Année 2011 Num d'ordre: 40556

Université Lille1

THÈSE

pour obtenir le grade de
Docteur,
spécialité Informatique



présentée et soutenue publiquement par

Hassen DRIRA

le 4 juillet 2011

Statistical Computing On Manifolds for 3D Face Analysis and Recognition

préparée au sein du laboratoire LIFL

sous la direction de

Mohamed DAOUDI
Boulbaba BEN AMOR

COMPOSITION DU JURY

M. Liming Chen	(Professeur, Ecole Centrale de Lyon, France)	Président
M. Remco Veltkamp	(Professeur, Université de Utrecht, Pays-bas)	Rapporteur
M. Pietro Pala	(Associate Professor, Université de Florence, Italie)	Rapporteur
M. Anuj Srivastava	(Professeur, Florida State University)	Examinateur
M. Mohamed Daoudi	Professeur TELECOM Lille1/LIFL	Directeur
M. Boulbaba Ben Amor	(Maître de conférences, TELECOM Lille1/LIFL)	Encadrant

Acknowledgements

First of all, I would like to thank my thesis advisors for their trust, and more specifically for having made me discover scientific research and the two research fields of top excitement that are Shape Analysis and Biometric. During these three years, Pr. Mohamed Daoudi has been a mentor to me, even more than at a pure scientific level. I would like to express my gratitude to Dr. Boulbaba Ben Amor for his outstanding human qualities and its unconditional encouragements which definitely helped me mutate into a grown-up researcher. I would also like to thank Pr. Srivastava, who has provided me with insights into the research process, and for inviting me during two mouths in his laboratory.

A special thanks is due my PhD committee members for taking the time to participate in this process, and especially the reviewers of the manuscript for having accepted this significant task: M. Pietro Pala (Associate professor at University of Florence) and M. Remco Veltkamp (Professor at University of Utrecht) and kindly accepted to sacrifice their precious time for the manuscript reviewing.

I also thank the committee members M. Liming Chen (Professor at Ecole Centrale Lyon) and M. Anuj Srivastava (Professor at Florida State University). All these people made me the honor to be present for this special day despite their highly charged agendas and Im sincerely grateful for that.

A special thank goes to my research buddies and especially master students Rim Slama and Eric Foe for their help.

I thank my family, my friends, people at TELECOM Lille 1 and LIFL for their never decreasing warm encouragements. Finally, I d like to thank my wife Emna and my parents Abdelaziz and Najet, whose support of my graduate career has helped me to endure the marathon of the PhD process.

Table of contents

Table of contents **8**

1 Introduction **15**

1.1 Problematic and Motivation 15

1.2 Goals and contributions 18

1.3 Organization of the thesis 20

2 State-of-the-art on 3D face recognition **21**

2.1 Introduction 21

2.2 3D Face acquisition and available databases 22

2.3 3D face representations 25

2.4 Challenges of 3D face recognition 27

2.5 Elements on geometry of facial surfaces 28

2.5.1 Algorithms for geodesic computation on surfaces 30

2.5.2 Facial expressions modeling 30

2.6 Geodesic-based approaches 32

2.7 Deformable template-based approaches 33

2.8 Local regions/ features approaches 35

2.9 Discussion 36

2.10 Conclusion 37

3 3D nose shapes analysis for partial biometrics 39

3.1 Introduction . . . 39

3.2 Automatic 3D scan preprocessing . . . 43

3.3 A Geometric Framework for nose shape analysis . . . 46

3.3.1 Nose curves . . . 47

3.3.2 Nose surfaces . . . 51

3.4 Experiments . . . 53

3.4.1 Preliminary results . . . 53

3.4.2 Large scale experiment . . . 54

3.4.3 Our approach vs. rigid-based face and nose matching . . . 57

3.5 Conclusions . . . 61

4 3D Face Recognition using Elastic Radial Curves 65

4.1 Introduction . . . 65

4.2 Introduction and overview of the proposed approach . . . 67

4.3 Automatic data preprocessing . . . 68

4.4 Facial shape representation using radial curves . . . 70

4.5 Elastic metric for expression invariant 3D face recognition . . . 71

4.5.1 Motivation for the elastic metric . . . 71

4.5.2 Elastic metric . . . 72

4.5.3 Square Root Velocity representation SRV . . . 74

4.5.4 Riemannian elastic metric on open curves . . . 75

4.5.5 Optimal Re-parametrization for curve matching . . . 80

4.5.5.1 Dynamic programming . . . 80

4.5.5.2 Computer implementation . . . 81

4.5.6 Riemannian metric on facial surfaces . . . 81

4.6 Robustness to pose and missing data . . . 83

4.7 Results and discussions . . . 85
4.7.1 Evaluation on GavabDB dataset . . . 85
4.7.2 Evaluation on FRGCv2 dataset . . . 89
4.7.3 Robustness to quality degradation: experimental setting . . . 90

5 Towards statistical computation on 3D facial shapes 99
5.1 Introduction . . . 99
5.2 Mean shape computation . . . 100
5.2.1 Intrinsic mean shape . . . 100
5.3 Hierarchical organization shapes . . . 102
5.3.1 Hierarchical shape database organization . . . 102
5.3.2 Hierarchical shape retrieval . . . 106
5.4 3D face restoration/ Recovery . . . 110
5.4.1 Learning statistical models on the shape space . . . 112
5.4.1.1 Training data collection . . . 112
5.4.1.2 Projection onto tangent space on mean shape . . . 113
5.4.1.3 Tangent PCA . . . 115
5.4.2 Curves restoration . . . 116
5.4.2.1 Projection of full faces . . . 116
5.4.2.2 Projection of faces with missing data . . . 118
5.4.3 3D face recognition under occlusion . . . 121
5.4.3.1 Overview of proposed approach for recognition under occlusion 121
5.4.3.2 Detection and Removal of Extraneous Parts . . . 122
5.4.3.3 Completion of Partially-Obscured Curves . . . 122
5.4.3.4 Experimental results . . . 123
5.5 Conclusion . . . 126

6 Conclusion and Future study 129

Bibliography **135**

List of figures

1.1 bloc diagram of the project tasks. 16

2.1 Output of the 3D sensor: range and color image at top row and 3D mesh with and without color mapped on in bottom row. 23

2.2 Example of face from FRGCv2 face. 3D data is shown at top and the texture is mapped on it at the bottom. 24

2.3 Examples of all 3D scans of the same subject from GAVAB dataset. 24

2.4 Examples of faces from the Bosphorus and illustration of the problems of incomplete data (due to the pose or occlusion) and deformations of the face (due to expressions). 25

2.5 3D face model and different representations: point cloud and 3D mesh. . . . 26

2.6 Deformations under facial expressions (a). Missing data due to self occlusion (b). 28

2.7 Euclidean path (in red) and geodesic one (in blue) on facial surface 29

2.8 Points correspondence (neutral face at left and expressive face at the right) 31

3.1 An illustration of challenges in facial shape in presence of large expressions. The level curves of geodesic distance to the nose tip lying in regions outside the nose are significantly distorted while the curves in nose region remain stable. 41

3.2 Histograms of distances (for some level curves) between different sessions of the same person 42
3.3 Automatic FRGC data preprocessing and nose curves extraction. 44
3.4 Automatic nose tip detection procedure. 45
3.5 Examples of pre processed data: nose region with geodesic level curves . . . 46
3.6 Illustration of the Pre-Shape Space $\mathcal{C}$ and geodesic in this pre-shape space. 48
3.7 Examples of geodesic between curves 50
3.8 Computing geodesics in the quotient space. 51
3.9 Geodesic path between source and target noses (a) First row: intra-class path, source and target with different expressions (b) Three last rows: inter-class path 52
3.10 Similarity matrices for (a) neutral vs. neutral comparsions, (b) neutral vs. small expressions, and (c) neutral vs. large expressions, top row includes marices for face-to-face matching and buttom row includes nose-to-nose matching. 54
3.11 Rank one recognition rates using 1 to k curves 55
3.12 Verification Rate at FAR=0.1% using (1 to k) curves 56
3.13 ROC curves using 1 to k curves 57
3.14 Receiver operating characteristic curves for our approach and ICP (baseline) 58
3.15 Cumulative Match Characteristic curves for our approach and ICP (baseline) 59
3.16 Examples of noses that are recognized by our approach and not by ICP algorithm on nose region 62
3.17 Examples of noses that are not recognized by our approach 63

4.1 Important challenges of 3D face recognition. (a) neutral face, (b) expression variations, (c) missing parts due to scan technology and pose variation, (d) pose and expression variations 66
4.2 Significant changes in both Euclidean and surface distances under different face expressions. 67

4.3 Overview of the proposed method. 69

4.4 Different steps of preprocessing: acquisition, filling holes, cropping and smoothing 70

4.5 Radial curves extraction : left image illustrates the intersection between the face surface and a plan to form two radial curves. The collection of radial curves is illustrated at the right image. 71

4.6 Illustration of elastic metric. In order to compare the two curves in (a), some combination of stretching and bending are needed. The elastic metric measures the amounts of these deformations. The optimal matching between the two curves is illustrated in (b). 72

4.7 An example of matching of two radial curves extracted from two faces. a) A curve on an open mouth, (c) a curve on closed mouth , b) change of parametrization before matching 76

4.8 Illustration of shape space and geodesic between its elements. 77

4.9 An example of matching and geodesic deforming radial curves extracted from two faces. a) A curve on an open mouth, (c) a curve on closed mouth ,(d) a geodesic path between curves (a) and (c), b) a result of matching using dynamic programming 79

4.10 Faces comparison by pairwise curves comparisons. 82

4.11 Examples of intra and inter-class geodesics in the shape space. 84

4.12 Examples of geodesics in shape space, pre-shape space and linearly interpolated path after ICP alignment. 85

4.13 Quality filter: examples of detection of broken and short curves (in red) and good curves (in blue). 86

4.14 Examples of facial scans with different expressions in GavabDB. (a) and (b) neutral, (c) smile, (d) laugh, (e) random gesture, (f) looking down, (g) looking up, (h) right profile, (i) left profile. 87

4.15 Examples of correct and incorrect matches. For each pair, the probe (on the left) and the ranked-first faces from the gallery (on the right) are reported. 87

4.16 CMC curves of our approach for the following scenario: neutral vs. neutral, neutral vs. expressions and neutral vs. all. 90
4.17 Examples of 3D faces with radial curves under different quality degradation. 92
4.18 DET curves for experiments on resolution changes. 94
4.19 DET curves for experiments on missing data. 95
4.20 DET curves for experiments on noisy data. 96
4.21 Recognition accuracy for different quality degradation experiments. 97
4.22 Variation in shapes of radial curves versus changes in locations of reference points. 97
4.23 Distributions of intra-class and inter-class distances. 98

5.1 An example of Karcher mean of faces, the face at the right is the karcher mean of the eight faces in the left. 101
5.2 Examples of nasal shapes and their means. 102
5.3 The result of hierarchical organization of gallery faces from FRGCv2 dataset and an example of path parsed by a query across the tree. 105
5.4 Path from top to bottom in the tree show increasing shape resolution. . . . 106
5.5 The tree resulting on hierarchical clustering. 107
5.6 Paths from top to buttom in the tree show increasing shape resolutions . . 108
5.7 Tree parsing for efficient search of a query 110
5.8 Retrieval of query shape (on the left) and its closest shape at each level of the tree 111
5.9 Radial curves collections for training step. 113
5.10 Illustration of mapping shapes onto the tangent space of μ, $T_\mu(\mathcal{S})$. Dashed arcs illustrate the projection of elements of the manifold to the tangent space. 115
5.11 Restoration of curves of different index. 117
5.12 Restoration of full face. The absolute and signed deviations 118

5.13 Restoration of curve with missing data. The probe curve is illustrated in (a) and the mean curve for correspondent index in (b). Together are illustrated in (c). In (d), the part of mean curve corresponding to missing data in probe curve. The result of restoration is illustrated in (e) and all curves are represented in (f). 118

5.14 Restoration of 3D face with missing data. Original face and its reconstruction from curves are respectively illustrated by (a) and (d). The reconstructed face and is illustrated in (e)The face with missing parts is illustrated in (b) and (c) illustrates radial curves. 120

5.15 Different steps of recognition under occlusion. 122

5.16 Gradual removal of occluding parts in a face scan. 123

5.17 (a) Faces with external occlusion, (b) faces after detection and removal of occluding parts, and (c) estimation of the occluded parts using a statistical model on shape spaces of curves. 124

5.18 Examples of faces from the Bosphorus database. 124

5.19 Recognition results on Bosphorus database and comparison with state-of-the-art approaches. 124

5.20 Examples of non recognized faces. Each row illustrates, from left to right, the occluded face, the result of occlusion removal and the result of restoration. 125

Table 1: **List of symbols and their definitions used in this thesis.**

Symbol	Definition /Explanation
S	a smooth facial surface.
$dist$	length of the shortest path on S between any two points
λ	a variable for the value of $dist$ from the tip of the nose
$\beta_\lambda(s)$	level curve of $dist$ on S at the level λ.
α	a variable for the value of the angle that makes radial curve with reference one.
$\beta_\alpha(s)$	radial curve of index α.
$q(s)$	the scaled velocity vector, $q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}}$.
$\langle a, b\rangle$	the Euclidean inner product in $\mathbb{R}^3$
$SO(3)$	group of rotation in $\mathbb{R}^3$.
Γ	group of re-parametrization in $[0, 1]$.
$[q]$	equivalence class of curves under rotation and re-parametrization.
$\mathcal{C}$	the set of all curves in $\mathbb{R}^3$, (closed curves in chapter 3 and open curves in chapter 4).
d_c	distance on $\mathcal{C}$, $d_c(q_1, q_2) = cos^{-1}(\langle q_1, q_2\rangle)$.
$\mathcal{S}$	the shape space: $\mathcal{S} \doteq \mathcal{C}/(SO(3) \times \Gamma)$.
d_s	distance on $\mathcal{S}$, $d_c([q_1], [q_2])$.
$T_v\mathcal{S}$	the space of all tangents to $\mathcal{S}$ at v
$\langle f_1, f_2\rangle$	the Riemannian metric on $\mathcal{C}$, $\int_0^1 f_1(s)f_2(s)ds$
$\psi_t(v_1, v_2)$	a geodesic path in $\mathcal{C}$, from v_1 to v_2, parameterized by $t \in [0, 1]$
$\mathcal{S}^{[0,\alpha_0]}$	indexed collection of radial curves (a face).
$\mathcal{S}^{[0,\lambda_0]}$	indexed collection of level curves (a nose).
μ	shape of mean face.
$T_\mu(\mathcal{S}^{[0,\alpha_0]})$	tangent space at μ on $\mathcal{S}^{[0,\alpha_0]}$

Chapter 1

Introduction

The present thesis was supported by the ANR under the project ANR-07-SESU-004. The project is known as Facial Analysis and Recognition Using 3D (FAR3D). The aim of the FAR3D project is to develop algorithms and techniques in 3D face recognition answering in particular the 3D challenges related to 3D acquisition, face expression variation and face recognition reliability/stability versus 3D face model quality (resolution and depth precision). Two face recognition scenarios will be studied within the project: on the one hand face recognition in symmetrical mode by comparing a probe 3D face model with the ones in the gallery, and on the other hand face recognition in asymmetrical mode by comparing an image or sequence of 2D facial images (probe) with the gallery ones. Figure 1.1 illustrates the bloc diagram of the project tasks, we are concerned by the geometric approach.The partners are Ecole Centrale de Lyon/LIRIS, Eurecom, THALES and USTL/LIFL which is the coordinator of the project.

1.1 Problematic and Motivation

The need for civil defense on the one hand and the fight against fraud, crime and terrorism on the other hand, make person's identity verification an ultimate aim for computer vision researchers. Person's identity can be checked in one of three ways: by something they have (e.g. an ID card); something they know (e.g. the answer to a security question); or

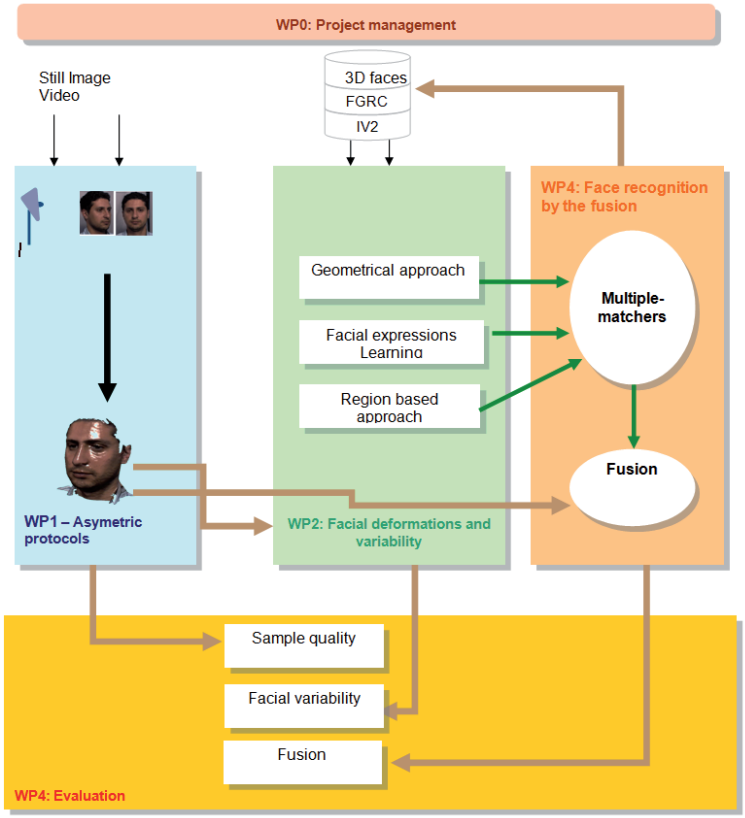


Figure 1.1: bloc diagram of the project tasks.

something they are (a biometric). Several special biometrics are suitable for differentiating persons from others. Biometrics such as faces, iris patterns, hand geometry, fingerprints, speech, retina, dynamic signature, and gait are among the most frequently used modalities. By human biometrics we mean the use of physiological characteristics, human body parts and their appearances, to pinpoint individual human beings in the course of daily activities. The appearances of body parts, especially in imaged data, have a large variability and are influenced by their shapes, colors, illumination environment, presence of other parts, and so on. Therefore, the biometrics researchers have focused on body parts and images that try to minimize this variability within class (subjects) and maximize it across classes. The use of shapes of facial surfaces (or 3D faces) is an important example of this idea. Actually, face recognition has many benefits over other biometric technologies due to the

natural, non-intrusive, and high throughput nature of face data acquisition. Thus, the techniques for face recognition have received a growing attention within the computer vision community over the past three decades. Since 2D (visible light) images of faces are greatly susceptible to variations in the imaging environments (camera pose, illumination patterns, etc), the researchers have argued for the need to use 3D face data, typically collected by laser scanners, for studying shapes of peoples' faces and using this shape analysis for biometrics. The output from laser scanners are minimally dependent on the external environmental factors and provide faithful measurements of shapes facial surfaces. It's the case the only remaining variability that is manifested within the same class, i.e. within the measurements of the same person, is the one introduced by changes in facial expressions, self occlusion (due to pose variation) and aging. Facial expressions, such as smile, serious, fear, and anger, are prime indicators of the emotional state of a person and, thus, are important in estimating mood of a person, for example in developing intelligent ambient systems, but may have a lesser role in biometric applications. In fact, change in facial expressions changes the shapes of facial surfaces to some extent and introduce a nuisance variability that has to be accounted for in shape-based 3D face recognition. We argue that the variability introduced by facial expressions has become one of the most important issue in 3D face recognition. The other important issue is related to data collection and imperfections introduced in that process. It is difficult to obtain a pristine, continuous facial surface, or a mesh representing such a surface, with the current laser technology. One typically gets holes in the scanned data in locations of eyes, lips, and outside regions. For instance, scans of people with open mouths result in holes in the mouth region. Moreover, when the subject is non cooperative and the acquisition phase is unconstrained, one gets variation in pose and unforeseen extraneous occlusions. Many studies treated pose and expression variation but few ones tried to focus on solving the problem of occlusion even if a face can easily be partially hidden by objects like glasses, hats, scarves or a hand, hair and beard. Occluded parts represent wrong information which can degrade recognition accuracy, so how to locate and remove occlusion on face quickly and automatically is a challenging task.

In the biometrics literature, recognition is a general term which includes identification and verification. However, recognition mostly refers to identification in many papers, and

authentication is used instead of verification.

One application of the verification task is access control where an authorized individual is seeking access to a secure facility and presents to the system his or her identity. Here, a one-to-one matching is performed : the 3D image for this individual is acquired, preprocessed and finally compared to an enrollment acquisition already incorporated in the system database. If the similarity is greater than a defined threshold the subject is granted access, otherwise access is denied. The identification scenario consists on determining identity of an authorized user in a database of many individuals. As far as verification scenario is concerned, it is important to examine the probability of correct verification as a function of the false acceptance rate (or Imposter access). This is often shown with a Receiver Operating Characteristic (ROC) curve that graphs the probability of correct verification versus the false acceptance rate. In the identification scenario, however, the results for facial identification are often displayed using a Cumulative Match Characteristic (CMC) curve. This curve displays the cumulative identification rates as a function of the rank distribution. This provides an indication of how close one may be to getting the correct match if the rank-one match was incorrect

1.2 Goals and contributions

In this thesis, we investigate shape analysis for 3D facial surfaces comparison, averaging and restoration. The contributions can be grouped according to the following categories:

- *Riemannian Analysis of 3D Nose Shape For Partial Human Biometrics.* The nose region is observed as the most stable region in the face during expression variations. We propose to study the contribution of this biometric in person identification. The main tool presented is the construction of geodesic paths between two arbitrary nasal surfaces. The length of a geodesic between any two nasal surfaces is computed as the geodesic length between a set of their nasal curves. In particular, we present results for computing geodesics, computing statistical means and stochastic clustering to perform hierarchical clustering. We demonstrate these ideas in two application contexts : the authentication and the identification biometric scenarios using nasal shapes on a large

dataset involving 2000 scans, and hierarchical organization of noses gallery to allow efficient searches.

- *Pose and Expression-Robust 3D Face Recognition Using Elastic Radial Curves.* Our contribution here is an extension of the previous elastic shape analysis framework [JKSJ07] of opened curves to surfaces which is able to model facial deformations due to expressions, even when the mouth is open. The key idea here is the representation of a face by its radial curves, and the comparisons of those radial curves using the elastic shape analysis. We demonstrate the effectiveness of our approach by comparing to the state of the art results. However, unlike previous works dealing with large facial expressions, especially when the mouth is open which require lips detection, our approach mitigates this problem without any lip detection. In [DBADS10], we test this algorithm on FRGC v2 dataset and organize hierarchically the face gallery to allow efficient searches. We illustrate this geometric analysis of 3D facial shapes in presence of both facial expressions and pose variations. We report results on GAVAB dataset and our result outperforms previous works on that dataset. We present the results of our approach in 3D face recognition designed to handle facial expression variations, pose variations and missing data between gallery and probe images. The method has several benefits that make it suited for 3D face recognition and retrieval in a large dataset. Firstly, to handle a pose variation and missing data, we propose a local representation by using a curve representation of 3D face and a quality filter of a curves. Secondly, to handle variations in facial expressions, we propose an elastic shape analysis of 3D face. Lastly, to accelerate the running time of recognition algorithm we propose hierarchical organization of the gallery to allow for efficient retrieval. Experiments that are performed on the FRGCv2 and GavabDB databases and within the SHREC'10 contest show high effectiveness also in the presence of large facial expressions, pose variations and missing data. For the identification experiment, we achieve the second best recognition rate of 97.7% on FRGCv2 and the best rank (96.99%) on GavabDB. The behavior of our algorithm in presence of different kind of quality degradation is analyzed through the next analysis. We simulate the presence of noise, missing data and decimation on a subset of FRGC v2 dataset, then we present

a comparative study of the results of our algorithm performed on these data.

- *3D face shape restoration: Application to 3D face recognition in presence of occlusions.* Our contribution here is a novel framework for 3D face shape restoration.
  3D face recognition in presence of missing data : the surface is first represented by radial curves (emanating from the nose tip for the example of the 3D face). Then a statistical computation on non linear manifold are proposed to built eigen-curves basis. Hence, the missing part of the curve can be restored by projecting the existing part of it on the base.

1.3 Organization of the thesis

This thesis is organized as follows: in chapter two we lay out challenges and the state of the art on 3D face recognition task. In chapter three, we study the contribution of the nose region, seen as the most stable region in the face under expression variations. Chapter four describes our approach for 3D face recognition. The same framework used for recognition is used to calculate statistics in chapter five. Finally we present concluding remarks and some perspectives of the current work in the last chapter.

Chapter 2

State-of-the-art on 3D face recognition

2.1 Introduction

The full 3D model of a human head can be acquired from multi-view stereo systems during the scanning process. However, multi-view are not practical for recognition scenarios. Therefore, static 3D sensors such as laser scanners or structured light based stereo systems produce the so-called 2.5D image (depth map or range image). 2.5D surface is usually defined as having at most one z-depth measurement from a given (x,y) coordinate. It is possible to generate full 3D face models by combining several 2.5D images. In the 3D face recognition literature, the term 3D is commonly used to denote 2.5D data. In terms of a modality for face imaging, a major advantage of 3D scans over 2D color imaging is that variations in illumination, pose and scaling have less influence on the 3D scans. Actually, three dimensional facial geometry represents the geometric structure of the face rather than its appearance influenced by environmental factors. For instance, 3D data is insensitive to illumination, head pose [BCF06b], and cosmetic [MS01]. Therefore, 3D face acquisition requires more sophisticated sensor than 2D one. The rest of this chapter in organized as follow. In section 2, we describe the acquisition process and survey available 3D face databases. Section 3 enumerates and describes the representations of a 3D face model. In

section 4, we summarize the challenges of 3D face recognition. Section 5 gives necessary elements to describe the geometry of a 3D face. Sections 6, 7 and 8 describe the previous approaches. In section 9 we discuss the previous approaches and we finally give concluding remarks in section 10.

2.2 3D Face acquisition and available databases

Many technical solutions have been proposed for the acquisition of 3D geometry from human face. Even if Computer Vision practitioners have made a significant effort to develop less expensive alternative techniques based on passive sensors, most 3D face databases are collected by active sensors. Active techniques are based on lighting up the 3D face. Most of laser scanners use an optical camera and laser rays to measure by triangulation the distance to sample target points on the faces' boundary to produce depth images. Depth images are 2D views of the face accompanied with the depth information (the distance from the face to the scanner) for each pixel. Laser-based sensors are able to produce a dense and precise reconstruction in about 0.5s for fast mode and 2.5s for high resolution mode. During acquisition, the subject should make no movement. However, they generate some noise and missing data due to laser absorption by dark area and self-occlusion as illustrated in Figure 2.6.b. Laser sensors are expensive as they need rotating mirrors and cylindrical lens to generate a plan of laser and scan the face. Figure 2.1 illustrates the output of the scanner as depth and color channels. The range and color image are illustrated at top row respectively by Figure 2.1.a and Figure 2.1.b. The 3D mesh is illustrated at bottom row in Figure 2.1.c, Figure 2.1.d shows the texture mapped on it.

Most popular 3D face datasets were collected using laser-based sensors. FRGCv1 and FRGCv2 were collected by researchers of the University of Notre Dame and contain 948 3D images for training (FRGCv1) and 4007 3D face for testing (FRGCv2) of 466 different persons belonging to different sex, different races and different ages [PFS$^+$05]. We illustrate an example of faces from this database in Figure 2.2, left image illustrates the 3D data and right image illustrates the color image mapped on the 3D one.

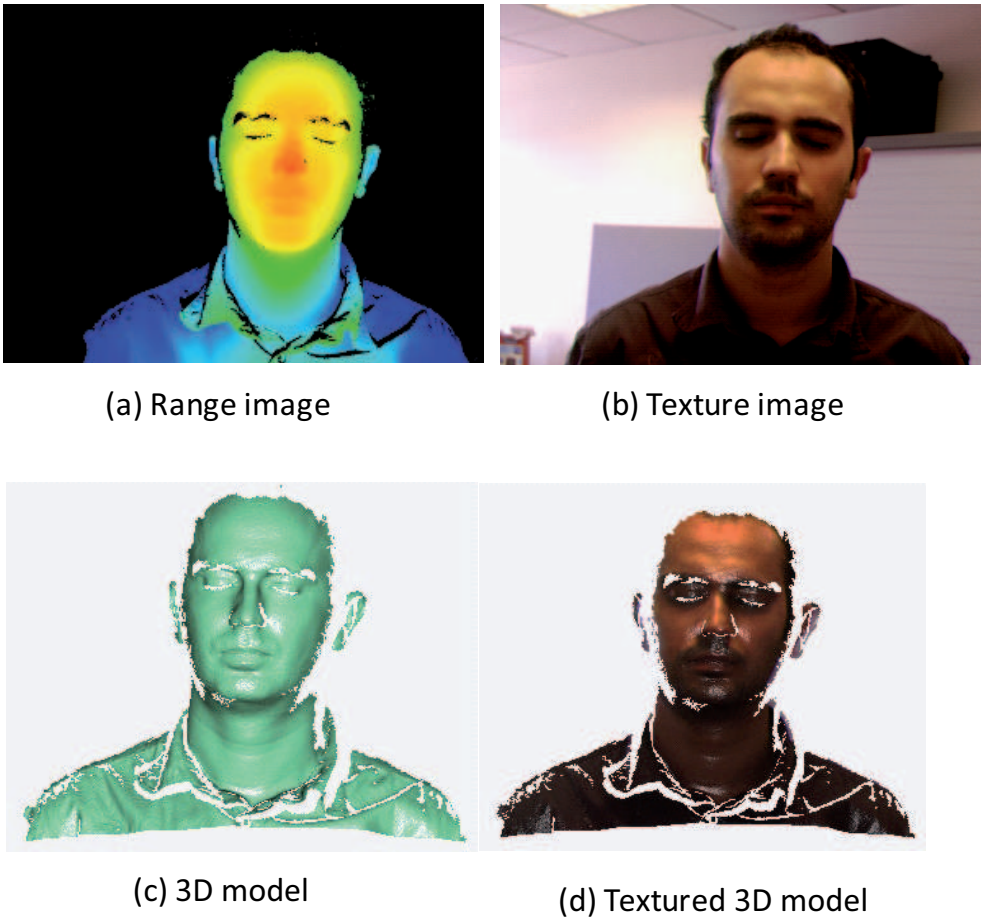


Figure 2.1: Output of the 3D sensor: range and color image at top row and 3D mesh with and without color mapped on in bottom row.

The GavabDB consists of Minolta Vi-700 laser range scans from 61 different subjects. The subjects, of which 45 are male and 16 are female, are all Caucasian. Each subject was scanned 9 times for different poses and expressions, namely six neutral expression scans and three scans with an expression. The neutral scans include two different frontal scans, one scan while looking up (+35 degree), one scan while looking down (-35 degree), one scan from the right side (+90 degree), and one from the left side (-90 degree). The expression scans include one with a smile, one with a pronounced laugh, and an arbitrary expression freely chosen by the subject [MS04]. Figure 2.3 shows some examples of faces in this dataset.



Figure 2.2: Example of face from FRGCv2 face. 3D data is shown at top and the texture is mapped on it at the bottom.

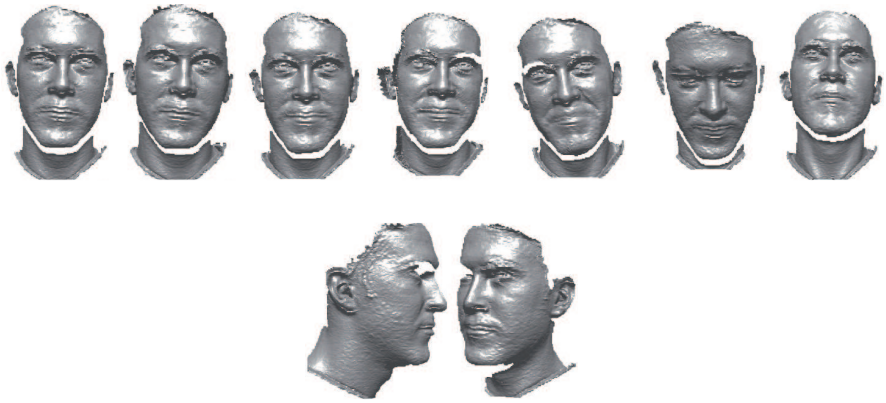


Figure 2.3: Examples of all 3D scans of the same subject from GAVAB dataset.

GavabDB is used by many researchers to compare their methods. However, the commonly used Benchmark for the task of 3D face recognition is FRGC. New challenges such as face occlusion and typical expression classification are not considered by these datasets. The typical expressions classification is not the main task in this manuscript, however, the face occlusion is a new challenge on 3D face recognition. BOSPHORUS database [SAD$^+$08] is the suitable 3D face database for that task; Actually, it contains facial models of 60 men and 45 women, 105 subjects in total, in various poses, expressions and in presence of occlusions (eyeglasses, hand, hair). The majority of the subjects are aged between 25 and 35. The number of total face scans is 4652, where for every subject 54 scans are available,

but 34 of them have 31 scans. For each subject, four occluded scans were gathered. These occlusion are i) mouth occlusion by hand, ii) eyeglasses, iii) occlusion of face with hair, and iv) occlusion of left eye and forehead regions by hands. Figure 5.18 shows sample images from the Bosphorus 3D database illustrating expression variations and typical occlusions.

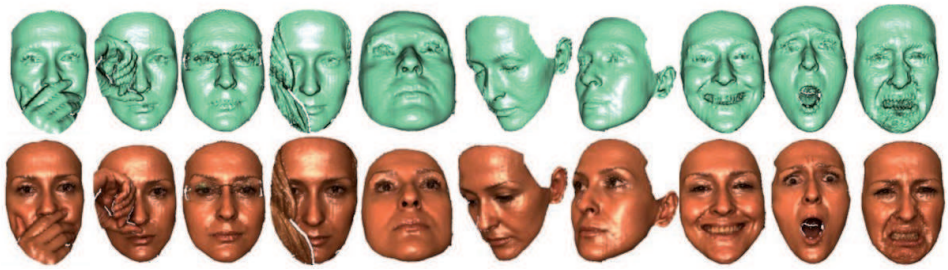


Figure 2.4: Examples of faces from the Bosphorus and illustration of the problems of incomplete data (due to the pose or occlusion) and deformations of the face (due to expressions).

We present in Table 2.1 a summary of the most known and available 3D face datasets and their characteristics.

Table 2.1: Popular 3D face datasets and their different characteristics.

	Sensor used	Number of subjects	Total scans	Expressions	Missing data	Occlusion
FRGCv1	Laser sensor	275	948	No	No	No
FRGCv2	Laser sensor	466	4007	Yes	No	No
GavabDB	Laser sensor	61	549	Yes	Yes	No
BosphorusDB	Structured light	105	4652	Yes	Yes	Yes

2.3 3D face representations

Despite all physical objects that surround us are continuous, digital computers can only work with discrete data. Therefore, surfaces are usually approximated using discrete representations. The most noticeable exception being represented by NURBS modeling. As far as 3D face is concerned, the output of 3D sensors is a range image accompanied with a color image. This specific representation is illustrated in Figure 2.1.a and Figure

2.1.b. The range image has the same resolution as the color one. The values of color level are replaced in that case by the depth value (distance to the sensor). Therefore, the neighborhood of points in 3D in known and the 3D model (illustrated in Figure 2.1.c can be constructed).

