Processing and analysis of 2.5D face models for non-rigid mapping based face recognition using differential geometry tools

Przemyslaw Szeptycki

THESE

pour obtenir le grade de
DOCTEUR DE L'ECOLE CENTRALE DE LYON
Spécialité : Informatique

dans le cadre de l'Ecole Doctorale InfoMaths
présentée et soutenue

Processing and analysis of 2.5D face models for non-rigid mapping based face recognition using differential geometry tools

---

par
PRZEMYSŁAW SZEPTYCKI

Juillet 2011

Directeur de thèse : Liming CHEN
et
Co-directeur de thèse : Mohsen ARDABILIAN
JURY

-	Prof. Georgy Kukharev	West Pomeranian University of Technology in Szczecin	Rapporteur
-	Prof. Faouzi Ghorbel	École Nationale des Sciences de l'Informatique de Tunis	Rapporteur
-	Prof. Jean-Marie Morvan	University Claude Bernard Lyon 1	Examinateur
-	Prof. Dimitris Samaras	Stony Brook University of New York	Examinateur
-	Dr Dijana Petrovska-Delacrétaz	TELECOM-SudParis	Examinateur
-	Prof. Liming Chen	Ecole Centrale de Lyon	Directeur de thèse
-	Dr Mohsen Ardabilian	Ecole Centrale de Lyon	Co-directeur de thèse

Abstract

This Ph.D thesis work is dedicated to 3D facial surface analysis, processing as well as to the newly proposed 3D face recognition modality, which is based on mapping techniques.

Facial surface processing and analysis is one of the most important steps for 3D face recognition algorithms. Automatic anthropometric facial features localization also plays an important role for face localization, face expression recognition, face registration ect., thus its automatic localization is a crucial step for 3D face processing algorithms. In this work we focused on precise and rotation invariant landmarks localization, which are later used directly for face recognition. The landmarks are localized combining local surface properties expressed in terms of differential geometry tools and global facial generic model, used for face validation. Since curvatures, which are differential geometry properties, are sensitive to surface noise, one of the main contributions of this thesis is a modification of curvatures calculation method. The modification incorporates the surface noise into the calculation method and helps to control smoothness of the curvatures. Therefore the main facial points can be reliably and precisely localized (100% nose tip localization using 8 mm precision) under the influence of rotations and surface noise. The modification of the curvatures calculation method was also tested under different face model resolutions, resulting in stable curvature values. Finally, since curvatures analysis leads to many facial landmark candidates, the validation of which is time consuming, facial landmarks localization based on learning technique was proposed. The learning technique helps to reject incorrect landmark candidates with a high probability, thus accelerating landmarks localization.

Face recognition using 3D models is a relatively new subject, which has been proposed to overcome shortcomings of 2D face recognition modality. However, 3D face recognition algorithms are likely more complicated. Additionally, since 3D face models describe facial surface geometry, they are more sensitive to facial expression changes. Our contribution is reducing dimensionality of the input data by mapping 3D facial models on to 2D domain using non-rigid, conformal mapping techniques. Having 2D images which represent facial models, all previously developed 2D face recognition algorithms can be used. In our work, conformal shape images of 3D facial surfaces were fed in to $2D^2$ PCA, achieving more than 86% recognition rate rank-one using the FRGC data set.

The effectiveness of all the methods has been evaluated using the FRGC and Bosphorus datasets.

Keywords : 3D face landmarking, 3D face recognition, curvatures classification, conformal mapping.

Résumé

Ce travail de thèse concerne l'analyse de surfaces faciales en 3D, ainsi que leur traitement, dans le récent cadre de la modalité de reconnaissance de visages en 3D, basé sur des techniques d'appariement.

Le traitement de la surface faciale et son analyse constituent une étape importante dans les algorithmes de reconnaissance de visage en 3D. La localisation de points d'intérêt anthropométriques du visage joue par ailleurs un rôle important dans les techniques de localisation du visage, de reconnaissance d'expression, de recalage, etc. Ainsi, leur localisation automatique joue un rôle crucial dans les algorithmes de traitement du visage 3D. Dans ce travail, nous avons mis l'accent sur la localisation précise et invariante en rotation des points d'intérêt, qui seront utilisés plus tard pour la reconnaissance de visages. Ces points d'intérêt sont localisés en combinant les propriétés locales de la surface faciale, exprimées en termes de géométrie différentielle, et un modèle global et générique du visage. Etant donné que la sensibilité des courbures, qui sont des propriétés de géométrie différentielle, au bruit, une des contributions de cette thèse est la modification d'une méthode de calcul de courbures. Cette modification incorpore le bruit de la surface dans la méthode de calcul, et permet de contrôler la progressivité des courbures. Par conséquent, nous pouvons localiser les points d'intérêt de la surface faciale avec précision et fiabilité (100% de bonnes localisation du bout du nez avec une erreur maximale de 8mm par exemple) y compris en présence de rotations et de bruit. La modification de la méthode de calcul de courbure a été également testée pour différentes résolutions de visage, présentant des valeurs de courbure stables. Enfin, étant donné que donné que l'analyse de courbures mène à de nombreux candidats de points d'intérêt du visage, dont la validation est coûteuse, nous proposons de localiser les points d'intérêt grâce à une méthode d'apprentissage. Cette méthode permet de rejeter précocement des faux candidats avec une grande confiance, accélérant d'autant la localisation des points d'intérêt.

La reconnaissance de visages à l'aide de modèles 3D est un sujet relativement nouveau, qui a été propose pour palier aux insuffisantes de la modalité de reconnaissance de visages en 2D. Cependant, les algorithmes de reconnaissance de visage en 3D sont généralement plus complexes. De plus, étant donné que les modèles de visage 3D décrivent la géométrie du visage, ils sont plus sensibles que les images 2D de texture aux expressions faciales. Notre contribution est de réduire la dimensionnalité des données de départ en appariant les modèles de visage 3D au domaine 2D à l'aide de méthodes, non rigides, d'appariement conformal. L'existence de modèles 2D représentant les visages permet alors d'utiliser les techniques précédemment développées dans le domaine de la reconnaissance de visages en 2D. Dans nos travaux, nous avons utilisé les cartes conformales de visages 3D en conjonction avec l'algorithme $2D^2$ PCA, atteignant le score de 86% en reconnaissance de rang 1 sur la base de données FRGC.

L'efficacité de toutes les méthodes a été évaluée sur les bases FRGC et Bosphorus.

Mots clés : Landmarking de visages 3D, reconnaissance de visage 3D, classification de courbures, appariement conformal.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Research Topic

From many proposed modalities for people identification and verification, face recognition is one of the most acceptable techniques by a society. Human face is also a valuable and rich physiological feature, used by humans to distinguish a person.

Face biometrics solutions have been developing for over 50 years, starting from semi-automated systems for face recognition developed in 1960s and ending on multimodal solutions used nowadays. It is also an important alternative for selecting and developing an optimal biometric system, without any sophisticated hardware and need of physical contact with the device. The face recognition modality has now matured into a science of sophisticated mathematical representations and matching processes, where to achieve better accuracy, instead of 2D images, 3D face model-based modality has been proposed.

Indeed, 3D face models contain more information which is very useful for processing, analysis and face-based people recognition. Unlike 2D modality, 3D face recognition has ability to overcome problems of face rotations as well as light variations, which are the main concerns with 2D facial images. Except texture, 3D models also contain 3D face geometry properties which can be described in differential geometry forms like Mean or Gaussian curvatures. Having more properties, 3D face models analysis is more complex, allows more variations and results in precise solutions. It also leads to more sophisticated and complex face recognition algorithms which gain better results than 2D counterparts, especially in case of lighting changes.

Among all aspects of face analysis, we mainly focused on precise face landmarks detection, without any change in the model's structure, which is one of the proposed contributions. Secondly, we used localized anchor points for one of 3D face recognition techniques, which relies on 3D surface mapping to 2D domain, where non-rigid mapping techniques are used. The mapping leads to creating 2D facial images, embedding non-rigid surface deformations. This kind of mapping enables all previously developed 2D face recognition algorithms to be used on 3D data.

1.2 Problems and Objective

Recently 3D face recognition modality has emerged as a major research direction for face recognition. The 3D face recognition modality is theoretically insensitive to illumination and rotations, which are the main factors imposing strong hurdles on

2D face recognition counterparts. Therefore, use of 3D face models, which contain ample information on both geometry and texture, can improve the effectiveness of face recognition systems.

While the problem of illumination changes in 3D face recognition can be solved using different 3D models properties, like curvatures, etc., the rotation problem can be fixed by pose normalization before 3D facial shape-based matching. Having many degrees of freedom in 3D, most of the existing 3D facial surface registration techniques rely on facial landmarks. The essence of the surface registration problem, is how to bring two models together into a common coordinate system. One of the solutions, is to find some key points and try to match them. In 3D, to fulfil all six degrees of freedom (three degrees in rotations and three degrees in translations), one needs at least three non-co-linear, labelled key points between two 3D models. On 3D facial models, the most marked-out facial points are the nose tip and two inner eyes cavities, therefore a reliable localization of these facial feature points is significant for further face surface registration/recognition algorithms. The key points can be also used as reference points for recognition algorithms, where generally the nose tip is used as the origin. Therefore the most important factor, influencing recognition algorithm, is precision of the localized point. One of the objectives of this thesis is to develop an algorithm which will be able to handle different face rotations and resolution, achieving precise main facial landmarks in each condition.

3D face models are widely thought of as a major solution in dealing with the problem of illumination changes as geometric features can be used instead of 2D texture features. Very popular features, which can be projected onto 3D models are features from differential geometry, for instance Gaussian or Mean curvatures which characterize the intrinsic properties of a surface. However these features, while simple for the computation, are also very sensitive to surface noise. To achieve smooth decomposition of those features on noisy models, noise filtering techniques like median or Gaussian filters can be used. The downside is that they also tend to remove fine details, which can be useful for people identification (like scars or marks). Therefore another objective of this thesis is to modify curvatures calculation technique to control smoothness of curvatures and achieve their smooth decomposition without filtering techniques.

While 3D face models capture accurately facial surfaces thereby offering pose and lighting condition invariant geometric features, they are likely more sensitive to facial expression changes as compared to their 2D counterparts. Higher dimensional data, makes the whole processing algorithm more complex and sophisticated. On the other hand, face recognition in 2D is well known and studied for many years. Connecting all these facts together, a third objective of this thesis is to create 2D images of 3D facial models, which will simplify the 3D recognition problem to 2D while embedding facial expressions, thereby making applicable all the techniques previously developed for face recognition in 2D.

1.3 Our Approaches and Contributions

Increasing quality of 3D face models is a common part of all 3D face preprocessing algorithms. Models suffer from holes and spikes which are the main quality issues in nowadays laser-based scanning techniques. Removal of those artefacts is a crucial step for further face analysis and recognition. Our first contribution is the preprocessing method, which increases models quality without removing fine details on facial surface. The method consists of adding simple threshold to the median filtering technique, which helps to remove certain noise from processed models.

Nowadays, the most popular method for 3D face models analysis and processing, relies on curvatures and their decomposition on facial models. While the processing of the curvatures allows facial expressions analysis, landmarks localization, face recognition, etc., their calculation method is very sensitive to the surface noise. One of the solutions is to smooth out processed models before curvatures calculation. This solution however influences also fine details on facial models, which can be further used for face recognition purpose. Therefore our second contribution refers to the curvatures calculation method. Simple change in the points neighborhood, used for the curvatures approximation, affects the smoothness of the curvatures decomposition. The smoothness of the curvatures relies on the size of the neighborhood used for the curvatures approximation. Additionally, expressing the size of the neighborhood in real units [mm], results in very stable curvature values across different models' resolutions and points density.

Subsequently the developed method was used for precise facial main points localization, which is our next contribution. In order to precisely localize the main facial points, namely: nose tip and inner corners of eyes, a large neighborhood in the Gaussian and Mean curvatures approximation method was used. Next, Gaussian and Mean curvatures classification was used to localize possible regions of the nose and eyes. To localize precise landmark points, the maximum Gaussian curvature points in each region was localized. Analysis of the curvatures, calculated with a large neighborhood, results in analysis of the main patches on faces, which refer to the nose and eye cavities. Therefore reliable, accurate and precise landmarking algorithm can be achieved. While the curvatures are insensitive to the rotations, their correct analysis and validation method makes the algorithm also rotation invariant.

This kind of analysis generally leads to many facial anthropometric candidate points. To identify the true main facial points, generally mean face model (expressed for example in geometry or appearance) and its fitting error was used. While highly effective, these kind of approaches are computationally expensive and suffer from the exponential number of configurations of the candidate points. Alternatively, the nose tip can also be localized using a data-driven approach, by creating point signatures and thereby enabling fast statistical learning. Our next contribution is in studies and comparison of four geometry based point properties namely Mean and Gaussian curvatures, Shape Index and Curvedness in terms of descriptiveness for the nose tip localization.

Face recognition is one of the most acceptable biometric identification modali-

ties. As such, face recognition based on 2D images has been studied for many years. Unfortunately, its performance drops drastically in case of inconsistent light conditions and facial rotations. Therefore face recognition in 3D has been proposed. However, the recognition algorithms are likely more complicated. Additionally, since 3D face models describe facial surface geometry, they are more sensitive to facial expression changes. Our next contribution, is reducing dimensionality of the input data by mapping 3D facial models onto 2D domain using non-rigid, conformal or quasi-conformal mapping techniques. Since facial expressions introduce non-rigid deformations of facial surface, those mapping techniques have the potential to deal with them, reducing its influence on 2D images. Having 2D images which represent facial models, all previously developed 2D face recognition algorithms can be used.

