

# Geodesic Distances and Hidden Markov Models for the 3D Face Recognition

Giuseppe Mastronardi<sup>1,2</sup>, Lucia Cariello<sup>1,2</sup>,  
Domenico Daleno<sup>1,2</sup> and Marcello Castellano<sup>1,2</sup>

<sup>1</sup>*Department of Electrotechnical and Electronics, Polytechnic of Bari,  
Via Orabona, 4 - 70125 Bari - Italy*

<sup>2</sup>*e.B.I.S. s.r.l. (electronic Business In Security), Spin-Off of Polytechnic of Bari,  
Via Pavoncelli, 139 - 70125 Bari - Italy*

## 1. Personal Identification Technologies

The citizen security is a problem that in the recent years, as the daily crime news show, has grown in importance. In particular, the world research in this area, has spent his energy into the producing of a personal identification system as most possible secure. In the last years the Biometrics methods have imposed itself more and more in personal recognition field.

The word "Biometrics" is used to refer to the study of the automatic methods of identification or authorization of people which entail the use of physiologic and behavioural characteristics (Ashbourn, 2000).

The main technologies used for the identification of people and based on biometric recognition (Cascetta & De Luccia, 2004) can be divided in

- personal identification founded on the biometric recognition of "static features": capturing and processing of human anatomic characteristics such as fingerprints, geometry and "footprint" of the hand vascular, facial, iris and retinal geometry;
- personal identification founded on the recognition of "dynamic features": vocal timbre, speech peculiarities (spectral analysis of the sound field), dynamic signature (pressure), digitization (pressure), gait (steps in walking).

There are also other recognition techniques among which DNA matching that are not widely employed due to their complexity and the impossibility to work in real time (Jain & Bolle et al., 1999)( Jain & Halici et al., 1999).

### 1.1 Biometric Identification Systems

The biometrical personal recognition techniques require the use of expert systems, neural networks, fuzzy logic systems and the development of sophisticated computing. These methods offer the main advantage over traditional ones to be able to remember and learn.

For a long time, the aim of international researchers and scientists has been to create machines and systems capable of imitating certain human abilities, among which the identification based on biometric recognition or the identification through the acquisition and subsequent processing of images.

The main areas of interest of biometric technologies are:

- direct authentication and verification of personal identity, proof of the identity stated by the individual;
- indirect identification of a person through the available biometric characteristics.

The main physiologic or behavioural characteristics that may be used for personal identification must meet the following essential requirements:

- universality (each individual must have the characteristics that have been defined);
- uniqueness (it is not possible for two people to be the same in terms of the characteristics);
- permanence/invariance (biometric characteristics must remain unchanged in time or their change must be very slow and independent from changes in weight of the subjects and from the acquisition process);
- collectability (in the sense that biometric characteristic can be measured quantitatively);
- acceptability (the acquisition should not result, in case of invasive methods, not tolerated from the subjects);
- reliability (detection must be reproducible);
- reducibility (the data must be compressed and organized into easily accessible files for following use);

In the follow list are described the main “intelligent” methods of biometric recognition (Prabhakar et al., 2003) (Zhang, 2000):

- a. Hand recognition: hand geometry recognition systems measure the physical characteristics (geometry) of an individual’s hand that is palm and fingers. The uniqueness and permanence feature are questioned as there may be many causes of insidious instability, and furthermore changes may intervene over time (age, diseases, accidents, etc). This method is acceptable (when excluding the hygiene factors for the hand positioning on the capture plate), requires easy reproducible acquisition methods, does not affect the privacy, but is used only when there are minimal identification requirements;
- b. Fingerprint recognition: the fingerprint is represented as a series of ‘ridges’ and ‘valleys’ with local discontinuities in the ridge flow pattern called minutiae. The map of ridges and minutiae is unique for each individual and may change only with the destruction of the skin;
- c. Vein recognition: the map of veins on the back of the hand varies for each individual, sufficiently to secure recognition from the comparison of the respective maps. The non-invasive method has encouraged the search for the realization of sensors for recognition based on this principle: the map of veins is acquired by an infrared optical systems by temperature sensors;
- d. Face recognition: the identification of an individual through the face is based on the extraction of physiognomic parameters and metric measurements. The methodologies differ greatly depending on the type of recognition: they can be fully automatic and delegated to the appropriate software and hardware, or involving the operator for off-line identification;
- e. Digitization (pressure): Gaines and his colleagues at the Rand Corporation have demonstrated the potential of dynamic typing in 1980 by measuring the amount of latency in the pressure of two consecutive letters in the text beaten by seven professional secretaries. The applicability of this methodology requires a large set of input data. A

research of 1986 showed that the typing of own name provides reliable performance and easier applicability. Alternative methods include the use of special keyboards that also measure the pressure exerted on the letters when typing;

- f. Retina recognition: is based on retinal scanning, a special scanner illuminates the retina through the pupil with an infrared (IR) light and memorises the information from the reflection of vascular contrast. The vascular map is stable and unambiguous. The biggest problem with this methodology relates to the difficulty of acquisition that requires extensive collaboration of the user. Moreover, the costs are high, so the method is confined in applications that require secure identification, such as research facilities, military and detention;
- g. Iris recognition: iris contains parameters of high power of discrimination for who does not use drugs, does not wear corrective contact lenses or colored and is not suffering from diseases (such as glaucoma). The acquisition is performed through scanning with high definition laser light. The method has limits of applicability similar to those mentioned for the retina recognition;
- h. Vocal timbre: these systems, based on the recognition of vocal timbre, are divided into two classes according to whether approval is given by depending on the text pronounced or independent of it. In the first case the personal recognition is carried out by asking to pronounce a "password" with which the various models are loaded. To avoid the use of recordings that can "circumvent" the system, it is preferred to use more words of which, the person must be recognized, they are asked to say a few. In the case of systems independent from the text, the system does not know in advance what will be pronounced and therefore for the enrollment must be loaded many words. For correct enrollment, therefore, it requires a large quantity of recordings.

## 1.2 Biometric Recognition Applications

In the traditional personal identification systems, the individual declares his identity through an personal identification code, numeric or alphanumeric code (PIN, login or userID). The weakness of these methods of personal identification is their fraudability due to the fact that the access code or password can be stolen or deducted, and then used fraudulently by others, obviously malicious. In fact, to limit and mitigate the fraudability, and then to improve security at physical access (of people) or in the remote systems (Internet, Intranet, corporate computer networks, etc..) is possible to replace alphanumeric access codes with the biometric features of the subject.

This Section briefly reviews the main biometric recognition applications (Buciu, 2008):

*Security in the control of the physical and informatic accesses* (banks, courts, judicial police and military facilities, strategic sectors of industry, patent offices, Research & Development, etc.): can be realized accesses with turnstiles or gates (sliding doors automatically opening), by inserting the magnetic badge in appropriate readers and by validating the identity through the extraction of a biometric characteristic (fingerprint, iris, hand geometry).

*Accreditation at institutions or services* (digital signature and biometric signature).

*Counterfeiting of identity documents*: the growing need for security has prompted many states, including Italy, to realize electronic identity documents (identity cards, passports, etc.). The smart cards can store, into the chip, much more information than traditional anagraphic data: the inclusion of fingerprints or images of iris/retina, making it very efficient and

secure the recognition of the subject, traditionally given to the photograph that accompanies the anagraphic data.

*Face identification* was widely used for identifying driver licenses, in immigration programs, passports, or welfare registration.

*Access control* deals with border – crossing, vehicle access, ATM, computer access, computer network access, online transaction access, online database access.

*Security* refers to the terrorist alert issue in airports, secure boarding systems, file encryption, intranet and internet security, or medical records. Many airports have been adopted face recognition technology for improving security.

*Surveillance* is another application area where face recognition plays a major part, including video surveillance, CCTV control or portal control.

*Multimedia management* deals with face-based searching information, face-based video segmentation and summarization or event detection. Human faces are frequently seen in news, sports, films, home video, and other multimedia content. Indexing this multimedia content by face detection, face tracking, face recognition, and face change detection is important to generate segments of coherent video content for video browsing, skimming, and summarization.