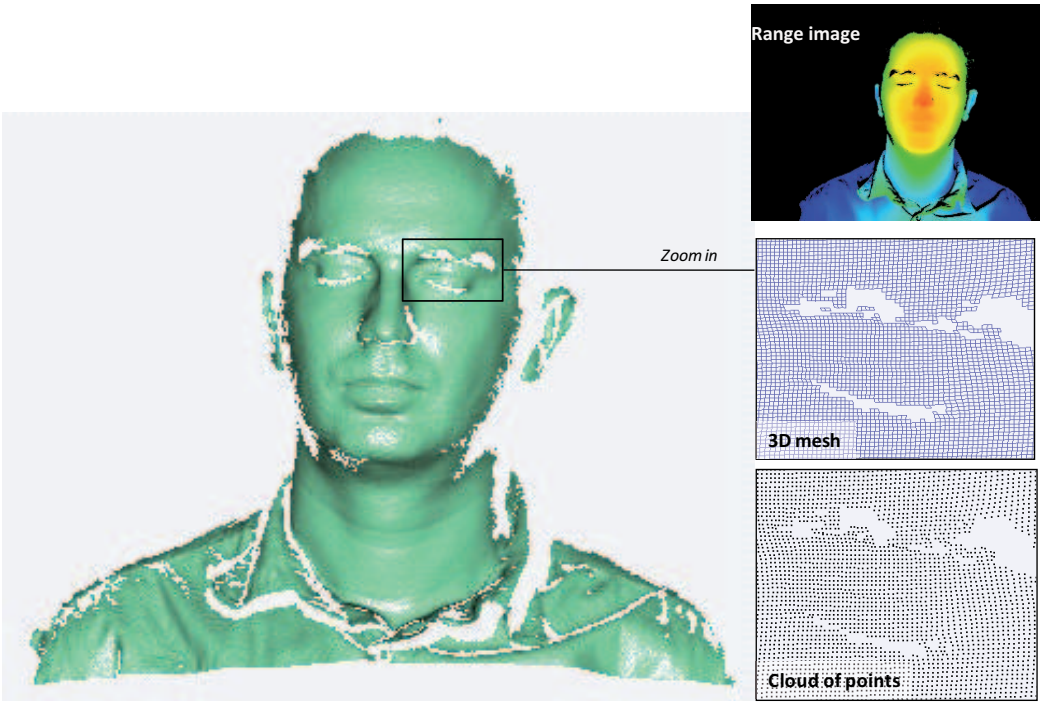


Figure 2.5: 3D face model and different representations: point cloud and 3D mesh.

Figure 2.5.a illustrates the 3D model reconstructed from range image. The different representation of the 3D face are illustrated in Figure 2.5.b. The top image illustrates the range image, the middle one illustrates a zoom on the resulting mesh and the bottom image illustrates a zoom on the cloud of points. The representation of the 3D face by range image allows the use of 2D image processing algorithms. The point cloud representation is used by several algorithms which do not need to know the neighborhood of points.

A surface mesh representation is given by the definition of an arbitrary number of sample points p of coordinates (x, y, z) lying on the surface. Then the surface is locally

approximated by polygon primitives (especially triangles) which connect the sample points into a graph structure, denoted as a polygon mesh M. The polygon primitives are denoted as faces f , the segments of faces as edges e and the extremities as vertices v (one vertex per sample point p):

$M = \{V, E, F\}$, $F = \{f_1, f_2, ..., f_N\}$, $E = \{e_1, e_2, ..., e_N\}$ and $V = \{v_1, v_2, ..., v_N\}$.

The automatic reconstruction of the mesh surface from point sets is an active research topic which is far beyond the scope of this manuscript. We defer the reader to (MM99) for a detailed survey and to (KBH06, SLSL06, BHGS06) for recent automatic techniques. In practice, interactive techniques based on the canonical Delaunay triangulation are mostly employed and can be considered as part of the acquisition operation.

2.4 Challenges of 3D face recognition

When acquired in non-controlled conditions, scan data often suffer from the problem of missing parts due to self-occlusions or laser-absorption by dark areas. Actually, the 3D face needs more than one scan to be fully acquired. Especially when the pose is not frontal as illustrated in Figure 2.6.b, the resulting scan is said 2.5D and not full 3D. However, this 2.5D scan is roughly approximated by 3D scan by 3D face recognition community researchers. Moreover, bottom row in Figure 2.6.b illustrates that dark area (hair or eyebrows) absorb the laser and generate missing data in the 3D mesh as illustrated at the top row of the same figure.

Additionally, variations in face data due to facial expressions cause deformations in the 3D mesh. Figure 2.6.a illustrates expressive faces at the bottom row (as 3D textured mesh). The top row illustrates the resulting 3D mesh with deformations.

Any 3D face recognition approach should successfully match face scans in presence of expression-based deformations and/or missing data (as illustrated respectively in Figure 2.6.a and Figure 2.6.b) to good quality, neutral, frontal 3D model. We note that generally the enrolled face scans are collected in controlled conditions and exhibit good data quality.

Past literature has tackled this fundamental issue with a varying degree of success as described in the survey paper [BCF06a].

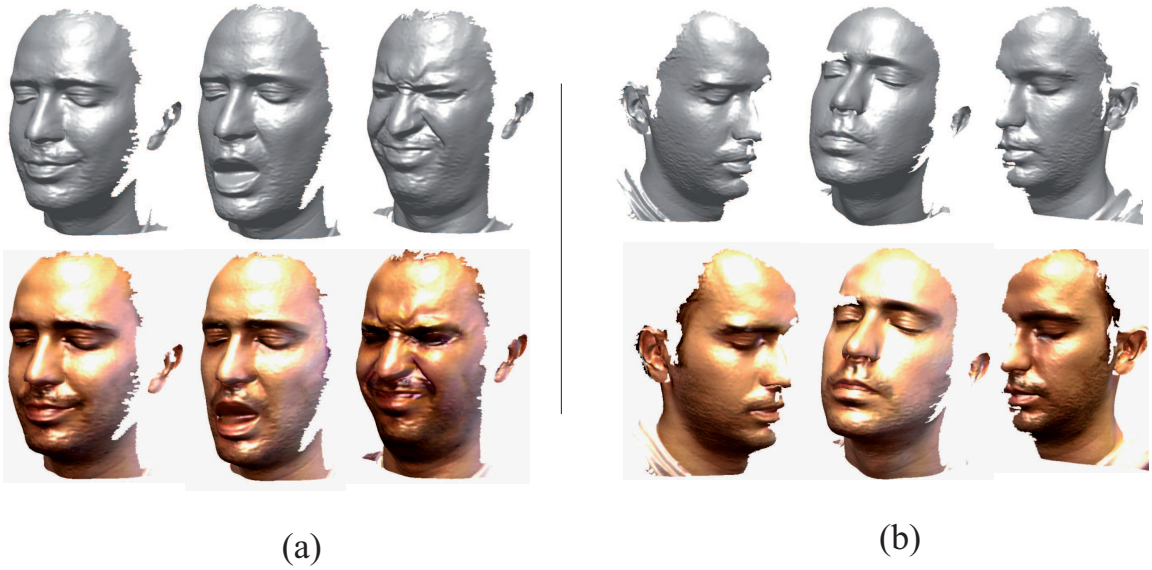


Figure 2.6: Deformations under facial expressions (a). Missing data due to self occlusion (b).

We start with modeling the geometry of human faces, next we will present most recent and elaborated approaches for that task.

2.5 Elements on geometry of facial surfaces

From geometric point view, a human facial surface can be modeled as a two-dimensional manifold, denoted by S. Roughly speaking, manifolds with finite dimension can be defined as topological spaces having locally similar structures to $\mathbb{R}^n$. Every element of a manifold should belong to neighborhood which is homomorphic to a neighborhood in $\mathbb{R}^n$. In topology, a diffeomorphism is an isomorphism in the category of smooth manifolds. It is an invertible function that maps one differentiable manifold to another, such that both the function and its inverse are smooth.. Therefore, the parametrized facial surface can be seen a differential application from an open set of $\mathbb{R}^2$ to an open set of $\mathbb{R}^3$. The parametrized surface is said regular if for each point $(u, v) \in U$, the vectors $\frac{\partial X}{\partial u}(u, v)$ and $\frac{\partial X}{\partial v}(u, v)$ are linearly

independent. In that case, the derivatives $\frac{\partial X}{\partial v}(u,v)$ and $\frac{\partial X}{\partial u}(u,v)$ constitute a local non-orthogonal coordinate system on S, and span an affine subspace of $\mathbb{R}^3$ called *the tangent space* and denoted by $T_x(S)$ for every $x \in S$. The restriction of the euclidean inner product to the tangent space $T_x(S)$ is called *first fundamental form*. This metric is intrinsic to the manifold S and allows to perform local measures independently of the coordinates of the embedding space[Kre91]. Using this metric, let d, the surface will be considered as metric space (S,d) or Riemannian manifold. In order to define an intrinsic distance, let $x, y \in S$ be two points on the surface and let $c : [0,1] \to S$ be a smooth curve in arc length parametrization connecting x and y on S. The length of c is defined by:

$$l[c] = \int_0^1 \| \frac{\mathrm{d}c(t)}{\mathrm{d}t} \| \, \mathrm{d}t. \tag{2.5.1}$$

Then the intrinsic distance on S between the two points is given by

$$d(x,y) = \inf_c l[c]. \tag{2.5.2}$$

The paths of minimum length is obtained by the minima of the function $l[c]$ and called minimal geodesics, and the resulting distance $d(x,y)$ is called geodesic distance. Figure 2.7 illustrates the euclidean and geodesic distances between two points on the face surface.

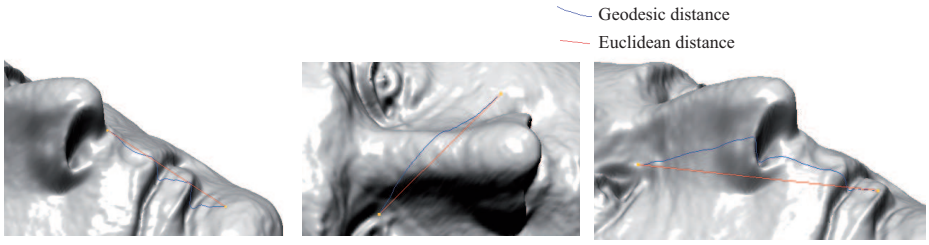


Figure 2.7: Euclidean path (in red) and geodesic one (in blue) on facial surface

There are two main algorithms in a 3D mesh processing literature that can be applied on the facial surface mesh to calculate the geodesic distance between any pair of points on it: the Dijkstra algorithm and the Fast marching algorithm. These are briefly described in

the following.

2.5.1 Algorithms for geodesic computation on surfaces

- Dijkstra algorithm: we start our discussion by considering the surface mesh as a graph, then we will extend it to triangular meshes. For a given source vertex (node) in the graph, the Dijkstra algorithm [Dij59] finds the path with lowest cost (i.e. the shortest path) between that vertex and every other vertex. It can also be used for finding costs of shortest paths from a single vertex to a single destination vertex by stopping the algorithm once the shortest path to the destination vertex has been determined. In our case, the edge path costs represent the length of the edge, Dijkstra's algorithm can be used to find the shortest route between two given vertices in the surface.

- Fast marching algorithm

  The fast marching algorithm [Set96] is characterized by flame propagation. The typical illustration of flame propagation in a forest helps to understand how fast marching works. Trees reached by the fire are consumed so the fire never propagates backward. To evacuate people, the firemen try to record front position at different positions in time. Hence, they can predict when the fire arrives to unburnt regions of the forest and order evacuation of people. The time of arrival of the fire front to a point in the forest is related to the shortest distance from that point to the source of the fire.

  Fast marching algorithm has the ability to find paths passing 'between' the graph edges and this can be of interest in case of low resolution meshes. In case of facial surface acquired with laser sensors, the resolution is quite hight and Graph-based algorithm, such as Dijkstra's, can be considered as one of best choices for finding shortest paths lengths.

2.5.2 Facial expressions modeling

Let (S_1, g) and (S_2, h) be two facial surfaces modeled as Riemannian manifolds. Particularly, in the same way as g induces a distance metric d_{S_1} on S_1, the tensor h induces a distance

metric d_{S_2} on S_2. The two faces belong to the same person (the author) but S_1 is a neutral face whereas S_2 is an expressive one. Let $f : S_1 \rightarrow S_2$ be a diffeomorphism modeling the expression (see Figure 2.8).

The diffeomorphism between two topological spaces is a continuous bijection, and the inverse application is continuous. Roughly speaking, a topological space is a geometric object, and the diffeomorphism is a continuous stretching and bending of the object into a new shape. Thus, a neutral face and an expressive one are diffeomorphic to each other. Notice that the holes (like open mouth) should be filled, otherwise this assumption is false. Some authors [BBK05] approximate this diffeomorphism by an isometry. In other words, they assume that f preserves the geodesic distance between every pair of points, that is,

$$d_{S_1}(x, y) = d_{S_2}(f(x), f(y)), \forall x, y \in S_1. \tag{2.5.3}$$

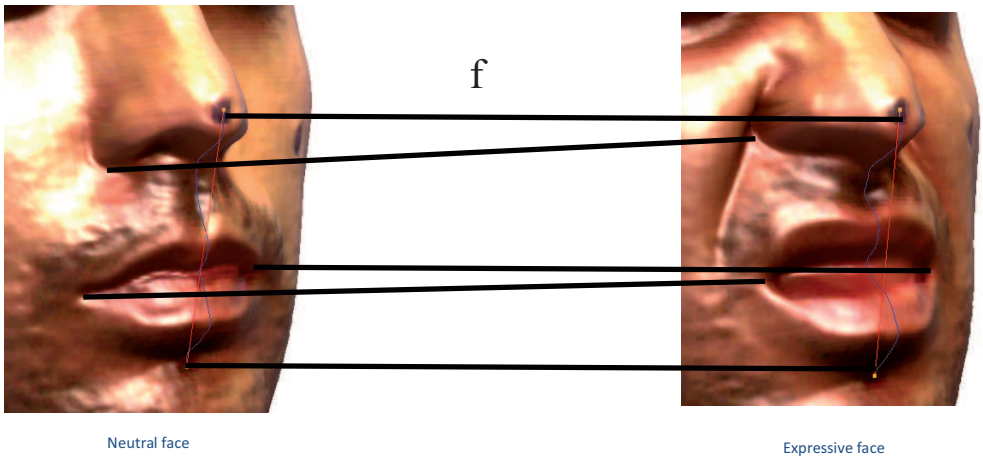


Figure 2.8: Points correspondence (neutral face at left and expressive face at the right)

Several approaches adopt this assumption and used geodesic distance on the facial surface to design expression-invariant approaches.

2.6 Geodesic-based approaches

In [BBK05], the authors presented an experimental validation of the isometric model. They placed 133 markers on a face and tracked the change of both euclidean and geodesic distances under facial expressions. The distribution of the absolute change of geodesic distances was closer to zero than the distribution of the change of euclidean distance. Therefore, the authors assumed that the change of geodesic distance is insignificant and conclude that geodesic distance remains unchanged under facial expression. Under this assumption, in [BBK05], they represent 3d faces as canonical forms. Canonical surfaces were obtained from face surfaces by warping according to a topology preserving transformation. Finally face models were represented with the geometric moments up to the fifth order computed for the 3D face canonical forms. However, while the effect of expressions was attenuated, a similar attenuation also occurred for discriminating features such as eye sockets and nose. This approach was improved in [BBK06] where the authors handled the challenge of missing parts. They embed the probe facial surface into that of the gallery. Faces belonging to the same subject are nearly isometric and thus result in low embedding error, whereas different subjects are expected to have different intrinsic geometry, and thus produce higher embedding error. The open mouth corrupts the isometric model. This problem was handled later by the authors in [BBK07] by using geodesic mask excluding the mouth region. The authors first detected and removed the lips, then the computation of geodesic distance geodesics is calculated on surface in presence of a hole corresponding to the removed part. This is done while avoiding passing in mouth area.

The assumption of the isometric model motivated several authors to use geodesic distance on facial surface. In [SSDK09], the geodesic distance to the tip of the nose were used as surface distance function. Differential geometry was used to compare 3D level curves of the surface distance function. This approach was an improvement of an earlier work of the authors [SSD06], were they used the level curves of the height function to define facial curves. The use of geodesic distance in [SSDK09] allows this approach to handle facial expressions. However, the open mouth corrupts the shape of some level curves and this parametrization did not address this problem, and the experiments were restricted to

a small subset of FRGCv2 database.

A similar geodesic polar parametrization of the face surface was proposed in [MMS07], but rather than studying the shape of curves, they studied local geometric attributes under this polar parameterization. To handle data with open mouths, they modified their geodesic polar parametrization by disconnecting the lips. Therefore, their approach required lips detection, as was the case in [BBK07].

In [BBP10], the authors used the geodesic distance on the face to extract iso-geodesic facial stripes. Equal-width iso-geodesic facial stripes were used as nodes of the graph and edges between nodes were labeled with descriptors, referred to as 3D Weighted Walkthroughs (3DWWs), that captured the mutual relative spatial displacement between all the pairs of points of the corresponding stripes. Face partitioning into iso-geodesic stripes and 3DWWs together provided an approximate representation of local morphology of faces that exhibits smooth variations for changes induced by facial expressions.

A common limitation of the previously described approaches is that they assume that the facial shape deforms isometrically, i.e. the surface distances between points are preserved, which is not valid in the case of large expressions. Actually, the movement of mimic muscles can stretch and/or shrink the face surface and not only bending it. The deformation caused by the expression in Figure 2.8 illustrates that it is not isometric. This will be discussed in more details in chapter 4.

2.7 Deformable template-based approaches

In recent years there has been focus on deforming surfaces, one into another, under a chosen criterion. Grenander's *deformable template* theory [Gre93] has been successfully applied to studying shapes of anatomical parts using medical images [MY01, GM98]. The set of non-rigid deformations can be subdivided into linear and nonlinear deformations. Nonlinear deformations imply local stretching, compression, and bending of surfaces to match each

other, and are also referred to as **elastic** deformations. Earlier attempts at elastic matching utilized graphs based on texture images of faces [KTP00].
Kakadiaris et al. [KPT+07] utilize an *annotated face model* to study geometrical variability across faces. The annotated face model is deformed elastically to fit each face thus allowing the annotation of its different anatomical areas such as the nose, eyes and mouth.

Points of an annotated 3D face reference model were shifted according to elastic constraints so as to match the corresponding points of 3D target models in a gallery. Similar morphing was performed for each query face. Then, face matching was performed by comparing the wavelet coefficients of the deformation images obtained from morphing. This approach is full automatic. Similar approaches were based on manually annotated model [MMS08, LJ06, LJ08].

In [LJ08], the authors presented an approach that is robust to self-occlusions (due to huge pose variations) and expressions. 3D deformations learned from a small control group was transferred to the 3D models with neutral expression in the gallery. The corresponding deformation was synthesized in the 3D neutral model to generate a deformed template. The matching was performed by fitting the deformable model to a given test scan, which was formulated as a minimization of a cost function.

In [tHV10], the authors propose a multi-resolution approach to semi-automatically build seven morphable expression models, and one morphable identity model from scratch. The proposed algorithm automatically selects the proper pose, identity, and expression such that the final model instance accurately fits the 3D face scan.

A strong limitation of these approaches is that the fiducial landmarks needed during expression learning have to be extracted manually for some approaches. They are usually semi-automatic and rarely full automatic.

2.8 Local regions/ features approaches

A different way proposed in the literature to handle expression variations is to match parts or regions of faces rather than the whole faces. Several notable local techniques were proposed in [Gor92, MSVD05] where the authors employed surface areas, curvatures around facial landmarks, distances and angles between them with a nearest neighbour classifier. In [LSY+05], based on ratios of distances and angles between eight fiducial points, the technique used a support vector machine classifier. Euclidean/geodesic distances between anthropometric fiducial points were employed as features in [GAMB07] along with linear discriminant analysis classifiers. However, a successful automated detection of fiducial points is critical here.

In [MFA08, MAM09], the authors presented low-level geometric features-based and reported results on neutral faces but the performance decreases when expressions variations were introduced. Using similar features, the authors in [LJZ09] proposed to design a feature pooling and ranking scheme in order to collect various types of low-level geometric features, such as curvatures, and rank them according to their sensitivities to facial expressions. They applied sparse representations to the collected low-level features and achieved good results on a challenging dataset (GavabDB [MS04]). This approach, however, required a training step.
Along similar lines, in [WLT10], a signed shape difference map (SSDM) was computed between two aligned 3D faces as an intermediate representation for the shape comparison. Based on the SSDMs, three kinds of features were used to encode both the local similarity and the change characteristics between facial shapes. The most discriminative local features were selected optimally by boosting and trained as weak classifiers for assembling three collective strong classifiers. The individual features were of the type: Haar-like, Gabor, and Local Binary Pattern (LBP).

In [MR10], the authors used 3D Fisherface region ensemble approach. After faces registration using ICP algorithm, the Fisherface approach seeks to generate a scatter matrix for improved classification by maximizing the ratio of the between scatter and within

scatter. Twenty-two regions were used as input for 3D Fisherface. In order to select most discriminative regions, *Sequential Forward Search* [STG08] was used.

In [HZA+10, HAWC11, HOA+11], the authors propose to use the Multi-scale LBP as a new representation for 3D face jointly to shape index. SIFT-based local features are then extracted. The matching involves also holistic constraint of facial component and configuration.

In [CMCS06], the Log-Gabor templates were used to exploit the multitude of information available in human faces to construct multiple observations of a subject which were classified independently and combined with score fusion. Gabor features were recently used in [MMJ+10] on automatically detected fiducial points.

In [CBF06, MBO07] the focus was on matching nose regions albeit using ICP. To avoid passing over deformable parts of the face encompassing discriminative information, the authors in [FBF08] proposed to use a set of 38 face regions that densely cover the face and fused the scores and decisions after performing ICP on each region.

In [QSBS10], the circular and elliptical areas around the nose were used together with forehead and the entire face region for authentication. Surface interpenetration measure (SIM) were used for the matching. Taking advantage of invariant face regions, an Simulated annealing approach was used to handle expressions.

In [AGD+08], the authors proposed to use the Average Region Models (ARMs) locally to handle the missing data and the expression-inducted deformation challenges. They manually divided the facial area into several meaningful components and registration of faces was carried out by separate dense alignments to relative ARMs. A strong limitation of this approach is the need for manual segmentation.

2.9 Discussion

The previous description of state of the art illustrates that recent research on 3D face recognition can be categorized into three classes. The geodesic-distance based approaches

are based on intrinsic distance and have the advantage to be independent of the embedding of the surface. In term of robustness to different challenges, these approaches are sensitive to noise, missing data and large expressions involving stretching and/or shrinking of the face and thus geodesic distance variation. These deformations due to large expressions are handled by deformable template-based approaches. These approaches are usually based on an annotated face model which deforms based on deformation of the landmarks. In term of accuracy, local approaches have shown best performances in general. Local approaches are usually based on matching stable regions together and reduce the effect of matching non stable regions.

In this thesis, we present a unified Riemannian framework that provides optimal matching, comparisons and deformations of faces using a single elastic metric. The proposed framework handles large expressions as well as missing parts due to large pose variation and occlusions. The first tool proposed in this thesis concerns partial biometric. That is, the construction of geodesic paths between arbitrary two nasal surfaces. The length of a geodesic between any two nasal surfaces is computed as the geodesic length between a set of their nasal curves. Facial surfaces are represented by collections of radial curves which allows local matching in order to handle pose variation, expressions and occlusion challenges. In addition, using the same framework, we are able to compute some statistics on 3D facial surfaces. This allows an hierarchical organization for efficient shape retrieval and missing parts restoration.

2.10 Conclusion

In this chapter, we first presented the 3D face acquisition process, the different representations of the 3D facial surface and the available benchmark datasets. The different representation of 3D facial surface. This allows the lay out of different challenges in 3D face recognition task. Second we present the geometric modeling of the 3D face, we present an early observation showing that the deformations of the face resulting from facial expressions can be modeled as isometries such that the intrinsic geometric properties of the facial surface, like geodesic distance, are expression-invariant. We presented the past approaches who were based on that assumption, even it is not valid in case of elastic deformation.

Then, we presented approaches based on face deformation. Finally we present the third kind of 3D face recognition approaches, based on local regions/features matching. At the end of the chapter, we discussed the presented approaches with their robustness to different challenges. The limitations of previous 3D face recognition approaches motivate further researches on several tasks as missing parts and non isometric deformations. We introduce our approach in this context.

Chapter 3

3D nose shapes analysis for partial biometrics

3.1 Introduction

In order to handle shape variability due to facial expressions and presence of holes in mouth, we advocate the use of *nose region* for biometric analysis. At the outset the shape of the nose seems like a bad choice of feature for biometrics. The shapes of noses seem very similar to a human observer but we will support this choice using real data and automated techniques for shape analysis. We do not assert that this framework will be sufficient for identifying human subjects across a vast population, but we argue for its role in shortlisting possible hypotheses so that a reduced hypothesis set can be evaluated using a more elaborate, multi-modal biometric system.

Why do we expect that shapes of facial curves are central to analyzing the shapes of facial surfaces? There is plenty of psychological evidence that certain facial parts, especially those around nose, lips and other prominent parts, can capture the essential features of a face. Our experiments support this idea in a mathematical way. We have computed geodesic distances between corresponding facial curves of different faces – same people different facial expressions and different people altogether. We have found that the distances are typically smaller for faces of the same people, despite different expressions, when compared to the

distances between facial curves of different people.

The stability of nose data collection, the efficiency of nasal shape analysis, and the invariance of nasal shape to changes in facial expressions make it an important region of the face.
To resume, the nose region has certain advantages over the face, for example :

1. The robustness to facial expressions of nose region. Indeed, its shape undergone less variation compared with other regions like mouth, eyes, etc..under facial expressions.
2. The nose region is immune to missing data (holes) due to the open mouth in case of frontal poses.
3. The pre-processing speed to extract features and informative region.
4. The efficiency of nasal shape analysis.

Our approach for analyzing shapes of the nasal surface is to represent such a surface by a dense collection of curves. These curves are defined such that they do not depend on the pose of the original face and have been used for full face recognition earlier. In the experimental results presented previously [ADB+09], we worked only with limited facial expressions, i.e. large facial expressions involving open mouths were avoided. Also, it was assumed that the holes are small enough to be filled by simple linear interpolations. The Figure 3.1 shows an example of extraction of facial curves for a 3D scan with large expression. Figure 3.2 illustrates histograms of distances for some level curves of geodesic distance to the tip of the nose. The three top histograms correspond to the distributions of distances calculated between curves lying on nasal region of the same person with different sessions. The distances have small values in contrast with distances shown in the down three histograms. The curves concerned in these histograms lie outside the nose region. As illustrated in the figure, the values of distances are quite high compared to the distances lying on the nose region. This experience advocates our assumption that a curve from the nose region is stable and captures the underlying true shape, while a curve from outside the nose region is significantly distorted due to the open mouth. If such a distorted curve is used to study the shape of the underlying facial surface, it will be difficult to distinguish people from one another.

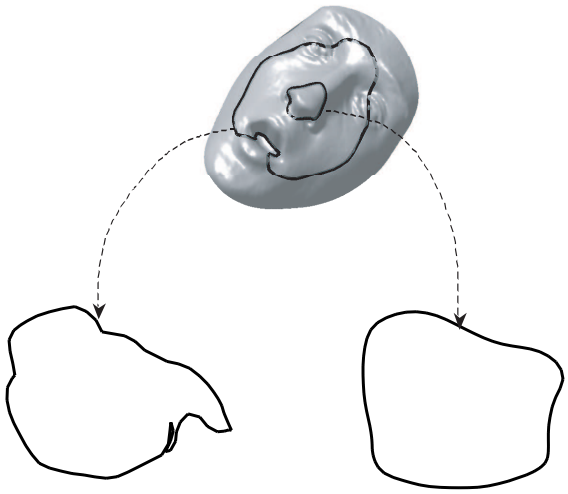


Figure 3.1: An illustration of challenges in facial shape in presence of large expressions. The level curves of geodesic distance to the nose tip lying in regions outside the nose are significantly distorted while the curves in nose region remain stable.

Our framework for quantifying shape differences between 3D surfaces is based on accumulating shape differences between the curves representing the two surfaces. This difference between shapes of curves, in turn, is computed by constructing geodesic paths under the elastic metric on a Riemannian space of closed curves in $\mathbb{R}^3$. On one hand these distances are useful in quantifying shape differences between objects, for example faces or parts of faces belonging to different people. On the other hand, they are also useful in studying variability in shapes of curves resulting from changes in facial expressions. For instance, we have computed distances between level curves on faces of sample people but under different expressions. Two types of curves were used in this experiment: the curves that lie in the nose region and the curves that lie outside the nose region.

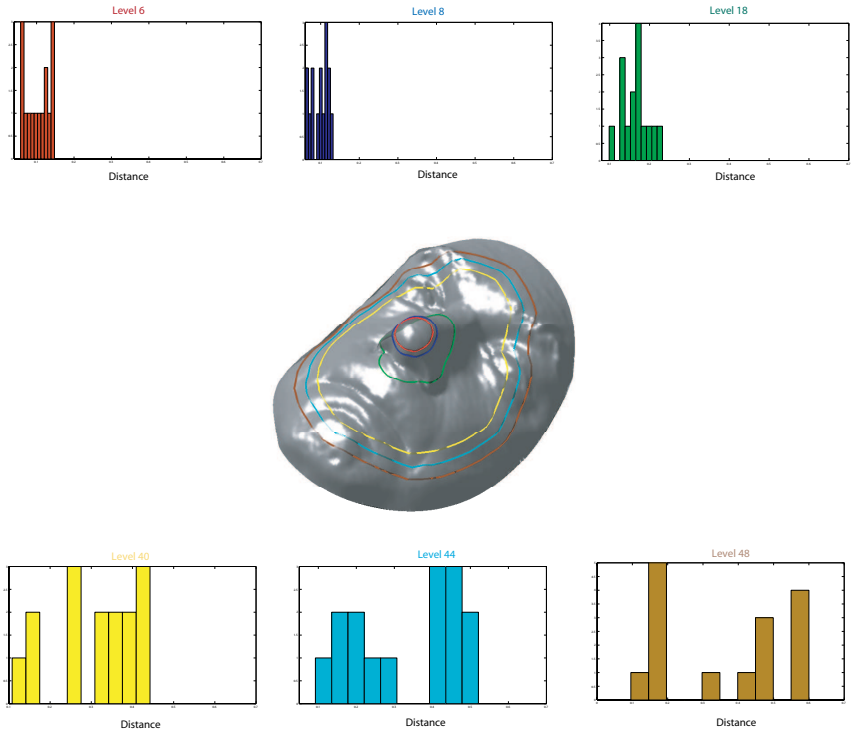


Figure 3.2: Histograms of distances (for some level curves) between different sessions of the same person

The rest of this chapter is organized as follows, section 2 gives a brief description of the preprocessing step. In section 3, we explain the differential-geometric framework of curves

and its extension to 3D surfaces that we used to analyze 3D shapes of nose region. Finally, in section 4, we show the experimental protocol, some preliminary results on a subset of FRGC v2 database containing expressive faces and large scale results.

3.2 Automatic 3D scan preprocessing

In order to assess the recognition performance of the proposed framework, we use 3D models of FRGCv2 dataset. This benchmark database contains sessions with both neutral and non-neutral expressions. Moreover, the laser-based 3D scanner used in the acquisition process introduces noise in the data. In fact, some of 3D face scans suffer from missing data (holes), spikes, artifacts specially in teeth region, occlusions caused by the hair, etc.

We focus in this work on designing a complete solution for 3D face analysis and recognition using only the nose region. For that purpose, it is crucial to begin by de-noising the data by removing spikes, filling holes and extract only the useful part of face and then nose from 3D scan. Figure 3.3 shows different steps of our preprocessing solution to overcome these problems. Starting from an original range image of a face, we firstly apply a 2D median filter in order to remove spikes while preserving edges. Secondly, using a smooth interpolation, we fill holes by adding points in parts where the laser has been completely absorbed (e.g. eyes, eyebrows and open mouth). This is done based on cubic interpolation (splines). On the triangular mesh, we apply the procedure shown in Figure 3.4 to localize the nose tip, necessary for cropping the useful region of the face and nose region segmentation. For this purpose, a sphere function having center the nose tip and radius $R = 100mm$ is constructed and the part inside the sphere is kept. Finally, a collection of geodesic level curves are extracted by locating iso-geodesic points from the reference vertex (nose tip) using the Dijkstra algorithm [Dij59]. In order to segment nose region, we consider the N first level curves ($N = 10$):

- **Nose tip detection** We are able to automatically detect the tip of the nose when the 3D face is frontal. This is verified in 3D models of FRGCv2 dataset. Through the center of the mass of the face, a first transverse plane (parallel to x-z plane) slices the facial surface. The intersection of the plane and the surface results in a horizontal

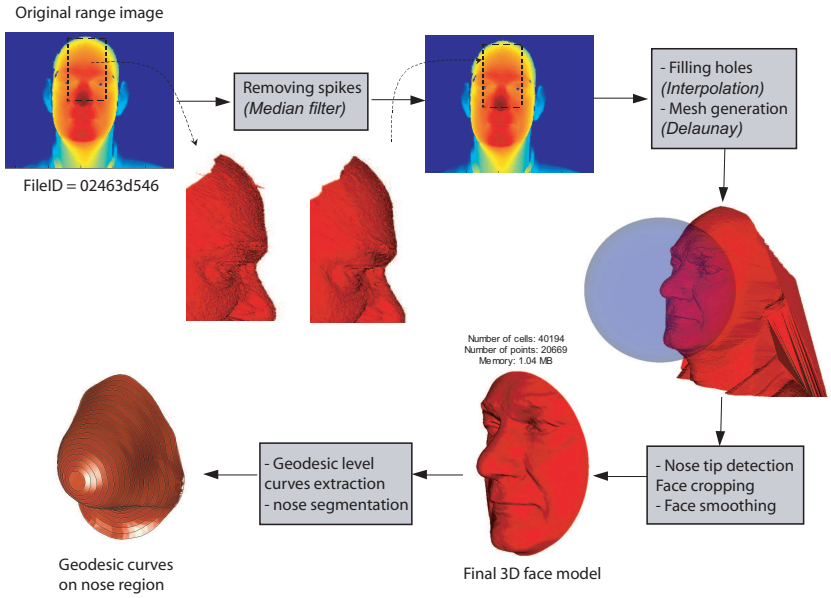


Figure 3.3: Automatic FRGC data preprocessing and nose curves extraction.

profile of the face. A second sagittal plane (parallel to y-z plane) passing through the maximum of the obtained horizontal profile, cuts the face on a vertical profile. The nose tip is located at the maximum of the middle part of this 2D curve as illustrated in Figure 3.4. This step is very important in our general framework as it is necessary to correctly crop faces and extract level curves from the surface.