To achieve comparable 2D images, non-rigid, conformal mapping techniques require the surface topology as well as the surface boundary to remain unchanged. The facial surface topology changes mainly due to large facial expressions, where open mouth causes appearance of additional vertices between lips. To preserve surface topology between expressions, open mouth detection and removal are necessary. To remove open mouth, one can rely on manual or automatic landmarks, yet automatic landmarking of the non-rigid part is a very challenging task. Additionally, removal of open mouth based on a few landmark points can be imprecise, causing incorrect topology preservation. Our another contribution, relies on open mouth detection based on curvatures analysis. Open mouth characterizes a large cavity which can be easily detected using curvatures decomposition analysis. Analysis of the principal curvatures permits to detect open mouth cavities. Additionally, the multi-scale curvatures analysis permits to detect open mouth cavities with varying size and position.

1.4 Organization of the Thesis

This thesis is organized as follows:

Chapter 2 focuses on face preprocessing and landmarking, where firstly we introduce the background of the topic, showing different modalities of 3D face landmarking techniques. Afterwards, we propose a modification of the spikes removal technique, which removes spikes without smoothing the whole 3D face models, thereby preserving the fine details. Subsequently, we propose a modification of the curvatures calculation method, which excludes facial surface noise. Next, in section 2.3 curvatures calculation method was used for 2.5D facial models landmarking algorithm. The proposed landmarking algorithm is able to precisely detect the main facial landmarks under facial rotations and different models resolution. Since validation of the point candidates given by the analysis of the curvatures is very time consuming, therefore in section 2.4 a data driven learning approach was proposed, which can exclude incorrect nose tip candidates with a high probability. Additionally in section 2.5 tests on stability of the modified curvatures calculation method using models with a different resolution were performed and demonstrated stability

of the curvatures values across different resolutions.

Chapter 3 covers 3D face recognition using non-rigid mapping techniques. First we introduce the related work on mapping-based 3D face recognition techniques. Section 3.4 shows dedicated models preprocessing for topology and boundary preservation, which are the conditions to achieve comparable face maps. The face surface topology is mainly affected by open mouth, therefore in this section we will focus on automatic open mouth detection and removal. The last section of the chapter (3.5) treats the problem of 3D face recognition based on 2D non-rigid mapping techniques, which are used to simplify the 3D face recognition problem to 2D, thereby making available all previously developed 2D face recognition algorithms. Additionally, since facial expression introduces non-rigid facial surface deformations, mapping techniques used are also non-rigid.

Finally chapter 4 concludes the thesis and gives hints and suggestions for the future work.

Chapter 2

2.5D Face Preprocessing, Analysis and Landmarking

2.1 Introduction

A problem of 2.5D[1] face preprocessing and analysis is common for all 3D face recognition algorithms. While 3D face models are theoretically insensitive to lighting conditions, the acquisition technique can be affected by the light [Bowyer *et al.* 2006]. The majority of scanning technologies are light based, where different, strong sources of light can affect the model's geometry (figure 2.1). Even under ideal illumination conditions, it is common for artifacts like spikes or holes to appear in specular, oily facial regions, the eyes and regions of facial hair (eyebrows, mustache or beard). Therefore a reliable preprocessing algorithm is required for models quality improvement, to exclude the impact of the quality on the recognition algorithm.

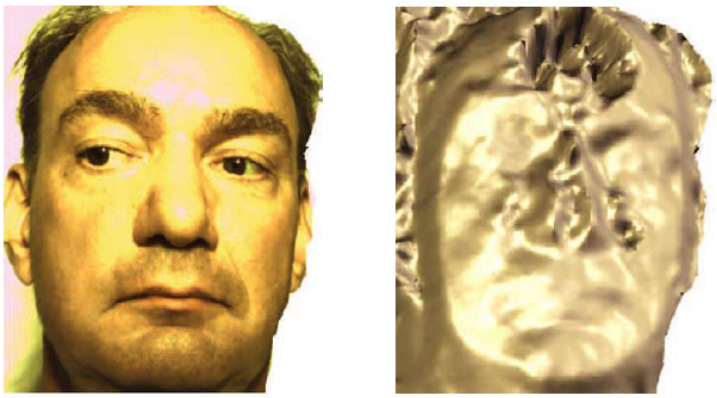

Figure 2.1: An example of incorrect shape caused by a strong source of light, presented in [Bowyer *et al.* 2006].

Rough 2.5D face models generally come without information about the face position or rotation. Therefore analysis of 3D facial surfaces is very important for 3D face recognition. Such an analysis gives the basic information about face position, size or rotation relative to the camera. These information in general is deduced

[1] A 2.5D model is a three-dimensional model with a constraint that the surface is given as a function $f(x, y)$, which means that every point in the (x, y) plane has at most one depth (z) value. 2.5D models are also called depth maps or depth images (figure 2.5).

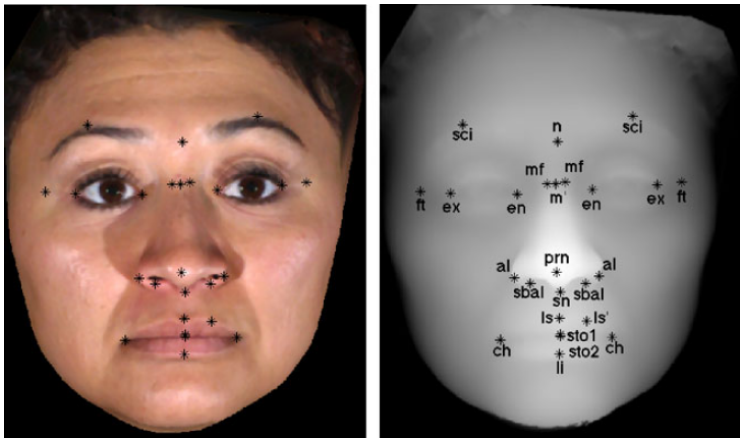

Figure 2.2: Example of anthropometric points used for face recognition (left: color image, right: range image).[Gupta *et al.* 2010]

from facial anchor points, which are fiducial points having consistent reproducibility between faces even in adverse conditions. These stable points generally include: nose tip, corners of eyes, corners of mouth, corners of nose, etc. (figure 2.2).

Accurately localized facial 3D landmarks, can be used for objectives like:

- mesh registration [Mian *et al.* 2007, Chang *et al.* 2006, Besl & McKay 1992, Shi *et al.* 2006, Lu *et al.* 2006],
- face recognition [Soltana *et al.* 2010, Gupta *et al.* 2010],
- facial expressions recognition [Zhao *et al.* 2010, Niese *et al.* 2008],
- face tracking [Sun & Yin 2008a],
- reference points for parametrization [Bronstein *et al.* 2005],
- facial regions cropping [Faltemier *et al.* 2008a, Chang *et al.* 2006],
- face localization [Colombo *et al.* 2006, Nair & Cavallaro 2009],
- face pose estimation [Sun & Yin 2008b, Malassiotis & Strintzis 2005].

Automatic facial landmarks localization generally meets the challenges of illumination, facial expressions, occlusions but also head pose and face models resolution. Additionally in the case of 2.5D scans, the head pose variations not only influence the facial appearance in the texture but also cause self-occlusions, where some landmark points can be hidden. The texture can be also affected by variations of lighting intensity or lighting source position, which change the facial appearance. Therefore approaches using only the texture, to localize facial points, are affected by its changes. Indeed, 3D facial models help to overcome the problems of lighting

variations and face position only while landmarks are localized by facial geometry analysis.

In this chapter, we focus on locating facial feature points on 2.5D face scans without any reinforcement from the texture side.

2.2 Related Work on 2.5D Models Quality Improvement and Landmarking

Preprocessing of 3D face meshes is a very important step in many applications, especially in the 3D face recognition task, where 3D mesh quality is influenced by noise, holes and movement of the subject. Like all classification techniques, the 3D face recognition requires the samples to be normalized and registered. The normalization as well as the face localization for many face recognition algorithms is based on the facial anthropometric points and is incorporated in the preprocessing step. Therefore a common preprocessing part for all 3D face recognition algorithms consists of: quality improvement, anchor points localization and face detection.

2.2.1 Quality improvement

As mentioned previously, actual scanning techniques are mainly based on the the light response analysis. Those techniques are sensitive to specular, oily regions, the eyes and regions of facial hair (eyebrows, mustache or beard), where in general some scanning artifacts can be observed [Bowyer *et al.* 2006, Boehnen & Flynn 2005]. Removal of those artifacts is a crucial step in the models quality improvement.

2.2.1.1 Noise removal

Noise appears on models often in the acquisition step by noisy sensors or insufficient environmental conditions. Such an impulse noise is called "spike" (figure 2.3) and in general, it is a point lying far from a surface created by its neighbors. To remove such a noise, its characteristic needs to be known. In 2.5D models, acquired by laser scanners, the impulse noise creates only in the z direction. Removal of the spikes artifacts can be divided in two groups: filtering and geometrical analysis.

Among the techniques frequently used to remove spike artifacts is the median filter [Zhao *et al.* 2009b, Zhao *et al.* 2009a, Faltemier *et al.* 2008b, Faltemier *et al.* 2008a, Faltemier *et al.* 2007, Mian *et al.* 2007]. Due to its simplicity, it has been extensively used for impulse noise removal. Simply calculating the median value from a neighborhood of a point and replacing its value by the calculated median, will lead to remove the spike noise from a processed model. Another filtering technique was used in [Wang *et al.* 2009, Colombo *et al.* 2006], where the Gaussian filter was used to remove the spikes and smooth data. Nevertheless these filtering techniques are simple and effective in the impulse noise removal, they also tend to remove fine details and change the whole model. Since the input models

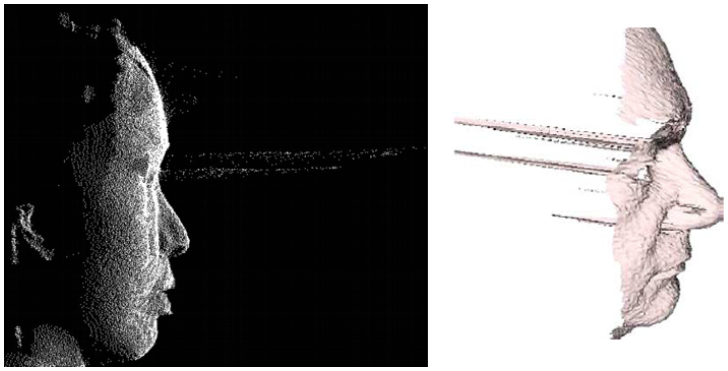

Figure 2.3: Example of spikes artefacts on 3D model acquired by laser based scanners. [Mian *et al.* 2007, Bowyer *et al.* 2006]

may differ in terms of resolution, the control of the kernel matrix size is necessary to adjust the strength of the filtering technique on the input model.

Another group of methods, are methods relying on geometrical properties analysis. Chang et al. in [Chang *et al.* 2006] analyzed angle between the optical axis and the point's local surface normal. If the angle was greater than a threshold value (80°), the point was considered as a spike. Another different technique was proposed by Bronstein et al. in [Bronstein *et al.* 2005], where the authors used discrete gradient norm; a vertex which has a high value of the norm, was removed as a potential spike. Mian et al. in [Mian *et al.* 2007, Mian *et al.* 2006] defined an outlier point as the one, whose distance is greater than a threshold d_t from any of its neighbors; d_t is calculated using the mean distance between neighboring points and its standard deviation.

Clearly it is hard to say which method has better accuracy in the noise removal. Ranking the methods according to the popularity, the most popular technique to remove the impulse noise is the median filter. The characteristic of the noise which appears only in the z direction, makes median filter a very effective technique to remove such artifacts.

2.2.1.2 Holes removal

As it was mentioned before, a general way to acquire 3D models is by contact-less laser techniques. The second type of artifacts, which are very often present on 3D facial models are holes. Holes, are mainly created in regions which absorb the light, like: hair, eyebrows, mustache or pupil (see figure 2.4). This occurs, because the eye is translucent and the light stripe is refracted instead of being reflected [Queirolo *et al.* 2010].

Discontinuities on the 3D model caused by errors in the enrollment stage, can be localized and filled by various methods. For those, who work with range images, holes are easy to find by scanning the range image for missing points which

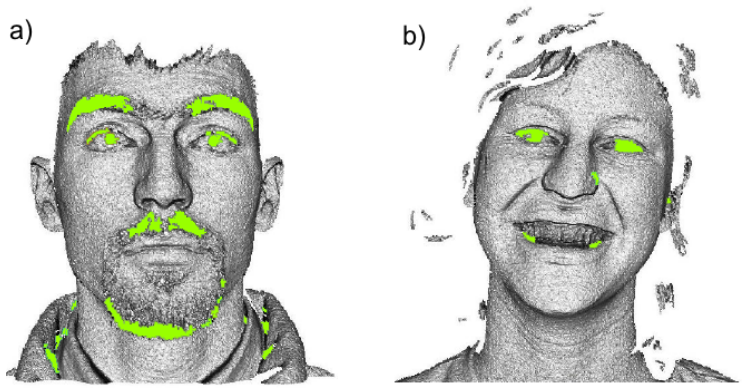

Figure 2.4: Example of holes on 3D models acquired by laser based scanner (marked in green, better seen in color).[Faltemier *et al.* 2008a]

are surrounded by the valid neighbors. Localized hole can be filled by simply interpolating x, y, z coordinates based on surrounding points [Faltemier *et al.* 2008b, Faltemier *et al.* 2008a, Faltemier *et al.* 2007]. Different method, developed to localize the holes on 2.5D models was presented in [Zhao *et al.* 2009b, Zhao *et al.* 2009a], where the authors used morphological reconstruction to locate the holes and cubic interpolation to fill them. The cubic interpolation has also been used in [Mian *et al.* 2007].