*Low enforcement* is closely related to suspect tracking and investigation, identifying cheats in casinos, criminal face retrieval and recognition.

*Human – computer interaction* refers to interactive gaming and proactive computing.

Table 1 lists some of the applications of face recognition that is the scgoal of this chapter (Zhao et al., 2003).

Areas	Specific Applications
Entertainment	Video Game, Virtual Reality, Training Programs
	Human-Robot-Interaction, Human-Computer-Interaction
Smart Cards	Drivers' Licenses, Entitlement Programs
	Immigration, National ID, Passports, Voter Registration
	Welfare Fraud
Information Security	TV Parent Control, Personal Device Logon, Desktop Logon
	Application Security, Database Security, File Encryption
	Intranet Security, Internet Access, Medical Records
	Secure Tracing Terminals
Law Enforcement and Surveillance	Advanced Video Surveillance, CCTV Control
	Portal Control, Post-Event Analysis
	Shoplifting, Suspect Tracking and Investigation

Table 1. Typical applications of face recognition

### 1.3 Face Recognition

Face recognition is an innate method used by humans to recognize one another. Face recognition techniques have an advantage over the other biometric techniques in that they are non invasive and require little or no cooperation from the (passive) individual subjected to recognition as they are not susceptible to behavioural modifications (voluntary or involuntary).

The main technologies used for facial recognition are:

- Principal Component Analysis (PCA);
- Local Feature Analysis (LFA);
- Neural Networks.

Recognition systems should be designed according to the type of application and to the attitude of the individual. The latter may be of three types:

1. Cooperative: the subject is motivated to use the system in order to be recognised so that s/he can access portals and gates to certain areas;
2. Uncooperative: the subject does not help or hinder the recognition process;
3. Hostile or reticent: the subject tries to avoid recognition and shows evasive behaviour.

The human face is made up of a "multidimensional set of images". From a biometric point of view, face recognition is not characterized by high permanence: the numerous facial expressions, age, the radical changes which can have certain features as hair, hairstyle, beard, moustache, etc; the presence of eye glasses, changes in the use of cosmetics or facial color, variation of the user pose and change the direction of the incident light, are examples of exterior features that can change in time and make facial recognition difficult. The 'impermanent' features of the face add to the complexity of the technical problems to solve. Nevertheless, successful techniques for personal identification have been developed at reasonable prices.

#### 1.4 3D Face Recognition Systems

The difficulty of two-dimensional system of recognition (photographs, videotapes, etc. ...), comes from the type of data used to verify the similarity between two faces. This is because these devices are working on a two-dimensional representation in a three-dimensional scene.

The 3D Face Recognition Systems have the ability to reproduce in 3D an image, capturing the smallest detail with outstanding precision. The promises of 3D face recognition are the high accuracy of recognition necessary for high-security applications, reducing the problems of pose and lighting, the best location of facial features.

The advantages of this technology are in fact:

- The 3D technology is more efficient than two dimensions because it is able to analyze much more information because you have access to new information;
- A system of identifying 3D is less sensitive to the illumination conditions;
- The problem of the pose can be solved with the realignment of the faces;
- The occlusion can easily be found through a process of segmentation;
- The automatic generation of synthetic facial expressions;

conversely, there is that:

- The 3D acquisition hardware is expensive (the cost increases with the precision of acquisition);
- Some 3D acquisition systems are invasive;
- Long acquisition times;
- Some scanner systems can even be dangerous for the retina (laser);
- Replace 2D devices (cameras, videocameras, etc..) with 3D with new equipment is a process that requires time and high costs.

A first important classification of the techniques used is derived directly from the type of images that are presented as input to the process of recognition by dividing it into two broad categories: still images and video clips. In particular, in general applications of the automatic recognition are generally used still facial images of the subject to recognize or identify; these pictures are taken in layers of gray, as the color, almost anything used in the algorithms, presented a decidedly unfavorable report between the information conveyed

and the additional computational load required for its preparation. In some special applications, however, requires the face recognition in a complex scene and time variable, usually by video sequences taken by cameras (which would be that of people in an airport). Face recognition systems currently available are based on following approaches: *Stereoscopy/photogrammetry* is the technique that detects the position, shape and size of an object starting from its photographs. The method performs the steps without coming into contact with the object, and is based on the principle of triangulation. The subject is photographed through a pair of stereoscopic cameras, at least two different angles and from each camera is projected a "line of view" to the points of the object. The intersection of these lines provides the 3D coordinates of points of interest. The basic concept is that knowing the location of the cameras in relation to the same points in two different images, you can calculate the 3D depth by triangulation. The reconstruction algorithms are divided in two main stage: matching (search for corresponding points between images) and reconstruction (from a match is possible to obtain the 3D point associated). As regards reconstruction, there are two modes: classic stereo (with two fixed cameras) and multiple views (with a camera in motion).

*Time of flight acquisition* is the technique that derive the distance to the object framed by measuring the time that a spot light takes to reach the subject framed and come back to the receiver placed in the device itself, from time of flight can be calculated immediately the distance traveled. In particular, they are optomechanic devices capable of emitting a laser, or an electromagnetic pulse that is distorted by the surface on which an impact, and receiving the reflected signal. From latter the time interval (time of flight) and then the distance between the instrument and the point found are measured.

*Laser scanner* is based on optical-mechanical devices capable of emitting an electromagnetic pulse (the laser) and receiving the reflected signal, measuring the time interval and therefore the distance between the instrument and the point found. This system operates by measuring thousands of points in the second form of "point clouds". For each measurement ( $x, y, z$ ), the system provides the intensity of the return signal describing the surface of the object scanned. These systems, also known as active methods, project on the object a specific pattern of light and obtain the depth image by analyzing the deformation that the pattern projected undergoes. The patterns of structured light that can be used are different and it is also possible to combine different patterns in the same device. The main patterns are: binary, n-ary, gray code and phase-shift. These are among the most used because they allow also the reconstruction of subject in the slight movement. The quality of the extracted model depends on the conditions of global illumination, which limits the use of these systems to environments with controlled lighting. These systems are conceptually very similar to those in time of flight, but with the important difference that the first image is acquired in a single pulse, while the scanners require several acquisitions to reconstruct the complete 3D model of the object.

## 2. Face Recognition Methods

The Face recognition has involved during the past years several research areas: psychology, pattern recognition, neural networks, computer vision and computer graphics. It is due to this fact that the literature on face recognition is vast and diverse. In fact, many methods of face recognition are based on different principles, algorithms and a mixture of techniques.

The usage of a mixture of techniques makes it difficult to classify these systems based purely on what types of techniques they use for feature representation or classification. We propose the following categorization (Zhao et al., 2003):

- **Holistic matching methods.** These methods use the whole face region as the raw input to a recognition system. One of the most widely used representations of the face region is eigenpictures, which are based on principal component analysis.
- **Feature based (structural) matching methods.** Typically, in these methods, local features such as the eyes, nose, and mouth are first extracted and their locations and local statistics (geometric, and/or appearance) are fed into a structural classifier.
- **Hybrid methods.** Just as the human perception system uses both local features and the whole face region to recognize a face, a machine recognition system should use both. One can argue that these methods could potentially offer the best of the two types of methods.

Table 2 lists some of the techniques of face recognition.