- **Level curves extraction** Our facial shape analysis algorithm operates using 3D curves extracted by computing geodesic length function along the face. This choice is motivated by the robustness of this function to facial deformations as described in [BBK05]. In addition to its intrinsic invariance to rigid transformations such as rotations and translations, this distance better preserves the separations of facial features as compared to the Euclidean distance in $\mathbb{R}^3$.

  Let S be a facial surface denoting a scanned face. Although in practice S is a triangulated mesh, start the discussion by assuming that it is a continuous surface. Let $dist : S \longrightarrow \mathbb{R}^+$ be a continuous geodesic map on S. Let c_λ denote the level set of

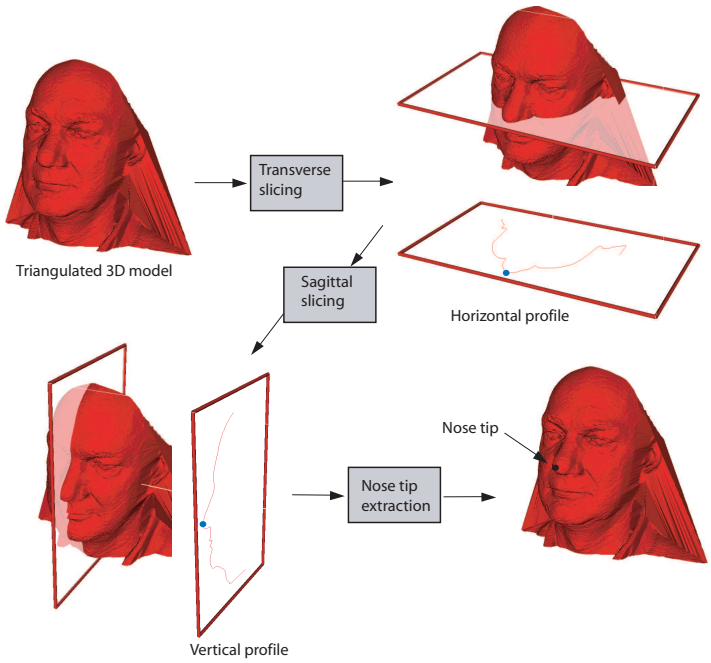


Figure 3.4: Automatic nose tip detection procedure.

dist, also called a **facial curve**, for the value $\lambda \in dist(S)$, i.e. $c_\lambda = \{p \in S | dist(P) = GD(r,p) = \lambda\} \subset S$ where r denotes the reference point (in our case the nose tip) and $GD(r,p)$ is the length of the shortest path from r to p on the mesh. We can reconstruct S from these level curves according to $= \bigcup_\lambda c_\lambda$.

- **Nose region segmentation** Detecting nose region is not easy in general. Fortunately, our approach deals with level curves and the nose tip is the reference point. So, it will be easy to delimit the nose region, in our case, by decreasing of value of λ and reconstruct the nose from level curves according to $N = \cup_\lambda c_\lambda$, for the value $\lambda \in D(N)$, i.e. $c_\lambda = \{p \in N | D(P) = GD(r,p) = \lambda\} \subset N$ where N denotes the nose region. The output of the preprocessing step is a nasal surface, as illustrated by Figure 3.5.

The preprocessing is one of the main issues in this such problems, so we have combined the preprocessing steps described above to develop an automatic algorithm. This algorithm

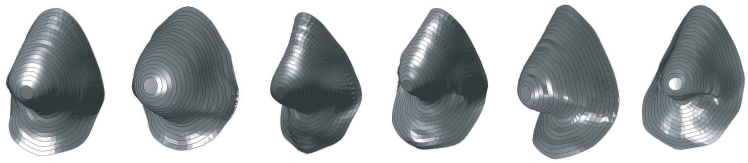


Figure 3.5: Examples of pre processed data: nose region with geodesic level curves

Table 3.1: Results of preprocessing procedure on FRGCv2 dataset.

	Original files	Success pre-processing	failed pre-processing	Success Rates (%)
Fall 2003	1893	1877	16	99.15
Spring 2004	2114	1994	20	98.99
FRGC v2	4007	3971	36	99.1

has successfully processed 3971 faces in FRGCv2, which means a success rate of 99.1% as described in the Table 3.1. This automatic preprocessing procedure failed for 16 faces taken in Fall 2003 and 20 faces taken in Spring 2004. Actually, it is the nose detection step that fails more than other steps. The main cause is the additive information which moves the mass center far from the face and the profiles extracted do no more match with the face profiles. For these faces, we have fixed manually the nose tip and so we have cleaned all the FRGC v2 faces, 99.1% automatically and 0.9% manually. The nose tip is considered correctly detected when the difference between the extracted coordinates of the nose tip against the coordinates provided with the dataset does not exceed 0.8 cm.

3.3 A Geometric Framework for nose shape analysis

As indicated earlier, our goal is to analyze shapes of nasal region surfaces using shapes of facial curves. In other words, we divide each surface into an indexed collection of simple, closed curves in $\mathbb{R}^3$ and the geometry of a surface is then studied using the geometries of the associated curves. Since these curves, previously called *facial curves*, have been defined as level curves of an intrinsic distance function on the surface, their geometries in turn

are invariant to the rigid transformation (rotation and translation) of the original surface. At least theoretically, these curves jointly contain all the information about the surface and one can go back-and-forth between the surface and the curves without any ambiguity. In practice, however, some information is lost when one works with a finite subset of these curves rather than the full set. Later, through experiments on real data, we will demonstrate that the choice of facial curves for studying shapes of facial surfaces is both natural and convenient.

In the following section, we will describe a differential-geometric approach for analyzing shapes of simple, closed curves in $\mathbb{R}^3$. In recent years, there have been several papers for studying shapes of continuous curves. The earlier papers, including [You98, KSMJ04, MM06, MSJ07a], were mainly concerned with curves in $\mathbb{R}^2$, while the curves in higher dimensions were studied later. In this chapter, we will follow the theory laid out by Joshi et al. [JKSJ07] for *elastic shape analysis* of continuous, closed curves in $\mathbb{R}^n$ and particularize it for facial curves in $\mathbb{R}^3$.

3.3.1 Nose curves

We start by considering a closed curve β in $\mathbb{R}^3$. Since it is a closed curve, it is natural to parametrize it using $\beta : \mathbb{S}^1 \to \mathbb{R}^3$. We will assume that the parametrization is non-singular, i.e. $\|\dot{\beta}(t)\| \neq 0$ for all t. The norm used here is the Euclidean norm in $\mathbb{R}^3$. Note that the parametrization is not assumed to be arc-length; we allow a larger class of parametrizations for improved analysis. To analyze the shape of β, we shall represent it mathematically using a *square-root velocity function* (SRVF), denoted by $q(t)$, according to:

$$q(t) \doteq \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}} \ . \tag{3.3.1}$$

$q(t)$ is a special function that captures the shape of β and is particularly convenient for shape analysis, as we describe next. Firstly, the squared $\mathbb{L}^2$-norm of q, given by:

$$\|q\|^2 = \int_{\mathbb{S}^1} \langle q(t), q(t) \rangle \, dt = \int_{\mathbb{S}^1} \|\dot{\beta}(t)\| dt \ ,$$

which is the length of β. Therefore, the $\mathbb{L}^2$-norm is convenient to analyze curves of specific lengths. Secondly, as shown in [JKSJ07], the classical elastic metric for comparing shapes of

curves becomes the $\mathbb{L}^2$-metric under the SRVF representation. This point is very important as it simplifies the calculus of elastic metric to the well-known calculus of functional analysis under the $\mathbb{L}^2$-metric.

In order to restrict our shape analysis to closed curves, we define the set:

$$\mathcal{C} = \{q : \mathbb{S}^1 \to \mathbb{R}^3 | \|q\| = 1, \int_{\mathbb{S}^1} q(t)\|q(t)\|dt = 0\} \subset \mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^3) \ . \tag{3.3.2}$$

Here $\mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^3)$ denotes the set of all functions from $\mathbb{S}^1$ to $\mathbb{R}^3$ that are square integrable. The quantity $\int_{\mathbb{S}^1} q(t)\|q(t)\|dt$ denotes the total displacement in $\mathbb{R}^3$ as one traverses along the curve from start to end. Setting it equal to zero is equivalent to having a closed curve. Therefore, $\mathcal{C}$ is the set of all closed curves in $\mathbb{R}^3$, each represented by its SRVF. Notice that the elements of $\mathcal{C}$ have the same length (equal to 1).

As illustrated by Figure 3.6, the set of these representations is a subset of a unit sphere of Hilbert space. Therefore the shortest path between elements in that space is projected to that subspace. It is not simply given by the greatest arc connecting the two points on the sphere.

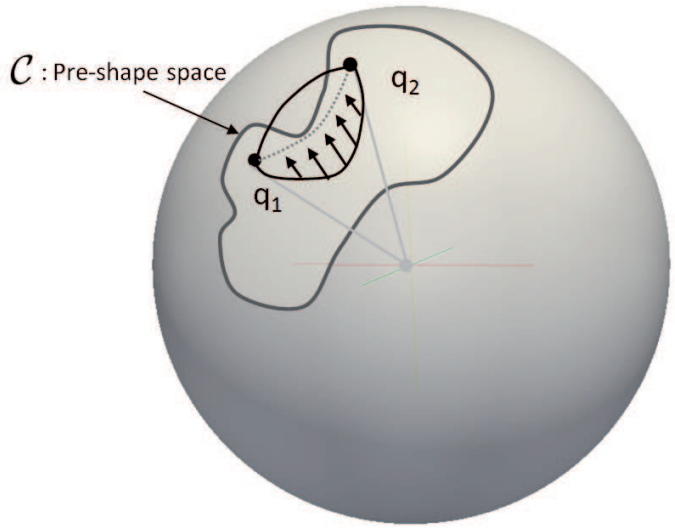


Figure 3.6: Illustration of the Pre-Shape Space $\mathcal{C}$ and geodesic in this pre-shape space.

Hence, due to a non-linear (closure) constraint on its elements, $\mathcal{C}$ is a non-linear manifold.

We can make it a Riemannian manifold by using the metric: for any $u, v \in T_q(\mathcal{C})$, we define:

$$\langle u, v \rangle = \int_{\mathbb{S}^1} \langle u(t), v(t) \rangle \, dt \ . \tag{3.3.3}$$

We have used the same notation for the Riemannian metric on $\mathcal{C}$ and the Euclidean metric in $\mathbb{R}^3$ hoping that the difference is made clear by the context. For instance, the metric on the left side is in $\mathcal{C}$ while the metric inside the integral on the right side is in $\mathbb{R}^3$. For any $q \in \mathcal{C}$, the tangent space:

$$T_q(\mathcal{C}) = \{v : \mathbb{S}^1 \to \mathbb{R}^3 | \, \langle v, w \rangle = 0, \ w \in N_q(\mathcal{C})\} \ ,$$

where $N_q(\mathcal{C})$, the space of normals at q is given by:

$$N_q(\mathcal{C}) = \text{span}\{\frac{q^1(t)}{\|q(t)\|}q(t) + \|q(t)\|\mathbf{e}^1, \frac{q^2(t)}{\|q(t)\|}q(t) + \|q(t)\|\mathbf{e}^2, \frac{q^3(t)}{\|q(t)\|}q(t) + \|q(t)\|\mathbf{e}^3\} \ ,$$

and where $\{\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3\}$ form an orthonormal basis of $\mathbb{R}^3$.

So far we have described a set of closed curves and have endowed it with a Riemannian structure. Next we consider the issue of representing the *shapes* of these curves. It is easy to see that several elements of $\mathcal{C}$ can represent curves with the same shape. For example, if we rotate a curve in $\mathbb{R}^3$, we get a different SRVF but its shape remains unchanged. Another similar situation arises when a curve is re-parametrized; a re-parameterization changes the SRVF of curve but not its shape. In order to handle this variability, we define orbits of the rotation group $SO(3)$ and the re-parameterization group Γ as the equivalence classes in $\mathcal{C}$. Here, Γ is the set of all orientation-preserving endomorphisms of $\mathbb{S}^1$ (to itself) and the elements of Γ are viewed as re-parameterization functions. For example, for a curve $\beta : \mathbb{S}^1 \to \mathbb{R}^3$ and a function $\gamma : \mathbb{S}^1 \to \mathbb{S}^1$, $\gamma \in \Gamma$, the curve $\beta(\gamma)$ is a re-parameterization of β. The corresponding SRVF changes according to $q(t) \mapsto \sqrt{\dot{\gamma}(t)}q(\gamma(t))$. We set the elements of the set:

$$[q] = \{\sqrt{\dot{\gamma}(t)}Oq(\gamma(t))|O \in SO(3), \ \ \gamma \in \Gamma\} \ ,$$

to be equivalent from the perspective of shape analysis. The set of such equivalence classes, denoted by $\mathcal{S} \doteq \mathcal{C}/(SO(3) \times \Gamma)$ is called the *shape space* of closed curves in $\mathbb{R}^3$. $\mathcal{S}$ inherits a Riemannian metric from the larger space $\mathcal{C}$ and is thus a Riemannian manifold itself.

The main ingredient in comparing and analyzing shapes of curves is the construction of a geodesic between any two elements of $\mathcal{S}$, under the Riemannian metric given in Eqn.

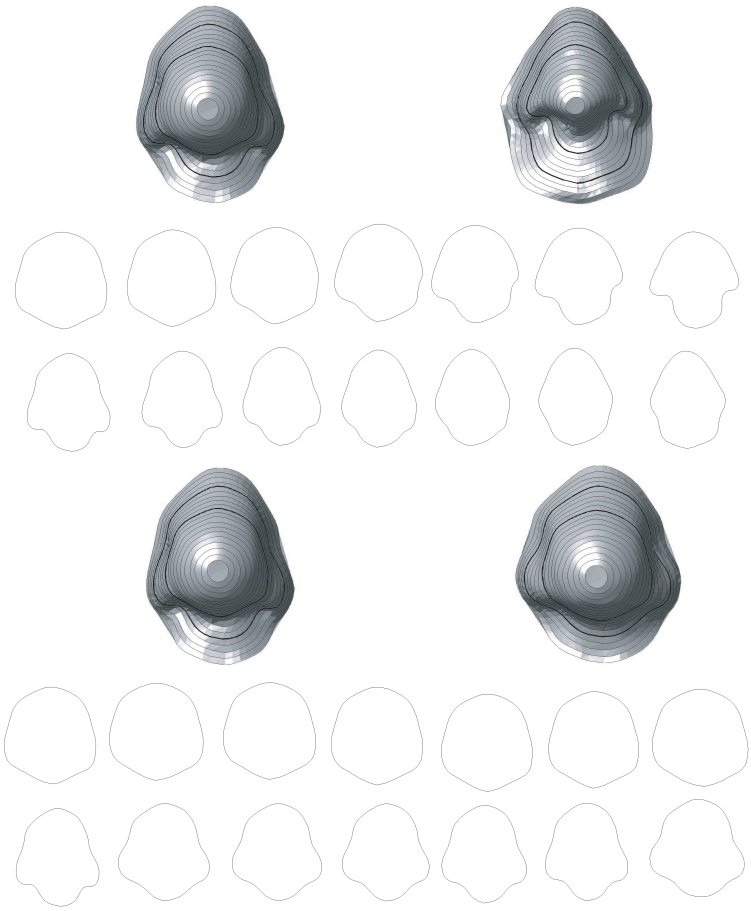


Figure 3.7: Examples of geodesic between curves

3.3.3. Given any two curves β_1 and β_2, represented by their SRVFs q_1 and q_2, we want to compute a geodesic path between the orbits $[q_1]$ and $[q_2]$ in the shape space $\mathcal{S}$. This task is accomplished using a *path straightening approach* which was introduced in [KS06]. The basic idea here is to connect the two points $[q_1]$ and $[q_2]$ by an arbitrary initial path ψ and to iteratively update this path using the negative gradient of an energy function $E[\psi] = \frac{1}{2}\int_s \left\langle \dot{\psi}(s), \dot{\psi}(s) \right\rangle ds$. The interesting part is that the gradient of E has been derived analytically and can be used directly for updating ψ. As shown in [KS06], the minimal values of E are actually geodesic paths in $\mathcal{S}$. Thus, this gradient-based update leads to a critical point of E which, in turn, is a geodesic path between the given points. This idea is

illustrated in Figure 3.8 where we show that the shortest distance between the two quotient spaces $[q_1]$ and $[q_2]$, is given by q_1 and q_2^*.

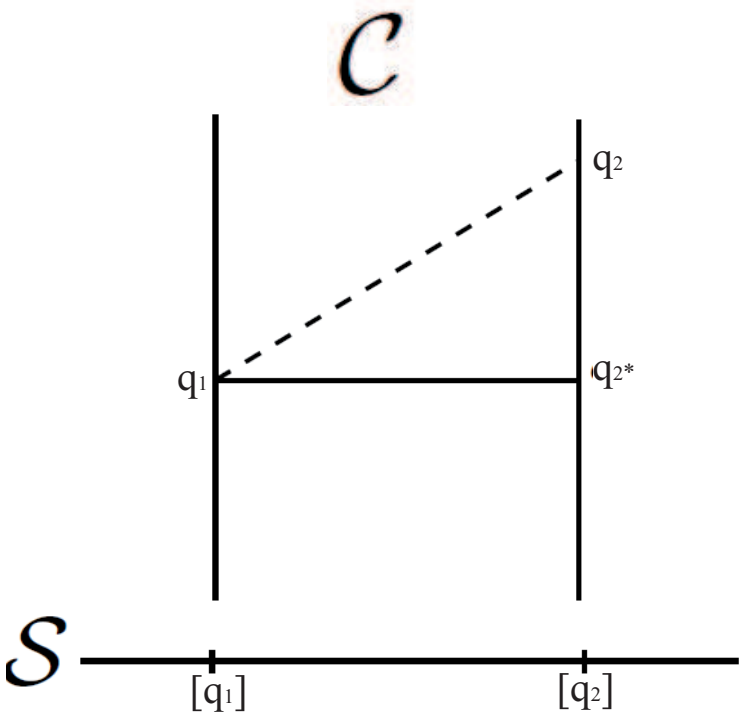


Figure 3.8: Computing geodesics in the quotient space.

3.3.2 Nose surfaces

Now we extend ideas developed in the previous section for analyzing shapes of nasal curves to the shapes of nose region surfaces. As mentioned earlier, we are going to represent a nose region surface S with an indexed collection of the level curves of the *dist* function.

That is,

$$N \leftrightarrow \{c_\lambda, \lambda \in [0, L]\} \ ,$$

where c_λ is the level set associated with $dist = \lambda$. Through this relation, each facial surface has been represented as an element of the set $\mathcal{S}^{[0,L]}$. In our framework, the shapes of any two noses are compared by comparing their corresponding nasal curves. Given any two nasal

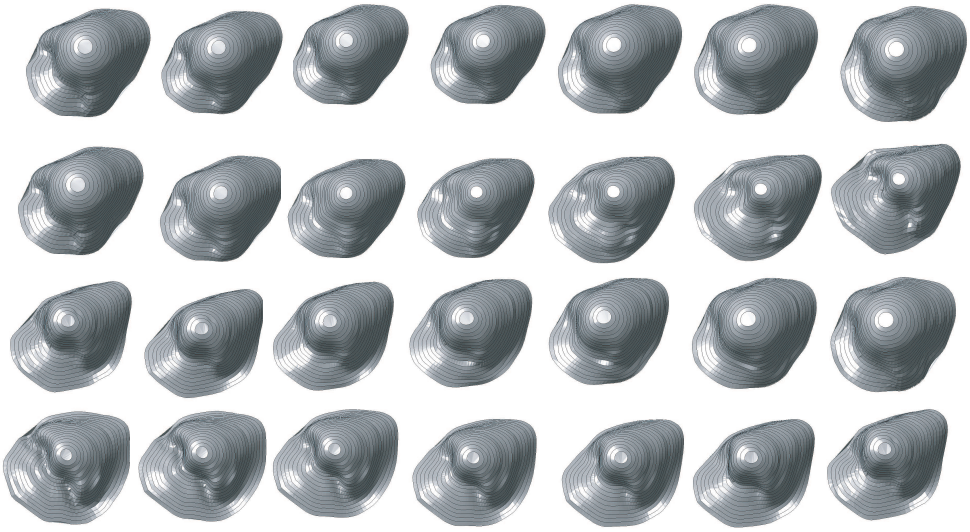


Figure 3.9: Geodesic path between source and target noses (a) First row: intra-class path, source and target with different expressions (b) Three last rows: inter-class path

surfaces N^1 and N^2, and their nasal curves $\{c^1_\lambda, \lambda \in [0, L]\}$ and $\{c^2_\lambda, \lambda \in [0, L]\}$, respectively, our idea is to compare the nasal curves c^1_λ and c^2_λ, and to accumulate these differences over all λ. More formally, we define a distance: $d_s : \mathcal{S}^{[0,L]} \times \mathcal{S}^{[0,L]} \to \mathbb{R}_{\geq 0}$, given by

$$d_s(N^1, N^2) = \int_0^L d(c^1_\lambda, c^2_\lambda) d\lambda \ . \tag{3.3.4}$$

Here, the distance inside the integral is the geodesic distance function between the shapes of any curves, described in the last section. It has been shown in [SSDK09] that this distance is actually a proper distance between the elements of the $\mathcal{S}^{[0,L]}$.

Given any two nasal surfaces N^1 and N^2, and their collection of curves $\{c^1_\lambda, \lambda \in [0, \lambda_0]\}$ and $\{c^2_\lambda, \lambda \in [0, \lambda_0]\}$, respectively, our idea is to compare the curves c^1_λ and c^2_λ, and to accumulate these distances over all λ. We define two possible metrics:

- **Arithmetic mean**: $d_a : \mathcal{S}^{[0,\lambda_0]} \times \mathcal{S}^{[0,\lambda_0]} \to \mathbb{R}_{\geq 0}$, given by $d_a(N^1, N^2) = \frac{1}{\lambda_0} \sum_{\lambda=1}^{\lambda_0} d(c^1_\lambda, c^2_\lambda)$.
- **Geometric mean**: $d_g : \mathcal{S}^{[0,\lambda_0]} \times \mathcal{S}^{[0,\lambda_0]} \to \mathbb{R}_{\geq 0}$, given by $d_g(N^1, N^2) = (\prod_{\lambda=1}^{L} d(c^1_\lambda, c^2_\lambda))^{(1/\lambda_0)}$.

In addition to the distance $d_s(N^1, N^2)$, which is useful in biometry and other classification experiments, we also have a geodesic path in $\mathcal{S}^{[0,L]}$ between the two points represented by N^1 and N^2. This geodesic corresponds to the optimal elastic deformations of nasal curves and, thus, nasal surfaces from one to other. Shown in Figure 3.9 are examples of such geodesic paths between given nasal surfaces – some examples involve faces of same people but with different facial expressions while the other examples show geodesics between faces that belong to different people.

3.4 Experiments

3.4.1 Preliminary results

Our framework is designed for computing geodesic path between nasal shapes via the shape of their curves and the distance d_s is used as the similarity score. Shown in Figure 3.10 are six examples of similarity matrices between noses and faces in different situations. Each of the six matrices in that figure shows similarity scores between 15×15 faces. The first case involves neutral expressions on the gallery and different neutral expressions on the probe, but of the same 15 people. The faces have been arranged in the same order so that the scores between the same person lies long the diagonal. Figure 3.10 illustrates these similarity matrices, the first row shows the results of the whole face and the second row gives the results on nose region. As expected, the values along the diagonal are in general much smaller (darker in the image) than the non-diagonal values.

The similarity matrices show that applying our approach on the nose region is better than applying it on the whole face when the probe faces represent large expressions. This result is expected as the large expressions do not affect the nose region.

One interesting target application of our framework is face authentication and recognition. In this section, we present some experimental evaluations produced on a subset of FRGCv2 database. We note that the main challenge in 3D face recognition is to handle deformations of faces caused by expressions or emotions.

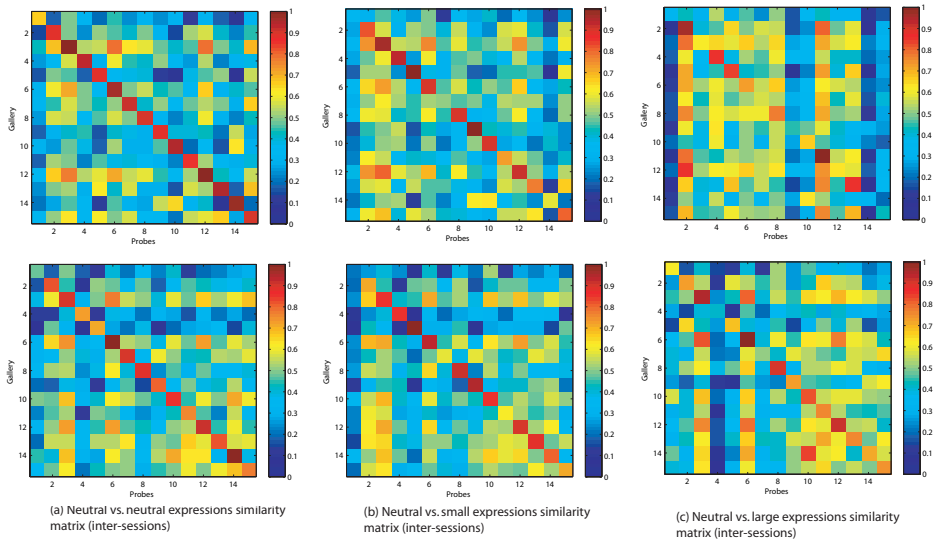


Figure 3.10: Similarity matrices for (a) neutral vs. neutral comparsions, (b) neutral vs. small expressions, and (c) neutral vs. large expressions, top row includes marices for face-to-face matching and buttom row includes nose-to-nose matching.

3.4.2 Large scale experiment

In this section we are going to present experiments on a large subset, the first 2000 scans, of FRGC v2. In order to produce results for both identification and authentication scenarios and to explore effect of the presence of facial expressions on performance, a similarity matrix between a gallery and a query datasets is computed. The gallery contains 2000 sessions for 209 different subjects taken from FRGC v2 dataset. The query set is identical to the gallery.

The effectiveness of d_a and d_g in biometric applications generally increases with the number of curves used. But the question is how to choose the curves which can give best results? Indeed, in the nasal region, there are some areas like nasal cavities which exhibit undesired variability. There are some tools and models, for instance feature selection, that explicitly conceived to target this goal. The results of identification and authentication scenarios will be presented separately in the following paragraphs.

Identification Scenario: The first type of experiment is an identification for which

performance is quoted as a rank-one recognition rate. In this experiment the gallery set consists of neutral faces of 209 subjects and the query set contains the remaining (2000−209) faces. Figure 3.11 shows the rank-one recognition rates obtained using both geometric and

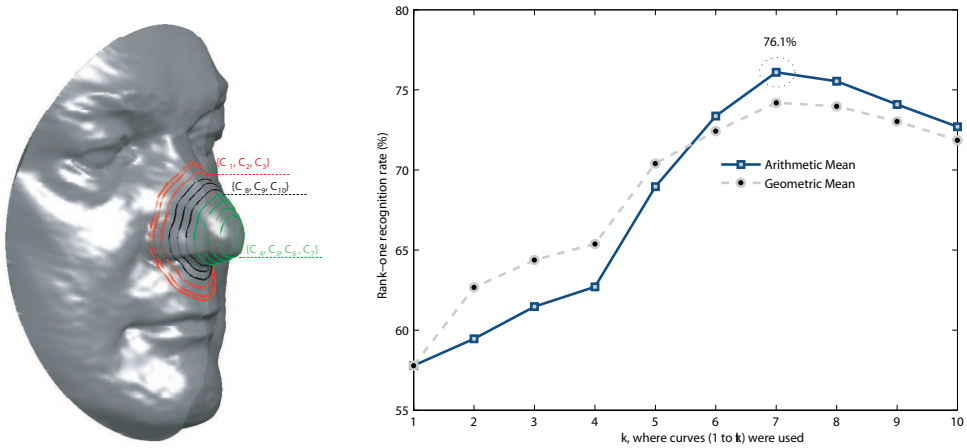


Figure 3.11: Rank one recognition rates using 1 to k curves

arithmetic mean (defined in section 3.3) in the right and the locations of these curves on a face in the left. We make the following remarks. First, we notice that the best recognition rate is obtained when using the arithmetic mean metric d_a. Actually, we obtain 76.1% using this metric using only seven curves. Second, we see that the recognition rate initially increases with the number of curves, reaches a peak and then starts decreasing, for both the metrics. The later curves which decrease the recognition rates are curves 8, 9 and 10 and, as we see in the Figure 3.11 (left), these curves pass through the nasal cavities. Clearly, shapes of this region (made up of cavities) are not reliable and can be excluded from the analysis. To compare our results with the literature, Faltemier et al. [FBF08] presented the recognition rates' results using many regions and on the whole FRGC v2 dataset; they report a 74.2% rank-one recognition rate in the region extracted with a sphere having the nose tip as center and 45mm radius. On a smaller region, extracted using a sphere with same center but 25mm radius, they reported the rate to be 80.5%. In contrast, our experiments

use first 2000 scans of FRGC and use a different region. Also, our approach is based on elastic deformations of curves while their approach is based on ICP. Our approach can handle deformations in surfaces more efficiently but fails when there are holes. Removal of curves from regions of nasal cavities improves the performance of our method.

Authentication Scenario: The second experiment involves authentication where the performance is measured using the verification rate (VR) for a given false acceptance rate (FAR) and with a receiver operating characteristic (ROC) curve that shows the trade-off between verification and false accept rates.

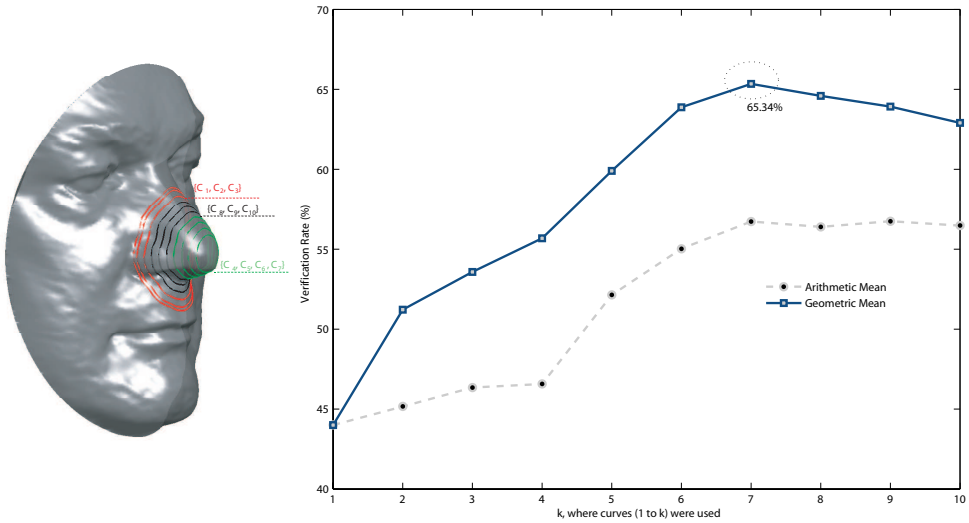


Figure 3.12: Verification Rate at FAR=0.1% using (1 to k) curves

Figure 3.12 (right) shows a plot of VR at FAR = 0.1%, computed using both the geometric and arithmetic means. The location of these curves on a face are shown in the left. The best verification rate is obtained using seven curves and the geometric mean metric; it is approximately 65.34%. As in the previous experiment, the verification rate increases with the number of used curves until $k = 7$ curves are used and then it decreases. The curves which decrease the recognition rates pass through the nasal cavities. Moreover, we notice that the verification rates using the geometric mean d_g are much better than those

using the arithmetic mean d_a. So we choose this metric for further experiments. Figure 3.13 shows the ROC curves for experiments involving different number k of curves. For $k < 7$, the ROC curves using curves 1 to k are higher for larger k. For $k > 7$, this trend is reversed. So, using a larger number of curves improves the authentication performance until the seventh curve and, after that, results get worse. Faltemier et al. [FBF08], for a fixed FAR = 0.1% using the region extracted with a sphere having the nose tip as center and 45mm radius on the whole FRGC v2 dataset and they reported a VR of 57.2%. On a smaller region, extracted using a sphere having always the nose tip as a center but a 25mm radius, they reported VR of 73.1% at a 0.1% FAR.

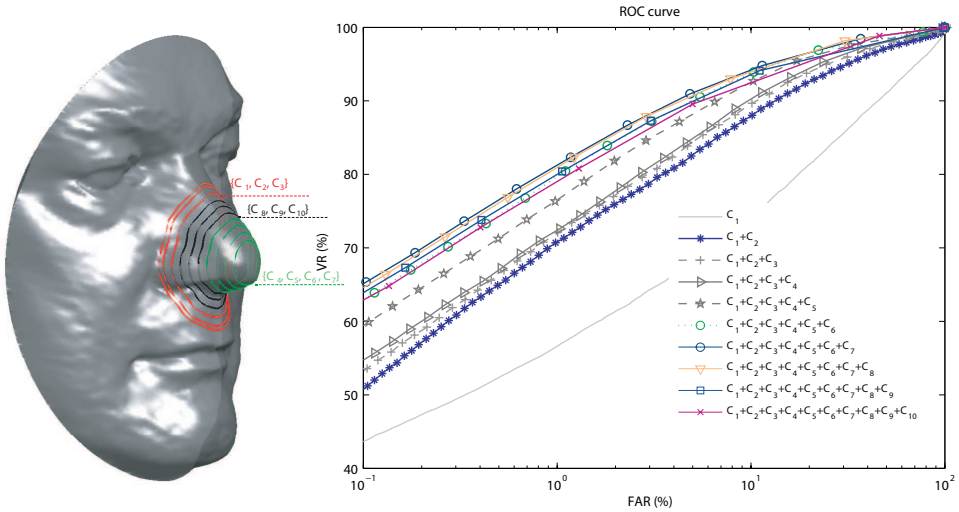


Figure 3.13: ROC curves using 1 to k curves

3.4.3 Our approach vs. rigid-based face and nose matching

In order to compare further our approach to rigid matching, a similarity matrix between a gallery and a probe dataset is computed. The gallery dataset contains 125 sessions for 125 different subjects acquired with neutral expressions selected from FRGCv2 dataset. The probe dataset includes completely different sessions of these subjects under non-neutral

facial expressions. Due to sensitivity of our algorithm to opened mouth, expressions in probe dataset include only scans with closed mouths.

In this matrix, the diagonal terms represent match scores (or Genuine Access) contrary to non-diagonal terms which represent Non-match scores (or Imposter Access). These scores allow us to produce the ROC and the CMC curves for this scenario.

We compare results of our algorithm with a standard implementation of ICP which is considered as a baseline in 3D face recognition. Baseline performance serves to demonstrate that a challenge problem can be executed, to provide a minimum level of performance, and to provide a set of controls for detailed studies. The same protocol was followed to compute similarity matrices for both the algorithms on the same preprocessed data (ICP was applied on nose regions of faces).