The problem of holes on 2.5D face models (range images) is not complex and can be resolved by simple interpolation. Holes can be localized by various methods starting from missing points localization on range images, analysis of number of connections between vertices or the number of triangles joining one edge.

2.2.2 Landmarks localization on 2.5D models

Face landmarking has gained increasing interest in resent years mainly because of its diverse applications, for example in: face normalization, face registration, face recognition etc.. To localize geometrically salient feature points such as a nose tip, eye corners, etc., most of the existing work relies on facial geometry analysis. The real-time solutions mainly base on point signatures, while precision on all landmarks is mainly gained by careful deformable models fitting. Thus all the 2.5D landmarking algorithms can be categorized by the tools used in facial points localization and are grouped into different categories in the following sections.

2.2.2.1 Curvatures analysis for the main points localization

Since the input of 3D face recognition algorithms are 3D models, the natural way to localize facial anthropometric points is by analysis of the facial surface. The most popular way to analyze a 3D surface is by a method called HK-classification, which in many cases, is the first step to reduce the number of candidate vertices for the

	The nose tip	The inner corners of the eyes
False	0.2%	0.2%
Non-existence	0%	5.71%
Correct	99.8%	94.09%

Table 2.1: Some of the facial regions appearance presented in [Moreno *et al.* 2003] using thresholded HK-classification, tested on 420 3D models.

anthropometric points.

Curvature analysis for the main landmark points localization and face detection was used in [Colombo *et al.* 2006, Moreno *et al.* 2003, Chang *et al.* 2006, Sun & Yin 2008b]. Usually, authors determine the three main anchor face points as: the nose tip and the inner corners of the eyes. The first step in the algorithm is to find the three main points candidates by a range image segmentation using the HK-classification. Since the second derivatives of a bi-quadratic polynomial surface approximation are very sensitive to noise, a smoothing filter on the range image was applied before the curvature calculation. To isolate high curvature regions authors used a thresholding process, discarding points with the low Mean and Gaussian curvature. After the thresholding process, in each highly convex or concave region the nose tip and the eyes candidates are chosen respectively. The appearance of the main facial regions after the thresholding was tested by Moreno et al. in [Moreno *et al.* 2003] and is presented in table 2.1 (tested on a set of 420 3D facial surfaces which contain different expressions and pose). The second, common part is the validation process. Having many nose tip and inner eye corners candidates, the authors in [Colombo *et al.* 2006] validated the correct three using PCA face reconstruction and reported 98.85% correct face localization using 140 face models. Different analysis is proposed by [Chang *et al.* 2006], where the authors analyzed the position of localized regions and reported 99.4% of successful localization of the eye cavities and the nose tip; the approach was tested on 4485 3D face images. To localize the inner eyes corners clusters, Sun et al. in [Sun & Yin 2008b] used classical decision tree method containing 5 steps. After the eyes regions are localized, the nose tip is found by fitting a reference plane using points from the eyes regions, the point which has the maximum distance to the reference plane became the nose tip. The authors tested the approach on BU-3DFE [Yin *et al.* 2006] database and achieved 92.1% correct face pose estimation using raw data.

Different curvature-based technique to localize expression-free facial points is presented in [Bronstein *et al.* 2005]. The feature points detector first locates the points corresponding to the local extrema of the Mean (H) and Gaussian (K) curvature. Then candidate points are specified: for the nose tip, both H and K must present a local maximum. Next the authors used a set of geometric relations to choose the best set of the candidate points and reported failure rate of the feature detector of less than 1%.

Curvatures were used also by Malassiotis et al. in [Malassiotis & Strintzis 2005], where the authors treated the problem of real-time 3D face head pose estimation. To estimate the pose, the authors looked for a reference point which was the nose tip. The nose tip was localized using principal curvatures k_1 and k_2. After face localization the authors sub-sampled the central part of the face and chose points which principal curvatures fell into the range $k_1 \in [0.1, 0.8], k_2 \in [0.12, 0.25]$. The range was experimentally determined by examining several faces corresponding to more than 20 individuals. The authors show that this procedure results in majority cases in a single point corresponding to the nose tip.

2.2.2.2 Shape Index and Curvedness Index analysis

Another, more complete studies on facial landmarks are presented by Nair et al. [Nair & Cavallaro 2009]. In this article the authors use curvatures, Shape Index, Curvedness and Point Distribution Model for a face detection, facial landmarks localization and registration. In order to characterize the curvature property of each vertex in the face mesh, authors computed two feature maps, namely the Shape Index and the Curvedness. The authors used Shape Index which describes subtle shape variations from concave to convex, providing a continuous scale between salient shapes. Because the Shape Index decomposition does not give an indication of the scale of the curvature at a vertex, the authors additionally calculated Curvedness, which describes how a surface is curved. The authors selected the correct three main points in two steps: in the first step, points which did not belong to any triangle with acute angles were removed. The second step was to validate the correct main three points by fitting a 3D points distribution model, the best fit of the model indicated the main facial points. The authors reported detection accuracy error by absolute mean distance in mm, for the nose tip: 8.83 mm, the left and right inner corner of the eyes, respectively: 12.11 mm and 11.89 mm.

Another interesting article, treating the main facial points localization for 3D face recognition, is the article of Lu et al. [Lu *et al.* 2006]. In this research work, the authors used Shape Index to localize three labeled feature points and to calculate the rigid transformation between facial models. Feature points were selected by determining the local Shape Index at each point within the 2.5D scan. Using a combination of the pose invariant Shape Index, the 3D coordinates and the texture image, the authors developed heuristics to locate a set of candidate feature points. The authors tested the algorithm on 113 facial scans and reported the error between the manually selected points and automatically derived features equals 10mm with a standard deviation of 17mm.

The Shape Index is one of the features which do not change while pose variations [Colbry *et al.* 2005]. This property was used by Colbry et al. in [Colbry *et al.* 2005] to design the fiducial anchor points detection algorithm, which will be rotation invariant. The authors localized the nose tip, which has a Shape Index value close to zero and consistent values around, using an averaging mask in the shape index space. The averaging mask identified the point with the largest average *cup* shaped

curvature. The outer corner of an eye was localized by the authors by following the rut that consistently runs along the bottom of the eye. Having a list of points candidates authors used relaxation algorithm to remove points which are not the main triple, where the relaxation matrix was based on constraints, derived by analyzing example scans. The authors tested the algorithm on 600 scans containing frontal, full profile and expression models. The authors reported the localization rate 99.0% for frontal faces with a neutral expression, 82.7% for profile faces with neutral expression and 75.0% for profile models with a smile expression. The authors also reported that approximately 90% of the scans are below 20mm error.

2.2.2.3 Symmetry plane localization and slice analysis

Human faces are characterized by the symmetry. This property can be used to find a symmetry plane going through the nose ridge and the nose tip.

This characteristic of the human face was used in the article of Faltemier et al. [Faltemier *et al.* 2008b]. The authors used rotated profile signatures for a robust 3D feature points detection. From each 2.5D face scan input the authors created 37 profile signatures, used later for the nose tip verification. To create a single profile signature, a 2.5D face was vertically rotated by a constant $\Delta_\Theta = 5$ in a range between 0° and 180°. At each rotation, the right intersection of the model with the background was taken as a profile signature which was compared later with two manually extracted reference profile signatures (one extracted with 0° pitch face rotation and the second with the 45° pitch face rotation). The authors tested the approach using NDOff2007 [Faltemier *et al.* 2008b] 3D face data set, which contains a total of 6911 non-frontal images and a single frontal neutral image for each of 406 distinct subjects. The authors reported almost 100% accuracy (in 10 mm precision) of the nose tip localization in 0° Yaw rotation and Pitch rotation in the range of $< -90°, 90° >$.

Another type of algorithms is the symmetry plane localization [Tang *et al.* 2008, Wang *et al.* 2008, Wang *et al.* 2009]. The main purpose of the algorithms was to localize the plane going through the nose ridge and the nose tip based on the symmetry assumption. One of the articles [Wang *et al.* 2008] used a posture localization algorithm, which was based on the facial central profile. The central profile is localized using mirrored copy of the model with respect to some plane. After the mirroring process, both the reference M model and the mirrored one M' are registered using Iterative Closest Point algorithm. After the fitting process, points from the model M became mirrored points on the model M'. This correspondence allows to calculate the facial symmetry plane. Since the article treats the face recognition problem, the main points localization is only a part of the whole pipeline and the accuracy of the main point localization is not reported.

Another technique, this time based on horizontal slices, for the nose tip localization is proposed by Mian et al. in [Mian *et al.* 2007, Mian *et al.* 2006, Mian *et al.* 2005], where the authors used the localized nose tip to place and crop the facial region. To localize the nose tip authors sliced a face horizontally at mul-

tiple steps d_v initially using a large value of d_v to improve the speed. After the nose tip was coarsely located the search is repeated in the neighboring region with a smaller value of d_v. Authors report the accuracy of the nose tip detection of 98.3%. Authors also marked that the algorithm assumes that the input data contains a frontal view of a single face with a small pose variation along the x and y axis.

2.2.3 Point signatures, local features

Another type of methods to identify the feature points, are methods which use points signatures. The signature can be understood as a description, propriety or characteristic of a point, which makes it distinguishable from other points.

A point signature has been also used by Xu et al. in [Xu *et al.* 2006], where authors combined local features for a robust nose location. In this article the authors used a cascade points filtering scheme. Combining rules to identify the nose tip position, the authors reported 99.3% of correct detection rate on 3DPEF database [Xu *et al.* 2006].

Another article, which treats feature points localization based on a points signature, is an article written by Conde et al. [Conde *et al.* 2005]. In this article, the authors localized landmark points on a facial 3D surface using Spin Images and an SVM classifier. The algorithm was tested on the FRAV3D data set [Conde *et al.* 2005], which contains different facial expressions and rotations. The authors reported the successful detection rate of the nose tip and the left and right eye inner corners at 98.65%.

A novel and pose invariant 3D shape descriptor was used by Pears in [Pears 2008]. In this article, the authors generated an implicit radial basis function (RBF) model of a facial surface for the nose tip localization. According to the article, the proposed descriptor is less affected by the noise than the Spin Image descriptor. The algorithm was tested on 1121 manually marked face models. The author reported a successful nose tip identification rate of 99.6% without mentioning about the precision of the localization.

2.2.4 Linear Shape and Appearance Models

The face global similarity allows to construct facial statistical models, which are non-linear, generative, parametric models of a certain visual phenomenon [Matthews & Baker 2004]. Those models can enclose any information about the points: position, texture, shape information, neighborhood appearance, movement, etc. Statistical models such as: Active Shape Models (ASM), Active Appearance Models (AAM), 3D Morphable Models (3DMM), Direct Appearance Model (DAM), Statistical Facial feAture Models (SFAM) [Zhao *et al.* 2009b], Active Blobs (AB) are extensively used for face analysis [Nair & Cavallaro 2009] and all of them belong to the closely related *linear shape and appearance models* [Matthews & Baker 2004].

One of the statistical models for the facial anchor points localization is presented by Nair et al. in [Nair & Cavallaro 2009]. In this paper the author proposed

a framework to accurately detect and segment 3D face as well as to localize facial landmarks. The proposed algorithm is based on fitting of a Point Distribution Model (PDM). The fitting was achieved through the transformation between the PDM and candidate vertices, where the candidate vertices were extracted by curvatures analysis. Other landmark points were localized by PDM energy function minimization. The authors reported absolute mean distance for the nose tip 8.83mm while for the inner corners of the eyes around 12mm.

Another article, which makes use of a statistical model is an article of Zhao et al. [Zhao *et al.* 2009b], where the authors applied a statistical model-based technique for 3D face landmarking. The statistical model was made from both: facial geometry and texture. Points around facial landmarks were sampled to create scale-free patches in range and texture. Using PCA algorithm the authors learned 2D shape variations and 2D texture variation of the created scale-free patches. To test the landmarks localization rate, the authors learned the models using half of the FRGCv1 data set and tested learned models on the second part of FRGCv1 data set as well as on randomly selected 1400 models from FRGCv2. The authors reported 100% of all 15 landmarks localizations in 20mm precision and 97% cases in precision 10mm.

Another method, presented by Dibeklioglu et al. [Dibeklioglu *et al.* 2008], used statistical landmarks localization. The local features were extracted using depth map and 7x7 neighborhood around each landmark to create 49-dimensional features. The statistical method, for non-expression points like the nose tip and the inner corner of the eye even in cross-database validation, achieved the accuracy close to 100%. In the case of the outer corner of the eye and the mouth corners the accuracy is respectively 90.09% and 87.16% for the FRGCv1 database, and 95.05%, 67.26% for the Bosphorus v1 database [Savran *et al.* 2008].