Approach	Representative Work
<b>Holistic methods</b>	
<i>Principal Component Analysis (PCA)</i>	
Eigenfaces	Direct application of PCA
Probabilistic Eigenfaces	Two-class problem with prob. measure
Fisherfaces/Subspace LDA	FLD on eigenspace
SVM	Two-class problem based on SVM
Evolution Pursuit	Enhanced GA learning
Feature Lines	Point-to-line distance based
ICA	ICA-based feature analysis
<i>Other Representations</i>	
LDA/FLD	FLD/LDA on raw image
PDBNN	Probabilistic decision based NN
<b>Feature-based methods</b>	
Pure geometry methods	Earlier methods; Recent methods
Dynamic Link Architecture	Graph matching methods
Hidden Markov Model	HMM methods
Convolution Neural Network	SOM learning based CNN methods
<b>Hybrid methods</b>	
Modular Eigenfaces	Eigenfaces and eigenmodules
Hybrid LFA	Local feature method
Shape-normalized	Flexible appearance models
Component-based	Face region and components

Table 2. Categorization of still face recognition techniques

### 2.1 Three - Dimensional Face Recognition Methods

In the recent years, different 3D face recognition methods have been proposed, a possible categorization can include the following four groups (Pan et al., 2005):

The 3D face recognition methods that can be categorized into four groups (Pan et al., 2005):

**Curvature analysis-based:** Curvature is the intrinsic local property of curved surface. The local shape could be determined by its primary curvature and direction (Trucco & Verri,

1998). Therefore, most of the early studies used curvature to analyze the features of 3D facial data (Lee & Milios, 1990; Gordon a & b, 1991; Yacoob & Davis, 1994). Gordon presented a template-based recognition system involving descriptors based on curvature calculations from range image data. The sensed surface regions were classified as convex, concave and saddle by calculating the minimum and maximum normal curvatures. Then locations of nose, eyes, mouth and other features were determined, which were used for depth template comparison. Yacoob proposed an approach to label the components of human faces. Qualitative reasoning about possible interpretations of the components was performed, followed by consistency of hypothesized interpretations. (Zhang et al., 2002) proposed an efficient method based on Gaussian curvature analysis. It consists of three major steps, Gaussian curvature estimation, boundary detection, and region growing. Moreover, (Song et al., 2006) presented a curvature analysis based method to estimate the principal freeform feature in a specified region of interest for template fitting. Based on the minimal, maximal, mean or Gaussian curvature computing, the geometric information is transferred to the curvature domain. Using a variant of Laplacian smoothing methods, the high frequency noises and interferences in the curvature domain are suppressed and the principal feature is addressed. By feeding back the extracted feature information to the geometric shape, the geometry of the template is estimated based on the feature analysis.

**Spatial matching based:** Recognizing was performed via matching facial surface or profile directly in 3D Euclidean space (Beumier & Acheroy, 2000; Pan et al., 2003; Wu et al., 2003; Pan & Wu, 2005; Lu et al., 2004). This kind of approaches generally assume that the facial surface is a rigid object so that are not competent for the recognition among the models with expressions.

**Shape descriptor based:** It exploits the shape representation or shape descriptor to achieve the recognition task in the representation domain. For instance, (Chua et al., 2000) used point signature - a representation for free-form surfaces for 3D face recognition, in which the rigid parts of the face of one person are extracted to deal with different facial expressions. (Wang et al., 2004) used a new shape representation called Local Shape Map for 3D face recognition. This kind of technique is simple and somewhat robust to small perturbations. However, its classification rate is considered to be low. (Berretti et al., 2008) presented an approach based on integral geometric shape information of a face. The method extracts salient face information by jointly considering metric and spatial properties between points of 3D face scans. First, the face surface is partitioned into a set of iso-geodesic surfaces, centered on the nose tip and characterized by increasing values of geodesic distance. Facial information captured by the iso-geodesic surfaces is then represented in a compact form by extracting their basic 3D shape and evaluating the spatial relationships between every pairs of surfaces. This is accomplished by a modeling technique capable to quantitatively measure the spatial relationships between 3D entities. Finally, surfaces and their relationships are cast to a graphlike representation, where graph nodes are the representations of the iso-geodesic surfaces, and graph edges are their spatial relationships. In this way, the similarity between two 3D face models can be estimated extracting their graph representation and combining distances between arc labels of the two graphs.

**Recover-and-synthesis based:** For this kind of methods, their probes still are 2D images but not 3D data. The partial 3D information was recovered from the 2D image, then the facial image in virtual view was synthesized for recognition (Zhao, 1999), (Lee & Ranganath, 2003), (Hu et al., 2004) or the recognition task was accomplished with the recovered

parameters (Blaiz et al., 2002). This type of approaches do not make full use of those online-available 3D information.

### 3. The Proposed System

In this chapter the authors present a 3D face recognition system to personal identification based on Pseudo 2D Hidden Markov Models training by expression-invariant representation (*Canonical Form*) of the databases faces.

In the section 1 has been emphasized that the main problems encountered in 3D face recognition are related to deformations of the face due to different facial expressions. Fortunately, the class of transformations that can suffer the surfaces representing the human faces are not arbitrary and can be modelled as isometric transformations that preserve the lengths. One way to find a representation which is the same for all isometric surfaces was proposed by Elad and Kimmel, named "*bending invariant canonical form*" (Elad & Kimmel, 2001). In fact, from the consideration that the identity of a person is associated to the intrinsic geometry of the surface of the face, while facial expressions are associated with extrinsic geometry, begin the idea underlying this method: to represent the intrinsic geometry of the surface of the face in a form that is identical for different postures of the face.

This invariant representation is an embedding of the intrinsic geodesic structure of the surface in a finite dimensional Euclidean space, in which geodesic distances are approximated by Euclidean ones: the *Canonical Forms*.

To construct a Canonical Form it is necessary, by first, a three-dimensional surface of face of subject that is in recognition phase, called Mesh; next, to measure the geodesic distances between points on the surfaces by Fast Marching on triangulated curved surfaces, and finally the Multi-Dimensional Scaling (MDS) algorithm is applied to extract a finite dimensional flat space in which geodesic distances are represented as Euclidean ones. The first phase of system is summarized in Fig. 1.

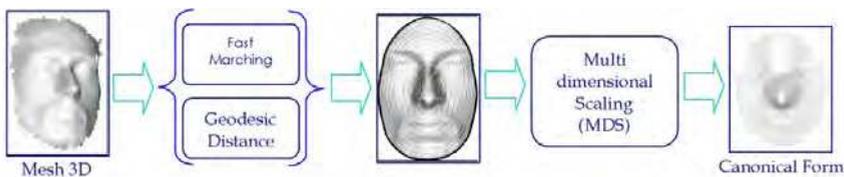


Fig. 1. Canonical Form Computation

In the second phase, the obtained expression-invariant representation of the face (*Canonical Form*) is putting into the Pseudo 2-D Hidden Markov Model with 3-6-6-6-3 structure to perform the facial recognition (see Fig. 2).

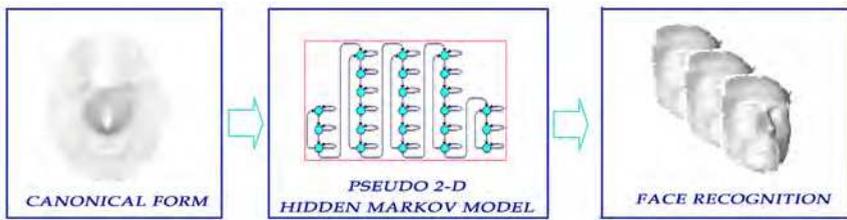


Fig. 2. Face Recognition Process

#### 4. Fast Marching Methods

The Fast Marching Method is a computational technique that approximate the solution of the "Eikonal equation" non-linear that is part of the broader class of Hamilton-Jacobi equations.

The Eikonal equations are in form:

$$|\nabla u(x)| = F(x) \text{ in } \Omega, \quad F(x) > 0 \quad (1)$$

$$u = g(x) \text{ on } \Gamma \quad (2)$$

where  $\Omega$  is a domain on  $\mathfrak{R}^n$ . The right side is generally known as it is the boundary condition that  $u$  equals a given function  $g(x)$  along the curve or surface  $\Gamma$  in  $\Omega$ .

One of the main difficulties in solving these equations is that the solution should not be differentiable even with smooth boundary data. The techniques based on Fast Marching Method are founded on two components. First, exploring the upwind "viscosity schemes", they automatically select solutions that include non-differentiability in natural ways. The second component matches the causality of these patterns with fast sorting methods borrowed from the problems of discrete networks, making the Fast Marching Method computationally efficient. Considering a domain  $\Omega$  with  $N$  total points, the complexity of these algorithms is the order of  $O(N \log N)$ .