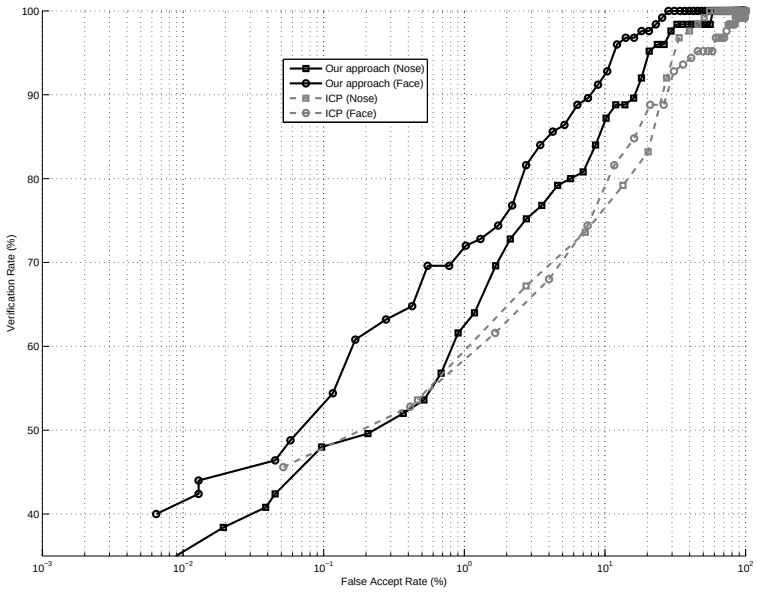


Figure 3.14: Receiver operating characteristic curves for our approach and ICP (baseline)

Figure 3.14 shows ROC curves for our approach on face, on nose region and ICP algorithm on nose and on face. As shown in 3.14, the ROC curves (for face and nose region)

of our approach are almost always above the ICP ones. Which means that our verification rates at each false accept rate are greater than ICP ones. This result is expected as our approach is more robust to face expressions.

As far as ICP is concerned, ROC curves for face and nose region are overlapped. Therefore, the nose region do not approve ROC curve of our approach applied on face is above the one of our approach itself applied using only nose region. We can say so that the verification rate at each false accept rate has decreased when using only nose region. ROC curve of our approach applied only on nose region is often above ICP ROC curves, applied on nose region and on the whole face.

Therefore, using only nose region in our approach gave a better verification results than ICP applied on the whole face. Although, using the whole face with our approach give always the better verification results.

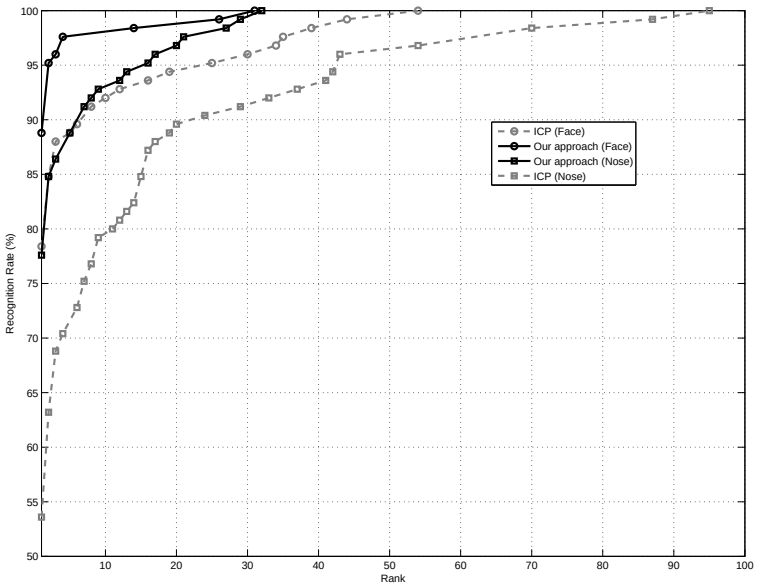


Figure 3.15: Cumulative Match Characteristic curves for our approach and ICP (baseline)

These observations are also confirmed in identification scenario which is illustrated with CMC curves of the two algorithms applied on nose region and on the whole face. As shown in (Figure 3.15), rank-one recognition rate given by our algorithm applied on the whole face is the best one, it's about 88.8%. Then results of rank-one recognition for our approach applied on nose region and ICP applied on the whole face are similar, about 78%. The worst result is given by ICP applied on nose region.

At the rank-four, our approach on the whole face is able to recognize 97.8% of the subjects in contrast with ICP on the whole face which gives a similar result with our approach on the nose region, it's about 89%. After this rank, the recognition rates given by our approach on the nose region are always better than the ones given by ICP applied on the whole face. So, using the whole face for our approach gives always the best recognition and authentication results. However, applying our approach only on the nose region gives better results than ICP applied on the whole face.

Figure 3.16 shows some examples of some noses (with their faces) which are recognized by our approach but not by the base line (ICP). This result can be explicated by the nature of the approaches themselves. Actually, as known, ICP algorithm search for the best rigid transformations to apply on a 3D points cloud (which represents the first face or nose) to be the closest to the second 3D points cloud (the second face or nose). As shown in the figure (Figure 3.16), 3D points positions have widely changed and it is difficult to ICP to find a smallest rigid transformation between these noses in probe and their correspondent in gallery. Although our approach successfully recognize them as it is based on shape comparison. As we see in the (Figure 3.16), the shape of the nose region looks the same although the non-rigid transformation between gallery and probe.

Unfortunately, our approach doesn't recognize some noses and we show in figure (Figure 3.17) some exemples of these noses with their faces and the noses that our algorithm ranks first. The result of our approach on the the whole face is better than his on the nose region. However, we have already with the nose region a better result than ICP. Moreover, the use of the nose region success to accelerate our approach, as shown is (table 3.2).

Table 3.2: Time consuming of the base line (ICP) and of our approach on the face and the nose region for preprocessing for extracting the face descriptors and recognition by matching face descriptors.

Method	**time (s)**
ICP	1
Our approach on the whole face	20
Our approach on the nose region	8

3.5 Conclusions

In this chapter, we studied the contribution of the shape of nose region in person identification. Our choice is based on plenty of psychological evidence that certain facial parts, especially those around nose can capture the essential features of a face. The main tool presented in this chapter is the construction of geodesic paths between arbitrary two nasal surfaces. The length of a geodesic between any two nasal surfaces is computed as the geodesic length between a set of their nasal curves. Our experiments support this idea in a mathematical way. We have computed geodesic distances between corresponding facial curves of different faces – same people different facial expressions and different people altogether. We have found that the distances are typically smaller for faces of the same people, although with different expressions, when compared to the distances between facial curves of different people. We have demonstrated these ideas in two application contexts : the authentication and the identification biometric scenarios using nasal shapes on a large dataset involving 2000 scans. We performed 76.1% as rank-1 recognition on this subset of FRGCv2 using 7 level curves. It is experimentally demonstrated that curves passing through nose holes decrease the performance.

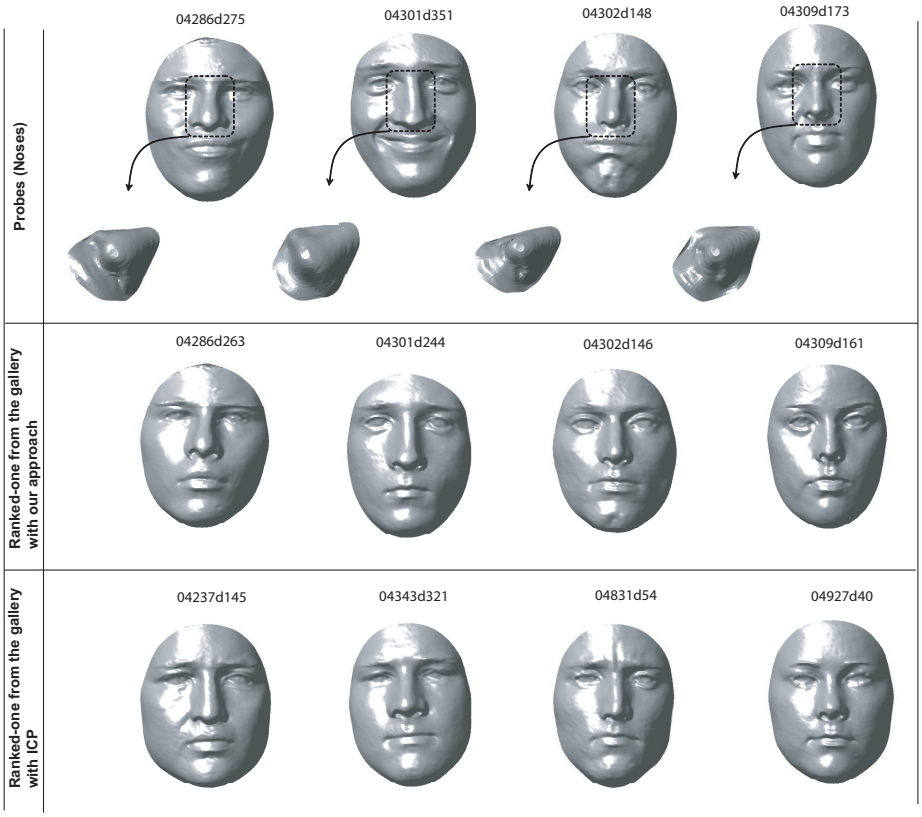


Figure 3.16: Examples of noses that are recognized by our approach and not by ICP algorithm on nose region

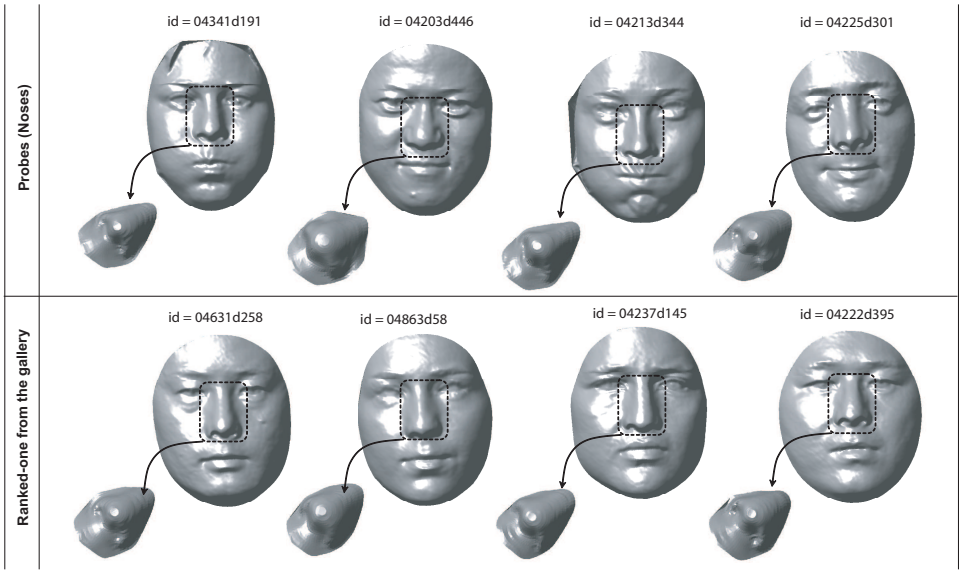


Figure 3.17: Examples of noses that are not recognized by our approach

Chapter 4

3D Face Recognition using Elastic Radial Curves

4.1 Introduction

The 3D face recognition problem can be considered as a special case of a more general 3D object recognition problem. An important distinction of human faces when compared to other 3D objects is that they all share common attributes and are similar to each other. Moreover, scan data often suffer from the problem of missing parts due to self-occlusions or imperfections in scanning technologies.

Additionally, variations in face data due to facial expressions, as shown in Figure 4.1.b, are challenging to 3D face recognition. In order to be useful in real-world applications, 3D face recognition approaches should be able to successfully recognize face scans even in the presence of large expression-based deformations and missing data due to occlusions and pose variation, as shown in Figure 4.1.c and Figure 4.1.d. Most recent research has been directed towards expression-invariant techniques and spent less effort to handle the missing parts problem. Few approaches handle the missing part problem but none has performed on a full database containing real missing data, they simulate some missing parts. In this chapter, we present a common framework handling both large expressions and missing parts due to large pose variation.

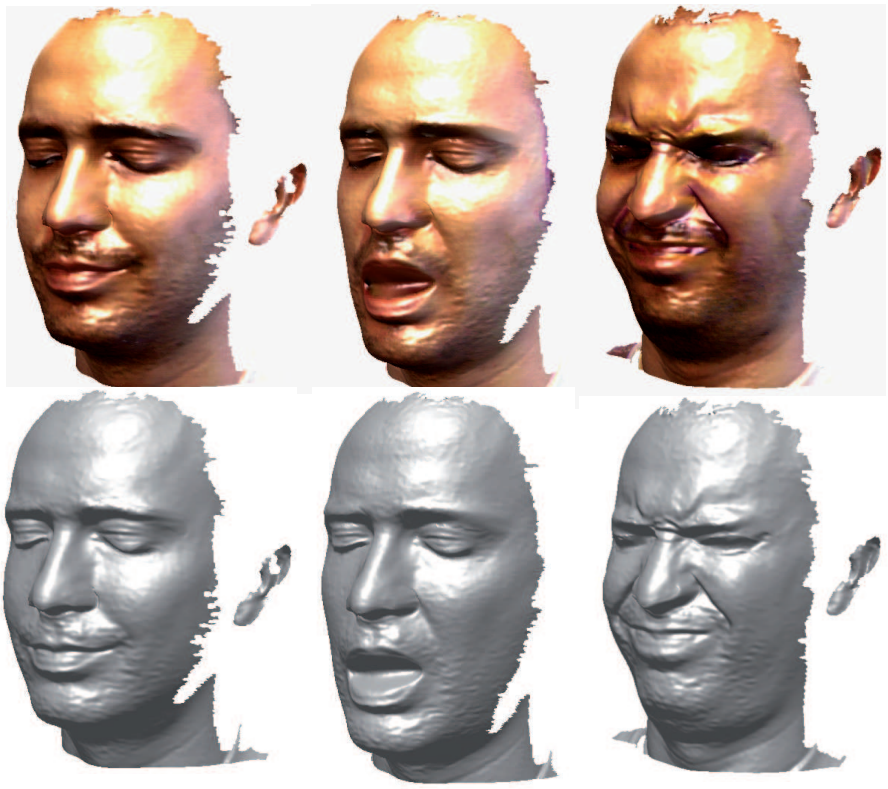


Figure 4.1: Important challenges of 3D face recognition. (a) neutral face, (b) expression variations, (c) missing parts due to scan technology and pose variation, (d) pose and expression variations

We propose a pose and expression robust 3D face recognition approach. Our approach is able to model facial deformations due to large expressions, even when the mouth is open. However, unlike some previous approaches [BBK07, MMS07], our approach deals with open mouths without any lip detection. Another contribution of this chapter is the robustness to missing data due to imperfect scan technology and huge pose variations. We automatically detect then discard curves passing through missing parts in order to avoid their influence. We demonstrate the effectiveness of our approach by comparing with the state-of-the-art results.

The rest of this chapter is organized as follows. In section 2, we present the limitation of the isometric model for faces under expressions and the overview of our proposed approach. Section 3 describes the pre-processing step. The representation of face by radial curves

is presented in section 4 and the curve elastic matching for large expression modeling is resumed in section 5. Section 6 presents how our approach deals with missing data due to huge pose variation. The experimental results and discussions are shown in section 7.

4.2 Introduction and overview of the proposed approach

Verifying quantitatively that facial expressions are indeed isometric is possible by tracking a set of feature points on the facial manifold and measuring how the distances between them change when varing facial expressions. Bronstein et al. [BBK05] placed 133 markers on a face and tracked how the distances between these points change due to facial expressions. The distribution of the absolute change of the geodesic and the Euclidean distances shows that geodesic distance is more stable than euclidean one under facial expression variations. This implies their assumption of the isometric model of facial expressions. This assumption is valid in the case of small expression. The large expression causes some combinations of shrinking and/or stretching. Under such variations, the geodesic distance is no more stable. In order to quantitatively verify this assumption, we placed 4 markers on a face and tracked the change of the geodesic and euclidean distances under large expression variations.

Figure 4.2: Significant changes in both Euclidean and surface distances under different face expressions.

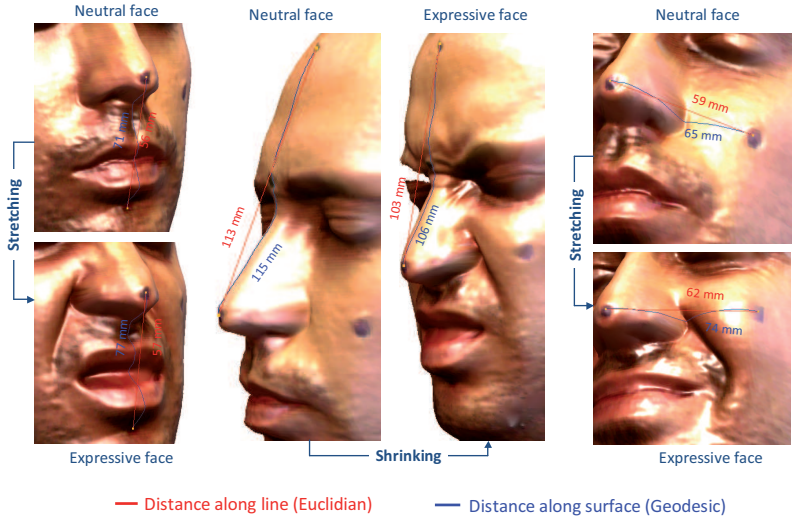


Figure 4.2 shows that a change in expression from neutral to non-neutral generally results in a shrinking or a stretching of the face shape, and both the euclidean and geodesic distances between points on the face are changed.

In one of the illustrated cases these distances decrease (from 113 mm to 103 mm for Euclidean distance, and from 115 mm to 106 mm for the Geodesic distance) while in the other two cases they increase. This clearly shows that large expression variations cause stretching and shrinking of face shape and under such elastic deformations neither the Euclidean distance nor the surface distance is preserved. Hence, the assumption of isometric deformation of the shape of face is not valid. One way to handle such elastic deformations of faces due to changes in expressions is to use an elastic shape analysis.

Geodesic level curves are no more significant for large expression variations. We propose to use radial curves with elastic matching in order to handle large expressions variations and pose variations. Radial curves were used by several authors [tHV09] for 3D face recognition. Radial and level curves are seen as features and best recognition rates are obtained with a set of 8 level curves.

Figure 4.3 provides an overview of the proposed 3D face recognition method. 3D faces of the probe **P** and the gallery **G** are firstly preprocessed. This step is essential to improve the quality of depth images and to extract informative parts of the face. Then, a coarse alignment is performed based on the translation vector formed by the tips of the noses. This step is followed by a finer alignment using ICP algorithm to normalize the pose. Next, we extract radial curves emanating from the tip of the nose and having different directions on the face. Within this step, a quality control module inspects the quality of each curve on both meshes and keeps only the *good* ones based on a pre-defined criterion.

4.3 Automatic data preprocessing

As illustrated in Figure 4.3, we start by preprocessing the input raw images to extract the relevant geometries and to improve data quality. Indeed, these images contain a number of imperfections such as holes, spikes, and include some undesired parts such as clothes, neck, ears and hair. The pre-processing step consists of the following pipeline of 3D mesh

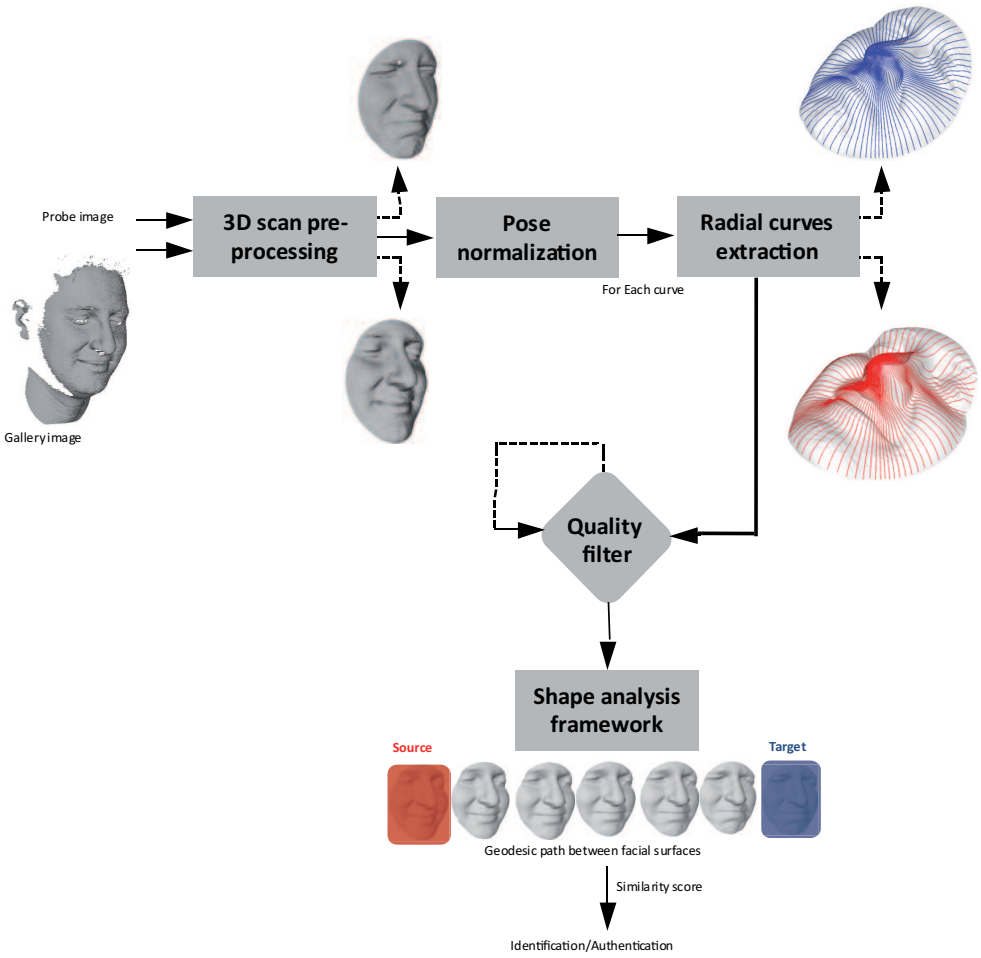


Figure 4.3: Overview of the proposed method.

processing filters as illustrated in Figure 4.4:

- *Filling holes filter* identifies and fills holes in input meshes. Holes are created either because of the absorption of laser in dark areas such as eyebrows and mustaches, self-occlusion or mouth opening. They are identified in the input mesh by locating boundary edges, linking them together into loops, and then triangulating the resulting loops.
- *A cropping filter* cuts and returns parts of the mesh inside a sphere centered at the nose tip of radius 75mm in order to discard as much hair as possible.

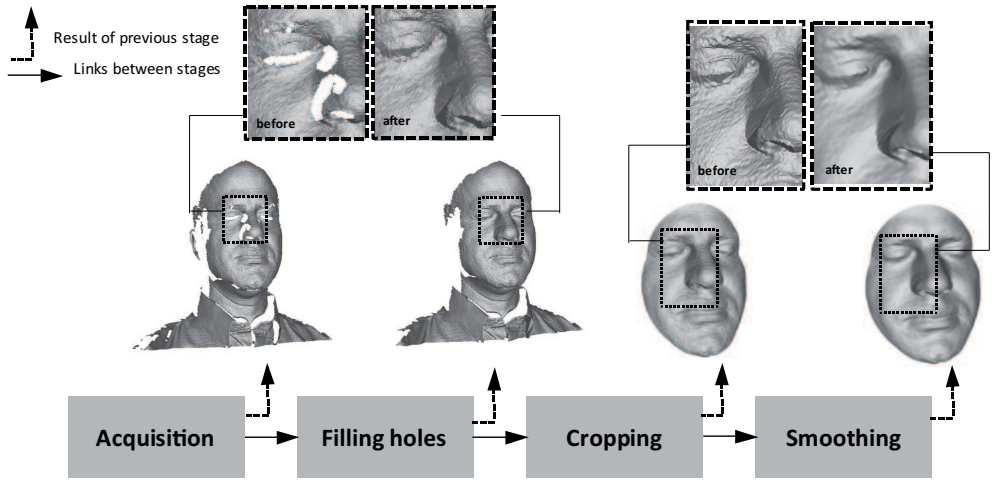


Figure 4.4: Different steps of preprocessing: acquisition, filling holes, cropping and smoothing

- *A smoothing filter* reduces high frequency information (spikes) in the geometry of the mesh, makes well-shaped cells, and distributes evenly the vertices.

Implementations of vtk library were used to apply the smoothing filter.

Our recognition approach is based on elastic matching of radial curves extracted from facial surfaces. After preprocessing and alignment stages, we represent each facial surface by an indexed collection of radial curves that are extracted as follows.

4.4 Facial shape representation using radial curves

Let S be a facial surface obtained from the output of the previous preprocessing step. Although S is a triangulated mesh, we start the discussion by assuming that it is a continuous surface. Let β_α denote the radial curve on S which makes an angle α with a reference radial curve. The reference curve is chosen to be the vertical curve once the face has been rotated to the upright position. In practice, each radial curve β_α is obtained by slicing the facial surface by a plane P_α that has the nose tip as its origin and makes an angle α with the plane containing the reference curve. That is, the intersection of P_α with S gives the radial curve β_α. We repeat this step to extract radial curves from the facial

surface at equal angular separation. Each curve is indexed by the angle α. Figure 4.5 shows an example of some radial curves.

If needed, we can approximately reconstruct S from these radial curves according to $S \approx \cup_\alpha \beta_\alpha = \cup_\alpha \{S \cap P_\alpha\}$ as illustrated in Figure 4.5. This indexed collection of radial curves captures the shape of a facial surface and forms its mathematical representation. In practice, we used 80 curves.

We have chosen to represent a surface with a collection of curves since we have better tools for elastically analyzing shapes of curves than we have for surfaces. More specifically, we are going to utilize an elastic method for studying shapes of curves that is especially suited for modelling deformations associated with changes in facial expressions.

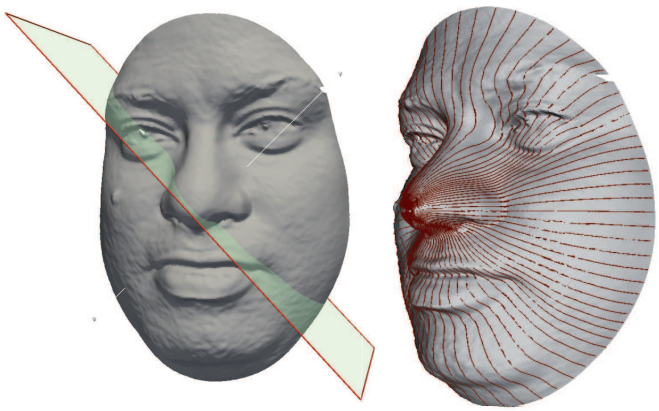


Figure 4.5: Radial curves extraction : left image illustrates the intersection between the face surface and a plan to form two radial curves. The collection of radial curves is illustrated at the right image.

4.5 Elastic metric for expression invariant 3D face recognition

4.5.1 Motivation for the elastic metric

Deformations on the facial surfaces are transferred to the radial curves. We need to quantify shape deformations of curves involving bending, stretching and shrinking. Consider the two

curves in Figure 4.6.a. Let us fix the parametrization of the top curve to be arc-length. That is we are going to traverse that curve with speed equal to one. In order to better match the curve with the down one, one should know at what rate we are going to move along the second curve so that points reached at the same time on two curves are as close as possible under some geometric criterion. In other words, peaks and valleys should be reached at the same time. Figure 4.6.b illustrates the matching where point 1 on the top curve matches to point 11 on the down curve. The part between the point 1 and 2 on the top curve shrinks on the curve 2. Therefore, the point 2 matches the point 22 on the second curve. An elastic metric is the measure of that shrinking. This is termed an *elastic matching* [SKJJ11].

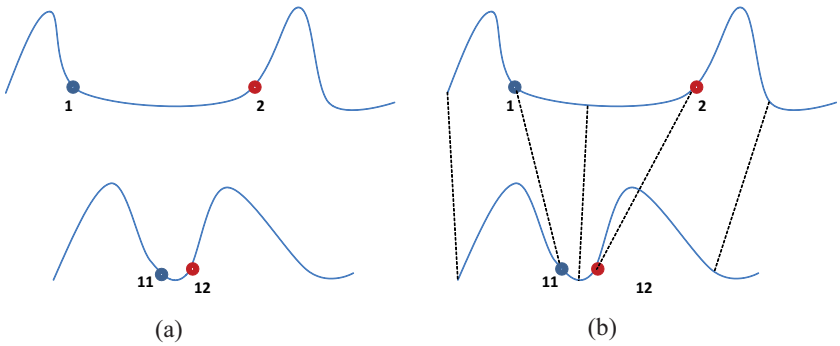


Figure 4.6: Illustration of elastic metric. In order to compare the two curves in (a), some combination of stretching and bending are needed. The elastic metric measures the amounts of these deformations. The optimal matching between the two curves is illustrated in (b).

4.5.2 Elastic metric

Let B the set of all parametrized curves in $\mathbb{R}^3$. $\beta : I \to \mathbb{R}^3$. β is supposed to be continuous and $\dot{\beta}(t)$ exists almost everywhere.

Let β an element of B whose derivative never vanishes: $\dot{\beta}(t) \neq 0, \forall t$. $\dot{\beta}$ can be seen as: $\dot{\beta}(t) = \exp(\phi(t))\upsilon(t)$, where ϕ represent the log-speed and $\upsilon(t)$ represent the direction vector as: $\phi(t) = log(\|\dot{\beta}(t)\|)$ and $\upsilon(t) = \frac{\dot{\beta}(t)}{\|\dot{\beta}(t)\|}$. Clearly $\upsilon(t)$ and ϕ completely specify $\dot{\beta}$ and the curve is seen as element in $\Phi \times \Upsilon$, where $\Phi = \{\phi : [0,1] \to \mathbb{R}\}$ and $\Upsilon = \{\upsilon : [0,1] \to \mathbb{S}^{n-1}\}$. Intuitively, ϕ tells us the (log of the) speed of traversal of the curve, while υ tells us direction

of the curve at each time t. In order to quantify the magnitudes of perturbations of β, a metric on $\Phi \times \Upsilon$ should be defined. First we note that the tangent space of $\Phi \times \Upsilon$ at any point (ϕ, υ) is given by

$$T_{(\phi,\upsilon)}(\Phi, \Upsilon) = \{(u, v) : u \in T_\phi\Phi, v \in T_\upsilon \mathbb{S}^{n-1}\} \tag{4.5.1}$$

with $T_\phi\Phi = \Phi$. as it is a linear space. Υ is a hypersphere in the Hilbert space $\mathbb{L}^2(I, \mathbb{R}^3)$ and its tangent space is given by:

$$T_\upsilon\Upsilon = \{f, f : [0, 1] \to \mathbb{R}^3 \ , \forall t, < f(t), \upsilon(t) >= 0\} \tag{4.5.2}$$

Suppose (u^1, f^1) and (u^2, f^2) are both elements of $T_{(\phi,\upsilon)}(\Phi, \Upsilon)$ and let a and b be positive numbers.

Definition [SKJJ11]***(Elastic Metric)***. For every point $(\phi, \upsilon) \in (\Phi \times \Upsilon)$, define an inner product on the tangent space $T_{(\phi,\upsilon)}(\Phi, \Upsilon)$ as:

$$< (u^1, f^1), (u^2, f^2) >= a^2 \int_0^1 u^1(t)u^2(t) \exp(\phi(t)) \, \mathrm{d}t + b^2 \int_0^1 < f^1(t), f^2(t) > \exp(\phi(t)) \, \mathrm{d}t \tag{4.5.3}$$

Note that $<, >$ in the second integral on the right denotes the standard dot product in $\mathbb{R}^n$. This elastic metric has the interpretation that the first integral measures the amount of *stretching*, since u^1 and u^2 are variations of the log speed ϕ of the curve, while the second integral measures the amount of *bending*, since f^1 and f^2 are variations of the direction υ of the curve [SKJJ11]. Therefore, the choice of weights a and b determines relative penalty on bending and stretching and a family of elastic metric is formed [MSJ07b]. The use of this metric to compare radial facial curves is motivated by the fact that the groups $SO(n)$ and Γ both act by isometries. Let $O \in SO(n)$ acts on a facial curve β by $(O, \beta)(t) = O\beta(t)$ and $\gamma \in \Gamma$ acts on β by $(\gamma, \beta)(t)$. $O \in SO(n)$ acts on (ϕ, υ) by $(O, (\phi, \upsilon)) = (\phi, O\upsilon)$ and $\gamma \in \Gamma$ acts on (ϕ, υ) by $(\gamma, (\phi, \upsilon)) = (\phi \circ \gamma + ln \circ \dot{\gamma}, \upsilon \circ \gamma)$. $O \in SO(3)$ acts by the restriction of a linear transformation on the tangent space of $\Phi \times \Upsilon$: $(O, (u, f)) = (u, Of)$, where $(u, f) \in T_{(\phi,\upsilon)}(\Phi, \Upsilon)$ and $(u, Of) \in T_{(\phi,O\upsilon)}(\Phi, \Upsilon)$.

The action of γ given in the above formula is affine linear, because of the term $ln \circ \gamma$. This make its action on the tangent space the same, but without this additive term:

$(\gamma,(u,f)) = (u\circ\gamma, \upsilon\circ\gamma)$, where $(u,f) \in T_{(\phi,\upsilon)}(\Phi,\Upsilon)$ and $(u\circ\gamma,\upsilon\circ\gamma) \in T_{(\gamma,(\phi,\upsilon))}(\Phi,\Upsilon)$. Combining the action of $SO(3)$ and Γ with the inner product presented in equation 4.5.3 on (Φ,Υ), it is easy to verify that these actions are by isometries, *ie*,
$< (O,(u_1,f_1)),(O,(u_2,f_2)) >_{(O,(\phi,\upsilon))} =< (u_1,f1),(u_2,f_2) >_{(\phi,\upsilon)}$
$< (\gamma,(u_1,f_1)),(\gamma,(u_2,f_2)) >_{(\gamma,(\phi,\upsilon))} =< (u_1,f1),(u_2,f_2) >_{(\phi,\upsilon)}$

We note that regardless of the values of a and b, both the groups $SO(3)$ and Γ act by isometries. An important question is: Is there some particular choice of weights a and b to make calculus easier? We propose to use the SRV representation already used in previous chapter for its simplicity of calculus. In this chapter the SRV representation finds its potential for elastically match facial curves.