2.2.5 Multi-decision

To achieve higher accuracy and trustworthy results a few techniques can be combined. Faltemier et al. in [Faltemier *et al.* 2007, Faltemier *et al.* 2008a] used consensus of three methods: first curvature and Shape Index were used to find nose tip candidates c_n. The next method aligned the input image to a template, using the ICP algorithm, where the position of the nose tip p_n is the highest Z value in the image. If the candidate nose tips was within 20mm distance, the average of p_n and c_n was reported as the final nose tip. The authors tested the approach on ND-2006 database [Faltemier *et al.* 2007] and achieved 98.4% of successful nose tip localization in 20mm precision.

Another multi-decision technique was presented by D'Hose et al. [D'Hose *et al.* 2007]. The authors used Gabor filter decompositions to amplify curvature information. The responses of the two orientations of the Gabor filters were merged to obtain an image integrating both horizontal and vertical curvature information. From these representations of the curvature, the authors extracted the vertical and horizontal lines of the maximum positive curvature.

The intersection between these two sets of lines resulted in a set of the nose tips candidates. The ICP algorithm and an average nose tip shape were used to find the nose tip. The authors tested the algorithm on the FRGCv1 data set and reported the nose tip localization 99.89% with 20mm precision, 99.37% with 10mm precision and 96.22% of eyes corners detection in 10mm precision.

2.2.6 Discussion

Face landmarking is a difficult problem, where the most important factor is the precision. Having many challenges like: face rotation, facial expressions, light variations, occlusions, resolution changes, partial models, numerous techniques were developed, starting from shape analysis through shape signatures, ending with sophisticated statistical models. Each method is designed to fulfil some of the challenges but yet none of them is able to meet all of the difficulties in automatic landmarking. To design real time algorithms, able for example to detect the nose tip, authors used some point signatures [Breitenstein *et al.* 2008, Xu *et al.* 2006]. Rotation invariance, which includes full profiles, can be handled by profile signatures, where the nose tip is a prominent point [Bronstein *et al.* 2005, Faltemier *et al.* 2008b]. Localization of the precise three main facial points (the nose tip and eyes inner corners) as well as the face validation can be done by curvatures analysis and a popular method called HK-classification [Colombo *et al.* 2006, Chang *et al.* 2006]. The precision of all landmark points is mainly achieved by careful statistical models fitting [Zhao *et al.* 2009a, Zhao *et al.* 2009b, Nair & Cavallaro 2007]. Thus a combination of the developed techniques is required to create a reliable landmarking system which will be able to localize landmark points even in adverse conditions (an example of such a system is presented in [Nair & Cavallaro 2007]).

Despite the increasing amount of related literature, current landmarking techniques need to improve both accuracy and robustness, especially in the presence of facial expressions, occlusions and rotation, thus the 3D face landmarking is still an open problem. In this chapter, we propose a general framework for 3D face landmarking which combines local surface properties expressed in terms of curvatures and global facial generic model, used for a face validation.

2.3 Curvature-based Anchor Points Localization on 2.5D Models

Curvatures, are important geometric properties that come from differential geometry and are widely applied in mesh processing. For 3D faces models, curvatures have been used in fields like: face localization [Colombo *et al.* 2006], landmarks localization [Chang *et al.* 2006, Nair & Cavallaro 2009], face segmentation [Chang *et al.* 2006, Nair & Cavallaro 2009], face recognition [Kakadiaris *et al.* 2007], point signatures calculation [Ceron *et al.* 2010], face pose estimation [Sun & Yin 2008b], etc. Curvatures are very popular because they

describe intrinsic as well as extrinsic surface properties. They are insensitive to rotations, simple to calculate and their decomposition over face surfaces is similar from one person to another. Curvatures describe local properties of a surface and beside of the advantages, are also very sensitive to a local surface noise and a change of a model resolution. In the following the curvatures calculation method, as well as our contribution, making the method resolution and noise insensitive, will be described. Next, the contribution in the curvatures calculation method will be examined in case: of facial landmarks localization, points signatures calculation, as well as stability in case of resolution changes.

2.3.1 Pre-processing of 2.5D face models

2.5D face models delivered so far by various scanning methods, are usually corrupted by an impulse noise or holes. To reduce the influence of the face model quality on a face landmarking algorithm, holes and spikes need first to be removed.

The most popular technique to remove an impulse noise (spikes) is the median filter. This method removes the impulse noise, but also tends to affect fine details in many cases. To avoid changes in correct vertices, in our work a decision-based median filtering technique is used. The median filter is applied only on vertices classified as potential spikes after a thresholding operation. This method can efficiently remove all the spikes, without touching properly scanned points.

Once all the spikes are removed, all holes are filled by fitting a mean square surface to a border of a hole. The border of a hole is located by searching for vertices having less than 8 neighbors. In comparison with linear or cubic interpolation, this method takes into account all the directions of the surface changes, which is more convenient for 3D models.

2.3.2 Curvatures calculation method on 2.5D models

Majority of 3D face datasets, recently used by the research community, are scanned using laser based techniques. The laser based scanning technology produces so called range maps (fig. 2.5), where the surface can be considered as a function $z = f(x, y)$ (2.5D models). Using this property of the scanned surface the Mean and Gaussian curvature can be directly computed.

In this section we are going to describe one of the most popular methods for the Mean and Gaussian curvature calculation at a certain point. The method is based on a bi-quadratic surface equation approximation and was used for example in [Colombo *et al.* 2006, Chang *et al.* 2006, Moreno *et al.* 2003, Sun & Yin 2008b]. There also exist different methods like: the Normal Cycle [Cohen-Steiner & Morvan 2003, Morvan 2008], Integral Invariants [Pottmann *et al.* 2007], Statistical Curvatures Estimation [Kalogerakis *et al.* 2007], Tensor Voting [Tong & Tang 2005] and many more [Gatzke & Grimm 2006].

The method for the Mean and Gaussian curvature calculation at a certain point is based on a bi-quadratic surface equation approximation [Colombo *et al.* 2006,

Chang *et al.* 2006, Moreno *et al.* 2003, Sun & Yin 2008b]. The surface equation is approximated using a local neighborhood of a point. Using the first and second derivatives of the approximated equation the Mean and Gaussian curvature can be calculated.

Let S be the surface defined by a C^2 real valued function $f : U^2 \rightarrow R$, defined on an open set $U \subseteq R^2$:

$$S = \{(x, y, z) | f(x, y) = z\}. \tag{2.1}$$

For every point $(x, y, f(x, y)) \in S$ two curvature measures can be computed, the Mean (H) and the Gaussian (K) curvature, which are given by equations [Toponogov & Rovenski 2006]:

$$H = \frac{EN + GL - 2FM}{2(EG - F^2)}, \tag{2.2}$$

$$K = \frac{LN - M^2}{EG - F^2}, \tag{2.3}$$

where E, F, G are coefficients of the first fundamental form of surface, L, M, N are coefficients of the second fundamental form [Trucco & Verri 1998, Toponogov & Rovenski 2006].

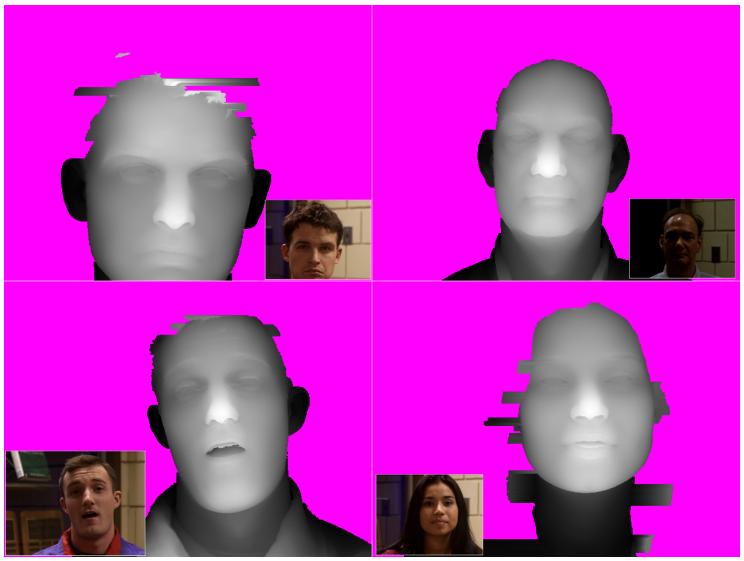

Figure 2.5: Range Maps created using models from the FRGC dataset, models after preprocessing step, removing spikes and filling holes (pink color indicates the background, photos show corresponding texture values).

The coefficients of the fundamental forms are given by following equations

[Toponogov & Rovenski 2006]:

$$\begin{aligned} E &= 1 + f_x^2, \\ F &= f_x f_y, \\ G &= 1 + f_y^2, \\ EG - F^2 &= 1 + f_x^2 + f_y^2, \end{aligned} \tag{2.4}$$

$$\begin{aligned} L &= \frac{f_{xx}}{\sqrt{1+f_x^2+f_y^2}}, \\ M &= \frac{f_{xy}}{\sqrt{1+f_x^2+f_y^2}}, \\ N &= \frac{f_{yy}}{\sqrt{1+f_x^2+f_y^2}}, \end{aligned} \tag{2.5}$$

where $f_x, f_y, f_{xy}, f_{xx}, f_{yy}$ are the first and the second derivatives of a function in a point (x, y).

Since surface representation is discreet, the partial derivatives must be estimated. For each point (x_i, y_i) on the range image we consider a bi-quadratic polynomial equation approximation of the surface:

$$f(x_i, y_i) = \alpha + \beta x + \gamma y + \delta x^2 + \varepsilon xy + \eta y^2, \tag{2.6}$$

where: $\alpha, \beta, \gamma, \delta, \varepsilon, \eta$ are equation coefficients, which can be localized by multiple linear regression algorithm using a local neighborhood of (x_i, y_i). Multiple linear regression attempts to model the relationship between two or more explanatory variables and a response variable by fitting a linear equation to observed data [Bishop 2006]. The system can be further rewritten in the matrix form: $JW = Y$ and the solution is given by: $W = (J^t J)^{-1} J^t Y$.

To create the system of equations at a point (x_i, y_i), a neighborhood needs to be defined. Each point within the neighborhood is used to create one equation, and further used for the estimation of the surface equation coefficients. One popular way to define such a neighborhood around a point (x_i, y_i) consists of using its closest vertices. Meanwhile, this solution is very sensitive to local noise. To solve this problem, noise removal techniques like: median filter, Gaussian filter, etc. were widely used. While those filtering methods exclude noise they also tend to remove fine details and affect the whole model.

The following section 2.3.2.1 describes our contribution into the calculation method, to make it: insensitive to local noise and stable across different model resolutions. The proposed modification incorporates the noise reduction technique into the surface approximation, which will exclude the need of smoothing and removing fine details on the models.

Having estimated the continuous surface equation of a point (x_i, y_i), the partial

derivatives can be computed:

$$\begin{aligned} f_x &= \beta + 2\delta x_i + \varepsilon y_i, \\ f_y &= \gamma + \varepsilon x_i + 2\eta y_i, \\ f_{xx} &= 2\delta, \\ f_{xy} &= \varepsilon, \\ f_{yy} &= 2\eta. \end{aligned} \tag{2.7}$$

Putting all equations together, the Gaussian and the mean curvatures (fig. 2.6) can be calculated using following formulas:

$$\begin{aligned} H(x_i, y_i) &= \frac{(1+f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1+f_x^2)f_{yy}}{2(1+f_x^2+f_y^2)^{\frac{3}{2}}}, \\ K(x_i, y_i) &= \frac{f_{xx}f_{yy} - f_{xy}^2}{(1+f_x^2+f_y^2)^2}, \end{aligned} \tag{2.8}$$

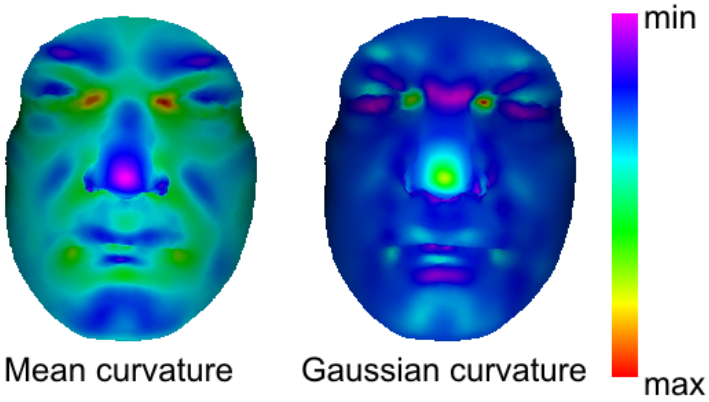

Figure 2.6: Decomposition of the Mean and Gaussian curvatures over facial model, computed using proposed method with 15mm neighborhood used for the surface approximation.

2.3.2.1 Modification of the curvatures calculation method

The curvatures calculation method, presented in the previous section, relies on a surface equation approximation. The method is very exposed to surface noise but also to resolution changes because direct, local point's neighborhood as well as the function derivatives are both sensitive to a noise [Colombo *et al.* 2006] and resolution changes. In this section, we define a geodesic distance-based neighborhood, for the computation of curvatures, which does not require smoothing the model and achieves better robustness in case of resolution changes.

One of the properties of 3D models is a real distance between vertices, which is expressed in [mm]. The distance stays approximately constant between certain

points under rigid deformations of a face as well as resolution changes. This property is no more guaranteed when using range images, where distance unit is undefined, thereby making it difficult to achieve consistent metric between models with a different resolution.