Many applications make use of the solution of non-linear Eikonal equation such as the problem of computing accurately the distance to a curve or a surface.

For example, consider the case of  $|\nabla u| = 1$  outside the unit circle  $u = 0$  on the unit circle it is verified that the function  $u(x,y) = (x^2 + y^2)^{1/2} - 1$  corresponding to the distance function from the unit circle solves the given Eikonal equation. By specifying the right side  $F(x) = 1$  and assuming zero the boundary condition, the solution is precisely the distance from the initial curve  $\Gamma$ . The solution is shown in Fig. 3, as a set of concentric circles, and you can verify that is differentiable everywhere.

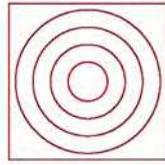


Fig. 3. Distance function to Eikonal equation  $|\nabla u|=1$

Another perspective is to imagine a disturbance propagating with unit speed away from the initial curve, and to calculate the propagation delay "first arrival time" at each point in the domain  $\Omega$ . The points where the solution is not differentiable are placed between two points on the boundary curve at the same distance. By implementing the same calculation on the surface and back (backtrack) in a way orthogonal to the curves of propagation (curves in which the propagation delay is equal for all points) we can calculate the shortest path on the manifold.

**4.1 Upwind Schemes**

To deduce the reason why to use the upwind schemes (forward) for approximating the gradient operator, consider the Eikonal equation given by

$$\begin{aligned} \sqrt{u_x^2} &= F(x) \quad u(0) = 0 \\ u(0) &= 0 \end{aligned} \tag{3}$$

The right side  $F(x) > 0$  is given, the objective is to calculate  $u(x)$  away from the boundary condition that  $u(0) = 0$ . It is possible to notice that the solution to this problem is not unique because if  $v(x)$  solves the problem, then does it also  $-v(x)$ , and then attention will focus on non-negative solutions  $u$ .

It is feasible to imagine building the solution "outwards" along the positive and negative x-axis from the origin by solving each problem separately. Considering the following ordinary differential equations:

$$\begin{aligned} \frac{du}{dx} &= F(x) \\ u(0) &= 0 \quad x \geq 0 \end{aligned} \tag{4}$$

$$\begin{aligned} \frac{du}{dx} &= -F(x) \\ u(0) &= 0 \quad x \leq 0 \end{aligned} \tag{5}$$

the right-hand side is only a function of  $x$ : a numerical quadrature is performing. Using the standard finite difference notion that  $u_i \approx u(i\Delta x)$  and  $F_i = F(i\Delta x)$ , you can approximate each of these two solutions using Euler's method,

$$\frac{u_{i+1} - u_i}{\Delta x} = F_i \quad i > 0 \quad (6)$$

$$\frac{u_i - u_{i-1}}{\Delta x} = -F_i \quad i \leq 0 \quad (7)$$

where  $u_0 = 0$ . This is an upwind scheme: it is possible to compute derivatives using points "upwind" or towards the boundary condition.

#### 4.2 The Fast Marching Method on Orthogonal Grids

In the previous section has been told that the Eikonal equation is non-differentiable and that the Fast Marching Method objective is to create an appropriate weak solution that arise from satisfying the entropy condition. The aim in this numerical scheme is the correct direction of the upwinding and treatment of sonic points. One particular upwind approximation to the gradient, is the following

$$\left[ \begin{array}{c} \max(D_{ij}^{-x}T, -D_{ij}^{+x}T, 0)^2 + \\ \max(D_{ij}^{-y}T, -D_{ij}^{+y}T, 0)^2 \end{array} \right]^{1/2} = F_{ij} \quad (8)$$

The Fast Marching Methods progress in an upwind wise to find the solution T. It is fundamental observe that in Eq. 8 the information propagates from smaller to larger values of T, so the algorithm is based on solving the Eq. 8 by building the solution outwards starting from the smallest value of T.

The Fast Marching algorithm is initialized as follows:

- The points are tagged as Alive
- All points one considered grid point away are tagged as Close
- All remaining grid points are tagged as Far

Then the algorithm proceeds as follow. The adjective "fast" is due to exploring of the points is done only for those places in a narrow strip outside of the front. The front is then moved forward in an upwind wise considering a set of points belonging to a band around the front stand; then it continues to move ahead on the points of the band; it freezes the values of existing points and, at last, incorporates some of those in front and calculate the new band around front (Chopp, 1993), (Malladi et al., 1993), (Adalsteinsson & Sethian, 1995). Ideas may be clarified by looking the main loop of method:

- 1) Points with the smallest value of T are tagged with Tial;
- 2) The Trial points becomes Alive and takes off from Close;
- 3) All Trial neighbors that are not Alive are tagged as Close and, if they are in Far, are removed and added to Close group;
- 4) The values of T for all the neighbors from the Eq. 8 are recalculated by solving the quadratic equation with only the values of the points that are Alive;
- 5) The cycle begins again from 1).

The algorithm works because the process of recalculation of the T values of upwind neighboring points produces a value less than those of accepted points (those in Alive). This feature, called monotone properties, allows to move forward by choosing the narrow strip of Trial points with the minimum value of T. In this way is possible to change neighbors without having to go back and to adjust the accepted values as is shown in Figure 4.

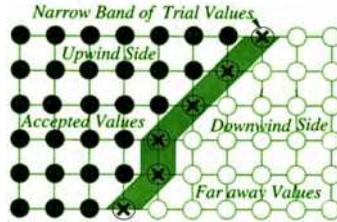


Fig. 4. Upwind construction of accepted values.

The Fast Marching methods are used to solve the Eq.10 by means of an update procedure, where it can imagine an uniform square grid and the goal is to update the value of T at the center point  $i,j$ . The Fast Marching methods, also, can be extended on a particular Triangulated Planar Domain (Kimmel & Sethian, 1998). To do it is necessary built a monotone update procedure on the Triangulated Mesh or, generally on an Arbitrary Triangulation (acute triangulation and obtuse triangulation) (Kimmel & Sethian, 1998).

**5. GEODESIC DISTANCE**

The Fast Marching Method algorithm can be used to compute distances on triangulated manifolds, and, hence, to construct minimal geodesics. Before proceeding, is useful to say a few words about a geodesic distances.

The geodesic distance between two pixels p and q is the length of the shortest path from p to q. Suppose  $P = \{p_1, p_2, \dots, p_n\}$  is a path between pixels  $p_1$  and  $p_n$ , i.e.  $p_i$  and  $p_{i+1}$  are connected neighbors for  $i \in \{1, 2, \dots, n-1\}$  and  $p_i$  belong to the domain for all i. The path length  $l(P)$  is defined as:

$$l(P) = \sum_{i=1}^{n-1} d_N(p_i, p_{i+1}) \tag{9}$$

the sum of the neighbors distances  $d_N$  between adjacent points in the path.

If the idea is to use the Fast Marching on triangulated manifolds first is necessary to solve the Eikonal equation on the triangulated surface with speed  $|\nabla T| = F = 1$  to calculate the distance from a source point; then you can then go back (backtrack) along the gradient by solving the ordinary differential equation

$$\frac{dX(s)}{ds} = -\nabla T \tag{10}$$

where at the source points the distance is zero and  $X(s)$  is the geodesic path. Successively, using the Heun's method of integration of the second order on Triangulated surface is possible to switch to a first order. Applying the Fast Marching method to triangulated domains requires a careful analysis of the update of one vertex in a triangle, while the  $T$  values at the other vertices are given: when the inner part of the triangle is integrated, three neighboring triangles are used to interpolate  $T$  with a polynomial second-order, whose six coefficients calculated from data values associated with  $T$  vertices. The fast marching on triangulated domains method can compute the geodesic distance between one vertex and the rest of the  $n$  surface vertices in  $O(n)$  operations. Repeating this computation for each vertex, we compute a geodesic distance matrix  $\Delta$  in  $O(n^2)$  operations. So we compute a geodesic distance matrix  $\Delta$ , where

$$\delta(p_i, p_j) = \delta_{ij} \quad (11)$$

Each  $ij$  entry of  $\Delta$  represents the square geodesic distance between the vertex  $i$  and the vertex  $j$ , that is  $\delta_{ij} = \text{GeodesicDistance}(\text{Vertex}_i; \text{Vertex}_j)$ , where

$$[\Delta] = \delta_{ij}^2 \quad (12)$$

Thereby, given a triangulated surface, we apply the fast marching procedure for each vertex in a subset of the vertices as a source point and obtain the geodesic distance matrix,  $\Delta$ .