4.5.3 Square Root Velocity representation SRV

In term of (ϕ,υ), SRV is given by $q(t) = \exp(\frac{1}{2}\phi(t))\upsilon(t)$. The tangent vector to $\mathbb{L}^2(I,\mathbb{R}^n)$ at q is given by a simple derivation calculus as: $h = \frac{1}{2}\exp(\frac{1}{2}\phi)u\upsilon + \exp(\frac{1}{2}\phi)f$. Let (u_1,f_1) and (u_2,f_2) denote two elements of $T_{(\phi,\upsilon)}(\Phi,\Upsilon)$, and let h_1 and h_2 denote the corresponding tangent vectors to $\mathbb{L}^2(I,\mathbb{R}^n)$ at q. The $\mathbb{L}^2$ inner product of h_1 and h_2 is given by:

$$< h_1,h_2 >= \int_0^1 < \frac{1}{2}\exp(\frac{1}{2}\phi)u_1\upsilon + \exp(\frac{1}{2}\upsilon)f_1, \frac{1}{2}\exp(\frac{1}{2}phi)u_2\upsilon + \exp(\frac{1}{2}\upsilon)f_2 > \,\mathrm{d}t \quad (4.5.4)$$

$$< h_1,h_2 >= \int_0^1 (\frac{1}{4}\exp(\phi)u_1u_2 + \exp(\phi) < f_1,f_2 >)\,\mathrm{d}t \quad (4.5.5)$$

In this computation, $\upsilon(t)$ is an element of the unit sphere hence the fact $< \upsilon(t),\upsilon(t) >= 1$ was used to reduce the formula. It was used also that $< \upsilon, f_i(t) >= 0$ since each $f_i(t)$ is a tangent vector to the unit sphere at $\upsilon(t)$.

This expression illustrates a particular elastic metric: for $a = \frac{1}{2}$ and $b = 1$. Therefore, the $\mathbb{L}^2$ metric on the shape of SRV representations corresponds to the elastic metric on $\Phi\times\Upsilon$ and this makes the calculus simpler. Actually, expressed in terms of SRV, the $\mathbb{L}^2$ metric does not depend on the point at which these tangent vectors are defined. Finally,

the inner product is simply given by:

$$< h_1, h_2 >= \int_0^1 < h_1(t), h_2(t) > \, \mathrm{d}t \tag{4.5.6}$$

In term of β, the SRV map is defined as: $SRV : B \to \mathbb{L}^2(I, \mathbb{R}^3)$

$$q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}} \ . \tag{4.5.7}$$

if $\dot{\beta}(t) \neq 0$ and 0 otherwise.

4.5.4 Riemannian elastic metric on open curves

Let $\beta : I = [0, 1] \to \mathbb{R}^3$, represent a radial curve on the face. To analyze the shape of β, we shall represent it mathematically using the *square-root velocity function* (SRVF), denoted by $q(t)$, according to:

$$q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}} \ . \tag{4.5.8}$$

$q(t)$ is a special function of β that we already used in previous chapter for its simplicity of calculus. Actually, the classical elastic metric for comparing shapes of curves becomes the $\mathbb{L}^2$-metric under the SRVF representation. This point is very important as it simplifies the calculus of elastic metric to the well-known calculus of functional analysis under the $\mathbb{L}^2$-metric. Hence, the SRV representation finds its potential for its ability for elastic matching. Actually, under $\mathbb{L}^2$-metric, the re-parametrization group acts by isometry on the manifold of q function (or SRV representation). This is not valid in the case of β. More formally, let β_1 and β_2 represent two radial curves and $\Gamma = \{\gamma : [0, 1] \to [0, 1], \gamma \text{ is a diffeomorphism } \}$ the set of all re-parametrizations.

$$\|\beta_1 - \beta_2\| \neq \|\beta_1 \circ \gamma - \beta_2 \circ \gamma\|. \tag{4.5.9}$$

The use of SRV representation allows the re-parametrization group to act by isometry on the manifold of SRV representations. This point is very important as the curve matching

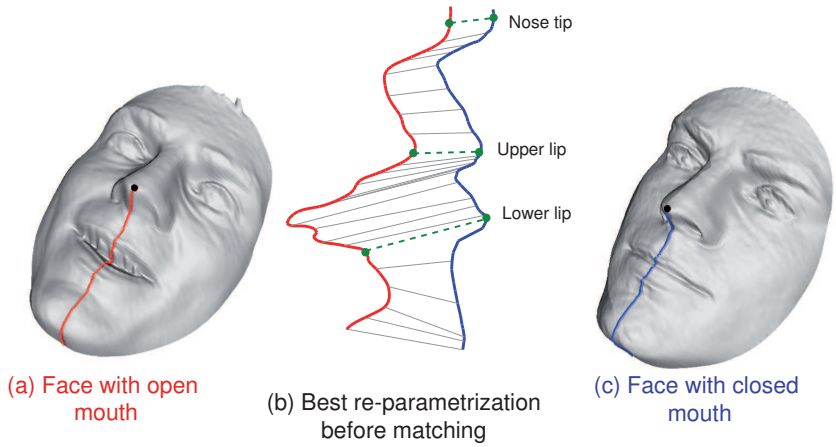


Figure 4.7: An example of matching of two radial curves extracted from two faces. a) A curve on an open mouth, (c) a curve on closed mouth , b) change of parametrization before matching

could be done after re-parametrization. The change of parametrization before the matching is able to reduce the effect of stretching and/or stretching of the curve.

This idea is illustrated in Figure 4.7. The task is to match two radial curves on two faces with two different expressions. The expression in the face at the left induces open mouth in contrast with the expression in face at the right. As shown in the middle panel, the anatomic points on the curves (upper and down lips) match together after re-parameterizing at least one curve. More formally, the elastic matching of the curves allows better matching of anatomical points on them.

We define the set (Pres-shape space):

$$\mathcal{C} = \{q : I \to \mathbb{R}^3, \|q\| = 1\} \subset \mathbb{L}^2(I, \mathbb{R}^3) \ . \tag{4.5.10}$$

The closure condition described in previous chapter is no more necessary in this case. This simplifies the calculus, with the $\mathbb{L}^2$ metric on its tangent spaces, $\mathcal{C}$ becomes a Riemannian manifold. In particular, since the elements of $\mathcal{C}$ have a unit $\mathbb{L}^2$ norm, $\mathcal{C}$ is a hypersphere in the Hilbert space $\mathbb{L}^2(I, \mathbb{R}^3)$. In order to compare the shapes of two radial

curves, we can compute the distance between them in $\mathcal{C}$ under the chosen metric. This distance is defined to be the length of a geodesic connecting the two points in $\mathcal{C}$. Since $\mathcal{C}$ is a sphere, the geodesic length between any two points $q_1, q_2 \in \mathcal{C}$ is given by:

$$d_c(q_1, q_2) = cos^{-1}(\langle q_1, q_2 \rangle) \ , \tag{4.5.11}$$

and the geodesic path $\psi : [0, 1] \to \mathcal{C}$, is given by:

$$\psi(\tau) = \frac{1}{\sin(\theta)} \left(\sin((1-\tau)\theta)q_1 + \sin(\theta\tau)q_2\right) \ ,$$

where $\theta = d_c(q_1, q_2)$. Figure 4.8 illustrates the space $\mathcal{C}$ and geodesic path between two elements of that space. As illustrated in Figure 4.8, the space of all curves is a sphere in Hilbert space. Thus, the geodesic on the space of curves is the arc of the great circle connecting the two curves seen as elements of this sphere.

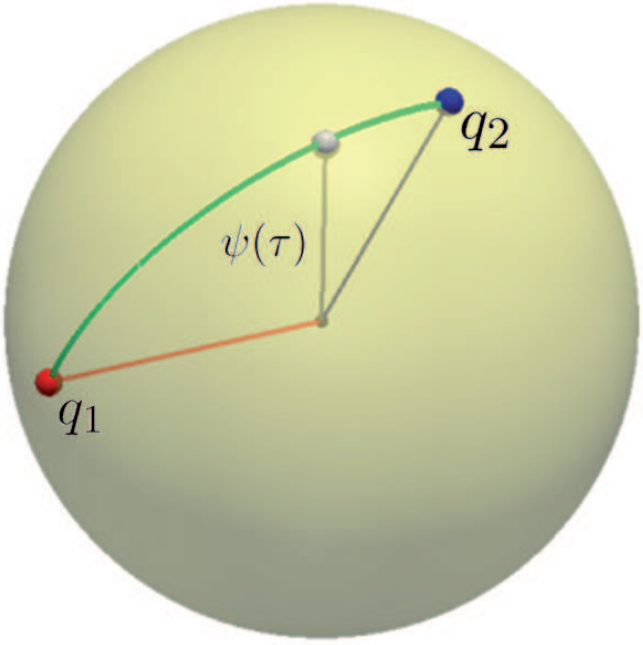


Figure 4.8: Illustration of shape space and geodesic between its elements.

As we did in last chapter, we define the equivalent class containing q as:

$$[q] = \{\sqrt{\dot{\gamma}(t)}Oq(\gamma(t))|O \in SO(3), \quad \gamma \in \Gamma\} \ ,$$

to be equivalent from the perspective of shape analysis. The set of such equivalence classes, denoted by $\mathcal{S} \doteq \mathcal{C}/(SO(3) \times \Gamma)$ is called the *shape space* of open curves in $\mathbb{R}^3$. $\mathcal{S}$ inherits a Riemannian metric from the larger space $\mathcal{C}$ due to the quotient structure [SKJJ11].

Thanks to SRV representation, the groups $\Gamma \times SO(3)$ act by isometries. This is a necessary condition to let the quotient space $\mathcal{S}$ inherit the metric from the pre-shape space $\mathcal{C}$.

To obtain geodesics and geodesic distances between elements of $\mathcal{S}$, one needs to solve the optimization problem:

$$(O^*, \gamma^*) = \arg .min_{(O,\gamma) \in SO(3) \times \Gamma} d_c(q_1, \sqrt{\dot{\gamma}}O(q_2 \circ \gamma)) \ .$$

For a fixed O in $SO(3)$, the optimization over Γ is done using Dynamic Programming. More description of this optimization method is given in the next section. Similarly, for a fixed $\gamma \in \Gamma$, the optimization over $SO(3)$ is performed using Singular Value Decomposition method.

By iterating between these two, we can reach a solution for the joint optimization problem. Let $q_2^*(t) = \sqrt{\dot{\gamma^*}(t)}O^*q_2(\gamma^*(t)))$ be the optimal element of $[q_2]$, associated with the optimal rotation O^* and re-parameterization γ^* of the second curve, then

$$d_s([q_1], [q_2]) \doteq d_c(q_1, q_2^*) \ , \tag{4.5.12}$$

and the shortest geodesic between $[q_1]$ and $[q_2]$ in $\mathcal{S}$ is given by:

$$\psi(\tau) = \frac{1}{\sin(\theta)} \left(\sin((1-\tau)\theta)q_1 + \sin(\theta\tau)q_2^*\right)$$

,

where θ is now $d_s([q_1], [q_2])$.

We recall the example of radial curves on open and closed mouth. The optimal matching for the two curves obtained using the dynamic programming is illustrated in Figure 4.9.b and it highlights the elastic nature of this framework. For the left curve, the mouth is open

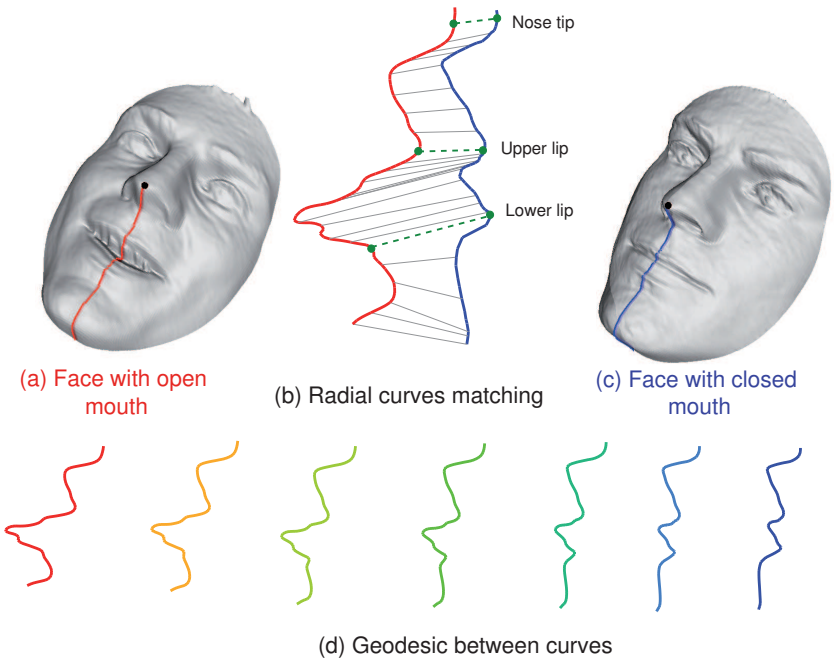


Figure 4.9: An example of matching and geodesic deforming radial curves extracted from two faces. a) A curve on an open mouth, (c) a curve on closed mouth ,(d) a geodesic path between curves (a) and (c), b) a result of matching using dynamic programming

and for the right curve, it is closed. Still the feature points (upper and bottom lips) match each other well. Figure 4.9.d shows the geodesic path between the two curves in the shape space $\mathcal{S}$ and this evolution looks very natural under the elastic matching.

4.5.5 Optimal Re-parametrization for curve matching

For the current rotation $O \in SO(3)$, let $\hat{q_2} = Oq_2$ and define a cost function $H : \Gamma \to \mathbb{R}_{>=0}$ by:

$$H(\gamma) = \int_0^1 \|q_1(t) - \sqrt{\dot{\gamma}(t)}\hat{q_2}(\gamma(t)\|^2 \, \mathrm{d}t \tag{4.5.13}$$

In order to find the optimal re-parametrization, we need to find a minimum of H in Γ. Several methods allow to do that like dynamic programming [BD62]. The cost function H is additive over the path $(t, \gamma(t))$.

4.5.5.1 Dynamic programming

Dynamic Programming algorithm (DP) [BD62], was first introduced in 1962 by Bellman and Dreyfus to solve matching problem. The key idea behind dynamic programming is quite simple. In general, to solve a given problem, we need to solve different parts of the problem (sub-problems), then combine the solutions of the sub-problems to reach an overall solution. Often, many of these sub-problems are really the same. The dynamic programming approach seeks to solve each sub-problem only once, thus saving a lot of computation. This is especially useful when the number of repeating sub-problems is exponentially large. This method is based on minimization of a certain type of cost function. This cost function has to be additive in time t, which the case in the cost function presented in equation 4.5.13. γ is seen as a graph from $(0, 0)$ to $(1, 1)$ in $\mathbb{R}^2$ such that the slope of this graph is always strictly between 0 and 90 degrees. To decompose the large problem into several sub-problems, define a partial cost function:

$$E(s, t; \gamma) = \int_s^t \|q_1(t) - \sqrt{\dot{\gamma}(\tau)}\hat{q_2}(\gamma(\tau)\|^2 \, \mathrm{d}\tau \tag{4.5.14}$$

so the original cost function defined in 4.5.13 is simply $E(0, 1; \gamma)$. The computer implementation details are provided in the next section.

4.5.5.2 Computer implementation

We remind that our goal is to find an optimal path from $(0,0)$ to $(1,1)$ in $\mathbb{R}^2$, corresponding to $(t,\gamma(t))$ that minimizes the cost function presented in equation 4.5.13. In order to use a numerical approach, the domain $[0,1]\times[0,1]$ is replaced with a finite grid and we restrict over search to that grid. The grid $G_n \times G_n$ is formed by uniform partition of G_n as $G_n = \{1/n, 2/n, ..., (n-1)/n, 1\}$. The search will be done over the set of all restrictions of γ to this grid. The total cost associated with the path is the sum of the costs associated with its linear segments. On an $n \times n$ grid there are only a finite number of paths, even less when we impose the slope constraint. Actually the path is never vertical or horizontal. However, this number of paths grows exponentially with n and we can not possibly search over all possible paths in an exhaustive fashion. Instead, the DP finds the optimal path in $O(n^2)$ time.

Denote a point on the grid $(i/n, j/n)$ by (i,j). Certain nodes are not allowed to go to (i,j) due to the slope constraint. Denote by N_{ij} the set of nodes that are allowed to go to (i,j). For instance $N(i,j) = \{(k,l)/0 \leq k < i; l < j \leq n\}$ is a valid set. Let $L(k,l;i,j)$ denote a straight line joining the nodes (k,l) and (i,j); for $(k,l) \in N_{ij}$ this is a line with slope strictly between 0 and 90 degrees. This sets up the iterative optimization problem:

$$(\hat{k},\hat{l}) = \arg.min_{(k,l)\in N_{ij}} L(k,l;i,j)\ . \tag{4.5.15}$$

4.5.6 Riemannian metric on facial surfaces

Now we extend the framework from radial curves to full facial surfaces. As mentioned earlier, we are going to represent a face surface S with an indexed collection of radial curves. That is, $S \leftrightarrow \{\beta_\alpha, \alpha \in [0,\alpha_0]\}$. Through this relation, each facial surface has been represented as an element of the set $\mathcal{S}^{[0,\alpha_0]}$. The indexing provides a correspondence between curves across faces. For example, a curve at an angle α on probe face is matched with the curve at the same angle on gallery face. Figure 4.10 illustrates an example of this correspondence. With this correspondence, we can compute pairwise geodesic paths and geodesic distances between the matched curves across faces. This computation has several interesting properties. Firstly, it provides a Riemannian distance between shapes of facial

surfaces by combining distances between the corresponding radial curves. The distance between two facial surfaces is given by:

$d : \mathcal{S}^{[0,\alpha_0]} \times \mathcal{S}^{[0,\alpha_0]} \to \mathbb{R}_{\geq 0}$, given by $d(S^1, S^2) = \frac{1}{\alpha_0} \sum_{\alpha=1}^{\alpha_0} d_s([q_\alpha^1], [q_\alpha^2])$. Here, q_α^i denotes the SRVF of the radial curve β_α^i on the i^{th}, $i = 1, 2$ facial surface and d_s is computed by the algorithm 1.

Algorithm 1: Geodesic distance computation algorithm between curves

Input: Two curves β^1 and β^2
Output: $d(\beta^1, \beta^2) = d_s([q^1], [q^2])$

1. For each $i = 1, 2$ compute the Square Root velocity Function $q^i(t) \doteq \frac{\dot{\beta}^i(t)}{\sqrt{\|\dot{\beta}^i(t)\|}}$
2. Compute the optimal rotation and re-parametrization O^* and γ^*:

$$(O^*, \gamma^*) = \arg .min_{(O,\gamma) \in SO(3) \times \Gamma} d_c(q^1, \sqrt{\dot{\gamma}} O(q^2 \circ \gamma)) \ .$$

3. $q_2^*(t) = \sqrt{\dot{\gamma}^*(t)} O^* q^2(\gamma^*(t)))$
4. Compute the geodesic distance between q^1 and q_2^*. $d_c(q^1, q_2^*) = cos^{-1}(\langle q^1, q_2^* \rangle)$.
5. $d_s([q^1], [q^2]) = d_c(q^1, q_2^*)$

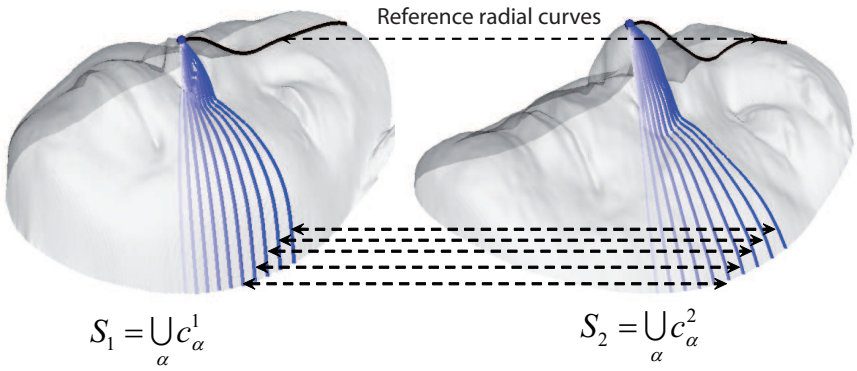


Figure 4.10: Faces comparison by pairwise curves comparisons.

Secondly, since we have geodesic paths denoting optimal deformations between individual curves, we can combine these deformations to obtain full deformations between faces. In fact, these full deformations are geodesic paths between faces when represented as elements

of $\mathcal{S}^{[0,\alpha_0]}$. Shown in Figure 4.11 are examples of such geodesic paths between source and target faces. The three top row illustrate paths between different subjects called inter-class geodesics, whereas the remaining rows illustrate paths between the same person under different expressions called intra-class geodesics.
In order to study the role of shape-preserving transformations, we compute the geodesics between surfaces in the **pre-shape space** $\mathcal{C}^{[0,\alpha_0]}$ and compare them with those in **the shape spaces** as illustrated in Figure 4.12. To illustrate the interest of our approach, we provide also a linear interpolation between the corresponding coordinates of two surfaces, after a rigid alignment using ICP, is also illustrated for comparison. The three rows show, respectively, a geodesic path in the shape space, the corresponding path in pre-shape space, and path which was linearly interpolated after ICP registration. Visually, it is clear that the geodesic in the shape space models exhibit the most natural deformation from the source to the target, especially in mouth region.

4.6 Robustness to pose and missing data

In situations involving non-frontal 3D scans, some curves are partially or totally hidden due to self occlusion during data collection. The use of these curves can severely affect the recognition performances. Therefore, these curves should be identified and discarded. The filter to detect such curves is applied after the extraction module and the aim of that filter is to discard curves that are of *bad* quality. To pass the quality filter, a curve should be one continuous piece and have a certain minimum of length. We retain the value 70 mm as minimum length, so a curve having less than 70 mm length or more than one piece will be detected then discarded by the quality filter. The dis-connectivity or a shortness of a curve results from the missing data on the face or presence of large noise. To illustrate this idea, we show an example of extracted curves on a face and then the result of quality filter on these curves in Figure 4.13. Once the good quality curves are retained, we can apply face comparison using available curves. We recall that during the pre-processing step, several holes are filled. The missing parts handled in this module are generally caused by pose variation. The range of allowable pose variations is discussed later in results sections.

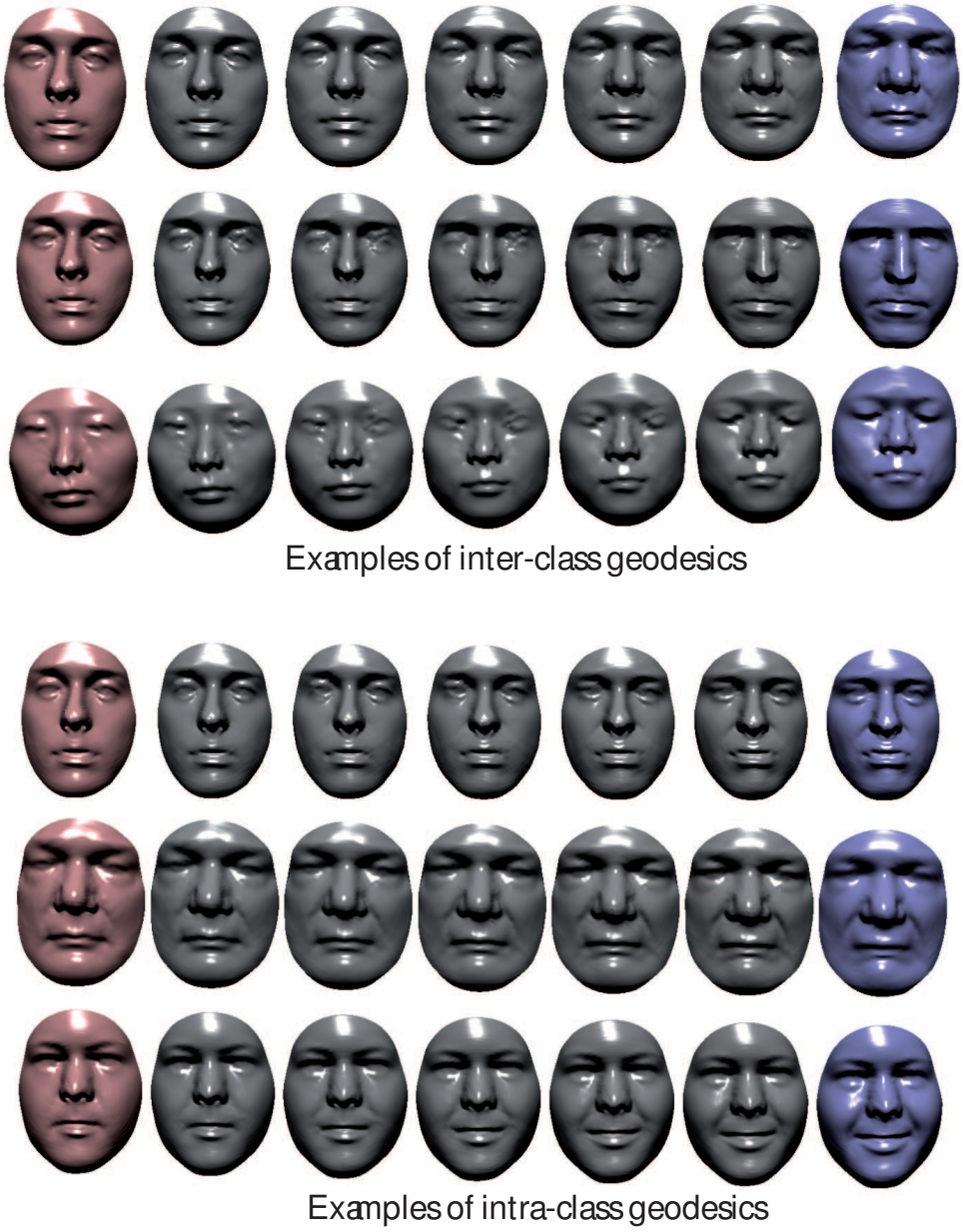


Figure 4.11: Examples of intra and inter-class geodesics in the shape space.



Figure 4.12: Examples of geodesics in shape space, pre-shape space and linearly interpolated path after ICP alignment.

4.7 Results and discussions

The proposed approach for representation and matching of faces has been evaluated using the GavabDB and FRGCv2 databases.

4.7.1 Evaluation on GavabDB dataset

GavabDB [MS04] is rich with expressive and noisy 3D face dataset currently available to the public. The GavabDB consists of Minolta Vi-700 laser range scans from 61 subjects, 45 male and 16 female, all Caucasian. Each subject was scanned nine times for different poses and facial expressions, six scans with neutral expression and three scans with a non-neutral expression. The neutral scans include two different frontal scans, one scan while looking up (+35 degree), one scan while looking down (-35 degree), one scan from the right side (+90 degree), and one from the left side (-90 degree). The non-neutral scans include one with a smile, one with a pronounced laugh, and an arbitrary expression freely chosen by the subject. Figure 4.14 shows some examples of scans from this dataset.

In order to assess the effectiveness of the proposed solution for face identification, we performed extensive experiments. In these experiments, one of the two frontal models with

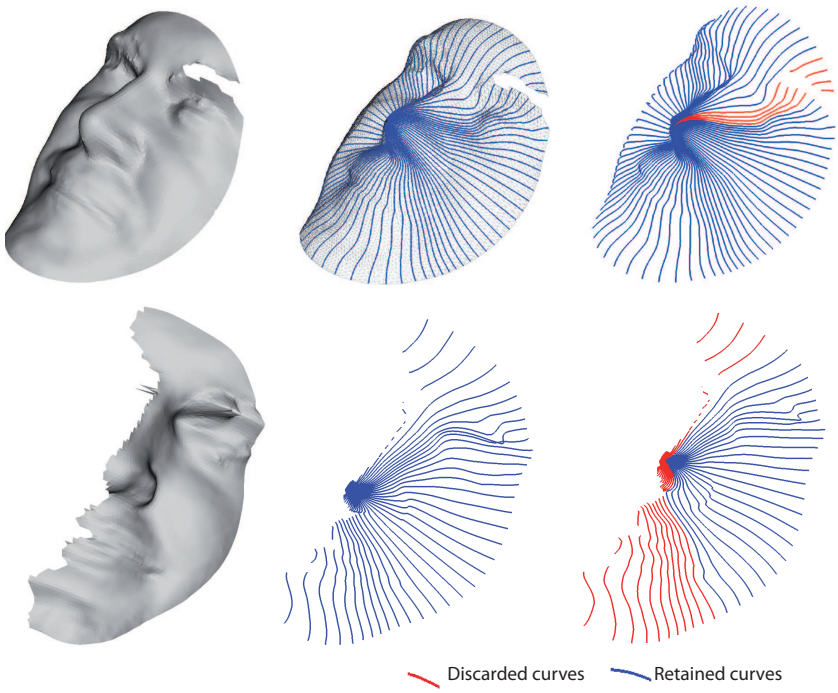


Figure 4.13: Quality filter: examples of detection of broken and short curves (in red) and good curves (in blue).

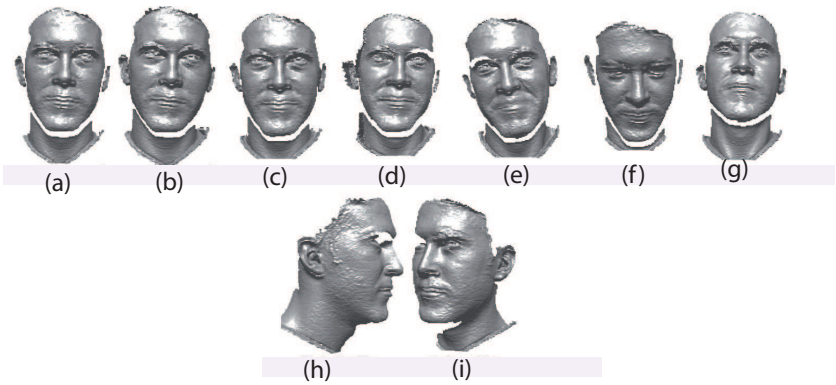


Figure 4.14: Examples of facial scans with different expressions in GavabDB. (a) and (b) neutral, (c) smile, (d) laugh, (e) random gesture, (f) looking down, (g) looking up, (h) right profile, (i) left profile.



Figure 4.15: Examples of correct and incorrect matches. For each pair, the probe (on the left) and the ranked-first faces from the gallery (on the right) are reported.

the neutral expression for each person is taken as a gallery model for the identification.

Table 4.1 summarizes the matching accuracy (rank-1 recognition rate) for different categories of probe faces. Each row in this table reports the accuracy of our method when 61 neutral frontal faces are used as gallery and 61 faces having the same expression and pose as probe. Different expressions and pose variation were already described and illustrated by Figure 4.14. Table 4.2 compares our results with previous approaches performed on GavabDB following the same protocol. We point out that our approach provides a high recognition rate for expressive faces (94.54%). This is due to both the face representation using radial curves and the elastic matching technique. Actually, each curve represents a

Table 4.1: Individual matching accuracy for different variations in the probe dataset.

Pose- expression of probe face	**Matching Accuracy**
frontal - neutral gesture	100%
frontal - smile gesture	98.36%
frontal - laugh gesture	95.08%
frontal - random gesture	90.16%
rotated looking down - neutral gesture	100%
rotated looking up - neutral gesture	98.36%
scans from right side - neutral gesture	70.49%
scans from left side - neutral gesture	86.89%

feature that characterizes a local region in the face. Besides, the elastic matching is able to establish a correspondence with guaranteed alignment among anatomical facial features. This was highlighted in Figure 4.9 that shows a satisfying matching lips despite a large deformation in one face due to the open mouth.

Figure 4.15 (top row) illustrates some examples of correct and incorrect matches of expressive and pose modified probe faces. In each case, the probe face is shown on the left, while the first ranked reference face (from the gallery) is shown on the right. These models also provide examples of the variability in terms of facial expression of face models included in the probe dataset.

Table 4.2: Recognition results comparison of different methods on GavabDB: (a) Neutral, (b) Expressive, (c) Neutral+expressive, (d) Rotated looking down, (e) Rotated looking up, (f) Overall, (g) Scans from right side, (h) Scans from left side.

	Li et al. [LJZ09]	Moreno et al. [MSVD05]	Mahoor et al. [MAM09]	Haar et al. [tHV10]	Mousavi et al. [MFA08]	Our method
a	96.67%	90.16%	-	-	-	100%
b	93.33%	77.9%	72%	-	-	94.54%
c	94.68%	-	78%	-	91%	95.9%
d	-	-	85.3%	-	-	100%
e	-	-	88.6%	-	-	98.36%
f	-	-	-	98%	81.67%	96.99%
g	-	-	-	-	-	70.49%
h	-	-	-	-	-	86.89%

As far as faces with the neutral expressions are concerned, the recognition accuracy sometimes depends on their pose. Actually, the accuracy is not affected when the pose is frontal or rotated looking down (the recognition rate is quite high). Figure 4.15 (top row) shows an example of a well matched face despite having a different pose (looking down). However, the accuracy remains very high (98.36%) for faces looking up.

For faces scanned from the right and the left sides, the accuracy is not as high because missing data becomes important. Figure 4.15 (lower row) illustrates some examples of failing matchings. For each pair, the probe (on the left) and the first closest face (on the left) are reported. As shown in table 4.2, our method outperforms majority of state-of-the-art approaches in terms of recognition rate. We note that no prior work included scans from left and right sides in their experimental protocol when using GavabDB dataset.

4.7.2 Evaluation on FRGCv2 dataset

We have also tested our method on the FRGCv2 dataset. In this dataset, the scans have been manually clustered into three categories: neutral expression, small expression, and large expression. We studied the following classification cases: neutral vs. neutral, neutral vs. non-neutral and neutral vs. all. The gallery consists of scans taken to be the first scan of each subject.Figure 4.16 shows the Cumulative Matching Curves of our method for the three cases.

Table 4.3: Recognition results comparison of different methods on FRGCv2 dataset (neutral vs. all scenario).

Approaches	Haar et al. [tHV10]	Berretti et al. [BBP10]	Queirolo et al. [QSBS10]	Faltemier et al. [FBF08]	Our approach
Rank-1	97%	94.1%	98.4%	97.2%	97.2%

Overall recognition results for our and state-of-the-art methods are presented in Table 4.3. We note that some rates were taken from Cumulative Matching Curve in the correspondent paper [BBP10]. We achieve a 97.2% rank-one recognition which is close to the highest published results on this dataset [QSBS10, FBF08]. The algorithm in [FBF08] is based on rigidly regions matching using ICP which needs to be applied more than 30 times.

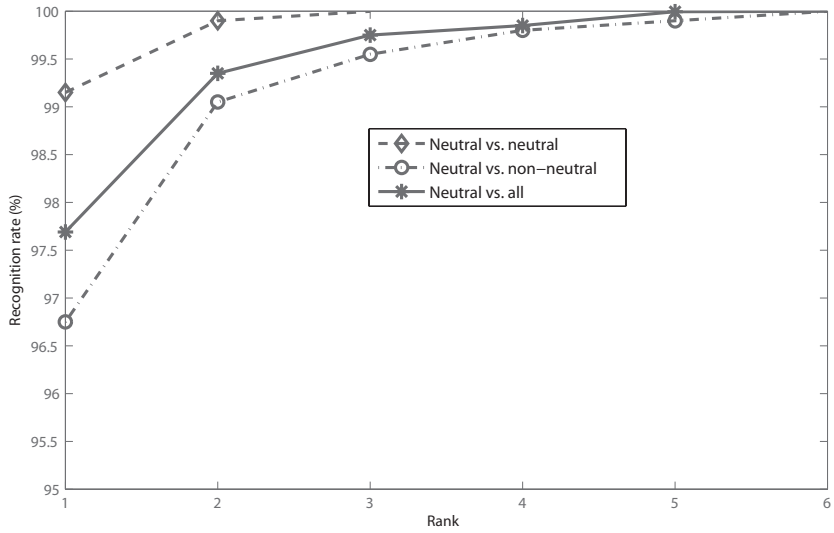


Figure 4.16: CMC curves of our approach for the following scenario: neutral vs. neutral, neutral vs. expressions and neutral vs. all.