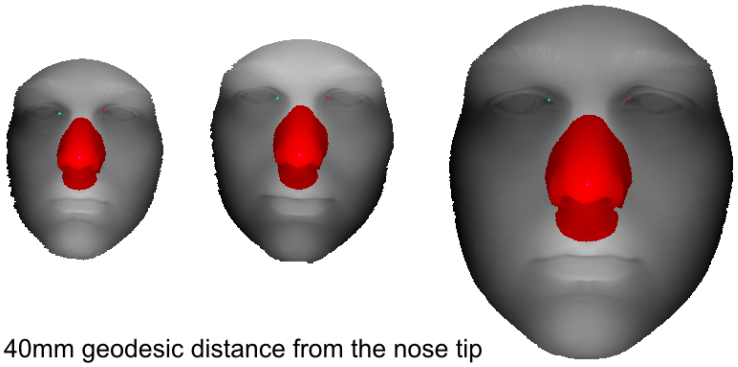

Figure 2.7: Models scanned from different distances from the scanner, resulting in different resolution and size of range images. The red area was marked using 40mm in geodesic distance from the nose tip, calculated on real x, y, z values of the models.

In our work we propose to keep 3D vertex coordinates and add the range coordinates as a UV parameterization for each point. Using real x, y, z values, the [mm] metric will be preserved even when the resolution will change. Using this metric one can find approximately the same area of the neighborhood in different resolutions (see fig. 2.7) because the size of the face will not change. The only difference is in the number of points enclosed in the neighborhood (see fig. 2.8).

To achieve smooth curvatures decomposition, insensitive to noise and resolution changes, we propose to use geodesic distance expressed in [mm] for the definition of a neighborhood instead of the first-neighbors. For each point on a surface the neighborhood will be taken using geodesic distance, which tends to preserve better surface topology than Euclidean one. In certain cases, Euclidean distance will mark incorrect points as those belonging to the neighborhood (see fig. 2.9). It happens due to the fact that Euclidean distance is not taking into account the surface topology and if the surface is rising very fast, for certain points the neighborhood can be incorrect, resulting in incorrect approximated surface. The geodesic distance can be computed using Dijkstra's algorithm, where for a given source vertex, in the graph (mesh), the algorithm finds the path with the lowest cost between the vertex and the other vertex. The cost can be considered as the geodesic distance when the weights between vertices are expressed in terms of Euclidean distance. The result of the curvatures calculation under different neighborhood can be seen on figure 2.10, where the Shape Index decomposition was used. The Shape Index is curvature description derived from the principal curvatures and will be further described in more details as a feature used for face recognition.

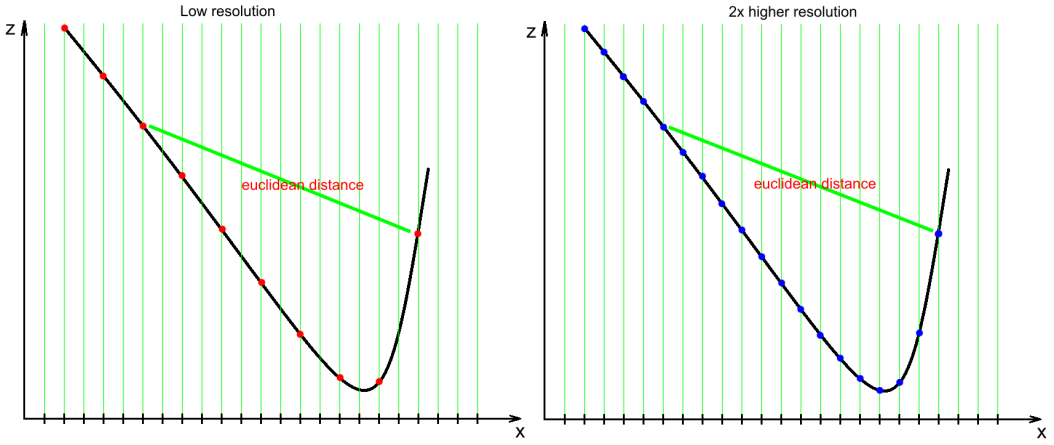

Figure 2.8: Difference in resolution, left image shows lower resolution sampling, right one shows 2x higher sampling. The difference in resolution can happen due to different distance from the scanner. As can be seen, between two certain points the euclidean distance stays approximately the same.

2.3.3 2.5D facial landmarking method

Our algorithm for automatic feature point localization is based on facial curvature analysis and makes use of a coarse-to-fine search strategy, consisting of two steps. At a coarse search step, candidate points for the three most salient facial feature points (nose tip and the two inner eye corners) are first identified based on curvature analysis and generic model fitting. As a fine search step, a generic face model is used to locate other feature points. But first of all, as 2.5D face scans are usually noisy they need to be cleaned up in a preprocessing step. All steps can be seen on figure 2.11.

2.3.3.1 Main points candidates - coarse search

The aim of this coarse search step is to localize feature point candidates on a 2.5D face model. In order to achieve rotation invariant feature points localization, we use Mean and Gaussian curvature analysis and classification [Besl & Jain 1986]. In this step we are targeting the three most salient feature points from the geometric perspective, namely the nose tip and the two inner eye corners. Indeed, the nose tip appears as a convex point on a facial surface while the two inner eye corners are concave points.

The range data of a 2.5D face model is thus first segmented into regions of homogeneous shapes according to HK classification. The HK classification labels each vertex into basic geometric shape classes (tab. 2.2, fig 2.12), using the sign of the mean (eq. 2.2) and the Gaussian (eq. 2.3) curvature [Trucco & Verri 1998] (the calculation method is described in the previous sections).

As it was introduced in section 2.3.2.1, noise sensitiveness can be reduced by

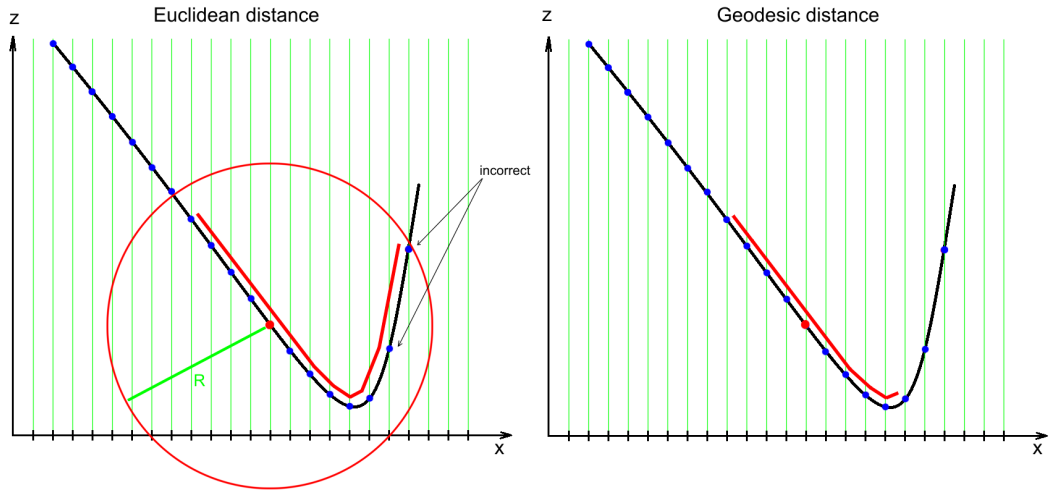

Figure 2.9: Difference in geodesic and euclidean distances. The images show the case when the point's (red point) neighborhood taken by geodesic distance is more appropriate than the neighborhood taken using euclidean distance.

	K < 0	K $= 0$	K > 0
H < 0	Hyperbolic	Cyl. convex	Ellip. convex
H $= 0$	Hyperbolic	Planar	Impossible
H > 0	Hyperbolic	Cyl. concave	Ellip. concave

Table 2.2: HK-Classification [Trucco & Verri 1998].

varying the extent of the neighborhood size. Figure 2.13 shows HK-Classification results with a different neighborhood size considered. As we can see, a large neighborhood used in a surface equation estimation hides the noise and also helps to find most marked out points.

In our experiments we fix the neighborhood size to 25 mm in the geodesic distance, which is more appropriate in such computations because it takes into account the facial surface topology. Recall that the geodesic distance between point A and B is defined as the sum of the Euclidean distances between points in the shortest path between points A and B on the underlying surface.

To localize the three most salient facial feature points candidates, namely the nose tip and the two eye inner corners, the concave and the convex regions of the nose and the eyes respectively are first searched. Each region can be localized according to the HK-Classification. To reduce the number of regions, the Gaussian curvature is thresholded with $K > 0.001$ for the nose and $K > 0.00005$ for the eyes regions (figure 2.14b).

In each located region (figure 2.14b), the most representative point in term of concavity or convexity is then identified. As can be seen in table 2.2, changes in

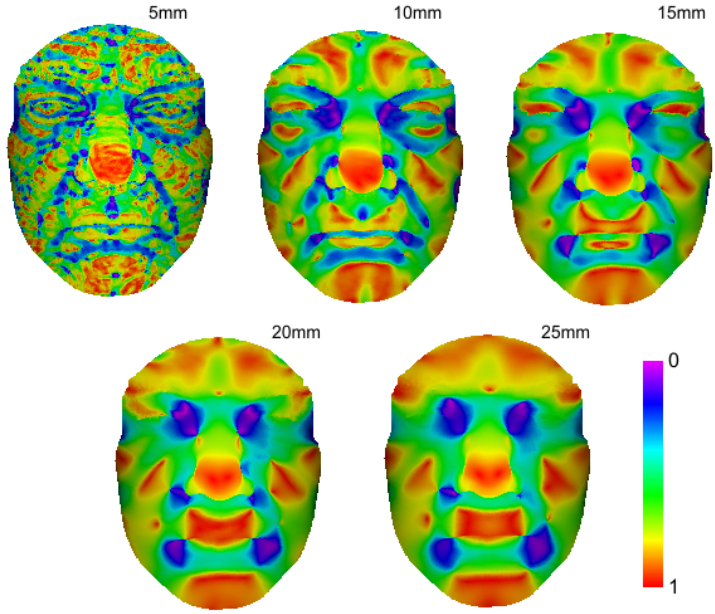

Figure 2.10: Example of Shape Index (normalized shape description, Shape Index scale represents numerous shapes starting from a spherical cup (0 value) ending on spherical cap (1 value)) decomposition, calculated from curvatures estimated using 5, 10, 15, 20, 25 mm in geodesic distance, note that estimation with a small neighborhood used to calculate curvatures is influenced by surface noise.

Gaussian curvature result in shape changes and maximum Gaussian curvature in each region gives the maximum convex or concave point. Such a point will be labelled as a **landmark candidate** in the convex regions for the nose tip and in the concave regions for the eye inner corners.

2.3.3.2 The three main point identification based on a Face Generic Model

To select true main three points from points candidates a selection step needs to be performed. The selection step is based on an error calculation between a generic model in certain position and a face surface.

Our generic face model (figure 2.15) is built from 40 models selected from the IV2 [Petrovska-Delacrétaz *et al.* 2008] dataset. The generic model is composed of 9 main face points whose positions have been calculated based on selected 2.5D facial models. These models were first manually landmarked for 9 feature points. Next, all models were translated and rotated to a frontal position having the nose tip as the origin. The fusion of all models relays on calculation of mean main points positions in the 3D space. The generic model is further normalized so that the distance between the two eye inner corners is 1 mm.

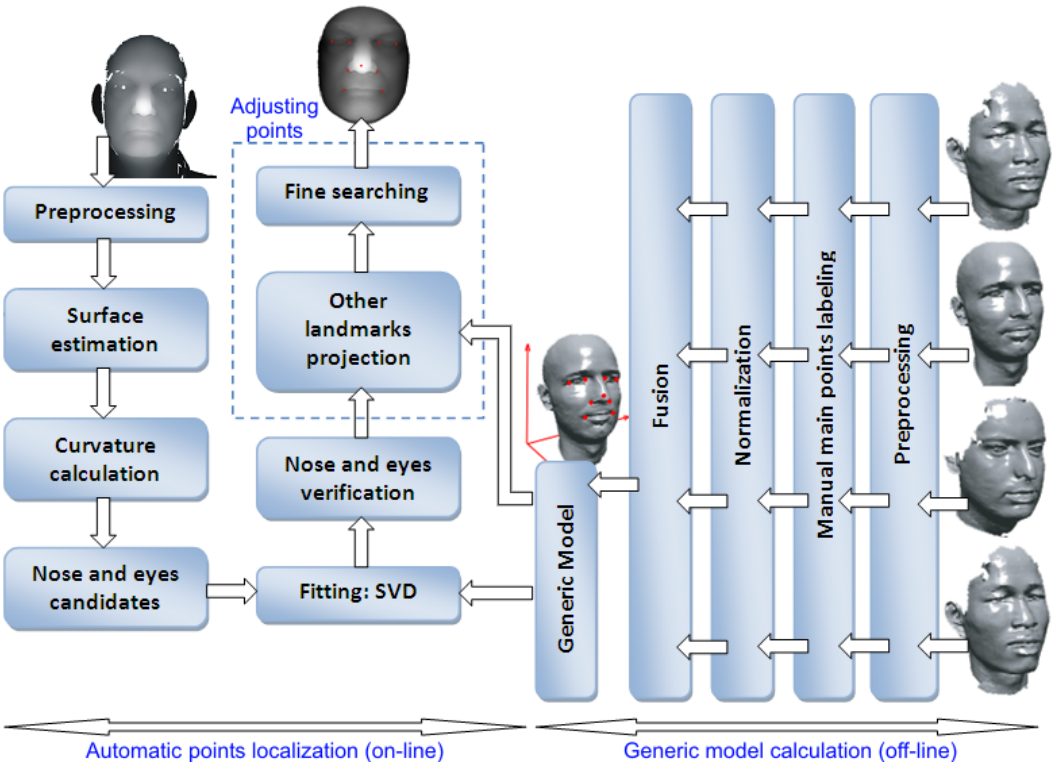

Figure 2.11: Algorithm schema for automatic main points localization.