## 6. Multidimensional Scaling

The main problem of human face is that it can not be considered a rigid object because it undergoes deformations resulting from facial expressions. So classical surface matching methods, aimed to find a Euclidean transformation of two surfaces which maximizes some shape similarity criterion, are more suitable for rigid objects than moving objects. The second problem is that facial surface has class of transformations that are not arbitrary, and in literature facial expressions can be modeled as isometric (or length-preserving) transformations that do not stretch and do not tear the surface, or more strictly, preserve the surface metric. When is studied a human face, the problem is implemented a deformable surface matching algorithm is to find a representation, which is the same for all isometric surfaces.

One of the algorithms most widely used for this purpose is the Multi-Dimensional Scaling (MDS). The idea is, preserving the relative distances between pairs of points, to transform a set of points in a high-dimensional space to a lower-dimensional one (Rohde, 2002). In most cases, in fact, is very complex to carry out results from studies handling high-dimensional vector spaces. So, if it is possible to obtain a set of vectors in a much lower-dimensional space, while preserving their similarity structure, the operations could be performed more efficiently. In other words multi-dimensional scaling (MDS) techniques are applied to extract a finite dimensional flat space in which geodesic distances are represented as Euclidean ones.

In the MDS method, the dissimilarity between pairs of objects of a collection are given as coefficients and then approximated as distances between corresponding pairs of entities in

the visual representation. The quality of this approximation is expressed as a loss function, which produces the best value to its minimum. It is possible to say that Richardson (Richardson, 1938) and Young and Householder (Young & Householder, 1938) may have officially initiated the multidimensional scaling literature, but frequent applications did not begin to appear until the papers by Torgerson (Torgerson, 1952) and, successively, by Shepard and Kruskal. Torgerson used a one-dimensional scaling technique to convert dissimilarity ratings to target distances and then attempted to find a set of points whose pairwise Euclidean distances best matched the target distances, according to mean-squared error. Though very effective, this technique is too complicated for many applications and a serious problem to another. An improvement occurred with Shepard (Shepard, 1962) that supposed that the goal of MDS should be to obtain a monotone relationship between the actual point distances. In literature Torgerson's method has spread as MDS metric. On the contrary, the technique introduced by Shepard is known as non-metric MDS. Kruskal (Kruskal, 1964) further developed the procedure and defined the function known as stress function, relating the pairwise distances and the ranking of dissimilarities. The basic technique developed by Shepard and Kruskal remained the standard for most applications of MDS.

Schwartz (Schwartz et al, 1989), were the first to use multidimensional scaling (MDS) as a tool for studying curved surfaces by planar models. In his work, he applied an MDS technique to flatten convoluted cortical surfaces of the brain, onto a plane, in order to study their functional architecture. Zigelman (Zigelman et al., 2002) and Grossman (Grossman et al., 2002) extended some of these ideas to the problem of texture mapping and voxel-based cortex flattening. A generalization of this approach was introduced in the recent work of Elad e Kimmel (Elad & Kimmel, 2001), that proposed an efficient algorithm to construct a signature for isometric surfaces.

The problem can be so defined: suppose you have a collection of  $n$  objects and a way to determine the differences between each pair

$$\Delta = [\delta_{ij} : i, j = 1, \dots, n] \tag{13}$$

The metric Multidimensional Scaling is a procedure to find a configuration of  $n$  points in a space of  $p$  size, usually Euclidean, so the point

$$x_i^* = (x_{i1}^*, \dots, x_{ip}^*)^T \tag{14}$$

represents uniquely the object  $i$ , and the Euclidean distance between points  $x_i^*$  and  $x_j^*$  is:

$$d_{ij}(X^*) = \|x_i^* - x_j^*\| = \sqrt{\sum_{a=1}^p (x_{ia}^* - x_{ja}^*)^2} \tag{15}$$

The corresponding dissimilarity  $\delta_{ij}$  for all pairs of objects  $(i, j)$  is so approximated:

$$\forall_{i < j} d_{ij}(X^*) \approx \delta_{ij} \tag{16}$$

Is usually sufficient to consider each pair of objects (i, j) only once with  $i < j$ , since differences are symmetrical. Asymmetric matrix elements  $\hat{\Delta}$  must be mediated:

$$\delta_{ij} = \delta_{ji} = (\hat{\delta}_{ij} + \hat{\delta}_{ji}) / 2 \quad (17)$$

For a given configuration  $X$ , the approximation error that is obtained to represent the dissimilarity between objects  $i$  and  $j$  can be defined as follows:

$$e_{ij} \stackrel{def}{=} |d_{ij}(X) - \delta_{ij}| \quad (18)$$

The MDS - least squares technique (least-squares) defines the loss function as a weighted sum of normalized errors and possibly on all pairs of objects (i, j), and penalizes the overall approximation error. The minimum of this function on  $X$  is found through numerical optimization to obtain the desired configuration  $X^*$ .

Each of the MDS algorithms is an instance of a heuristic minimization function well known. This heuristic provides a theoretical basis and is usually the result of a large number of options and variations derived from operations research professionals and other fields. Some of the choices are determined simply by the application domain, others may be based on theoretical or empirical evidence found in the literature.

## 7. Hidden Markov Models

Hidden Markov models (HMMs) are a well-developed technology for classification of multivariate data that have been used extensively in speech recognition, handwriting recognition and even sign recognition. They consist of states, possible transitions between states (and the probability of those transitions being taken) and a probability that in a particular state, a particular observation is made (see Fig. 5). The "Hidden" mean that the state of the HMM can not, in general, be known by looking at the observations. Then, they are also "Markov" because the probability of observing an output depends only on the current state and not on previous states. By looking at the observations, using an algorithm known as the Viterbi Algorithm, an estimate of the probability that a particular instance or stream observed was generated by that HMM can be computed (Rabiner, 1989).

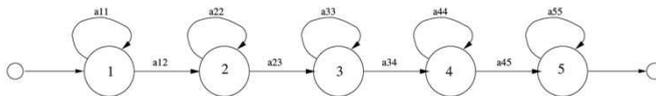


Fig. 5. Schema of Hidden Markov Models

For every class of interest, is possible apply HMMs, to calculate the probability returned by each of them and choose the most probable one.

### 7.1 One - Dimensional Hidden Markov Models

The main idea of this structure is to design a multi-state system which outputs this sequence while being in a state  $q_t$  at a given time  $t$  (Samaria & Young, 1994). At each state, the system is designed to output a certain observation from the vocabulary with a likelihood given by the output probabilities. At each time step the system will switch from its current state to the next (possibly stay in the same state) with a transition probability. The given of the transition probabilities will therefore implicitly design the (hidden) structure of the system. The state of the system at time  $t = 1$  is given by its initial state probabilities.