The authors perform their algorithm on faces with missing data, but they simulate the missing data and have not reported any results on a database with missing data. As far as the computational cost is concerned, it is very high (they apply 38 times ICP) compared to our method which takes less than 8s, as described in Table 4.4. The authors in [QSBS10] use similar approach based on multi-region rigid matching. They use interpeneration measure to handle deformations due to expressions. They do not present any results on datasets with less quality level such as GAVABdb. The missing data problem due to pose variation is not handled; all models in FRGCv2 are frontal.
This table presents the computational cost of different steps of our algorithm on a desktop PC with a 3 GHz Core 2 Duo processor with 3 GB memory.

4.7.3 Robustness to quality degradation: experimental setting

In order to investigate the impact of quality degradation in 3D face recognition task, we design and follow series of experiments setting which aim to highlight robustness or weaknesses of recognition algorithms. First, two subsets from the well known FRGCv2

Table 4.4: Consuming time to match two facial scans using our approach.

Step	**Time consumed (s)**
Mesh Preprocessing	0.375 (per face)
Mesh Alignment	0.125
Curves extraction	0.485 (all curves, 2 faces)
Quality control	0.000001875 (per curve)
Curve preprocessing	0.137 (per curve)
Curve comparison	0.0318 (per curve)
Overall	7.7

database are extracted. The first subset contains 410 neutral earlier scans for each subject having a second session in the dataset. Whereas the second subset includes, for each subject an arbitrary session from the remaining sessions. This subset called probe subset has also 410 neutral or non-neutral sessions. These subsets allow first simulations of different quality degradations and then to compute similarity matrices between probe and gallery datasets. Below are details about the quality degradation subsets:

- ***One reference set (Gallery dataset)***, contains the earlier neutral sessions from FRGCv2 database. Gallery images are preprocessed with a segmentation by a 100 mm radius sphere centered on the nose tip, and a median filtering over a window width of 5 points in order to remove spikes.

- ***Three sets of noisy data***, this degradation is a simple simulation of the noise (or measurement errors) in the 3D acquisition process. We have considered that the errors affect mainly the distance coordinate (the range). The injected noise is a normal distribution of error added to the Z-coordinate of the 3D models, with standard deviation of respectively 0.2, 0.4, 0.8 mm as illustrated in figures 4.17.c and 4.20. These sets allow to discuss the influence of noisy data on recognition algorithm and to study the algorithm behaviour depending on the strength of this injected noise.

- ***Three sets with missing data***, the acquisition process frequently generates files with missing data, particularly due to shadows effects, and lack of reflectivity of some parts of the

face. It is not easy to simulate in a controlled way real defects. Instead, we have represented the voids in the scanned surfaces by cuts with a 10 mm radius sphere, randomly centred on the face surface. We only have avoided to affect the nose tip area, which is critical for algorithms. The sets correspond respectively to 1, 2 and 3 voids on the sample as illustrated in figures 4.17.b and 4.19. These sets allow to investigate the presence of missing data in probe samples.

- ***Three down-sampled sets***, the decimation is done by a reduction of the number of vertices from the original mesh by a factor of respectively 2, 4 and 8, and decimated mesh filtering as illustrated in figures 4.17.d and 4.18. These sets allow to study the effect of low resolution mesh in recognition algorithm. Figure 4.17 shows examples of some radial curves extraction under different quality degradation.

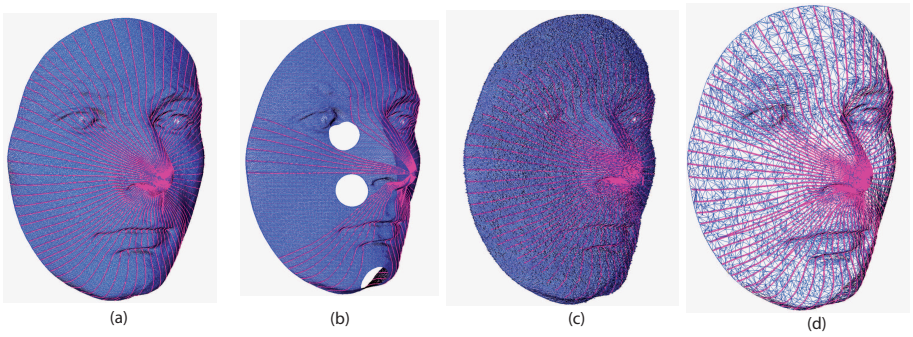


Figure 4.17: Examples of 3D faces with radial curves under different quality degradation.

After quality degradation simulations done on probe subsets, we conduct experiments in order to evaluate algorithm performances under all that quality degradation conditions. We are interested in both identification and authentication scenarios.

The results of identification and authentication scenarios will be presented separately.

- *Authentication Scenario* : The first experiment involves authentication where the performance is measured using a Detection Error Trade-off (DET) curve that shows the trade-off between the false reject and the false accept rates. As illustrated in Figure 4.18, Figure 4.19 and Figure4.20, the DET curves of tests including quality de-gradation are almost always above DET curve of tests on original data. This means that the performance

of the algorithm are in general decreased in the presence of quality degradation. However, the decrease of performance differs depending on the noise type and this induces more robustness of our approach to some kind of noises than others. In case of decimation for instance, accuracy are slightly affected. As illustrated in figure 4.18, when the number of points are reduced by a factor of 2, 4 and 8, the performance is very close to results on original data. The reduction by a factor of 8 decreases slightly the performance and the false reject rate at false accept rate = 0.01% increases from 25% to 30%. In the case of missing data, the performance are slightly decreased when only one hole is generated on each 3D facial surface as illustrated in figure 4.19. Increasing the number of holes decreases the performance and the false reject rate at false accept rate = 0.01% increases from 25% to 35%. As far as the noise is considered, as illustrated in figure 4.20, an injection of a normal distribution of error added to the z coordinate of the 3D files, with standard deviation of 0.2 mm does not affect seriously the performance. With standard deviation of 0.4 mm, the performance is decreased and the false reject rate at false accept rate = 0.01% increases from 28% to about 38%. However, an error with standard deviation 0.8 mm affects seriously the accuracy and the false reject rate at false accept rate = 0.01% increases from 38% to more than 75%.

- *Identification Scenario* : The second type of experiment is for identification for which the performance is quoted as a rank-one recognition rate. As far as decimation is concerned, the rank-one recognition rate remains almost unchanged when the number of points is reduced by a factor of 2 or 4 (90.9%) (It was 93.9% on the original data). Due to the reduction of points number by a factor of 8, the recognition rate becomes 87.3%. In the case of missing data, the rank-one recognition rate is more seriously affected and decreases from 93.9% for the original data to 83.4% (with one hole), 81.2% (with two holes) and 80% (with three holes). The accuracy is more seriously affected when a noise error is injected on the data. Actually, with only an error of 0.2 mm as standard deviation, the rank-one recognition rate is already 82.4%. When the error is more important (0.8 mm as standard deviation), the rank-one recognition rate is only 66.5%.

To summarize, our method is robust to decimation caused by reduction of the number

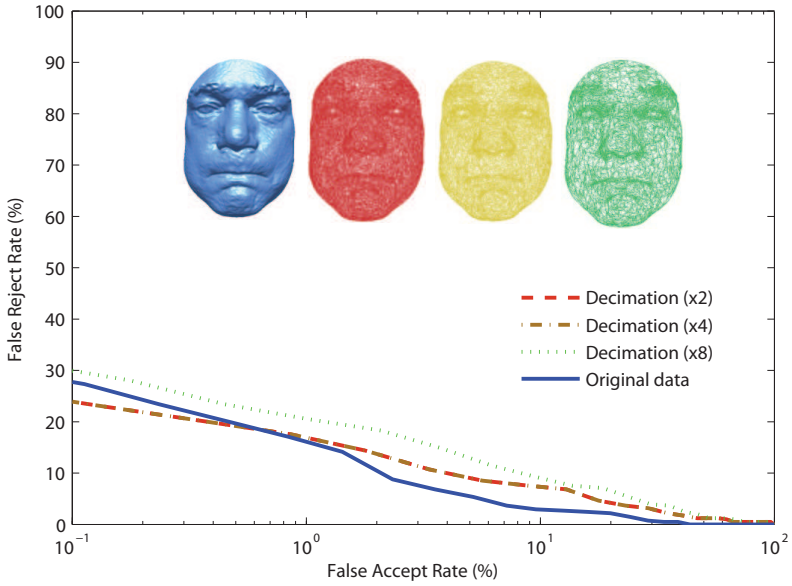


Figure 4.18: DET curves for experiments on resolution changes.

of points by a factor of 2, 4 and 8. It is also robust to missing data. However, the noisy data affects the performance of our algorithm. Actually, the algorithm is based on analyzing the shape of radial curve, which is strongly affected by the noise.

Since the radial curves rely critically on the location of the reference point (tip of the nose), and there is a possibility of numerical error in estimating this point, how stable is the framework presented here to such numerical errors? We have studied this issue experimentally and some results are presented in Figure 4.22 similarly to [SSJD09]. The figure illustrates the mean variations in shapes of radial curves versus changes in locations of reference points. X axis corresponds to the radius of the sphere where the tip of the nose lies and Y axis corresponds to the distance between the two facial surfaces. For example, 0 in X axis means that the nose tip is not moved. This point corresponds to 0 in Y axis which means that the distance stays 0. Hence the plot passes through $(0, 0)$ point. Figure 4.23 illustrates the distribution of distance calculated in quality experiment. As illustrated in this Figure, the intra-class distance distribution can be approximated to a normal distribution with mean equal to 0.155 with standard deviation of 0.3. Whereas, the inter-class distance

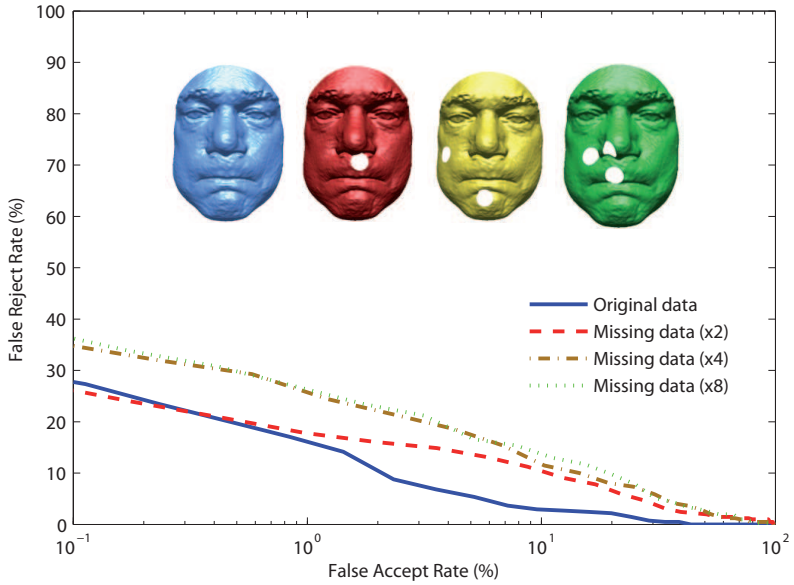


Figure 4.19: DET curves for experiments on missing data.

distribution can be approximated to a normal distribution with mean equal to 0.185 with standard deviation of 0.3. Based on this observation, a distance between two facial surfaces is considered intra-class distance if it is less than 0.14. We demonstrate through Figure 4.23 that when the nose tip lies in a sphere of radius 1cm, the distance is less than 0.1 and corresponds to intra-class distance. This implies a stability of shape analysis with respect to reference point location.

Conclusion

In this chapter, we have presented the results of our approach in 3D face recognition designed to handle facial variations, pose variations and missing data between gallery and probe images. The method has several benefits that make it suited for 3D face recognition. Firstly, to handle a pose variation and missing data, we have proposed a local representation by using a curve representation of 3D face and a quality filter of a curves. Secondly, to handle variations in facial expressions, we proposed an elastic shape analysis of 3D face.

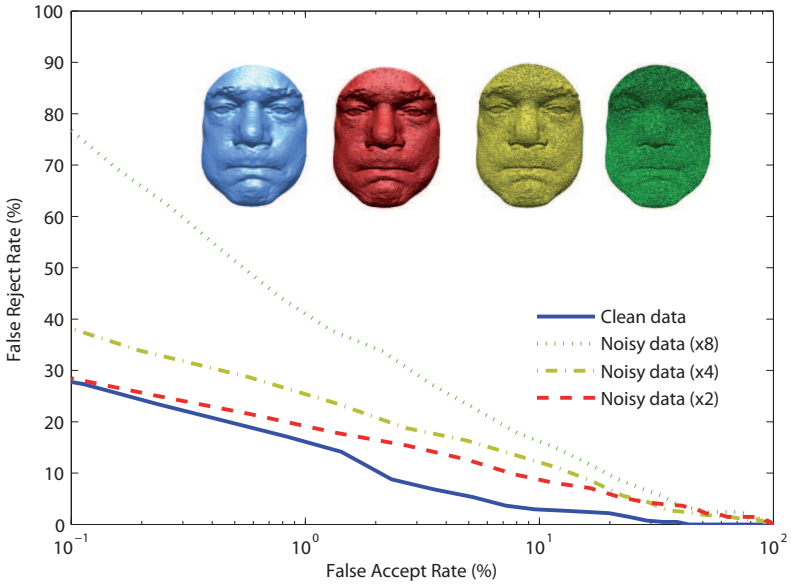


Figure 4.20: DET curves for experiments on noisy data.

Experiments that have been performed on the FRGCv2 and GavabDB databases have shown high effectiveness also in the presence of large facial expressions. However, the identification scenario costs a lot in terms of time computation. In our case, to retrieve a shape in a dataset of 410 facial surfaces, we need about 40 minutes. A hierarchical database organization can be performed in order to accelerate the identification process. In next chapter, we study an approach to cluster the gallery dataset using our framework. By the other hand, our solution to handle missing data needs a complete scan in the gallery in order to ensure minimum number of common curves for comparison. Some application need to match two scans with missing data. Our actual solution is no more sufficient, so we propose to restorate missing part of 3D face before comparison. This idea is proposed in the next chapter.

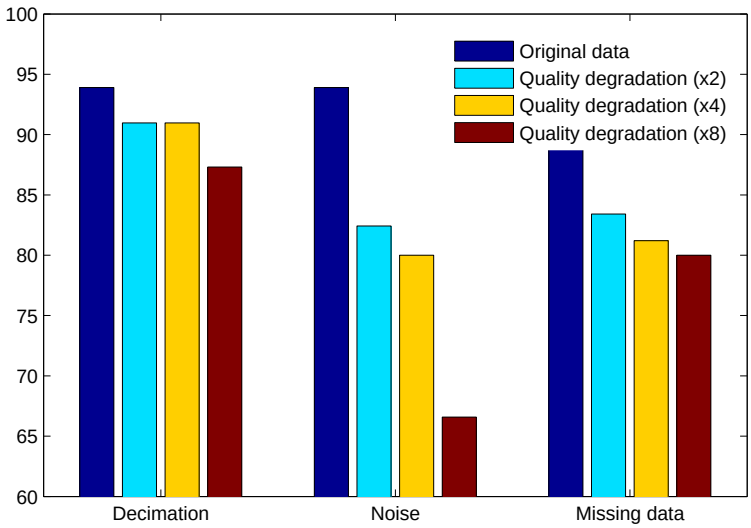


Figure 4.21: Recognition accuracy for different quality degradation experiments.

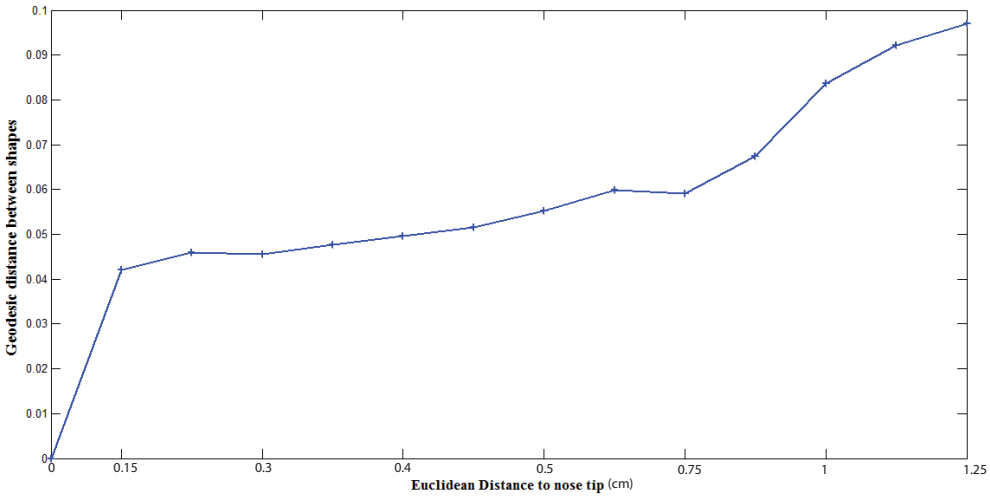


Figure 4.22: Variation in shapes of radial curves versus changes in locations of reference points.

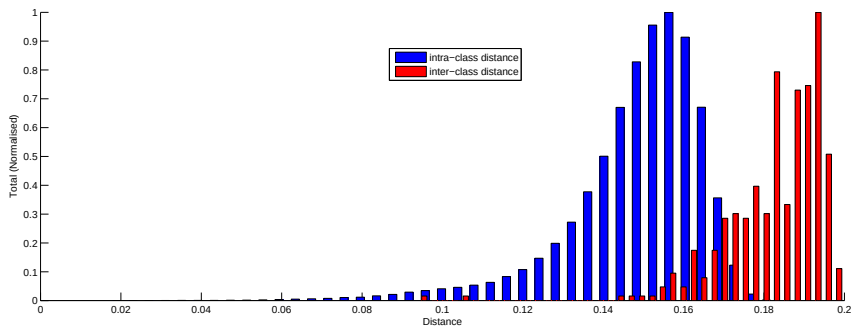


Figure 4.23: Distributions of intra-class and inter-class distances.

Chapter 5

Towards statistical computation on 3D facial shapes

5.1 Introduction

We presented in the previous two chapters tools for comparing facial and nasal surfaces. We demonstrated the effectiveness of the proposed approaches in both biometric scenario (identification and authentication) by performing our algorithms on several Benchmarks and comparing results with state of the art. Our framework provides the ability to discriminate between different shape of faces even in presence of facial expressions and partial data. In order to accelerate the shape retrieval, one need to build statistical shape priors. Moreover, 3D face with missing data can be recovered when one uses statistical-driven methods. We recall that missing data are due to occlusions of faces by other objects such as hair, hands, accessories, etc. This chapter outlines methods for intrinsic statistical analysis on 3D face shape space. We pursue the statistical approach with two main goals:

1. Hierarchical organization of shape galleries to allow efficient retrieval: We will consider the problem of clustering nasal or facial surfaces, according to the shapes of their curves. To improve efficiency, we will investigate a hierarchy in which the mean shapes are recursively clustered. This can significantly improve database searches in systems with shape-based queries. For instance, in identification scenario, testing a

query against prototypes of different clusters, to select a cluster, and then testing against shapes only in that cluster, is much efficient than the exhaustive testing.

2. Face recovery: 3D missing data restoration by learning shape models: Given 3d facial shape with missing data, we propose to automatically reconstruct the missing remaining data, yielding the most plausible reconstruction. Our approach is based on learning shape variability in order to recover queries with missing parts.

The rest of this chapter is organized as follows. Section 2 illustrates the intrinsic mean calculation of several facial surfaces. In section 3, we demonstrate the effectiveness of our approach in efficient search by hierarchical organization of face gallery database. We present in section 4 tools and algorithms for recovery of missing data in 3D face. Finally, section 5 provides some concluding remarks.

5.2 Mean shape computation

Statistics on shapes was widely studied in computer vision community with several applications such as clustering and learning. Previous researches on shape analysis [SJML05] proposed a hierarchical clustering of imaged objects according to the shapes of their boundaries. The nonlinear nature of shape space prevents the definitions of sample statistics, such as means and covariances in a conventional manner. Previous researchers on shape analysis suggest the use of Karcher mean [Kar77] to define average shapes by locally minimizing the average variance of a collection of shapes. We adopt that strategy for curve means computation and thus facial surface means computation.

5.2.1 Intrinsic mean shape

Let $\mathcal{C}$ denote *the pre-shape space* of open curves as described in previous chapter. We recall that $\mathcal{S} \doteq \mathcal{C}/(SO(3) \times \Gamma)$ is called the *shape space* of open curves in $\mathbb{R}^3$. That quotient space inherits a Riemannian metric from the larger space $\mathcal{C}$ due to the quotient structure [SKJJ11].

As described in last chapter, the distance between two facial surfaces is given by:

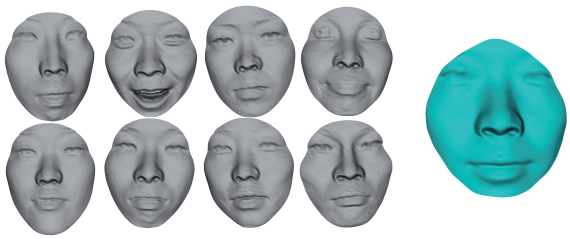


Figure 5.1: An example of Karcher mean of faces, the face at the right is the karcher mean of the eight faces in the left.

$d_s : \mathcal{S}^{[0,\alpha_0]} \times \mathcal{S}^{[0,\alpha_0]} \to \mathbb{R}_{\geq 0}$, given by $d_s(S^1, S^2) = \frac{1}{\alpha_0} \sum_{\alpha=1}^{\alpha_0} d_s([q_\alpha^1], [q_\alpha^2])$. Here, q_α^i denotes the SRVF of the radial curve β_α^i on the i^{th}, $i = 1, 2$ facial surface.

To calculate the Karcher mean of facial surfaces $\{S^1, ..., S^n\}$ in $\mathcal{S}^{[0,\alpha_0]}$, we define the variance function:

$$\mathcal{V} : \mathcal{S}^{[0,\alpha_0]} \to \mathbb{R}, \mathcal{V}(S) = \sum_{i=1}^{n} d_s(S^i, S^j)^2 \tag{5.2.1}$$

The Karcher mean is then defined by:

$$\overline{S} = \arg \min_{\mu \in \mathcal{S}^{[0,\alpha_0]}} \mathcal{V}(\mu) \tag{5.2.2}$$

The intrinsic mean may not be unique, i.e. there may be a set of points in $\mathcal{S}^{[0,\alpha_0]}$ for which the minimizer of $\mathcal{V}$ is obtained. To interpret geometrically, $\overline{S}$ is an element of $\mathcal{S}^{[0,\alpha_0]}$, that has the smallest total deformation from all given facial surfaces $\{S^1, ..., S^n\}$.

In order to compute facial mean surfaces, we use the Karcher mean algorithm. This approach, presented in Algorithm 1, uses the gradient of $\mathcal{V}$, in the space $T_\mu(\mathcal{S}^{[0,\alpha_0]})$, to iteratively update the current mean μ.

Algorithm 2: Karcher mean algorithm

Gradient search

Set k = 0. Choose some time increment $\epsilon \leq \frac{1}{n}$. Choose a point $\mu_0 \in \mathcal{S}^{[0,\alpha_0]}$ as an initial guess of the mean. (For example, one could just take $\mu_0 = S1$.) 1. For each $i = 1, ..., n$ choose the tangent vector $f_i \in T_{\mu_k}(\mathcal{S}^{[0,\alpha_0]})$ which is tangent to the geodesic from μ_k to N^i. The vector $g = \sum_{i=1}^{i=n} f_i$ is proportional to the gradient at μ_k of the function $\mathcal{V}$. 2. Flow for time ϵ along the geodesic which starts at μ_k and has velocity vector g. Call the point where you end up μ_{k+1}. 3. Set $k = k + 1$ and go to step 1.

Since this is a gradient approach, it only ensures a local minimizer of the variance function $\mathcal{V}$.

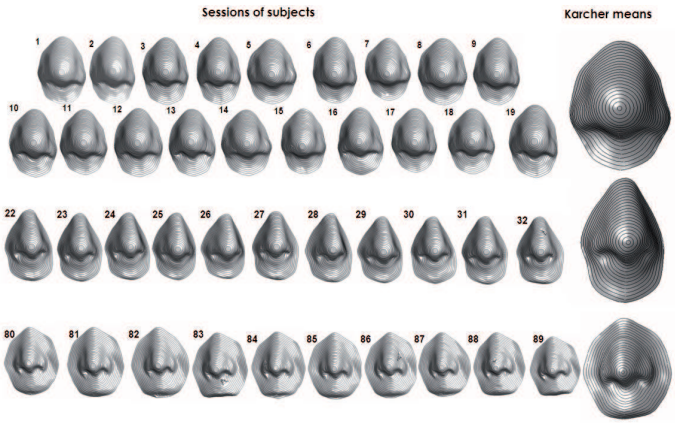


Figure 5.2: Examples of nasal shapes and their means.

An example of using the Karcher mean to compute average faces is shown in Figure 5.1. This Figure illustrates the mean face of faces belonging to different persons. Several examples of using the Karcher mean to compute average noses are shown in Figures 5.2.

5.3 Hierarchical organization shapes

One of the main goals for studying shapes of nose/face region is to conduct biometric search where query is often compared to a set of gallery shapes. This comparison can be made more efficient if we can organize the gallery elements in form of a hierarchical database, i.e. a tree, where the comparisons are performed only at the nodes. To construct such a shape tree we need to be able to cluster similar shapes, and that is the problem we study next.

5.3.1 Hierarchical shape database organization

Let $\mathcal{S}^{[0,\alpha_0]}$ denotes the set of facial shapes. Consider the problem of clustering n shapes (in $\mathcal{S}^{[0,\alpha_0]}$) into k clusters. A general approach is to form clusters in such a way that they minimize total "within-cluster" variance. Let a configuration C consists of clusters denoted

by $C_1, C_2, \ldots, C_k$, and let μ_is be the mean shapes in C_is and n_is be the sizes of C_is. There are several cost functions used for clustering, e.g the sum of traces of covariances within clusters. However, the computation of means μ_is of large shape clusters, and therefore their variances, is computationally expensive, especially when they are updated at every iteration. As a solution, one often uses a variation, called *pairwise clustering* [HB97], where the variance of a cluster is replaced by a scaled sum of distances (squared) between its elements:

$$Q(C) = \sum_{i=1}^{k} \frac{2}{n_i} \left(\sum_{S^a \in C_i} \sum_{b<a, \theta_b \in C_i} d(S^a, S^b)^2 \right). \tag{5.3.1}$$

We seek configurations that minimize Q, i.e., $C^* = \operatorname{argmin} Q(C)$. Notice that the metric used is the arithmetic mean d_a. We will minimize the clustering cost using a Markov chain search process on the configuration space. The basic idea is to start with a configuration of k clusters and to reduce Q by re-arranging shapes amongst the clusters. The re-arrangement is performed in a stochastic fashion using two kinds of moves. These moves are performed with probability proportional to the negative exponential of the Q-value of the resulting configuration. The two types of moves are following. (1) **Move a shape**: Here we select a shape randomly and re-assign it to another cluster. Let $Q_j^{(i)}$ be the clustering cost when a shape θ_j is re-assigned to the cluster C_i keeping all other clusters fixed. If θ_j is not a singleton, i.e. not the only element in its cluster, then the transfer of θ_j to cluster C_i is performed with probability: $P_M(j, i; T) = \frac{\exp(-Q_j^{(i)}/T)}{\sum_{i=1}^{k} \exp(-Q_j^{(i)}/T)}$ $i = 1, 2, \ldots, k$. Here T plays a role similar to temperature in simulated annealing. If θ_j is a singleton, then moving it is not allowed in order to fix the number of clusters at k. (2). **Swap two shapes**: Here we select two shapes randomly from two different clusters and swap them. Let $Q^{(1)}$ and $Q^{(2)}$ be the Q-values of the original configuration (before swapping) and the new configuration (after swapping), respectively. Then, swapping is performed with probability: $P_S(T) = \frac{\exp(-Q^{(2)}/T)}{\sum_{i=1}^{2} \exp(-Q^{(i)}/T)}$.

In order to seek global optimization, we have adopted a simulated annealing approach. Although simulated annealing and the random nature of the search help in avoiding local minima, the convergence to a global minimum is difficult to establish. Algorithm

3 illustrates the steps of the clustering algorithm.

Algorithm 3: Statistical shape clustering

For n shapes and k clusters, initialize by randomly distributing n shapes among k clusters. Set a high initial temperature T. 1. Compute pairwise geodesic distances between all n shapes. This requires $n(n-1)/2$ geodesic computations. 2. With equal probabilities pick one of the two moves:

- **Move a shape**: Pick a shape θ_j randomly. If it is not a singleton in its cluster, then compute $Q_j^{(i)}$ for all $i = 1, 2, \ldots, k$. Compute the probability $P_M(j, i; T)$ for all $i = 1, \ldots, k$ and re-assign θ_j to a cluster chosen according to the probability P_M.
- **Swap two shapes**: Select two clusters randomly, and select a shape from each. Compute the probability $P_S(T)$ and swap the two shapes according to that probability.

3. Update the temperature using $T = T/\beta$ and return to Step 2. We have used $\beta = 1.0001$.

It is important to note that once the pairwise distances are computed, they are not computed again in the iterations. Secondly, unlike k-mean clustering, the mean shapes are never calculated in this clustering. The algorithms for computing Karcher mean and clustering can be applied repeatedly for organizing a large database of human faces into a hierarchy that allows efficient searches during biometrics. Although there are 466 faces in the gallery in the FRGCv2 experimentation, we use only those 410 faces that have corresponding face in the probe.

Figure 5.3 shows a hierarchical organization of these shapes all the way up to the top. At the bottom level (level D in the Figure), these 410 are clustered in 29 clusters. Computing the means of each of these clusters, we obtain faces that are to be clustered at the next level (level C in the Figure 5.3). Repeating the clustering on the mean faces at level C, we obtain the next level (level B) containing 3 mean faces representing the clusters of the previous level. In level A, we just calculate the mean of the three mean faces in level B. This face represent the coarsest face along the tree.

If we follow a path from top to bottom of the tree, we can see the shapes getting more detailed structurally and leading up to individual faces, as illustrated in Figure 5.4.

Another example, we start with approximately 500 nose scans corresponding to 50

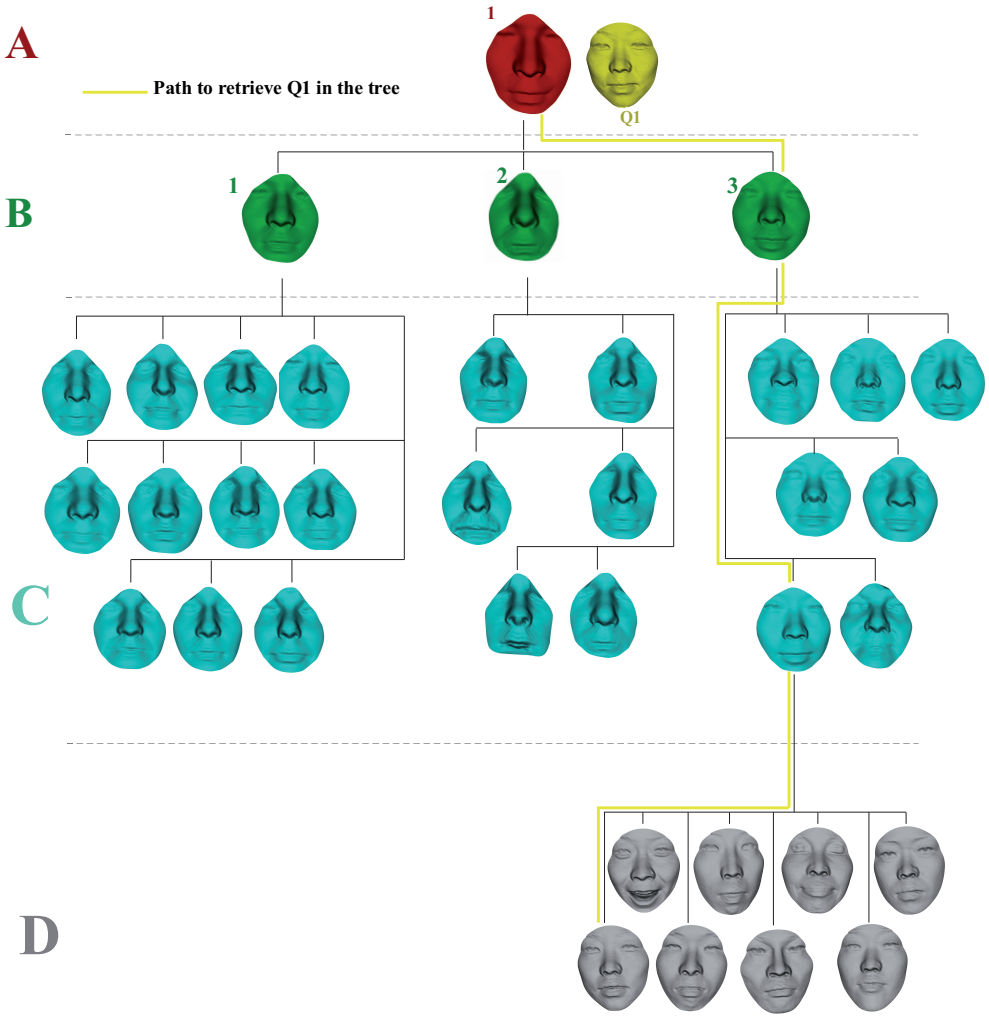


Figure 5.3: The result of hierarchical organization of gallery faces from FRGCv2 dataset and an example of path parsed by a query across the tree.

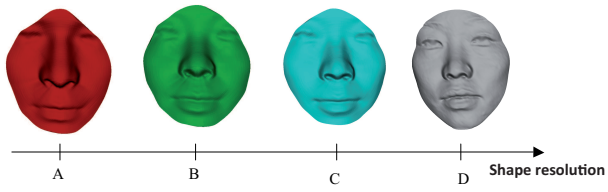


Figure 5.4: Path from top to bottom in the tree show increasing shape resolution.

distinct subjects. These noses form the bottom layer of the hierarchy, called level E in Figure 5.5. Then, we compute Karcher mean shapes for each person to obtain shapes at level D. These shapes are further clustered together and a Karcher mean is computed for each cluster. These mean shapes form the level C of the hierarchy. Repeating this idea a few times, we reach the top of the tree. If we follow a path from top to bottom of the tree, we can see the shapes getting more detailed structurally and leading up to individual faces, as illustrated in Figure 5.6.