The calculated generic face model is now used to sort out the true three main feature points (the nose tip and the two inner eye corners) from the set of candidate landmarks, resulting from the curvature analysis. As our 2.5D face models can be in an arbitrary position and we do not have any prior information about it, the basic idea is to consider all combinations of any three landmark candidates (the nose tip candidate and the inner corners of the eyes candidates). To calculate the error between the generic face model and a facial surface, for each combination of landmarks candidates (two points from the eyes candidates and one from the nose candidates are always taken under consideration), we are moving the whole generic face model above the face surface and calculating the sum of distances for all generic face model points to the closest points on the face surface. The movement is based on the rotation and the translation founded by SVD algorithm based on the three main points from the generic face model (red points on fig. 2.15) and landmarks candidates under consideration. Singular Value Decomposition algorithm [Umeyama 1991, Eggert *et al.* 1997, Wen *et al.* 2006] is a matrix decomposition algorithm which has been used iteratively in the ICP (Iterative Closest Point) algorithm. The SVD algorithm allows us to find fine translation and rotation between objects based on the covariance matrix between them. To be invariant to the scale, the generic model has been scaled based on the distance between concave points

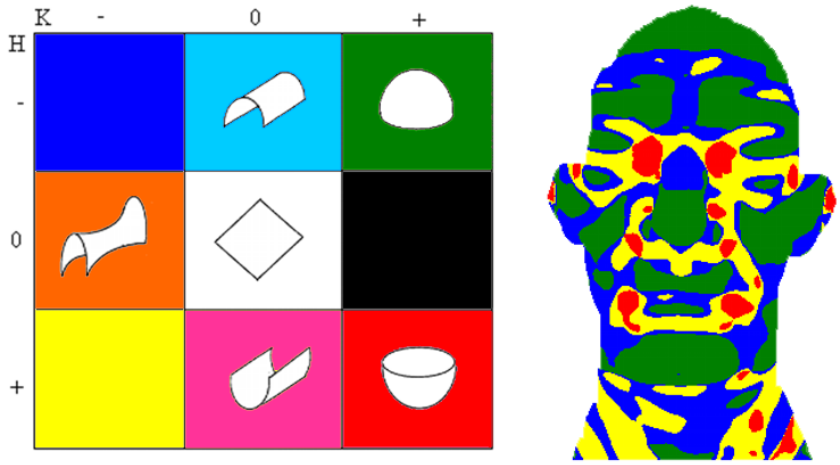

Figure 2.12: HK-Classification dictionary: elliptical concave: red, elliptical convex: green, hyperbolic concave: yellow, hyperbolic convex: blue) [Trucco & Verri 1998]

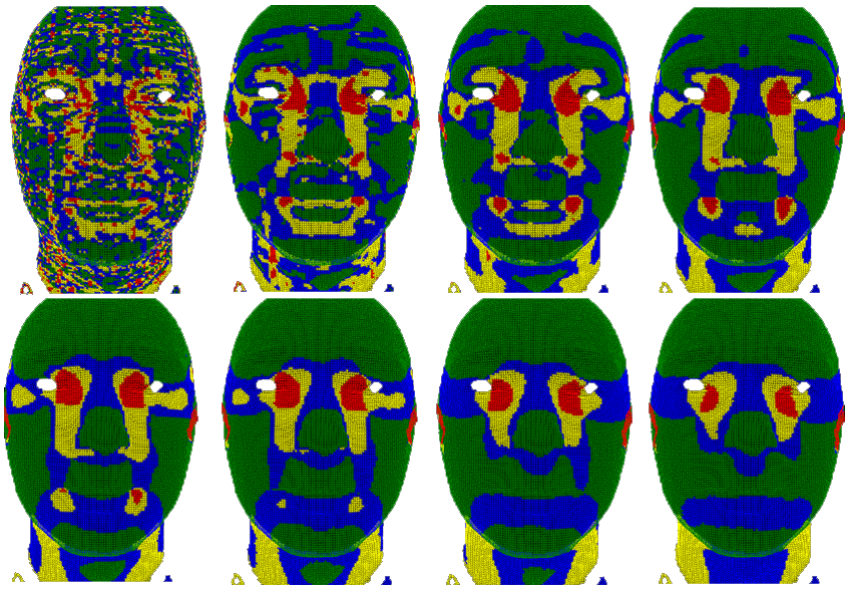

Figure 2.13: The HK-Classification with different neighborhood in the curvature calculation between 5 mm(left top) and 40 mm(right bottom) (elliptical concave: red, elliptical convex: green, hyperbolic concave: yellow, hyperbolic convex: blue).

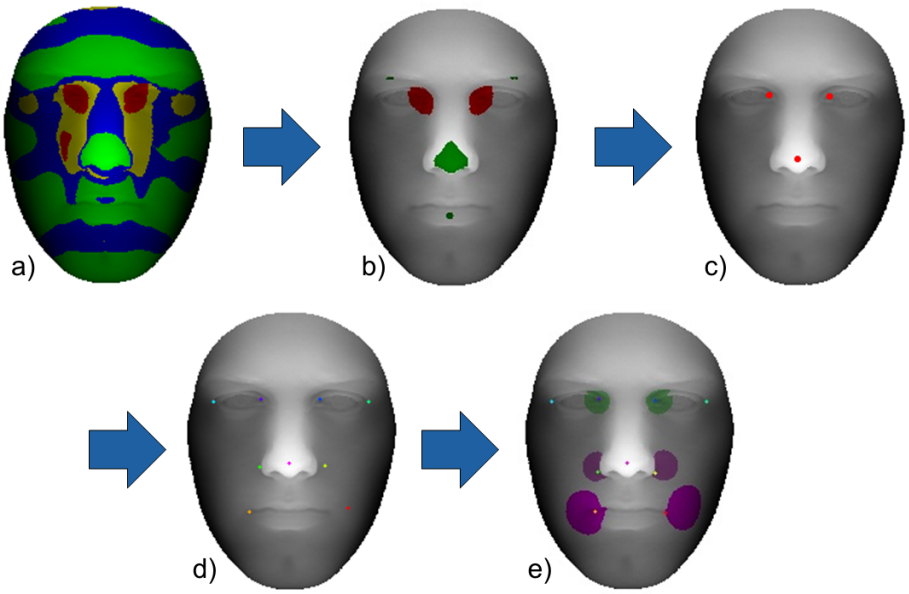

Figure 2.14: Main points localization algorithm: a) HK-Classification, b) nose and eyes regions, c) (coarse localization) the nose tip point and the inner corners of eyes points, d) generic model aligment, e) fine adjusting of points

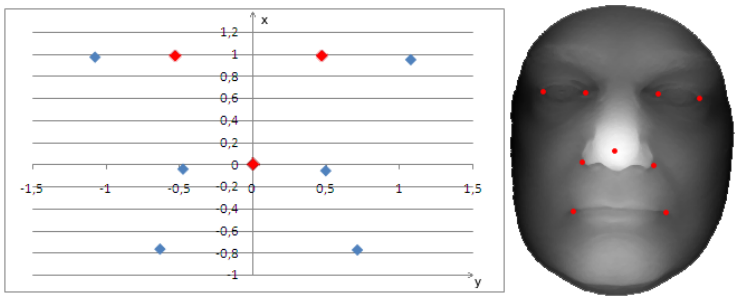

Figure 2.15: Generic model made from 40 models from IV2 data set (x,y projection, red points - main three points - inner corners of the eyes and the nose tip).

(the eyes candidates) under the consideration.

The smallest error between the generic face model and a face surface, under specific position, identifies the true main feature points thanks to the associated manually labelled landmarks on the generic face model (fig. 2.14c).

2.3.3.3 Generic Face Model-based fine analysis

Once the three main feature points have been located, we proceed to localize the other feature points from the generic face model (as showed on fig. 2.15). For this purpose, we project these manually labeled feature points from the generic face model onto a processed 2.5D face model, using the same rotation and translation computed previously by SVD. The closest points on the 2.5D face model to the generic model points will become the landmark candidates and obtain their labels (figure 2.14d).

A coarse-to-fine search strategy is executed by a local search, to deliver the accurate and precise location of the projected feature points. Indeed, the two corners of the lips, the two outer nose corners and the inner eye corners can be characterized as concave points within a certain neighborhood resolution. As we can see in figure 2.13, smaller neighborhood uncovers details on the surface like the lips corners or the nose corners. To localize them precisely, we calculate curvatures using a smaller neighborhood around these points in the surface approximation. In our work, 15 mm neighborhood size is chosen for the lips while 10 mm neighborhood size is chosen for the eyes and the nose, figure 2.14e. The vertex having its maximum Gaussian curvature gives us the most concave point in the concave region and is labelled as the final anchor point.

2.3.4 Experimental results on landmarks localization

In this section the experimental setting and results will be introduced. In order to assess the robustness of our method two face databases were used. From each database the precision of localized points will be expressed as the mean localization error expressed in the 3D Euclidean distance between a feature point automatically located and its corresponding manually labelled landmark.

2.3.4.1 Datasets

In our experiment, we made use of the FRGC ver. 1.0 and ver. 2.0 datasets [Phillips *et al.* 2005] as well as the Bosphorus dataset [Savran *et al.* 2008].

The first version of the FRGC dataset contains 953 face scans from 275 people, captured under controlled illumination conditions, generally neutral expressions and slight head pose and scale variations. The second version of the FRGC dataset contains 4007 face scans from 466 persons. These 2.5D face scans were captured under different illumination conditions and contain various facial expressions, including happiness, surprise, etc.. Since the database does not have manual landmarks, all faces have been manually landmarked by our research group with 15 feature points.

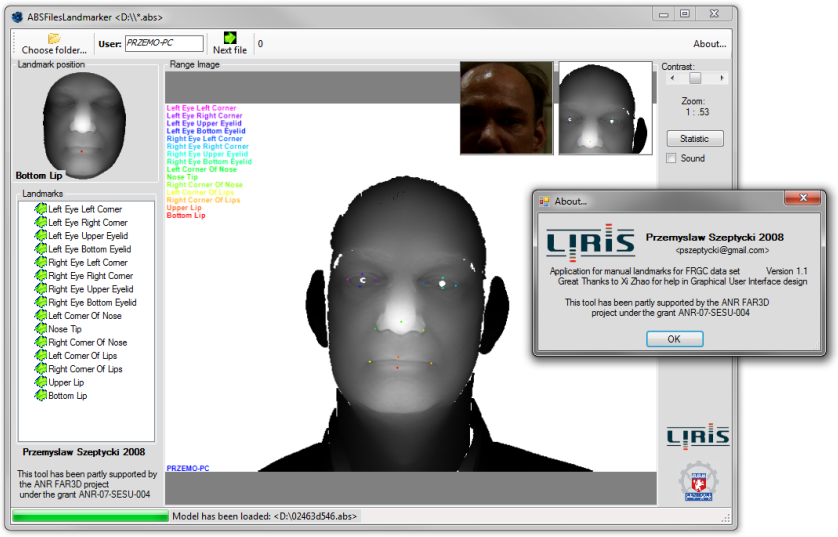

Figure 2.16: Manual main points landmarking application used for landmarking of the FRGC dataset.

As we can see in figure 2.17, these manually labelled anchor points include the eye and lips corners, the nose corners and its tip, and also upper and lower points in the eyelid and lips middle for future investigation. The landmarking application is illustrated in figure 2.16.

The second database used for landmarks evaluation was the Bosphorus dataset [Savran *et al.* 2008]. The Bosphorus Dataset is a dataset of 3D faces which includes a rich set of expressions, systematic variations of poses and different types of occlusions. The dataset contains 105 subjects in various poses, expressions and occlusion conditions with 24 labelled facial landmarks each (figure 2.17). Facial data are acquired using Inspeck Mega Capturor II 3D, which is a commercial structured-light based 3D digitizer device. The sensor resolution in x, y and z (depth) dimensions are 0.3mm, 0.3mm and 0.4mm respectively, and color texture images are high resolution (1600x1200 pixels). The facial expressions are composed of judiciously selected subset of Action Units as well as the six basic emotions, and many actors/actresses are incorporated to obtain more realistic expression data.

2.3.4.2 Experimental results

Our 3D face landmarking solution was benchmarked on a significant subset from FRGC ver. 1.0 and 2.0 as well as on a subset of the Bosphorus dataset. In the first experiment we have chosen 9 prominent feature points from the curvature viewpoint to assess our automatic 3D face landmarking solution. For this purpose, 1618 of models from the FRGC datasets were randomly selected.