The elements that characterised a HMMs are:

- $N = |S|$  is the number of states of the model. If  $S$  is the set of states, then  $S = \{s_1, s_2, \dots, s_n\}$ .  $s_i$  is one of the states that can be employed by the model. To observe the system are used  $T$  observation sequences, where  $T$  is the number of observations. The state of the model at time  $t$  is given by  $q_t$  in  $S$ ,  $1 < t < T$ ;
- $M = |V|$  is the number of different observation symbols. If  $V$  is the set of all possible observation symbols (also called the codebook of the model), then  $V = \{v_1, v_2, \dots, v_M\}$ ;
- $A = \{a_{ij}\}$  is the state transition probability matrix, where  $a_{ij}$  is the probability that the state  $i$  became the state  $j$ :

$$a_{ij} = p(q_t = s_j \mid q_{t-1} = s_i) \quad (19)$$

where  $1 \leq i, j \leq N$ , with constraint  $0 \leq a_{ij} \leq 1$ , and  $\sum_{j=1}^N a_{ij} = 1$ ,  $1 \leq i \leq N$

- $B = \{b_j(k)\}$  the observation symbol probability matrix,  $b_j(k)$  is the probability to have the observation  $k$  when the state is  $j$ :

$$b_j(k) = p(O_t = v_k \mid q_t = S_j) \quad (20)$$

where  $1 \leq j \leq N$ ;  $1 \leq k \leq M$ ; and  $O_t$  is the observation symbol at time  $t$ .

- $\Pi = \{\pi_1, \pi_2, \dots, \pi_N\}$  is the initial state distribution:

$$\pi_i = p(q_1 = S_i) \quad (21)$$

where  $1 \leq j \leq N$ .

Using a shorthand notation, a HMM is defined by the following expression:

$$\lambda = (A, B, \pi). \quad (22)$$

The training of the model, given a set of sequences  $\{O_i\}$ , is usually performed using the standard Baum-Welch re-estimation, which determines the parameters  $(A, B, \pi)$  that maximize the probability  $P(\{O_i\} \mid \lambda)$ .

**7.2 Two - Dimensional Hidden Markov Models**

Pseudo Two-Dimensional Hidden Markov Models (P2D-HMMs) can be considered how a generalization of the One-Dimensional HMMs, helpful to represent two-Dimensional data. A word “Pseudo” is used because of the state alignments of consecutive columns are calculated independently of each other (Werner & Rigoll, 2001).

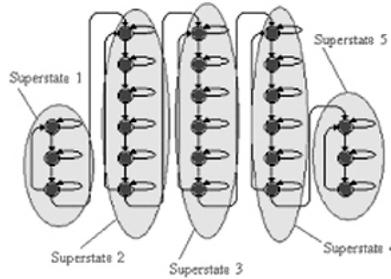


Fig. 6. Schema of Pseudo 2Dimensional Hidden Markov Model

In Fig. 6, P2D-HMMs are organized in a two-dimensional way: the horizontal states are named superstates; each of them consists of a One-dimensional HMMs in vertical direction. For recognition process a P2D-HMM can be transformed into an equivalent one-dimensional HMM.

The HMMs can be trained by the standard Baum-Welch algorithm and the recognition step can be carried out using the standard Viterbi algorithm.

The elements of an P2D-HMM are [Nefian & Hayes III, 1999] :

- $N_0$  are the number of superstates, and  $S_0 = \{S_{0,i} \mid 1 \leq i \leq N_0\}$  the set of superstates.
- The initial superstate distribution,  $\Pi_0 = \{\pi_{0,i}\}$ , where  $\pi_{0,i}$  are the probabilities of being in super state  $i$  at time zero.
- The super state transition probability matrix,

$$A_0 = \{a_{0,ij}\} \tag{23}$$

where  $a_{0,ij}$  is the probability of transitioning from super state  $i$  to superstate  $j$ .

- The parameters of the P2D-HMMs, which include:
  - ✓ The number of embedded states in the  $k$ th superstate,  $N_1^{(k)}$ , and the set of embedded states,  $S_1^{(k)} \{S_{1^{(k)},i}\}$ .
  - ✓ The initial state distribution,  $\Pi_1^{(k)} = \{\pi_{1^{(k)},i}\}$ , where  $\pi_{1^{(k)},i}$  are the probabilities of being in state  $i$  of super state  $k$  at time zero.
  - ✓ The state transition probability of transitioning from state  $k$  to state  $j$ .
- Finally, there is the state probability matrix,

$$B^{(k)} = \{b_{1^{(k)}}(O_{t_0, t_1})\} \tag{24}$$

for the set of observations where  $O_{t_0, t_1}$  represent the observation vector at row  $t_0$  and column  $t_1$ . In a *continuous density* HMM, the states are characterized by continuous observation density functions. The probability density function that is typically used, is a finite mixture of the form

$$b_i^{(k)}(O_{i_0,i_1}) = \sum_{m=1}^M c_{im}^{(k)} N(O_{i_0,i_1}, \mu_{im}^{(k)}, U_{im}^{(k)}) \tag{25}$$

where  $1 \leq i \leq N_1^{(k)}$ ,  $c_{im}^{(k)}$  is the mixture coefficient for the  $m$ th mixture state  $i$  of super state  $k$ .

$N(O_{i_0,i_1}, \mu_{im}^{(k)}, U_{im}^{(k)})$  is a Gaussian with mean vector  $\mu_{im}^{(k)}$  and covariance matrix  $U_{im}^{(k)}$ . Let

$$\Lambda^{(k)} = \{\Pi_1^{(k)}, A_1^{(k)}, B_1^{(k)}\} \tag{26}$$

be the set of parameters that define the  $k$ th super state. Using a shorthand notation, an embedded HMM is defined as the triplet:

$$\lambda = (\Pi_0, A_0, \Lambda) \tag{27}$$

where:

$$\Lambda = \{\Lambda^{(1)}, \Lambda^{(2)}, \dots, \Lambda^{(N_0)}\} \tag{28}$$

Although more complex than a one-dimensional HMM, a P2D-HMM is more appropriate for data that are two-dimensional, and has a complexity proportional to the sum of the squares of the number of states:

$$\sum_{k=1}^{N_0} (N_1^{(k)})^2 \tag{29}$$

In Fig. 3 it's shown an example of a P2D-HMM with a structure 3-6-6-6-3: the superstates 1 and 5 are constituted by a left to right 1D-HMM with 3 states; instead the superstates 2, 3 and 4 are constituted by a left to right 1D-HMM with 6 states. This is a particular structure used in the face recognition system realized by the authors and proposed in the followed sections.

### 8. The Mesh and GAVADB 3D face database

The Mesh, three-dimensional surfaces of face of subject, is in a ASE format, chosen for its excellent readability. A file ASE, in fact, is perfectly compatible with a text file, it opens up with a common notepad to show the content. The 3D image is a graphical model with a surface constructed from polygons. The polygons are described by the graphics system as solid faces, rather than as hollow polygons, as is the case with wireframe models. Separate portions of mesh that make up the model are called polygon mesh and quadrilateral mesh. The mesh is stored as three-coordinates points  $x$ ,  $y$  and  $z$  of each point of the mesh and the triangles defined by the points themselves. The number of a three-coordinates points are variable, in this system a Mesh have 3000 points.

To test the proposed system the GAVADB database (Moreno & Sanchez, 2004) is used. It is a 3D face database. It contains 549 three-dimensional images of facial surfaces. This database has 9 images for each of the 61 different individuals with 45 male and 16 female. The total of the individuals are Caucasian and their age is between 18 and 40 years old. The database

provides systematic variations with respect to the pose and the facial expression. In particular there are: 2 neutral expressions with frontal views; 2 neutral expressions with views x-rotated with  $\pm 30^\circ$ , looking up and looking down respectively; 2 neutral expressions with views y-rotated  $\pm 90^\circ$ , left and right profiles respectively; 3 frontal gesture images laugh, smile and a random gesture chosen by the user, respectively.

### 8.1 The Implementation

Given a polyhedral approximation (Mesh) of the facial surface,  $S$ , the fast marching on triangulated domains method computes the geodesic distance between one vertex and the rest of the  $n$  surface vertices. Repeating this computation for each vertex, we compute a geodesic distance matrix  $\Delta$ , where for each  $ij$  entry of  $\Delta$  represents the square geodesic distance between the vertex  $i$  and the vertex  $j$ , that is  $\delta_{ij} = \text{GeodesicDistance}(\text{Vertex}_i; \text{Vertex}_j)$ , with  $[\Delta] = \delta_{ij}^2$  (see Eq. 11, Eq. 12). One of the crucial steps in the construction of the canonical form of a given surface, is an efficient algorithm for the computation of  $\delta_{ij}$ , that is the geodesic distances.