5.3.2 Hierarchical shape retrieval

Using the trees constructed, one can conduct efficient biometric searches, especially in the identification scenario where one needs a comparison of the query face to the whole gallery dataset. In view of this structure, a natural way is to start at the top, compare the query with the faces at each level, and proceed down the branch that leads to the closest face. At any level of the tree, there is a number, say h, of possible shapes and our goal is to find the shape that best matches the query.

Figure 5.3 illustrates a retrieval example using the hierarchical organization of the dataset. For Q1 for example, the query facial shape (at the top) is first compared to the shapes of level B. As it is closer to shape B3, we proceed down the corresponding branch. There, the query proceeds down the branch of the shape C3 as this shape is the closest one in this level to the query. The decision of the retrieval is given after comparison with shapes at level D. In this case the query is matched to the shape D246. The last match decides to which cluster the query belongs. According to our tree, time computation for shape retrieval is approximately 10-20 times faster than exhaustive comparison for the 3541

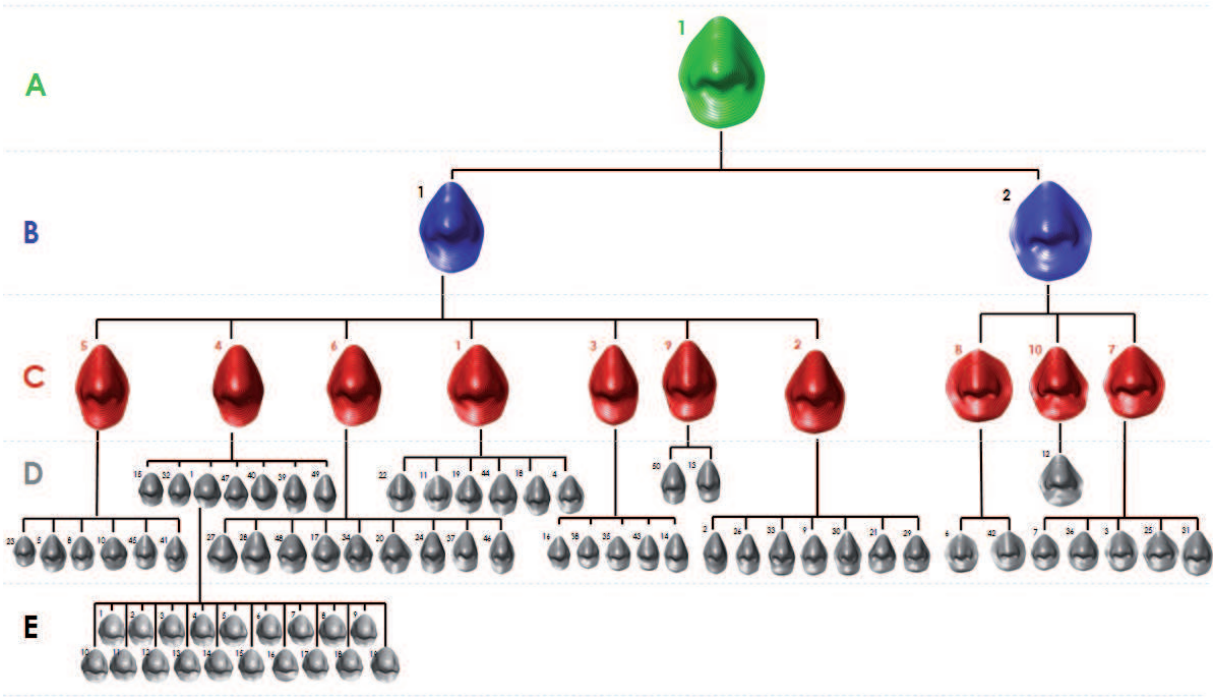


Figure 5.5: The tree resulting on hierarchical clustering.

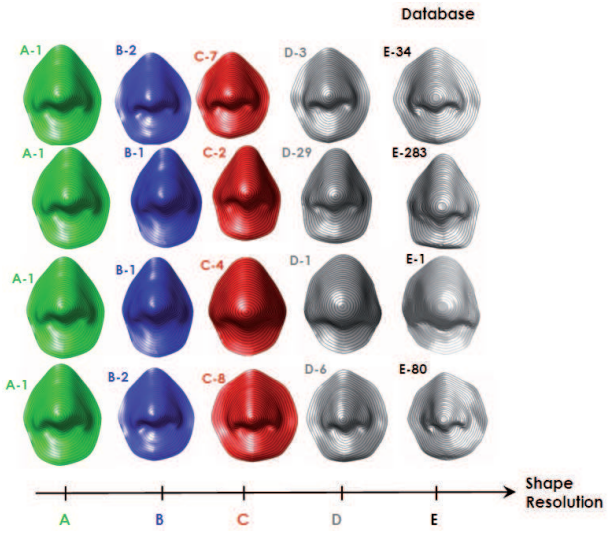


Figure 5.6: Paths from top to buttom in the tree show increasing shape resolutions

queries of FRGCv2 used in recognition scenario in previous chapter. Once the gallery is hierarchically organized, the face recognition becomes more efficient.

Such hierarchical organization greatly improves the retrieval times for large databases, although sometimes at a small loss in retrieval performance. This way one can conduct efficient biometric recognition, especially in the identification scenario, where one needs a comparison of the query face to the whole gallery. One starts at the top, compares the query with the faces at that level, and proceeds down the branch that leads to the closest face. Fig. 5.3 illustrates a retrieval example using the hierarchical organization of the dataset. The query facial shape Q1 (at the top) is first compared to the shapes of level B and is matched to B3. Then, at the next stage, the query best matched shape C23 and taking one more step, the retrieval decision is generated at level D. In this example, the query is matched to the shape D246. In our experiments, time taken for shape retrieval after hierarchical organization is approximately 10-20 times faster than exhaustive comparison for the 3541 queries of FRGCv2 used in recognition scenario above.

How is the classification performed in this scheme. One can classify a given query up to level B, C, and D. A query is said correctly classified at any level if it is in the same

cluster as the correct face class at that level. The full recognition is achieved at level D. The classification rate for 3541 queries at different levels are: 99.5% correct classification at level B, 92.4% at level C, and approximately 90% (354/3541 are incorrect) at level D. In contrast, the classification rate using an exhaustive search at the D level results in a performance of 97.7%. The average time for hierarchical retrieval is 3.5 mins while that for the exhaustive search is 52 mins. The same idea is also tested on nasal shapes. Figure 5.7 illustrates an example of retrieval using the hierarchical organization of the dataset. The query nasal shape (at the top) is first compared to the shapes of level B. As it is closer to shape B_1, we proceed down the correspondent branch. There, the query proceeds down the branch of the shape C_2 as this shape is the closest one in this level to the query. The decision of the retrieval is given after comparison with shapes at level D. In this case the query is matched to the shape D_2. The person is recognized as the last match is an intra class one. Actually, the shape D_2 is the mean shape of nasal shapes of the same person of the query. Notice that nasal shapes of each person at level E are represented by their mean at level D.

Figure 5.8 shows the query and shapes that match best at each level. The last match decides whether the person is recognized or not. In the first row, the query (in the left) is closest at level A to shape A_1. Then it is closest to C_1 at level C. Finally the closest shape to the query at level D will decide whether it is recognized or not.

Table 5.1: Time consuming and retrieval results

Query	Exhaustive	Hierarchical	decision
Q1	75 s	24 s	recognized
Q2	75 s	10.5 s	recognized

According to our tree, time consuming for shape retrieval is always smaller than 18 times time consuming between two shapes. Unlike the exhaustive retrieval, which needs 50 times time consuming between two shapes.

Statistical tools such as mean shape allow us to make comparison of a query to a set of gallery shapes more efficiently. This by organizing the gallery elements in form of a hierarchical database, i.e. a tree, where the comparisons are performed only at the nodes.

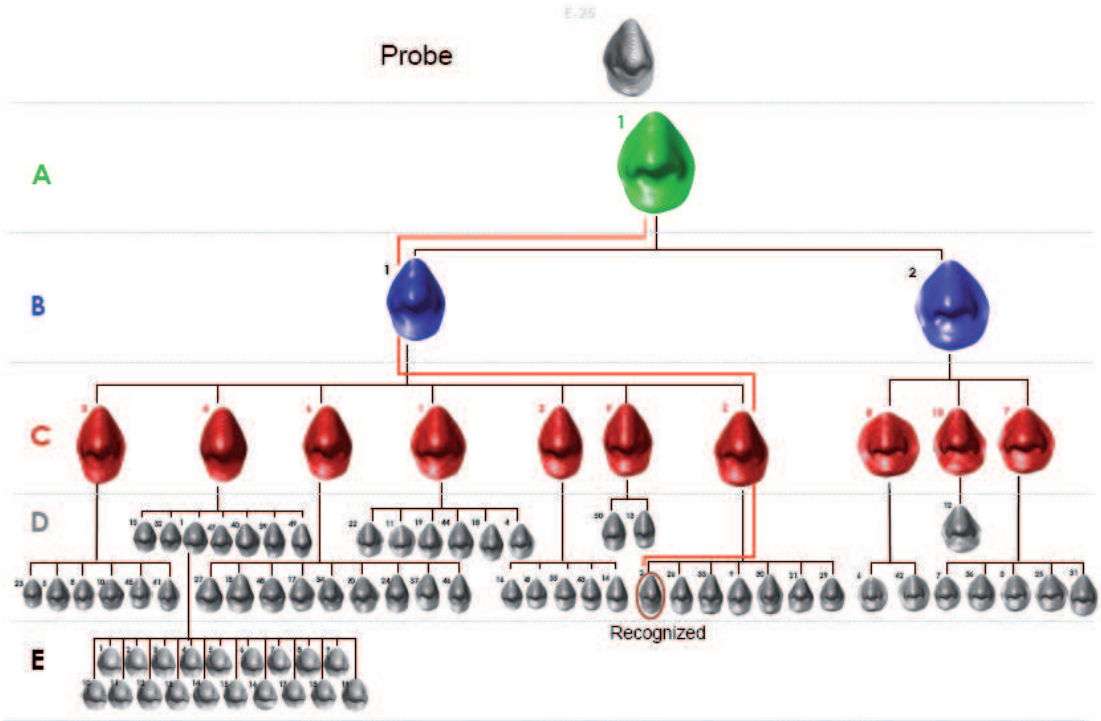


Figure 5.7: Tree parsing for efficient search of a query

The comparison is still performed as described in previous chapter. In case of missing data, we used to discard curves with missing data. Missing data can be recovered using statistical tools. Our contribution in this section consists on the restoration of the face. This restoration is local since we restore only invalid curves which are not complete. To perform this task, the main idea consists on constructing an eigencurves basis for each index of curve using PCA on the tangent space of the manifold $\mathcal{S}$. Then, we represent a restored curve as a linear combination of the best K eigencurves.

5.4 3D face restoration/ Recovery

The task of statistical learning on the space $\mathcal{S}$ is difficult. Actually, it is nonlinear space and therefore linear statistical approaches, associated with the vector space, can not be applied

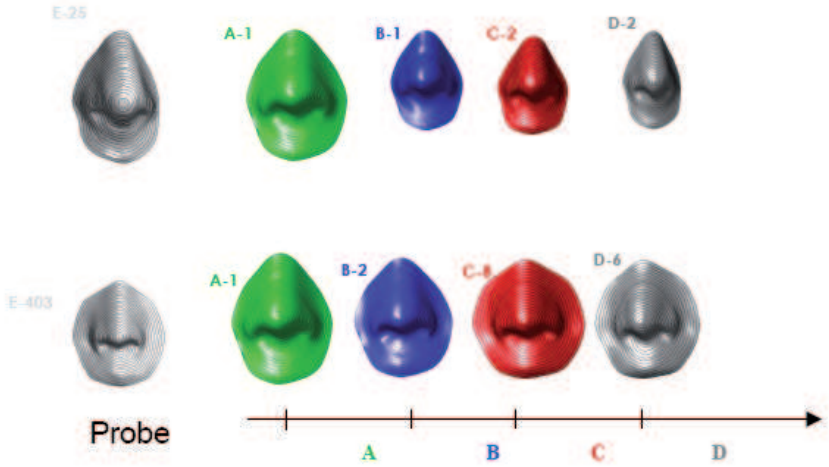


Figure 5.8: Retrieval of query shape (on the left) and its closest shape at each level of the tree

directly. Several authors used the range image representation that allows to apply linear tools such as PCA in order to recover faces with missing data. In [CCS09], the authors restored missing data in 3D face due to occlusion removal. They use GPCA (Gappy PCA) to reconstruct 3D damaged face, where only un-occluded pixels are used in analysis phase. The reconstruction is applied on the full face even original data are restored.

Our approach is to locally recover missed data in 3D face, we keep our face shape representation by indexed collection of radial curves and recover only curves involving missing data. That is, we project observed shapes (without missing data) onto the tangent spaces at sample means, and further reduce their dimensions using PCA, similarly to approaches presented in [DM98, SJML05, FLPJ]. Thus, we obtain a low-dimensional representations of shapes called TPCA. Using tangent principal components (TPCs) of observed shapes, we recover only missing parts in 3D face. Our solution involves making two approximations. First, we project elements of $\mathcal{S}$ on $T_\mu(\mathcal{S})$: the tangent space of their mean shape μ. The tangent space is a vector space and therefore, better suited for statistical modeling. This is performed using a specific map, called inverse exponential

map, this application maps elements on the shape space to the tangent space of a given shape. Second we perform dimension reduction in $T_\mu(\mathcal{S})$ using PCA. The resulting vectors are representative of initial shapes and will be used to restore missing data of query shapes. The first two steps are done off-line and are called Learning statistical models on the shape space.

5.4.1 Learning statistical models on the shape space

This step is done off-line. Training dataset contains several 3D faces, let N be the size of training set, without any missing data. 3D training faces should be first preprocessed, then dressed to upright position. Finally radial curves are extracted as described in previous chapter. The N radial curves of each index represent the input of learning model. Notice that the learning is done separately for each index.

5.4.1.1 Training data collection

Complete 3D faces in the gallery dataset are needed for the training. Each face is first preprocessed using the algorithm described in previous chapter, then dressed to the upright position. In practice we register all faces to a reference model in order to dress them to a frontal position by performing few iterations of the ICP algorithm. Collections of curves are constructed by extracting the curves of same index from all faces. Let K be the number of curves in each face and N the number of training 3D faces. Figures 5.9.a, 5.9.b and 5.9.c illustrate extracting three different collections i, j, and k. The collection C_i includes the N curves of index i from the N faces.

The training step will be done independently for each collection. Next, we will describe the training step for one collection of curves. As mentioned earlier, we project the shapes onto the tangent space of the mean shape. The mean shape is calculated as described earlier in this chapter.

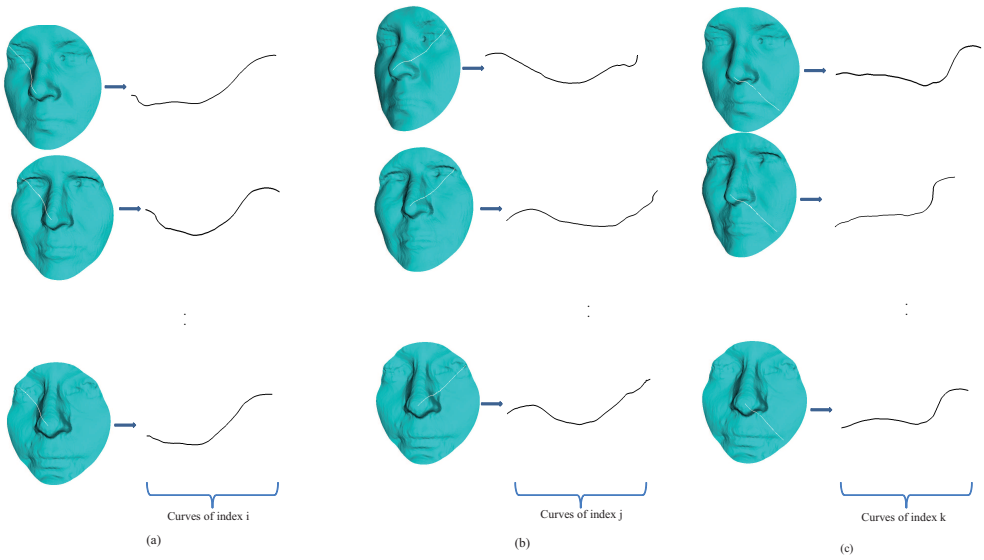


Figure 5.9: Radial curves collections for training step.

5.4.1.2 Projection onto tangent space on mean shape

First, we represent a curve by the SRV function as we did in previous chapter (q function). Therefore, a curve can be seen as element of infinite-dimensional non linear Riemannian manifold. Still, the tangent space $T_q(\mathcal{S})$ is a vector space. One can view $T_q(\mathcal{S})$ as a locally flat approximation of $\mathcal{S}$. It is local since this approximation is valid only in some neighborhood of q.

Hence, it is assumed that at least a set of complete curves are available (those acquired for the training step). Having $q_1, q_2, , ..., q_N$ elements of the manifold $\mathcal{S}$, we want to look for a subspace that best represents these elements. To construct the basis the idea is to work on the tangent space of the mean of the elements q_{mean}. That's why first of all, we project manifold's elements (each q_i) on the tangent space q_{mean}. Where q_{mean} is the SRV representation of the mean curve, which is the result of application of Karcher mean on the set $\{q_1, q_2, , ..., q_N\}$, calculated as described previously in this chapter. Next, let μ denotes the mean shape ($\mu = q_{mean}$).

In order to project shapes defined as elements of Riemannian manifold, we will use a

specific mathematical tool called exponential map, that enables to transfer points back and forth between the manifold $\mathcal{S}$ and the tangent space $T_\mu(\mathcal{S})$. We first need to recall the next theorem.

Theorem 5.4.1 *Given a point $q \in \mathcal{S}$ and a tangent vector $v \in T_q(\mathcal{S})$, there exists a unique constant-speed parametrized geodesic $\alpha_v : [-\epsilon, \epsilon] \to \mathcal{S}, \epsilon > 0$, such that $\alpha_v(0) = p$ and $\dot{\alpha}_v(0) = v$.*

Definition 5.4.1 *If $\mathcal{S}$ is a Riemannian manifold and $q \in \mathcal{S}$, the exponential map $exp_q : U \subset T_q(\mathcal{S}) \to \mathcal{S}$, is defined by $exp_q(v) = \alpha_v(1)$ where α_v is as defined above in the theorem.*

In our case, the exponential mapping *exp* maps a vector $v \in T_\mu(\mathcal{S})$ to a point of $\mathcal{S}$. In other words, to reach the point $exp_\mu(v)$, one starts at μ, and then moves for time 1 along the unique constant speed geodesic whose velocity vector at q is v. The inverse of an exponential map takes a point q_i on the manifold $\mathcal{S}$ and maps it to an element (or multiple elements) of the tangent space $T_\mu(\mathcal{S})$. A vector v is said to be the inverse exponential map of a point $q_i \in \mathcal{S}$, at the point $\mu \in \mathcal{S}$, if $exp_\mu(v) = q_i$. It is denoted by $v = exp_\mu^{-1}(q_i)$.

We notice that the manifold represents the unit sphere of Hilbert space as described in previous chapter.

As a result, the exponential map, $exp : T_\mu(\mathcal{S}) \to \mathcal{S}$, has also a simple expression. Let v be a vector $v \in T\mu(\mathcal{S})$, the exponential mapping of v gives an element of the manifold $\mathcal{S}$ as:

$$exp_\mu(v) = \cos(\|v\|)\mu + \sin(\|v\|)\frac{v}{\|v\|} \ ,$$

The exponential map is a bijection if we restrict the Euclidean norm $\|v\|$ so that $\|v\| \in [0, \pi)$. For a point $q_i \in \mathcal{S}$, such that $(q_i \neq \mu)$, the inverse exponential map $v = exp_\mu^{-1}(q_i)$ projects q_i on the tangent space of μ as:

$$v = \frac{\theta}{\sin(\theta)}(q_i - \cos(\theta)\mu) \ ,$$

where $\theta = cos^{-1}(< \mu, q_i >)$, we recall that $<,>$ refers to the inner product of Hilbert space $\mathbb{L}^2(I, \mathbb{R}^3)$. The result of this projection is elements on the tangent space $\{v_1, v_2, , ..., v_N\}$. It is possible to do traditional operations and work like in an Euclidean space using the

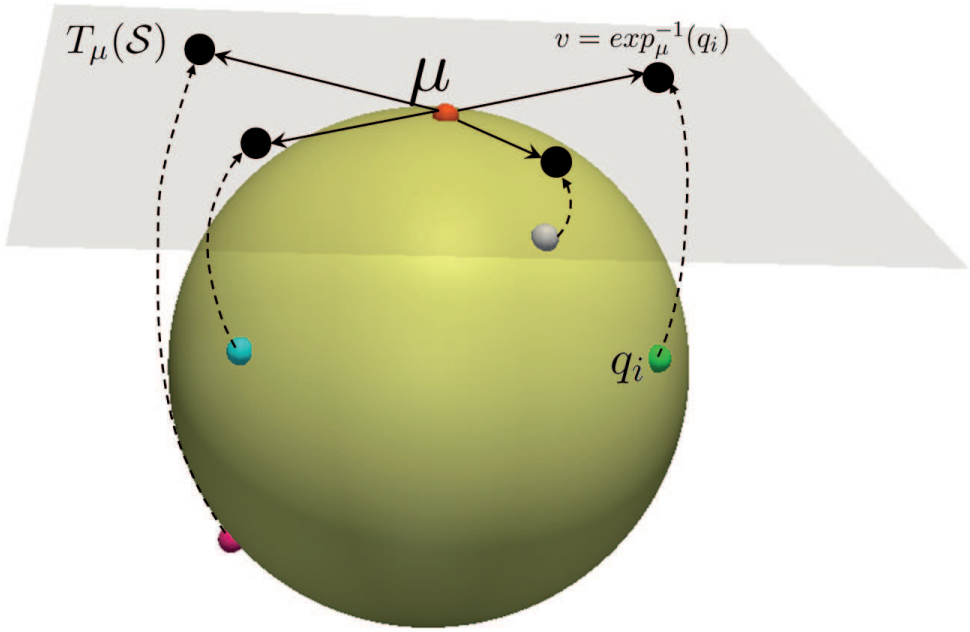


Figure 5.10: Illustration of mapping shapes onto the tangent space of μ, $T_\mu(\mathcal{S})$. Dashed arcs illustrate the projection of elements of the manifold to the tangent space.

projected elements. Figure 5.10 illustrates a mapping of elements of the shape space onto the tangent space of the manifold on the mean shape μ. PCA will be performed on these elements.

5.4.1.3 Tangent PCA

Next we perform PCA in $\mathbb{R}^n$ using the projected observations $\{v_1, v_2, , ..., v_N\}$. The covariance matrix $C \in \mathbb{R}^{n \times n}$ is first calculated. The resulting eigenvectors form a basis in the tangent space of the manifold on the mean shape. We retain 90% of the informative data and represent the resulting basis as $B = \{w_i^1, w_i^2, ..., w_i^m\}$ ($m \leq N$ and $m \leq n$) .

This basis represents a new space with smaller dimension for curves. Roughly speaking, a curve can be projected in that new space and has new coordinates. The shape of the curve is 90% conserved. The main idea of that basis is to have a new basis of projection.

We resume the off-line step in algorithm 4. In the next step, curves with missing data can be projected on that basis and hence recovered. We recall SRVF the function that enables to calculate the Square Root Velocity Function as described in previous chapter.

Algorithm 4: Eigenvectors computation.

Input: Training faces (without missing data) $G = \{y_i\}_{1 \leq i \leq N_G}$

Output: $B = \{v_{k_j}\}$: Eigenvectors

K : number of curves in each face;

for $k \leftarrow 1$ **to** K **do**

  μ_k = intrinsic mean of $SRVF(y_{k_i})$ (Karcher mean).

  for $i \leftarrow 1$ **to** N_G **do**

    $\beta_{k_i} = exp_{\mu_k}^{-1}(SRVF(y_{k_i}))$

  $S_k = \sum_{i=1}^{N_G} \beta_i \beta_i^T$

  $\{v_{k_j}\}$ = eigenvectors of S_k

Notice that the number of eigenvectors differs from one index to another. Hence we can not define any eigenface. Roughly speaking, we call the i^{th} eigenface the collection of the the i^{ths} eigencurves for each level k. Some i^{th} eigencurves are not able to provide a complete a face.

5.4.2 Curves restoration

In order to validate the new basis, we propose first to project curves without missing data and compare the resulting curve with the original one. Repeating that for a curve collection of a face, we obtain a projection of a 3D face. In the second experiment, we propose to handle missing data. That is, only curves with missing data will be projected and recovered.

5.4.2.1 Projection of full faces

This experiment illustrates a projection of curves of full faces on corresponding basis. The face is represented by an indexed collection of radial curves. For each index k, the training step is done as described above. The curve of each index k is next projected onto the correspondent basis. That is, the curve is first projected as v_k onto the tangent space

of the mean curve calculated in training step, then projected on the correspondent basis $B_{k_j}, k = 1..K$. The projection of v_k is a linear combination of vectors of $B_{k_j}, k = 1..K$. The coefficients are calculated using the Euclidean inner product $\langle c_k, v_{k_j} \rangle$. Algorithm 5 resumes the projection step.

Algorithm 5: Projection facial curves on eigen-curves.

Input: P: face without missing data, Eigenvectors $B_{k_j}, k = 1..K$

Output: $P\_restored$

K : number of curves in each face;

for $k \leftarrow 1$ **to** K **do**

$q_k = SRVF(c_k)$;

$v_k = exp^{-1}_{\mu_k}(q_k)$;

$q_k = exp_{\mu_k}(\sum_{j=1}^{N_{B_k}} \langle c_k, v_{k_j} \rangle v_{k_j})$;

$c_k = SRVF(q_k)$;

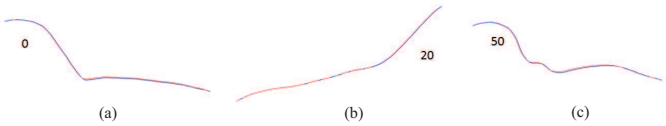


Figure 5.11: Restoration of curves of different index.

Figure 5.11 illustrates a restoration of three different curves. As illustrated respectively in Figure 5.11.a, Figure 5.11.b and Figure 5.11.c, the original curve and its restoration are very similar, almost the same. This validates our model which represents curves in a new basis while keeping 90% of information. The same idea is followed to recover 3D faces: curves of different index are restored. The collection of the restored curves represents the 3D face. Notice that the number of eigenvectors differs from one level to another. Figure 5.12 illustrates the face projection. The first row illustrates the original face (at left) and the restored one (at the right). In the second row, we see at the middle the original and restored face together. The left and right image illustrates respectively the signed and absolute deviation between the original and the restored face. The average absolute deviation is $0.21mm$ with $0.28mm$ as standard deviation. These observations demonstrate

on the closeness of restored face to the original one. Therefore, we propose next to restore curves with missing data and keep complete ones.

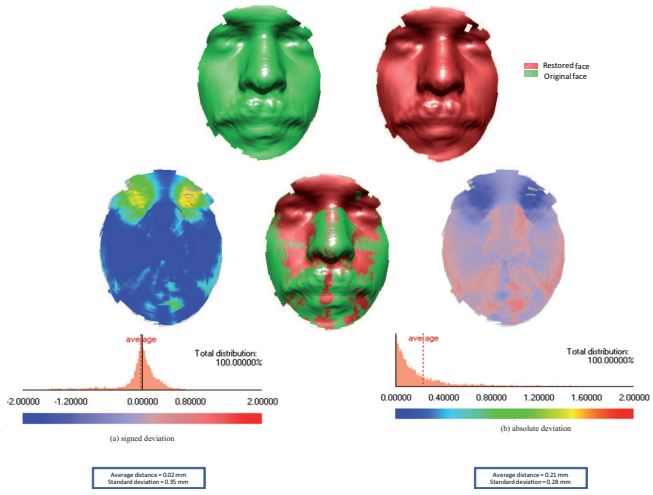


Figure 5.12: Restoration of full face. The absolute and signed deviations

5.4.2.2 Projection of faces with missing data

The restoration of 3D face with missing data is the goal of the actual section. We presented a learning step where we build a projection basis. The projection is tested for shapes (curves and then faces) without missing data. The resulting shape is very close to original data. In case of missing data, we pursue similar strategy. Several changes on the algorithms are needed in case of missing data however.

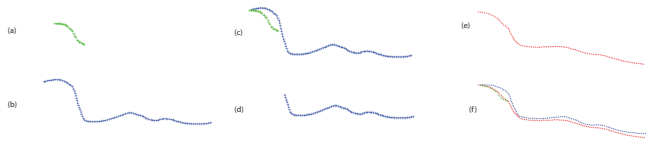


Figure 5.13: Restoration of curve with missing data. The probe curve is illustrated in (a) and the mean curve for correspondent index in (b). Together are illustrated in (c). In (d), the part of mean curve corresponding to missing data in probe curve. The result of restoration is illustrated in (e) and all curves are represented in (f).

Given a curve with missing data, the first step is to detect the missing part, comparing to the mean curve in the same index. Figure 5.13.a illustrates the curve of the index i in the mean face. The corresponding curve in the probe face is illustrated in Figure 5.13.b, this curve represents a missing part as illustrated in Figure 5.13.c. This missing part has to be recovered after projecting the curve on the basis generated by PCA. For that issue, we project the missing-part curve onto the tangent space of the mean shape, then the resulting vector is projected on the PCA-generated basis. Notice that during these steps, the calculus is restricted on the common parts. The part of the mean curve corresponding to missing data in the probe one is illustrated in Figure 5.13.d. The application of inverse exponential map is done on the common part of probe curve and the mean one. More formally, the inverse exponential map is a function of the inner product of q_i and μ. This inner product is restricted to common part of the two curves. This restriction is also preserved in the projection on PCA-basis. The restoration is done only for curves with missing data, the quality module described in previous chapter inspects whether a curve presents missing data. Therefore the restoration is local and majority of data is usually completely kept. We resume the restoration of a 3D face with missing part in algorithm 6. Figure 5.17 illustrates a restoration of a 3D face with missing data. Original face and its reconstruction from curves are respectively illustrated by (a) and (d). The reconstructed face is illustrated in (e). The face with missing parts is illustrated in (b) and (c) illustrates radial curves. The restoration preserves the shape of a 3D face and tries to better suit the original data.

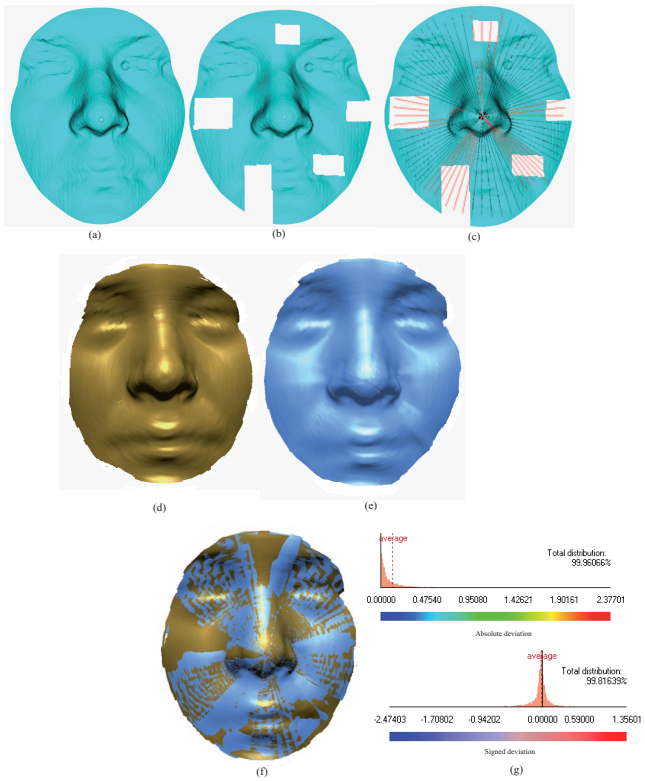


Figure 5.14: Restoration of 3D face with missing data. Original face and its reconstruction from curves are respectively illustrated by (a) and (d). The reconstructed face and is illustrated in (e)The face with missing parts is illustrated in (b) and (c) illustrates radial curves.

Algorithm 6: Restoration.

Input: P: face with missing data, Eigenvectors $B_{k_j}, k = 1..K$

Output: $P\_restored$

K : number of curves in each face; n: number of points in a complete curve. **for** $k \leftarrow 1$ **to** K **do**

  if *quality*$(c_k = 0)$ **then**

    for $i \leftarrow 1$ **to** n **do**

      $flag_k[i] = 1$;

      if *i corresponds to missing part* **then**

        $flag[i] = 0$;

    $q_k = SRVF(c_k)$;

    $v_k = exp^{-1}_{\mu_k}(q_k)$; for $flag[i] = 1$

    $q_k = exp_{\mu_k}(\sum_{j=1}^{N_{B_k}} \langle c_k, v_{k_j} \rangle v_{k_j})$; for $flag[i] = 1$

In the previous chapter, we illustrated a filtering step that finds and removes curves with missing parts. Although this step is effective in handling smaller missing parts, it may not be sufficient when larger parts of a face are missing due to external occlusions, such as glasses and hair. This problem is more challenging than self-occlusion because in addition to the missing face part, we also have parts of the occluding object(s) in the scan. In a non cooperative situation where the acquisition is unconstrained there is a high probability of this kind of occlusion.

We propose to apply the restoration step in order to improve face recognition in presence of occlusions: we will handle this problem in two steps. First, for a given scan, we detect what points belong to the face and what points belong to the occluding object, and remove the occluding object. Second, using the face part and the training data, we use the described statistical model in the shape space of radial curves estimate the parts of face that have been occluded, in order to recover the full face.

5.4.3 3D face recognition under occlusion

We will handle this problem in two steps. First, for a given scan, we detect what points belong to the face and what points belong to the occluding object, and remove the occluding object. Second, using the face part and the training data, we use a statistical model in the shape space of radial curves estimate the parts of face that have been occluded, in order to recover the full face as described in previous section.

5.4.3.1 Overview of proposed approach for recognition under occlusion

As illustrated in Figure 5.15, once the face is preprocessed, an occlusion removal module is applied. This module detects the presence of occlusion and discard additional parts. Then we extract the radial curves emanating from the tip of the nose. The quality module, similarly to previous chapter, inspects the quality of curves and detects short curves (with missing data). These curves are restored as described in previous section. The restored curves together with ones that passed the quality module, constitute the face for comparison.

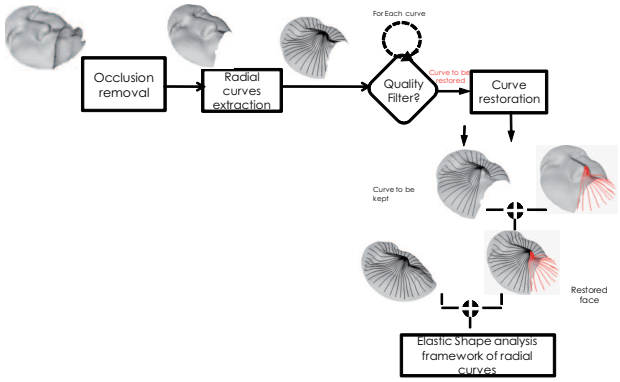


Figure 5.15: Different steps of recognition under occlusion.