While many of these face models have facial expressions (FRGCv2), all of them are in roughly frontal position. In order to test robustness of our solution as com-

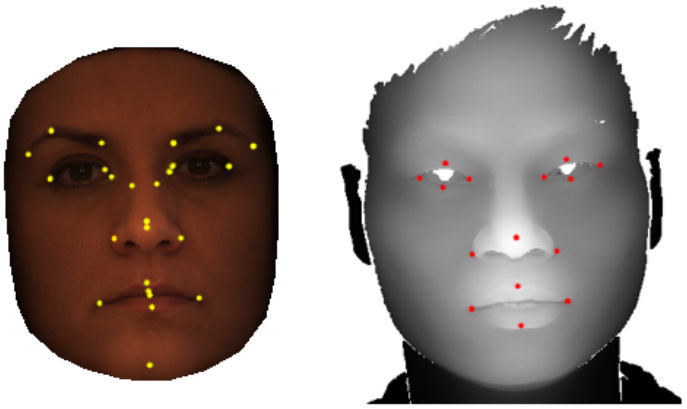

Figure 2.17: Manual landmark points on the Bosphorus (left) and FRGC (right) datasets.

pared to rotation, each selected 3D face model was rotated randomly in yaw (from -90 to 90 degrees), pitch (from -45 to 4 5degrees) and roll (from -30 to 30 degrees) before applying our automatic 3D face landmarking.

Figure 2.18 shows the localization results by our algorithm. Accumulative precision is displayed together with localization accuracy rate given in mm precision.

For comparison purpose, we also applied to the testing set Statistical Facial feAture Model (SFAM) method, developed within our team [Zhao *et al.* 2009b]. The method was applied on 1400 frontal models for the FRGC v2 dataset. However, the SFAM method was designed to achieve numerous facial feature points with high precision, the method is not invariant to facial rotations. The rotation problem in the SFAM is caused by the sampling method used to achieve patterns around each landmark for the learning and testing stages. The method was designed to localize 15 landmark points, since only 9 landmark points can be used for comparison between the two techniques since other facial feature points do not have prominent curvature properties. Cumulative precision curves of the SFAM method are shown in figure 2.19.

To accomplish tests and prove the method independence of the data, the algorithm was also tested on the Bosphorus data set. For this purpose, the first 46 subjects were selected from the Bosphorus data set. Since our landmarking method is distinguished by high precision in the main facial points localization, in this test only the main three points precision was examined (figure 2.20).

The whole results curves can be seen in figure 2.18. The best localization result was achieved for the nose tip, where 100% accuracy was achieved in 8 mm precision while the eye inner corners were localized with a 100% accuracy for left inner eye corner in 12 mm precision and 13 mm precision for the right inner eye corner. Therefore, the inner eye corners were located with more than 99% accuracy in a 10 mm precision. With an accuracy of 88.75% for left lips corner and 87.45% for

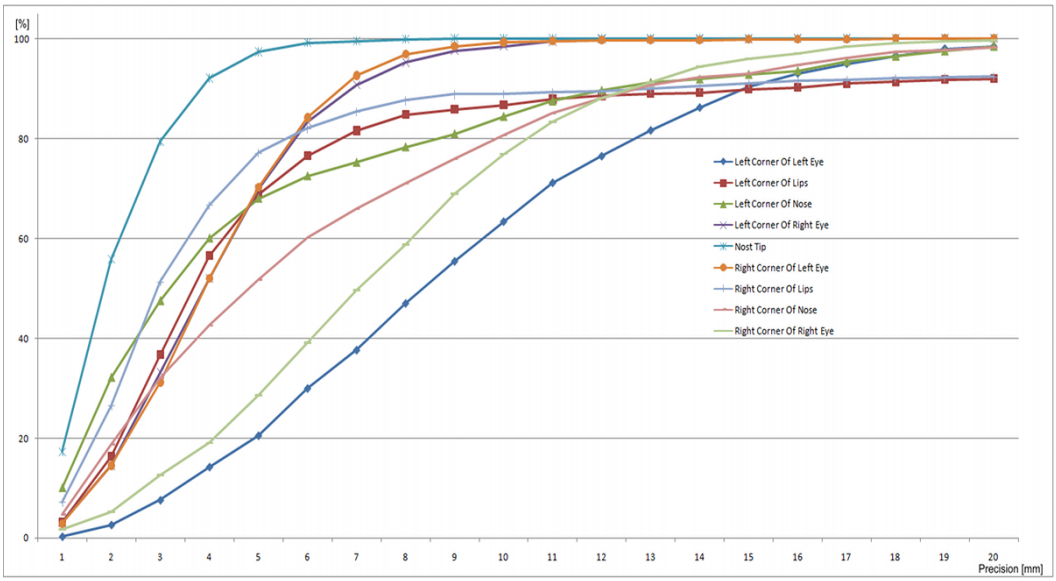

Figure 2.18: Precision curves for the 9 points on the FRGC dataset, localized by our method. Horizontal axis represents landmarks precision while vertical axis represents cumulative error distribution (better seen in color).

right lips corner with 20m precision, our algorithm displays the worst localization results which are mainly caused by facial expressions, mouth movement, etc. Fair localization is achieved for other feature points with respectively 99.62% and 99.87% accuracies for the left and right eye outer corners in 20 mm precision, and 98.2% and 99.0% accuracies for the left and right nose corners.

Figure 2.19 shows the precision curves achieved by the SFAM method [Zhao *et al.* 2009b]. As we can see from the figure, the nose tip and inner corners of eyes, having each prominent geometric feature, are better localized by the curvature analysis-based method. Since the SFAM method takes under the consideration geometrical relationship, local topology and texture of each landmarks, the method displays better precision in all other landmarks being also more complex.

2.3.4.3 Discussion

The major application of 3D face landmarking is 3D face registration and normalization. Thus rotation robustness of a 3D face landmarking solution is important as it relaxes the constraints on the input 2.5D or 3D face model, making 3D-based face processing techniques closer to realistic application requirements. In our work, such a rotation invariance was made possible thanks to the curvature-based analysis and the use of a generic 3D Face Model. As compared to [Faltemier *et al.* 2008b, D'Hose *et al.* 2007, Zhao *et al.* 2009b], our approach achieves higher precision in the main facial points localization while automatically providing up to nine 3D face landmarks.

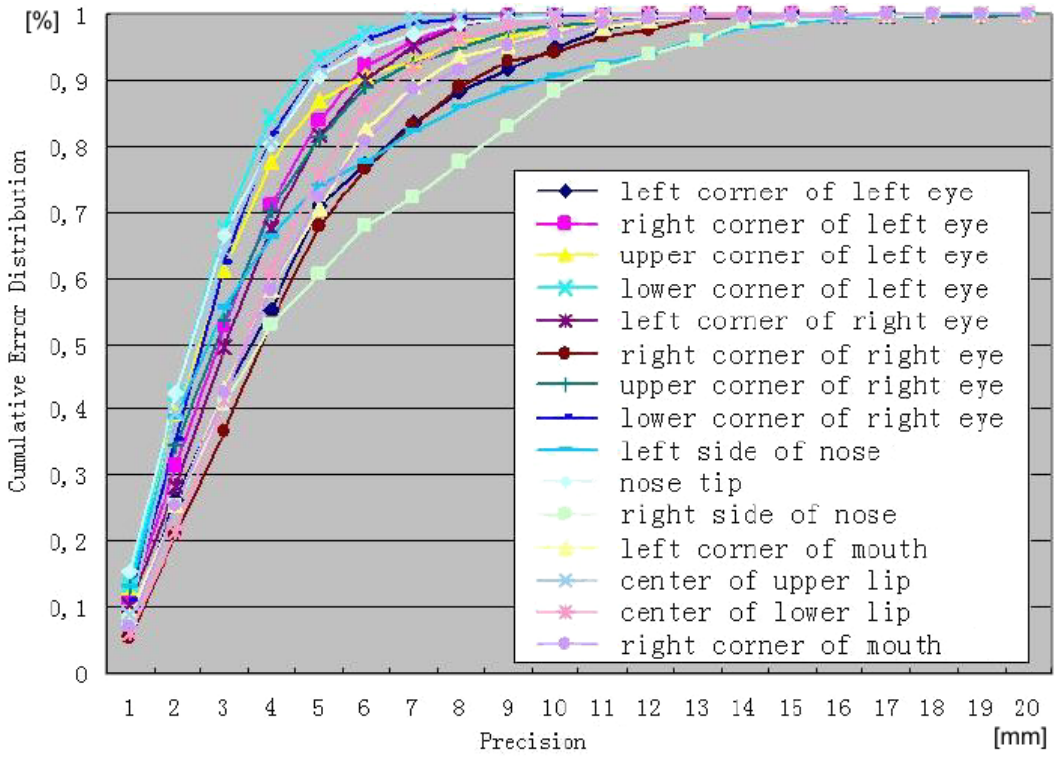

Figure 2.19: Precision curves for 15 manual landmarks localized by [Zhao *et al.* 2009b] on the FRGC dataset.

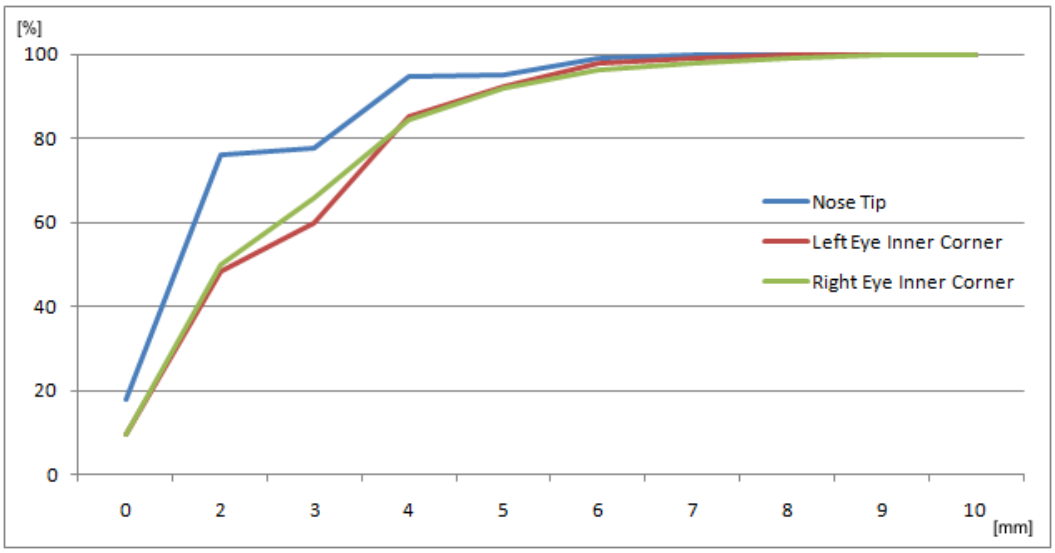

Figure 2.20: Precision curves for the three main points created based on first 46 subjects from the Bosphorus dataset. Points localized by our method. Horizontal axis represents landmarks precision while vertical axis represents cumulative error distribution (better seen in color).

Another analysis using our method, showing the frequency of localized landmarks within each precision interval (displayed on the figure 2.21). As we can see, the majority of localized nose tips are within 0-3 mm precision, the majority of eye inner corners are localized within 2-6 mm precision. The reason, why the inner corners of eyes are localized with smaller precision is the shape of nose and eyes cavities. The maximum Gaussian curvature finds the maximum convexity of concavity points. While the maximum convexity point always denotes the nose tip, the inner corners of eyes are not always characterized by this property. To achieve higher precision in the inner corners of eyes, further analysis of the neighborhood around localized inner corners of the eyes is necessary.

While the method localizes precisely the main three facial landmarks, the other landmarks, without prominent curvature properties are localized with low precision. To localize the other landmarks precisely, more sophisticated analysis of the local neighborhood of each landmark for example based on learning techniques can be more helpful. Therefore, the two methods developed within our team are complementary. The Curvature analysis can be used to localize the three facial main points, the face can be then normalized, subsequently the SFAM method [Zhao *et al.* 2009b] can be used to localize other facial landmarks.

To relate the results to humans' ability to manually localize facial landmarks, ten people were asked to manually label the previously defined anchor points using 10 randomly selected 2.5D models from 10 subjects from the FRGC dataset. Next, mean error and standard deviation were computed for each landmark which

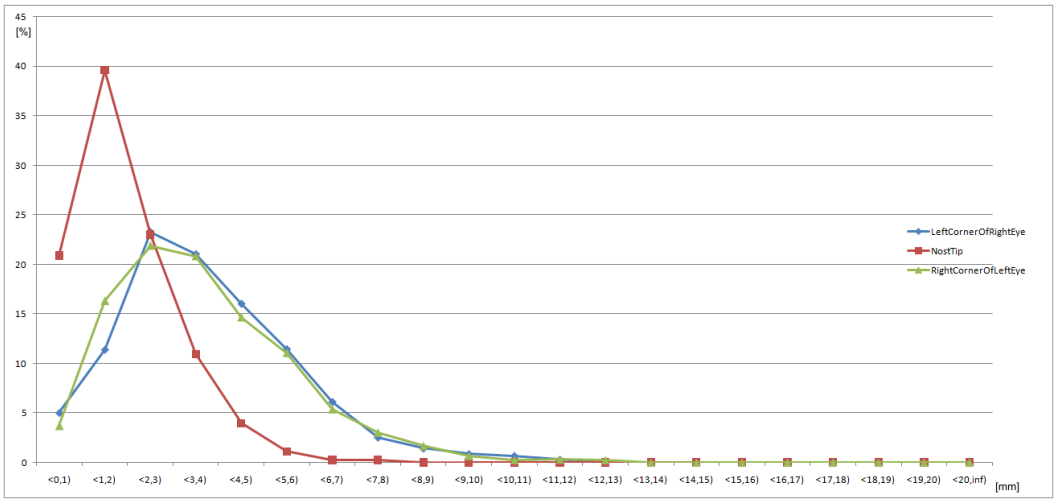

Figure 2.21: Landmarks frequency for the 3 main points on the FRGC dataset. Horizontal axis shows landmarks precision while vertical axis shows localization rate (better seen in color)

is summarized in table 2.3. As we can see from the table, the biggest manual errors as expected were made on landmarks not precisely defined such as right and left corners of nose while nose tip was among the anchor points labeled with the least errors. This experiment shows that each feature point for different person does not locate accurately at the same place, therefore an anchor point on a 3D face model should be considered more as a region than an exact point.