The resulting matrix  $\Delta$  is invariant under isometric surface deformations, but is not a unique representation of isometric surfaces since it depends on arbitrary ordering and the selection of the surface points. Treating the squared mutual distances as a particular case of *dissimilarities*, it is possible to apply a dimensionality-reduction technique, that is *multidimensional scaling* (MDS), in order to embed the surface into a low-dimensional Euclidean space  $\mathfrak{R}^m$ . This is equivalent to finding a mapping between two metric spaces,

$$\begin{aligned} \varphi : (S, \delta) &\rightarrow (\mathfrak{R}^m, d) \\ \varphi(p_i) &= x_i \end{aligned} \quad (30)$$

that minimizes the embedding error,

$$\begin{aligned} \varepsilon &= f\left(\left|\delta_{ij} - d_{ij}\right|\right) \\ d_{ij} &= \|x_i - x_j\|_2 \end{aligned} \quad (31)$$

for some monotone function  $f$  that sums over all  $ij$ . The obtained  $m$ -dimensional representation is a set of points  $x_i \in \mathfrak{R}^m (i = 1, \dots, n)$ , corresponding to the surface points  $p_i$ . Different MDS methods can be derived using different embedding error criteria (Borg et al. 1997). A particular case is the *classical scaling*, introduced by Young and Householder (Young et al. 1938), and that is a method here applied. The embedding in  $\mathfrak{R}^m$  is performed by double-centering the matrix  $\Delta$

$$B = -\frac{1}{2} J \Delta J \quad (32)$$

(here  $J = I - \frac{1}{n} U$ ;  $I$  is a  $n \times n$  identity matrix, and  $U$  is a matrix consisting entirely of ones). The first  $m$  eigenvectors  $e_i$ , corresponding to the  $m$  largest eigenvalues of  $B$ , are used as the embedding coordinates

$$\begin{aligned} x_i^j &= e_i^j \\ i &= 1, \dots, m; j = 1, \dots, n \end{aligned} \quad (33)$$

where  $x_i^j$  denotes the  $j$ -th coordinate of the vector  $x_i$ . We refer to the set of points  $x_i$  obtained by the MDS as the *canonical form* of the surface. In the described work it is obtained a surface representation with  $m=20$ .

The autovectors values obtained by MDS that give the best characterization of a face surface, calculated Canonical Form, are saved in the binary format HTK (Hidden Markov Model ToolKit) in order to use the Hidden Markov Model to perform the face recognition. The binary file has an extension .bin, and contains the data saved in the format Big Endian, which is commonly used by Motorola processors (family 680x0), IBM and Sun as opposed to Intel and AMD's x86 family.

The Hidden Markov Model Toolkit (HTK) (Young and Young, 1994) is a toolkit for building and manipulating Hidden Markov Models. HTK is primarily used for speech recognition research although it has been used for numerous other applications such as the face recognition. HTK consists of a set of library modules and tools to train and to test the HMM and to run a recognition. In the proposed system is used the Pseudo 2-D HMM with topology 3-6-6-6-3 shown in Fig. 7.

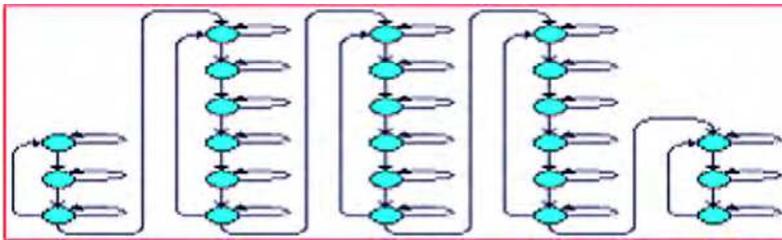


Fig. 7. Schema of used Pseudo 2D HMM

## 9. Experimental Results

After training of P2D-HMM on canonical forms constructed from GAVADB database, the authors performed several experiments and the results were encouraging, indeed the system achieved a rate of recognition equal to 98%. The robustness of the system was put to the test by the presence of various noise situations: incompleteness of mesh in the form of holes and occlusions. Moreover, comparing these outcomes with others different from other methods proposed in the literature and applied on the same 3D database, the proposed method results outperform all of them (see Table 2).

System	Authors	Results
Fishersurface (LDA)	(Heseltine et al., 2004a)	88.7%
Eigensurface (PCA)	(Heseltine et al., 2004b)	87.3%
Fishersurface (LDA)	(Heseltine et al., 2004c)	88.4%
3D matching + gray level	(Beumier & Acheroy, 2000)	98%
Morphable model [19]	(Blanz et al., 2002)	89%
3D eigenfaces [20]	(Xu et al, 2004)	71.1%
<b>The presented system</b>		<b>98%</b>

Table 2. Comparative results

Observing the Table 2, it is possible to notice that the proposed 3D facial recognition system obtained training a 3-6-6-6-3 P2D-HMM by expression-invariant representation (*Canonical Form*) of the databases faces, has reached reasonably good results, encouraging the authors to continue in this area for future developments in several directions such as: increasing the numbers of mesh points, that now are 3000; improving the Multi-Dimensional Scaling algorithm; reducing implementation time necessary to calculate the Canonical Form.

## 10. Reference

- Adalsteinsson, D. & Sethian, J. A. (1995). A fast level set method for propagating interfaces. *Journal of Computational Physics*, Vol. 118, No. 2, 269-277
- Ashbourn, J. (2000). *Biometrics: Advanced Identity Verification, The Complete Guide*, Springer, London
- Berretti, S.; Del Bimbo, A. & Pala, P. (2008). SHREC'08 Entry: 3D Face Recognition using Integral Shape Information, *Proceedings of IEEE International Conference on Shape Modeling and Applications*, Stony Brook, New York, USA, pp. 255-256, June 2008
- Beumier, C. & Acheroy, M. (2000a) Automatic Face Verification from 3D and Grey Level Clues, *Proceedings of RECPAD2000, 11th Portuguese conference on pattern recognition*, pp. 95-101, Porto, Portugal, May 11-12, 2000
- Beumier, C. & Acheroy, M. (2000b). Automatic 3D Face Authentication. *Image and Vision Computing*, Vol. 18, 315-321
- Blanz, V.; Romdhani, S. & Vetter, T. (2002). Face Identification across Different Poses and Illumination with a 3D Morphable Model, *Proceedings of IEEE International Automatic Face and Gesture Recognition (FGR'02)*, pp. 202-207, 2002
- Borg, I. & Groenen, P. (1997). *Modern multidimensional scaling - theory and applications*, Springer-Verlag, ISBN 0-3879-4845-7, New York
- Buciu, I. (2008). Overview of Face Recognition Techniques. *IEEE 2008*, 173-176
- Cascetta, F. & De Luccia, F. (2004). Personal Identification Systems. *Mondo Digitale*, No. 9
- Chopp, D. L. (1993). Computing Minimal Surfaces via Level Set Curvature Flow. *Journal of Computational Physics*, Vol. 106, 77-91
- Chua, C.S.; Han, F. & Ho, Y.K. (2000). 3D human face recognition using point signature, *Proceedings of IEEE International Automatic Face and Gesture Recognition (FRG'00)*, pp. 233-238, 2000
- Elad, A. & Kimmel, R. (2001). Bending Invariant Representations for Surfaces, *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'01)*, Vol. 1, pp. 168-174
- Gordon, G.G. (1991). Face Recognition based on Depth Maps and Surface Curvature. *Geometric Methods in Computer Vision, SPIE Proceedings*, Vol. 1570, 234-247
- Grossman, R.; Kiryati, N. & Kimmel, R. (2002). Computational surface flattening: a voxel-based approach, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 4, 433-441, ISSN 0162-8828
- Heseltine, T.; Pears, N. & Austin, J. (2004a). Three-Dimensional Face Recognition: A Fishersurface Approach, *Proceedings of International Conference on Image Analysis and Recognition (ICIAR 2004)*, pp. 684-691