5.4.3.2 Detection and Removal of Extraneous Parts

The first problem we face in externally occluded faces is that detection of the external object parts. We accomplish this by comparing the given scan with a template scan, where a template scan is developed using an average of training scans that are complete, frontal and have neutral expressions. The basic matching procedure between a template and a given scan is recursive ICP, implemented as follows. In each iteration, match the current face scan with the template using ICP and remove those points on the scan that are more than a certain threshold from their corresponding points on the template. This threshold has been determined using experimentation and is fixed for all faces. In each iteration, additional points that are considered extraneous are removed and the alignment (with the template) based on the remaining points is further improved. This, in turn, helps remove more extraneous points in the next iteration. Fig. 5.16 shows an example of this implementation. From left to right, each face shows an increasing alignment of the test face with the template, with the aligned parts shown in magenta, and also an increasing set of points labeled as extraneous, drawn in pink. The final result, the original scan minus the extraneous parts, is shown in green at the end.

5.4.3.3 Completion of Partially-Obscured Curves

The next problem we face is that the recovery of points on face that have been occluded by the extraneous object. The core of this problem, in our representation of facial surfaces by

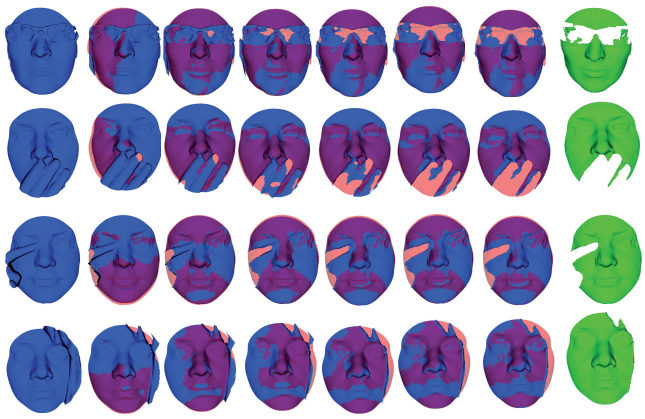


Figure 5.16: Gradual removal of occluding parts in a face scan.

curves, is to take a partial facial curve and predict its completion. The sources of information available for this prediction are: (1) the partially observed curve and (2) several training curves at the same angle that are extracted from full faces. The restoration step is applied on curves with missing data as described in previous chapter.

We present three examples of this procedure in Fig. 5.17, starting with three faces with external occlusion in (a). The detection and discarding of occluded parts is performed as described in the previous section, and the result of that step is shown in (b). Finally, the curves passing through the missing parts are restored and shown in (c).

5.4.3.4 Experimental results

In case of faces with external occlusion, we first restore them and then apply the recognition procedure. That is, the occluded part is detected and removed, and then recovered, resulting in a full face that can be compared with a gallery face using $d_{\mathcal{S}}$. In order to evaluate our approach, we perform this automatic procedure on BOSPHORUS database. BOSPHORUS database [SAD$^+$08] is suitable for this evaluation as it contains 3D scans of 60 men and 45 women, 105 subjects in total, in various poses, expressions and in presence of occlusions (eyeglasses, hand, hair). The majority of the subjects are aged between 25 and 35. The number of total face scans is 4652, where for most subjects 54 scans are available, but for 34 there are only 31 scans. The interesting part is that for each subject, there

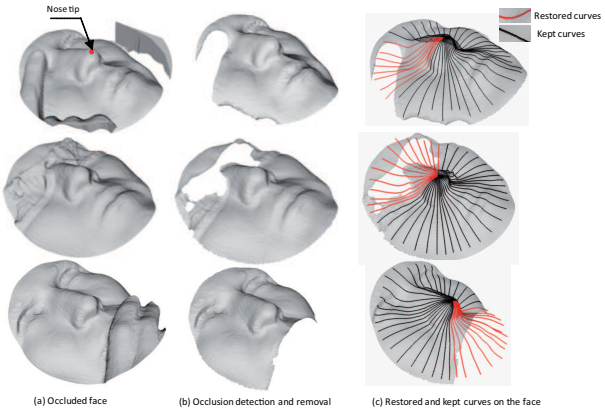


Figure 5.17: (a) Faces with external occlusion, (b) faces after detection and removal of occluding parts, and (c) estimation of the occluded parts using a statistical model on shape spaces of curves.



Figure 5.18: Examples of faces from the Bosphorus database.

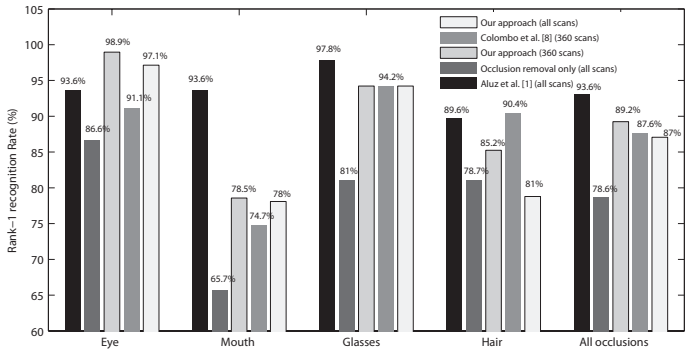


Figure 5.19: Recognition results on Bosphorus database and comparison with state-of-the-art approaches.

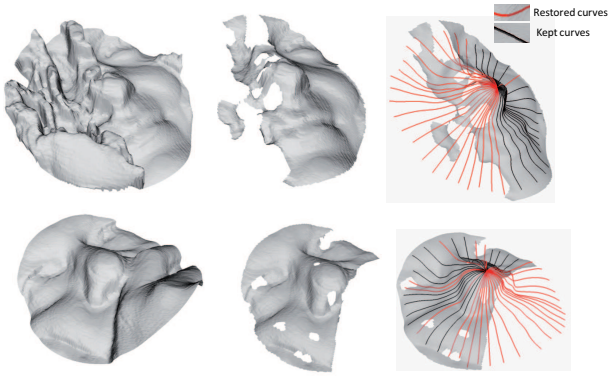


Figure 5.20: Examples of non recognized faces. Each row illustrates, from left to right, the occluded face, the result of occlusion removal and the result of restoration.

are four scans with occluded parts. These occlusion are i) mouth occlusion by hand, ii) eyeglasses, iii) occlusion of face with hair, and iv) occlusion of left eye and forehead regions by hands. Fig. 5.18 shows sample images from the Bosphorus 3D database illustrating expression variations and typical occlusions.

In order to collect training data, we consider 105 neutral faces with complete data of the 105 different persons present in the BOSPHORUS database. Similarly to previous approaches, we pursued the following protocol: a neutral face of each of the 105 persons is taken to form a gallery dataset with size 105. The probe contains scans with the 381 scans involving occlusions. The rank-one recognition rate is reported in Fig. 5.19 for different approaches per type of occlusion. As these results show, the process of restoring occluded parts increases the accuracy of recognition. The rank-one recognition rate is 78.63% when we delete the occluded parts and apply recognition algorithm using the remaining parts, as described in Section 6. However, if we perform restoration, the recognition rate improves to 87.06%. The results using the ICP algorithms are very low when compared with our approch. Examples of 3d faces recognized by our approach are shown in Fig. 5.17, with different steps of the algorithm. Compared to state-of-the-art approaches, Aluz et al [AGA08] reported 93.69% rank-one recognition rate as overall occlusion. While this reported performance is very good, the process has manual components. Actually, the authors partition the face manually and fuse the scores for matching differents parts of the

face together. In order to compare with Colombo et al. [CCS09], we reduce the probe dataset to 360 by discarding bad quality scans as Colombo et al. [CCS09] did. Our method outperforms this in the global results, with a performance of 89.25%, although individually our performance is smaller in the case of occlusion by hair. According this particular occlusion, the additional part is hardly completely discarded. Therefore, during the restoration step, our algorithm tries to keep non-removed parts. This leads to deformation in the shape of curves and hence affects the accuracy. We present some examples of non recognized faces in case of hair occlusion in Fig. 5.20. For these examples, our occlusion detection step fails. In that case, discarding curves passing through occlusion is better than restoring them as illustrated in the Fig. 5.19.

5.5 Conclusion

Building on a differential geometric representation of shapes, and geodesic lengths as shape metrics, we have presented in this chapter tools that enable a statistical analysis of 3D shape.Despite the non-linearity of the manifolds, we are able to calculate statistics using appropriate methods such as the Karcher mean algorithm for facial surfaces averaging. This presents the main step to allow hierarchical organization (i.e. a tree) of gallery of faces and perform efficient identification (by parsing the tree). We validated this idea using the gallery dataset of FRGCv2 including 410 subjects and achieved about 90% of identification rate (compared to 97.7% using exhaustive search). Time consuming for probe identification in that dataset gallery is reduced by factor of 20. Another important contribution presented in this chapter is the missing data recovery via a learning model proposed within the same manifold. We proposed to restrict to recover only the missing data on the 3D scan by considering the same representation (3D face scan as a collection of radial curves), by projecting observed shapes of radial curves located in the missing data onto the tangent space at sample mean (actually a vector space). That is, we construct a low dimensional subspace Eigencurves for each element of the indexed collection of curves, and then represent a curve (with missing date) as a linear combination of these Eigencurves. In order to quantatively validate the restoration, we propose a simple way to detect and discard occlusion then we apply our restoration algorithm on BOSPHORUS. Recognition

results match or improve the state-of-art of methods on BOSPHORUS database which poses the challenge of occlusion.

Chapter 6

Conclusion and Future study

Three-dimensional face recognition has become a major trend and attractive biometric modality, trying to use 3D geometry (shape) of the face in order to improve face recognition systems accuracies. While traditional two-dimensional face recognition methods suffer from sensitivity to external factor variations, such as lighting conditions, head pose, and are also sensitive to the use of cosmetics, 3D methods appear to be more robust to these variations. Yet, the problems of facial expressions, on one side, since the geometry of the face significantly changes, and missing data, on other side, are major challenges in 3D face recognition.

In this thesis we presented contributions to the 3D face analysis and recognition tasks by proposing a unified geometric framework for deforming, comparing, matching, averaging and clustering facial surfaces. We demonstrated and justified experimentally the effectiveness of our novel method in biometric recognition. The main ingredient of the tools introduced to perform such geometric analysis is the elastic matching of 3D faces. In fact, several previous researches proposed to model facial expressions as isometries of the facial surface. Using this model, the problem of expression-invariant face comparison returns to comparison of nearly-isometric non-rigid surfaces. However, this assumption seems to be sated in the case of large deformations because the facial shapes are affected by transformations that exceed surface bending and includes elastic transformation (stretching/shrinking) of the

facial surface, in addition to the open mouth problem. To mitigate the influence of facial expressions in face shapes comparison, we proposed a representation of the facial surfaces by a dense collection of radial curves and extended existing elastic metric for comparing open curves in $\mathbb{R}^3$ to compare facial surfaces. The task of 3D face comparison boils down to the comparison of pairwise radials curves. The elasticity in the proposed framework allows lessening of the expression deformation effects by matching better radial curves before their comparison. This curve-to-curve comparison allows also handling missing parts challenge by discarding bad quality curves in the visible part of the face scan. In fact, in real-word applications one to compare partial data, this is because current 3D sensors are able to capture depth maps that contains holes due to self-occlusions or imperfections of the range finder sensors. We experimentally validate the proposed approach using the well-known benchmarks (FRGCv2 and GavabDB) and compare performances to state-of-the-art ones. We reported the second best recognition rates on GavabDB, and provided satisfying results on both facial expressions and significant pose variations. Our approach achieved the third best recognition rate on FRGCv2 dataset.

Computational time for processing and comparing the geometry of 3D faces led us to think about this problem and propose efficient solutions. Despite the non-linearity of the manifolds, we are able to calculate statistics using appropriate methods as the Karcher mean algorithm. This presents the main step to allow hierarchical organization (i.e. a tree) of gallery of faces and perform efficient identification (by parsing the tree). We validated this idea using the gallery dataset of FRGCv2 including 410 subjects and achieved about 90% of identification rate (compared to 97.7% using exhaustive search). Time consuming for probe identification in that dataset gallery is reduced by factor of 20. Another important aspect presented in this thesis is the missing data recovery via a learning model proposed within the same manifold. This task is not-trivial due to the non-linearity of the manifold, therefore linear statistical approaches, associated with the vector space, could not be applied directly. We propose to restrict to recover only the missing data on the 3D scan by considering the same representation (3D face scan as a collection of radial curves), by projecting observed shapes of radial curves located in the missing data onto the tangent space at sample mean (actually a vector space), and further reduce their dimensions using T-PCA approach (PCA

on the tangent space of the manifold). It returns to construct a low dimensional subspace Eigencurves for each element of the indexed collection of curves, and then represent a curve (with missing date) as a linear combination of these Eigencurves. This approach may be beneficial in real-word application were facial scans could be damaged by the presence of occlusions (eyeglasses, hands, hairs, etc.). The problems of occlusion removal and partially matching facial surfaces are rarely addressed by the community and present a new important issue.

Future work

The issues that we hope to address in the future are of different nature. First, the presented shape analysis framework could be improved by introducing the texture information in addition to the facial geometry. In fact, 3D face scans consist of two channels; the texture image mapped on the 3D geometry of the faces.
In addition, the SRV representation and the elastic framework are designed for curves in $\mathbb{R}^n$ and could be easily extended to include this information. The introduction of color information of the face in this representation allows 3D shape matching more accurately because of the richness of the face parts with different textures (lips, eyes, cavities of the nose, eyebrows, etc.).

Second, the problems of occlusion removal then missing part recovery using the statistical learning method (described in chapter 5) could be experimentally validated on an appropriate dataset such as Bosphorus [SAD+08] that includes different occlusions by eyeglasses, hand or hair. With the availability of dynamic 3D sensors able to produce real time 3D reconstruction of the scene, these challenges are present in real-word like applications. Other face analysis-based applications could be addressed using geometric analysis tools developed in this thesis such as soft biometric identification (gender and ethnicity classification) in addition to Human-Machine Interface applications as 3D facial expression recognition that become one of the important research topic.

More fundamentally, the geometric shape analysis of parameterized surfaces (boundary

of 3D objects) issue is still an open problem and presents the main target goal of several researches especially when having such challenges (non-rigid/elastic deformations of the shape and missing data). While 2D shapes analysis has received a growing attention and a wide range of approaches is proposed, very few ones studying the shapes of surfaces directly, in a way that is analogous to shape analysis of curves, exist in the literature. The case of surfaces is more difficult than curves because curves are naturally parameterized (points on curves are ordered) while the points on surfaces are not (surface meshes). Certain approaches proposed to see the surface as a collection of curves and bring the problem to analyzing shapes of these curves which is relatively easier then return to surfaces. Riemannian analysis is a promising way since it offers powerful tools for removing all shape-preserving transformations from the pre-shape space (representation space) using the notion of quotient space and provides theory and algorithms to compute statistics (mean sample, covariances, etc.) that allows hierarchical organization of large datasets. This idea is widely illustrated and experimentally justified in this work in biometric context based on facial surfaces and leads both scientific and real-word challenges. Despite the benefits of curve-based representation and analysis to overcome missing data, an interesting future direction will be to work directly on surfaces. Recently, Kurtek et al. [KKDS10] proposed Riemannian framework for shape analysis of (continuous) surfaces, both closed and open, in $\mathbb{R}^3$ that are invariant to rotation, translation, scaling, and re-parameterization. The authors are concerned with explicit parametric representations of surfaces (not their embeddings in volumes or level functions, nor their restrictions to sets of landmarks).

Publications of the author

International journals

[ADB+09] B. BENAMOR, **H. DRIRA**, L. BALLIHI, A. SRIVASTAVA, M. DAOUDI An Experimental Illustration of 3D Facial Shape Analysis under Facial Expressions, *Annals of telecommunications, 2009.*

H. DRIRA, B. BENAMOR, M. DAOUDI, A. SRIVASTAVA, R. SLAMA, Pose and expression-robust 3d face recognition using elastic radial curves, *(IEEE Transactions on Pattern Analysis and Machine Intelligence, Undergoing major revision after the first round of review, 2011.).*

International conferences

[DASD09] **H. DRIRA**, B. BENAMOR, A. SRIVASTAVA, M. DAOUDI, A Riemannian Analysis of 3D Nose Shapes For Partial Human Biometrics, pp. 2050-2057, *in IEEE International Conference on Computer Vision, (ICCV 2009).*

[DBADS09]**H. DRIRA**, B. BENAMOR, M. DAOUDI, A. SRIVASTAVA, Nasal Region Contribution in 3D Face Biometrics Using Shape Analysis Framework, *Third International Conference on Biometrics, (ICB 2009).*

[DBDS10]**H. DRIRA**, B. BENAMOR, M. DAOUDI, A. SRIVASTAVA, Pose and Expression-Invariant 3D Face Recognition using Elastic Radial Curves, *Proceedings of the British Machine Vision Conference, (BMVC 2010).*

[DBADS10] **H. DRIRA**, B. BENAMOR, M. DAOUDI, A. SRIVASTAVA, Elastic Radial Curves to Model 3D Facial Deformations, *(3-D Object Retrieval ACM Multimedia Workshop 2010).*

[VvJD+11] VELTKAMP, R.C. and VAN JOLE, S. and **DRIRA, H.** and BEN AMOR, B. and DAOUDI, M. and LI, H. and CHEN, L. and CLAES, P., SMEETS, D. and

HERMANS, J. and VANDERMEULEN, D. and SUETENS, P., SHREC'2011 Track: 3D Face Models Retrieval, *(3-D Object Retrieval Eurographics Workshop 2011)*.

National conferences

[DASD] **H. DRIRA**, B. BENAMOR, A. SRIVASTAVA, M. DAOUDI, Analyse de forme des régions nasales pour la reconnaissance de visages 3D, *(Coresa 2009)*.

Technical reports

[DADa]**H. DRIRA**, B. BENAMOR, M. DAOUDI, 3D face matching algorithm, *Technical report ANR, FAR3D project, 2010)*

[DADb]**H. DRIRA**, B. BENAMOR, M. DAOUDI, Mathematical Definition of Faces Space, *Technical report ANR, FAR3D project, 2009)*

Bibliography

[ADB+09] Amor B.B., Drira H., Ballihi L., Srivastava A., Daoudi M., An Experimental Illustration of 3D Facial Shape Analysis under Facial Expressions , *Annals of telecommunications*, 64(5-6):369–379. 2009.

[AGA08] Alyuz N., Gokberk B., Akarun L., A 3D Face Recognition System for Expression and Occlusion Invariance , in *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*. 29 2008.

[AGD+08] Alyüz N., Gökberk B., Dibeklioglu H., Savran A., Salah A.A., Akarun L., Sankur B., 3D Face Recognition Benchmarks on the Bosphorus Database with Focus on Facial Expressions , in *BIOID*, pp. 57–66. 2008.

[BBK05] Bronstein A.M., Bronstein M.M., Kimmel R., Three-Dimensional Face Recognition , *International Journal of Computer Vision*, 64(1):5–30. 2005.

[BBK06] Bronstein E.M., Bronstein M.M., Kimmel R., Robust expression-invariant face recognition from partially missing data , in *in Proc. ECCV. 2006, Lecture Notes on Computer Science*, pp. 396–408, Springer. 2006.

[BBK07] Bronstein A.M., Bronstein M.M., Kimmel R., Expression-Invariant Representations of Faces , *IEEE Transactions on Image Processing*, 16(1):188–197. 2007.

[BBP10] Berretti S., Bimbo A.D., Pala P., 3D Face Recognition Using Isogeodesic Stripes , *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(12):2162–2177. 2010.

[BCF06a] Bowyer K.W., Chang K., Flynn P., A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition , *Comput. Vis. Image Underst.*, 101(1):1–15, ISSN 1077-3142. 2006.

[BCF06b] Bowyer K.W., Chang K.I., Flynn P.J., A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition , *Computer Vision and Image Understanding*, 101(1):1–15. 2006.

[BD62] Bellman R., Dreyfus S., *Applied Dynamic Programming*, RAND Corporation. 1962.

[CBF06] Chang K., Bowyer W., Flynn P., Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression , *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(10):1695 –1700, ISSN 0162-8828. oct. 2006.

[CCS09] Colombo A., Cusano C., Schettini R., Gappy PCA Classification for Occlusion Tolerant 3D Face Detection , *J. Math. Imaging Vis.*, 35(3):193–207, ISSN 0924-9907. 2009.

[CMCS06] Cook J., McCool C., Chandran V., Sridharan S., Combined 2D/3D Face Recognition Using Log-Gabor Templates , in *AVSS*, p. 83. 2006.

[DADa] Drira H., Amor B.B., Daoudi M., 3D face matching algorithm, ANR FAR3D project, 2010 , Technical report, University of Lille1.

[DADb] Drira H., Amor B.B., Daoudi M., Mathematical Definition of Faces Space, ANR FAR3D project, 2009 , Technical report, University of Lille1.

[DASD] Drira H., Amor B.B., Srivastava A., Daoudi M., Analyse de forme des régions nasales pour la reconnaissance de visages 3D , in *3-D Object Retrieval Eurographics Workshop*, Toulouse, France. mars 19.

[DASD09] Drira H., Amor B.B., Srivastava A., Daoudi M., A Riemannian Analysis of 3D Nose Shapes For Partial Human Biometrics , in *IEEE International Conference on Computer Vision*, pp. 2050–2057. 2009.

[DBADS09] Drira H., Ben Amor B., Daoudi M., Srivastava A., Nasal Region Contribution in 3D Face Biometrics Using Shape Analysis Framework , in *Third International Conference on Biometrics*, pp. 357–366, Alghero, Italy. June 2-5 2009.

[DBADS10] Drira H., Ben Amor B., Daoudi M., Srivastava A., Elastic Radial Curves to Model 3D Facial Deformations , in *3-D Object Retrieval ACM Multimedia Workshop*, pp. 75–80, Florence, Italy. October 25 2010.

[DBDS10] Drira H., Benamor B., Daoudi M., Srivastava A., Pose and Expression-Invariant 3D Face Recognition using Elastic Radial Curves , in *Proceedings of the British Machine Vision Conference*, pp. 90.1–90.11, BMVA Press, doi:10.5244/C.24.90. 2010.

[Dij59] Dijkstra E.W., A Note on Two Problems in Connection with Graphs , *Numerische Math.*, 1:269–271. 1959.

[DM98] Dryden I., Mardia K., *Statistical Shape Analysis*, John Wiley & Son. 1998.

[FBF08] Faltemier T.C., Bowyer K.W., Flynn P.J., A Region Ensemble for 3-D Face Recognition , *IEEE Transactions on Information Forensics and Security*, 3(1):62–73. 2008.

[FLPJ] Fletcher P.T., Lu C., Pizer S.M., Joshi S., Principal Geodesic Analysis for the Study of Nonlinear Statistics of Shape , *To appear, IEEE Transactions on Medical Imaging 2011*.

[GAMB07] Gupta S., Aggarwal J.K., Markey M.K., Bovik A.C., 3D Face Recognition Founded on the Structural Diversity of Human Faces , in *Proceeding of Computer Vision and Pattern Recognition, CVPR*. 2007.

[GM98] Grenander U., Miller M.I., Computational Anatomy: An Emerging Discipline , *Quarterly of Applied Mathematics*, LVI(4):617–694. 1998.

[Gor92] Gordon G., Face recognition based on depth and curvature features , in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 108–110. 1992.

[Gre93] Grenander U., *General Pattern Theory*, Oxford University Press. 1993.

[HAWC11] Huang D., Ardabilian M., Wang Y., Chen L., A novel geometric facial representation based on multi-scale extended local binary patterns , in *FG*, pp. 1–7. 2011.

[HB97] Hofmann T., Buhmann J.M., Pairwise Data Clustering by Deterministic Annealing , *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(1):1–14. 1997.

[HOA+11] Huang D., Ouji K., Ardabilian M., Wang Y., Chen L., 3D Face Recognition Based on Local Shape Patterns and Sparse Representation Classifier , in *MMM (1)*, pp. 206–216. 2011.

[HZA+10] Huang D., Zhang G., Ardabilian M., Wang, Chen L., 3D face recognition using distinctiveness enhanced facial representations and local feature hybrid matching , in *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, pp. 1–7. 2010.

[JKSJ07] Joshi S.H., Klassen E., Srivastava A., Jermyn I., A Novel Representation for Riemannian Analysis of Elastic Curves in $\mathbb{R}^n$, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR*. 2007.

[Kar77] Karcher H., Riemannian center of mass and mollifier smoothing , *Communications on Pure and Applied Mathematics*, 30:509–541. 1977.

[KKDS10] Kurtek S., Klassen E., Ding Z., Srivastava A., A novel riemannian framework for shape analysis of 3D objects , in *Proceedings of Conference on Computer Vision and Pattern Recognition, CVPR*, pp. 1625–1632. 2010.

[KPT+07] Kakadiaris I.A., Passalis G., Toderici G., Murtuza M.N., Lu Y., Karampatziakis N., Theoharis T., Three-Dimensional Face Recognition

in the Presence of Facial Expressions: An Annotated Deformable Model Approach , *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(4):640–649. 2007.

[Kre91] Kreyszig E., *Differential Geometry*, Dover Publications Inc. 1991.

[KS06] Klassen E., Srivastava A., Geodesics Between 3D Closed Curves Using Path-Straightening , in *Proceedings of European Conference on Computer Vision, ECCV*, pp. I: 95–106. 2006.

[KSMJ04] Klassen E., Srivastava A., Mio W., Joshi S., Analysis of Planar Shapes Using Geodesic Paths on Shape Spaces , *IEEE Pattern Analysis and Machine Intelligence*, 26(3):372–383. March, 2004.

[KTP00] Kotropoulos C., Tefas A., Pitas I., Frontal Face Authentication Using Morphological Elastic Graph Matching , *IEEE Transactions on Image Processing*, 9(4):555–560. April 2000.

[LJ06] Lu X., Jain A.K., Deformation Modeling for Robust 3D Face Matching , in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2006*, pp. 1377–1383. 2006.

[LJ08] Lu X., Jain A., Deformation Modeling for Robust 3D Face Matching , *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(8):1346 –1357, ISSN 0162-8828. aug. 2008.

[LJZ09] Li X., Jia T., Zhang H., Expression-insensitive 3D face recognition using sparse representation , *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 0:2575–2582. 2009.

[LSY+05] Lee Y., Song H., Yang U., Shin H., Sohn K., Local Feature Based 3D Face Recognition , in *Proceedings of Audio- and Video-Based Biometric Person Authentication, AVBPA*, pp. 909–918. 2005.

[MAM09] Mahoor M.H., Abdel-Mottaleb M., Face recognition based on 3D ridge images obtained from range data , *Pattern Recognition*, 42(3):445–451. 2009.

[MBO07] Mian A., Bennamoun M., Owens R., An Efficient Multimodal 2D-3D Hybrid Approach to Automatic Face Recognition , *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(11):1927 –1943, ISSN 0162-8828. nov. 2007.

[MFA08] Mousavi M.H., Faez K., Asghari A., Three Dimensional Face Recognition Using SVM Classifier , in *ICIS '08: Proceedings of the Seventh IEEE/ACIS International Conference on Computer and Information Science*, pp. 208–213, Washington, DC, USA. 2008.

[MM06] Michor P.W., Mumford D., Riemannian Geometries on Spaces of Plane Curves , *Journal of the European Mathematical Society*, 8:1–48. 2006.

[MMJ+10] Moorthy A., Mittal A., Jahanbin S., Grauman K., Bovik A., 3D Facial Similarity: automatic assessment versus perceptual judgments , in *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, pp. 1–7. 2010.

[MMS07] Mpiperis I., Malassiotis S., Strintzis M.G., 3-D Face Recognition With the Geodesic Polar Representation. , *IEEE Transactions on Information Forensics and Security*, 2(3-2):537–547. 2007.

[MMS08] Mpiperis I., Malassiotis S., Strintzis M., Bilinear Models for 3-D Face and Facial Expression Recognition , *Information Forensics and Security, IEEE Transactions on*, 3(3):498 –511, ISSN 1556-6013. sep. 2008.

[MR10] McKeon R., Russ T., Employing region ensembles in a statistical learning framework for robust 3D facial recognition , in *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, pp. 1–7. 2010.

[MS01] Mavridis N. F.T.D.P.S.M., Strintzis M.G., The HISCORE face recognition application: Affordable desktop face recognition based on a novel 3D camera , in *International Conference Augmented Virtual Environments and 3D Imaging*, pp. 157–160. 2001.

[MS04] Moreno A.B., Sanchez A., GavabDB: A 3D face database , in *Workshop on Biometrics on the Internet*, pp. 77–85. 2004.

[MSJ07a] Mio W., Srivastava A., Joshi S., On Shape of Plane Elastic Curves , *International Journal of Computer Vision*, 73(3):307–324. 2007.

[MSJ07b] Mio W., Srivastava A., Joshi S., On Shape of Plane Elastic Curves , *International Journal of Computer Vision*, 73:307–324, ISSN 0920-5691, 10.1007/s11263-006-9968-0. 2007.

[MSVD05] Moreno A.B., Sanchez A., Velez J.F., Díaz F.J., Face recognition using 3D local geometrical features: PCA vs. SVM , in *Int. Symp. on Image and Signal Processing and Analysis*. 2005.

[MY01] Miller M.I., Younes L., Group Actions, Homeomorphisms, and Matching: A General Framework , *Intl. Journal of Computer Vision*, 41(1/2):61–84. 2001.

[PFS+05] Phillips P.J., Flynn P.J., Scruggs W.T., Bowyer K.W., Chang J., Hoffman K.J., Marques J., Min J., Worek W.J., Overview of the Face Recognition Grand Challenge , in *CVPR (1)*, pp. 947–954. 2005.

[QSBS10] Queirolo C.C., Silva L., Bellon O.R., Segundo M.P., 3D Face Recognition Using Simulated Annealing and the Surface Interpenetration Measure , *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32:206–219, ISSN 0162-8828. 2010.

[SAD+08] Savran A., Alyüz N., Dibeklioğlu H., Çeliktutan O., Gökberk B., Sankur B., Akarun L., Biometrics and Identity Management , chapitre Bosphorus Database for 3D Face Analysis, pp. 47–56, Berlin, Heidelberg: Springer-Verlag. 2008.

[Set96] Sethian J., A fast marching level set method for monotonically advancing fronts , in *The National Academy of Sciences of States of America, PNAS*, volume 93, pp. p1591–1595. 1996.

[SJML05] Srivastava A., Joshi S.H., Mio W., Liu X., Statistical Shape Analysis: Clustering, Learning, and Testing , *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(4):590–602. 2005.

[SKJJ11] Srivastava A., Klassen E., Joshi S., Jermyn I., Shape Analysis of Elastic Curves in Euclidean Spaces , *To appear, IEEE Transactions on, Pattern Analysis and Machine Intelligence*. 2011.

[SSD06] Samir C., Srivastava A., Daoudi M., Three-Dimensional Face Recognition Using Shapes of Facial Curves , *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28:1858–1863, ISSN 0162-8828. 2006.

[SSDK09] Samir C., Srivastava A., Daoudi M., Klassen E., An Intrinsic Framework for Analysis of Facial Surfaces , *International Journal of Computer Vision*, 82(1):80–95. 2009.

[SSJD09] Srivastava A., Samir C., Joshi S.H., Daoudi M., Elastic Shape Models for Face Analysis Using Curvilinear Coordinates , *Journal of Mathematical Imaging and Vision*, 33(2):253–265. 2009.

[STG08] Sun Y., Todorovic S., Goodison S., A feature selection algorithm capable of handling extremely large data dimensionality , in *In Proceedings of the 8th SIAM International Conference on Data Mining*, pp. 530–540. 2008.

[tHV09] ter Haar F.B., Veltkamp R.C., A 3D face matching framework for facial curves , *Graphical Models*, 71(2):77–91. 2009.

[tHV10] ter Haar F., Velkamp R.C., Expression modeling for expression-invariant face recognition , *Computers and Graphics*, 34(3):231–241. 2010.

[VvJD+11] Veltkamp R.C., van Jole S., Drira H., Amor B.B., Daoudi M., Li H., Chen L., Claes P., Smeets D., Hermans J., Vandermeulen D., Suetens P., SHREC '11 Track: 3D Face Models Retrieval , in *3DOR*, pp. 89–95. 2011.

[WLT10] Wang Y., Liu J., Tang X., Robust 3D Face Recognition by Local Shape Difference Boosting , *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32:1858–1870, ISSN 0162-8828. 2010.

[You98] Younes L., Computable elastic distance between shapes , *SIAM Journal of Applied Mathematics*, 58:565–586. 1998.

Abstract

We propose, in this thesis, a unified Riemannian framework for comparing, deforming, averaging and hierarchically organizing facial surfaces. This framework is applied within the 3D face recognition problem where facial expressions, pose variations, and occlusions are the main challenges of this topic. The facial surfaces are represented by collections of level curves and radial ones. The set of closed curves (level curves) constitute an infinite dimensional sub-manifold and is used to represent the nasal region, the most stable part of the face. The facial surface is represented by an indexed collection of radial curves. In this case, the calculus is simpler and the space of open curves shape is simply the hypersphere of Hilbert space. The comparison in this shape space is done via an "elastic" metric in order to handle non-isometric deformations of facial surfaces. We propose algorithms for computing means and eigenvectors in these nonlinear manifolds and hence algorithms for estimation of missing parts of 3D facial surfaces. Comparison with competitor approaches using a common experimental setting on the FRGCv2, GAVAB, BOSPHORUS databases, shows that our solution is able to obtain, and outperform in some scenarios, the state-of-the-art results.

Résumé

Dans cette thése, nous proposons un cadre Riemannien pour comparer, déformer, calculer des statistiques et organiser de maniére hiérarchique des surfaces faciales. Nous appliquons ce cadre á la biométrie faciale 3D oú les défis sont les expressions faciales, les variations de la pose et les occultations du visage par des objets externes. Les surfaces faciales sont représentées par un ensemble de courbes de niveaux et de courbes radiales. L'ensemble des courbes fermées (de niveau) constitue une sous-variété non-linéaire de dimension infinie et est utilisé pour représenter le nez, la partie la plus stable du visage. La surface faciale est présentée, par ailleurs, par une collection indexée de courbes radiales. Dans ce cas, le calcul se simplifie et l'espace des formes des courbes ouvertes se raméne à une hyper sphère de l'espace de Hilbert. La comparaison dans l'espace des formes se fait via une métrique élastique afin de faire face aux déformations non-isométriques (ne conservant pas les longueurs) des surfaces faciales. Nous proposons des algorithmes pour calculer les moyennes, les vecteurs propres dans ces variétés non-linéaires et l'estimation des parties manquantes des surfaces faciales 3D. L'approche présentée dans cette thése a été validée sur des Benchmarks connus (FRGCv2, GAVAB, BOSPHORUS) et obtenu des résultats compétitifs par rapport aux méthodes de l'état de l'art.