulation of aging effects on face images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24: 442–455.

Lee, J. C. M. E. (1990). Matching range images of human faces, *Third International Conference on Computer Vision*, pp. 722–726.

Lee, Y., Song, H., Yang, U., Shin, H. & Sohn, K. (2005). Local feature based 3d face recoginition, *Lecture notes in Computer Science* 3546: 909–918.

Mahoor, M. H. & Abdel-Mottaleb, M. (2009). Face recognition based on 3d ridge images obtained from range data, *Pattern Recognition* 42(3): 445 – 451.

Moreno, A. B. & Sánchez, A. (2004). GavabDB: a 3D Face Database, *Workshop on Biometrics on the Internet*, Vigo, pp. 77–85.

Nagamine, T., Uemura, T. & Masuda, I. (1992). 3d facial image analysis for human identification, *Pattern Recognition, 1992. Vol.I. Conference A: Computer Vision and Applications, Proceedings., 11th IAPR International Conference on*, pp. 324 –327.

O'Toole, A., Phillips, P., Jiang, F., Ayyad, J., Penard, N. & Abdi, H. (2007). Face recognition algorithms surpass humans matching faces over changes in illumination, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 29(9): 1642 –1646.

Pan, G., Wang, Y., Qi, Y. & Wu, Z. (2006). Finding symmetry plane of 3d face shape, *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, Vol. 3, pp. 1143–1146.

Phillips, P., Grother, P., Micheals, R., Blackburn, D., Tabassi, E. & Bone, M. (2003). Face recognition vendor test 2002, *Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on*, p. 44.

Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W. & Worek, W. (2006). Preliminary face recognition grand challenge results, *in Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on*, pp. 15–24.

Sinha, P., Balas, B., Ostrovsky, Y. & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about, *Proceedings of the IEEE* 94(11): 1948 –1962.

Su, Y., Shan, S., Chen, X. & Gao, W. (2009). Hierarchical ensemble of global and local classifiers for face recognition, *Image Processing, IEEE Transactions on* 18(8): 1885 –1896.

Sun, C. & Sherrah, J. (1997). 3d symmetry detection using the extended gaussian image, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 19(2): 164 –168.

Turk, M. & Pentland, A. (1991). Eigenfaces for recognition, *Journal of cognitive Neuroscience* 3(1): 71–86.

Y., C. J., Lapreste, J. T. & M., R. (1989). Face authentification or recognition by profile extraction from range images, *IEEE Computer Society Workshop on Interpretation of 3D Scenes*, pp. 194–199.

Zhang, L., Razdan, A., Farin, G., Femiani, J., Bae, M. & Lockwood, C. (2006). 3d face authentication and recognition based on bilateral symmetry analysis, *The Visual Computer* 22(1): 43–55.

Zhang, X. & Gao, Y. (2009). Face recognition across pose: A review, *Pattern Recognition* 42(11): 2876 – 2896.

Zhao, W., Chellappa, R., Phillips, P. J. & Rosenfeld, A. (2003). Face recognition: A literature survey, *ACM Comput. Surv.* 35(4): 399–458.

Zhiliang, W., Yaofeng, L. & Xiao, J. (2008). The research of the humanoid robot with facial expressions for emotional interaction, *Intelligent Networks and Intelligent Systems, 2008. ICINIS '08. First International Conference on*, pp. 416 –420.

New Approaches to Characterization and Recognition of Faces
Edited by Dr. Peter Corcoran

ISBN 978-953-307-515-0
Hard cover, 252 pages
Publisher InTech
Published online 01, August, 2011
Published in print edition August, 2011

As a baby, one of our earliest stimuli is that of human faces. We rapidly learn to identify, characterize and eventually distinguish those who are near and dear to us. We accept face recognition later as an everyday ability. We realize the complexity of the underlying problem only when we attempt to duplicate this skill in a computer vision system. This book is arranged around a number of clustered themes covering different aspects of face recognition. The first section presents an architecture for face recognition based on Hidden Markov Models; it is followed by an article on coding methods. The next section is devoted to 3D methods of face recognition and is followed by a section covering various aspects and techniques in video. Next short section is devoted to the characterization and detection of features in faces. Finally, you can find an article on the human perception of faces and how different neurological or psychological disorders can affect this.

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

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

IntechOpen