- Heseltine, T.; Pears, N. & Austin, J. (2004b). Three-Dimensional Face Recognition: An Eigensurface Approach, *Proceedings of International Conference on Image Processing (ICIP 2004)*, pp. 1421-1424, ISBN 0-7803-8554-3, IEEE, New York
- Heseltine, T.; Pears, N. & Austin, J. (2004c). Three-Dimensional Face Recognition Using Surface Space Combinations, *Proceedings of British Machine Vision Conference, 2004*
- Hu, Y.; Jiang, D.; Yan, S.; Zhang, L. & Zhang, H. (2004). Automatic 3D Reconstruction for Face Recognition, *Proceedings of IEEE International Automatic Face and Gesture Recognition (FGR'04)*, pp. 17-19, Seoul, Korea, 2004
- Jain, A.; Bolle, R. & Pankanti, S. (1999). *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Press, Boston
- Jain, L.C.; Halici, U.; Hayashi, I.; Lee, S.B. & Tsutsui, S. (1999). *Intelligent biometric techniques in fingerprint and face recognition*, CRC Press, Boca Raton, Florida, USA
- Kimmel, R. & Sethian, J. A. (1998). Computing geodesic paths on manifolds, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 95, No. 15, pp. 8431-8435
- Kruskal, J. B. (1964). Multidimensional scaling by optimising goodness of fit to a nonmetric hypothesis. *Psychometrika*, Vol. 29, 1-27
- Lee, J.C. & Milios, E. (1990). Matching Range Images of Human Faces, *Proceedings of IEEE 3rd International Conference on Computer Vision (ICCV 1990)*, pp. 722-726
- Lee, M.W. & Ranganath, S. (2003). Pose-invariant face recognition using a 3D deformable model. *Pattern Recognition*, Vol. 36, No. 8, 1835-1846
- Lu, X.; Colbry, D. & Jain, A.K. (2004). Three-Dimensional Model Based Face Recognition, *Proceedings of 17th International Conference on Pattern Recognition*, pp. 362-366, 2004
- Malladi, R.; Sethian, J. A. & Vemuri, B. C. (1995). Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 2, 158 - 175
- Moreno, A.B. & Sanchez, A. (2004). GavabDB: A 3D Face Database, *Proceedings of 2nd COST Workshop on Biometrics on the Internet: Fundamentals, Advances and Applications*, pp. 75-80, Vigo, Spain, March 2004
- Nefian, A.V. & Hayes, M.H. III. (1999). An embedded HMM-based approach for face detection and recognition. *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP Proceedings.*, pp. 3553-3556 vol.6, 1999
- Pan, G.; Wu, Z. & Pan, Y. (2003). Automatic 3D Face Verification from Range Data. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)*, Vol. 3 (2003), 193-196
- Pan, G.; Han, S.; Wu, Z. & Wang, Y. (2005). 3D Face Recognition using Mapped Depth Images, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 175-181
- Pan, G. & Wu, Z. (2005). 3D Face Recognition from Range Data. *International Journal of Image and Graphics*, Vol. 5, No. 3, 573-593
- Prabhakar, S.; Pankanti, S. & Jain A. K. (2003). Biometric Recognition: Security and Privacy Concerns. *IEEE Security & Privacy*, pp. 33-42, March-April 2003
- Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, pp. 257 - 286 Vol. 77 Issue 2, Feb 1989
- Richardson, M. W. (1938). Multidimensional psychophysics, In: *Psychological Bulletin*, Vol. 35, 659-660

- Rohde, D.L.T. (2002). Methods for binary multidimensional scaling. *Neural Computation*, Vol. 14, No. 5, 1195-1232, ISSN 0899-7667
- Samaria, F. & Young, S. (1994). HMM based architecture for face identification. *Image and computer Vision*, pp.537-543 vol. 12, October, 1994
- Schwartz, E.L.; Shaw, A. & Wolfson, E. (1989). A numerical solution to the generalized mapmaker's problem: flattening nonconvex polyhedral surfaces. *IEEE Transactions on Pattern Analysis and machine Intelligence*, Vol. 2, No. 9, Sept. 1989, 1005-1008
- Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika*, Vol. 27, 125-139, 219-246
- Song, Y.; Vergeest, J.S.M.; Wigers, T. & Langerak, T.R. (2006). Estimating principal deformable freeform features by curvature analysis, *Proceedings of 7th International Conference on Computer-Aided Industrial Design and Conceptual Design (CAIDCD '06)*, pp. 1-6, November 2006
- Torgerson, W. S. (1952). Multidimensional scaling: I. Theory and method. *Psychometrika*, Vol. 17, 401-419
- Trucco, E. & Verri, A. (1998). *Introductory Techniques for 3-D Computer Vision*, Prent. Hall Inc.
- Wang, Y.; Pan, G. & Wu, Z. (2004). 3D Face Recognition using Local Shape Map, *Proceedings of IEEE International Conference on Image Processing (ICIP'04)*, 2004
- Werner, S. & Rigoll, G. (2001). Pseudo 2-Dimensional Hidden Markov Models in Speech Recognition. *IEEE Workshop on Automatic Speech Recognition and Understanding*, pp. 441-444
- Wu, Y.; Pan, G. & Wu, Z. (2003). Face Authentication based on Multiple Profiles Extracted from Range Data. *AVBPA'03, LNCS*, pp. 515-522 vol. 2688
- Xu, C.; Wang, Y.; Tan, T. & Quan, L. (2004). A new attempt to face recognition using eigenfaces, *Proceedings of the 6th Asian Conf. on Computer Vision*, Vol. 2, pp. 884-889
- Yacoob, Y. & Davis, L.S. (1994). Labeling of Human Face Components from Range Data. *CVGIP: Image Understanding*, Vol. 60, No.2, 168-178
- Young, G. & Householder, G. S. (1938). Discussion of a set of points in terms of their mutual distances. *Psychometrika* Vol. 3, No. 1, Roger E. Millsap, (Ed.), 19-22
- Young, S.J. (1994). The HTK Hidden Markov Model Toolkit: Design and Philosophy. *Entropic Cambridge Research Laboratory, Ltd*, Vol., 2, 2-44
- Zhang, D. (2000). *Automated Biometrics Technologies and Systems*, Kluwer Academic Publishers, Boston
- Zigelman, G.; Kimmel, R. & Kiryati, N. (2002). Texture mapping using surface flattening via multi-dimensional scaling. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 8, No. 2, 198-207



## Face Recognition

Edited by Milos Oravec

ISBN 978-953-307-060-5

Hard cover, 404 pages

**Publisher** InTech

**Published online** 01, April, 2010

**Published in print edition** April, 2010

This book aims to bring together selected recent advances, applications and original results in the area of biometric face recognition. They can be useful for researchers, engineers, graduate and postgraduate students, experts in this area and hopefully also for people interested generally in computer science, security, machine learning and artificial intelligence. Various methods, approaches and algorithms for recognition of human faces are used by authors of the chapters of this book, e.g. PCA, LDA, artificial neural networks, wavelets, curvelets, kernel methods, Gabor filters, active appearance models, 2D and 3D representations, optical correlation, hidden Markov models and others. Also a broad range of problems is covered: feature extraction and dimensionality reduction (chapters 1-4), 2D face recognition from the point of view of full system proposal (chapters 5-10), illumination and pose problems (chapters 11-13), eye movement (chapter 14), 3D face recognition (chapters 15-19) and hardware issues (chapters 19-20).

### How to reference

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

Giuseppe Mastronardi, Lucia Cariello, Domenico Daleno and Marcello Castellano (2010). Geodesic Distances and Hidden Markov Models for the 3D Face Recognition, Face Recognition, Milos Oravec (Ed.), ISBN: 978-953-307-060-5, InTech, Available from: <http://www.intechopen.com/books/face-recognition/geodesic-distances-and-hidden-markov-models-for-the-3d-face-recognition>

# INTECH

open science | open minds

### InTech Europe

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

### InTech China

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

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.