2.3.5 Conclusion on landmarks localization using curvatures decomposition

In this section we have presented our curvature-based method for automatic 2.5D face landmarking. The proposed method consists of curvature calculation with coarse-to-fine analysis. The coarse analysis is based on curvatures calculated using large neighborhood to display only the main patches on the facial surface, thus making the coarse analysis easier. The coarse analysis is performed to localize the main three points on facial surface which later are used as a projection-base for the other points from the Generic Facial Model. The method was tested on more than 1600 2.5D facial models from the FRGC dataset and achieved, from our knowledge, the highest precision for the nose tip and inner corners of the eyes localization among all know nowadays 2.5D methods. The method is insensitive to the lighting conditions as well robust to facial rotations since it uses geometrical surface properties.

Since the other facial points, except the main: nose tip, inner corners of eyes, are not characterized by prominent curvature properties, those cannot be localized accurately using curvatures analysis method. To localize other facial landmarks

Anchor point	Mean error	Standard deviation
Left Eye Left Corner	2.9531	1.4849
Left Eye Right Corner	2.4194	1.0597
Left Eye Upper Eyelid	2.0387	1.3744
Left Eye Bottom Eyelid	1.9424	0.8507
Right Eye Left Corner	2.0473	1.077
Right Eye Right Corner	2.7559	1.5802
Right Eye Upper Eyelid	2.108	1.6449
Right Eye Bottom Eyelid	1.8362	0.8105
Left Corner of Nose	3.8023	1.9839
Nose Tip	1.9014	1.0474
Right Corner of Nose	4.4974	2.1489
Left Corner of Lips	1.9804	1.1045
Right Corner of Lips	1.9891	1.1905
Upper Lip	3.0414	1.5292
Bottom Lip	2.0628	1.3052

Table 2.3: Mean error and Standard Deviation of manual anchor points based on 10 models and 10 samples per model (expressed in mm).

other technique, for example the SFAM technique [Zhao *et al.* 2009b], can be used. Since SFAM method is not invariant to the facial rotations the both methods are complementary. Curvatures analysis can be used for the main points localization and thus for facial surface registration, afterwards the SFAM method can be used for the other landmarks localization.

2.4 Reducing Complexity of Landmarks Localization by a Learning Technique

Nose tip localization is often the basic step for 2.5D face registration and further 3D face processing and as such appears as a side problem of most research work on 2.5D or 3D face recognition. In this section, we carry out a comprehensive study of four popular rotation invariant differential geometric properties, namely Mean and Gaussian curvature, Shape Index and Curvedness, for the purpose of nose tip localization.

Localization of nose tip often appears in the literature as a side work for further 3D faces processing and generally makes use of a priori knowledge of nose shape through convexity/concavity analysis of facial surface [Colombo *et al.* 2006, Nair & Cavallaro 2009, Colbry *et al.* 2005]. However such an analysis generally leads to many candidate points. To identify the true nose tip from the set of candidate points, a mean face model fitting was used in [Nair & Cavallaro 2009], face reconstruction error was considered in [Colombo *et al.* 2006]. While highly

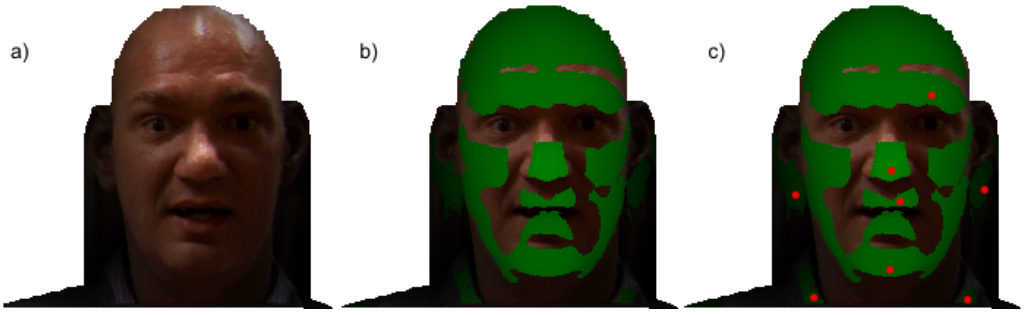

Figure 2.22: The nose tip localization, a) a face range image, b) the nose tip regions (elliptical convex regions), c) the maximum Gaussian curvature points in each nose tip candidate region.

effective, these approaches are computationally expensive and suffer from the exponential number of candidate point configurations. Alternatively nose tip can also be localized using a data-driven approach, by creating some point signatures and thereby enabling fast statistical learning and classification [Breitenstein *et al.* 2008]. Interesting point signatures so far reported in the literature include curvature related descriptors [Ceron *et al.* 2010], Effective Energy, Spin Images, etc. and record detection rates of 99.3% in [Xu *et al.* 2006], 98.65% in [Conde *et al.* 2005] or 99.6% in [Pears 2008].

In this section we also investigate a data-driven approach and study the relevance of using four rotation invariant curvature-based shape descriptors as a point signature for the task of nose tip localization. For this purpose, we studied and compared four differential geometry based point properties, namely Mean and Gaussian curvatures, Shape Index and Curvedness in terms of descriptiveness for the nose tip localization.

2.4.1 Detection of nose tip candidates

The nose tip is one of the most marked out points in the curvature space, with similar curvature characteristics for all faces. This property can be used to localize its candidate points. For this purpose, the input 2.5D face models are first preprocessed in order to remove holes and large spikes. The main steps for locating nose tip candidate points are summarized in figure 2.22. We first locate convex regions thanks to the well known HK-Classification (table 2.2). Within each of the located convex regions, a nose tip candidate is searched as a point having the maximum Gaussian curvature. All these nose tip candidate points are then checked out using an SVM classifier.

Having a 3D facial surface, the localization of local patches can be carried out through a curvature analysis using a threshold [Trucco & Verri 1998, Colombo *et al.* 2006, Chang *et al.* 2006]. While intuitive and simple, thresholding can also lead to loss of the nose tip region which happens in case of a flat nose (for

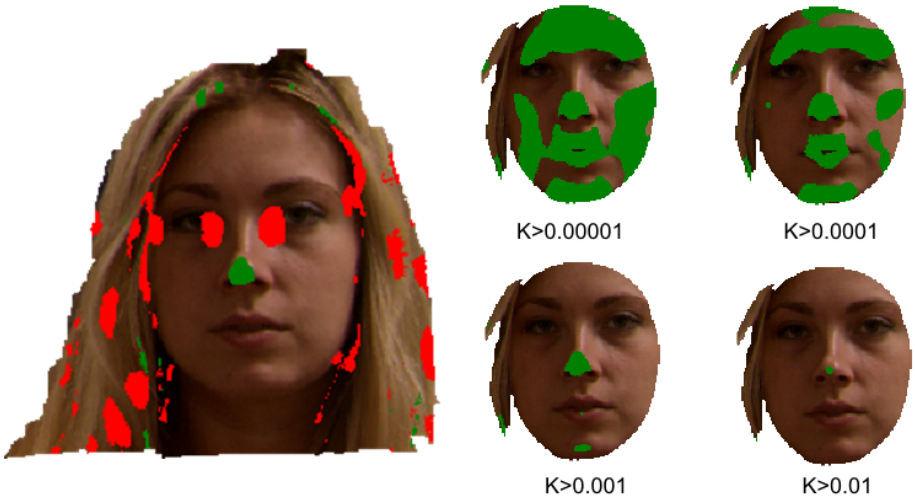

Figure 2.23: Left image: candidate regions for the nose tip (green) and inner corners of the eyes (red); Right image: Vanishing nose tip region after thresholding HK-Classification with different threshold used for the Gaussian curvature.

example Asian people), figure 2.23 shows how the nose tip region vanishes while different thresholds are used.

In this work, we do not make use of curvature thresholding (in section 2.3, the curvature thresholding was used to delimit the number of regions under consideration). Instead, we look for all convex regions thanks to the HK-classification (figure 2.12) which permits to partition a range data into regions of homogeneous shapes, called homogeneous surface patches based on the signs of Mean (H) and Gaussian (K) curvatures [Trucco & Verri 1998]. As the approximation of the curvatures over a range data generally is sensitive to noises [Colombo *et al.* 2006, Chang *et al.* 2006], we used the method described in section 2.3.2 which incorporates the Gaussian noise rejection in the curvature calculation method.

In this work, as illustrated in figure 2.22b, all the elliptical convex patches, thus with $H < 0$ and $K > 0$, are considered as the nose tip candidate regions. In each of these regions, the point having the maximum Gaussian curvature is then selected and considered as the nose tip candidate point (figure 2.22c) for which we compute a point signature which is further checked by the SVM classifier. Decomposition of the Gaussian curvature can be seen on figure 2.24, where the nose tip is clearly seen.

2.4.2 Point signatures

Once the set of the candidate points for the nose tip is collected, as described in the previous section, the SVM is invoked to further identify the true nose tip. For this purpose, we simply compute a point signature which gathers several state-of-the-art curvature-based descriptors [Colombo *et al.* 2006, Ceron *et al.* 2010], namely the

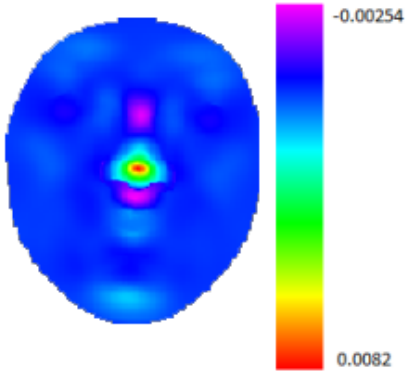

Figure 2.24: Gaussian curvature decomposition on 3D face model, calculated with 25 mm neighborhood in surface approximation, showing the appearance of the nose tip (max value - red, min value - pink).

Mean, Gaussian Curvatures, the Shape Index, Curvedness and study their discriminative power for the task of the nose tip localization when they are used alone or in combinations.

Curvatures for each nose tip candidate point are computed by a bi-quadratic surface equation approximation using a neighborhood of 25mm (please refer to section 2.3.2, where the curvatures calculation method is described), which proves to be a reasonable compromise in removing noise while keeping the rough shape of the local region.

2.4.2.1 Principal curvatures

Principal curvatures $k_1(p)$ and $k_2(p)$ (figure 2.25), can be extracted from equations [Trucco & Verri 1998]:

$$k_1(p) = H(p) + \sqrt{H(p)^2 - K(p)}, \tag{2.9}$$

$$k_2(p) = H(p) - \sqrt{H(p)^2 - K(p)}, \tag{2.10}$$

where H and K are Gaussian and mean curvature at point p. The Principal Curvatures measure, how the surface bends by different amounts in different directions at that point.

2.4.2.2 Shape Index

Shape Index (figure 2.26) is a normalized curvature representation in a certain point of a surface within 2.5D image, proposed for example by Dorai and Jain in 1995 [Dorai & Jain 1995]. The shape index (SI) of the face image is the quantitative measure of the shape of a surface at a point. This local curvature information about a point is independent of the coordinate system. The Shape index

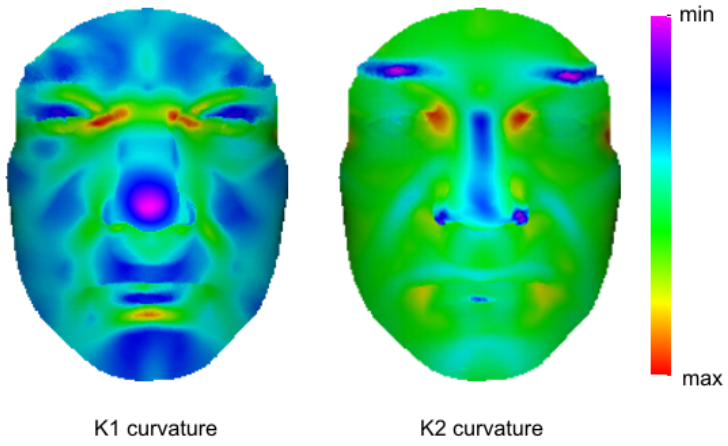

Figure 2.25: Principal Curvatures decomposed on facial model.

S at point p is calculated using maximum (k_1) and minimum (k_2) local curvature [Nair & Cavallaro 2009]:

$$S(p) = \frac{1}{2} - \frac{1}{\pi} \operatorname{tg}^{-1} \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)}, \tag{2.11}$$

where $k_1(p)$ and $k_2(p)$ are principal curvatures at point p of the surface S.

2.4.2.3 Curvedness

The curvedness (figure 2.27) of a surface at a vertex (v) can be calculated based on equation [Nair & Cavallaro 2009]:

$$Curv(v) = \frac{\sqrt{k_1^2(v) + k_2^2(v)}}{2}, \tag{2.12}$$

where $k_1(p)$ and $k_2(p)$ are principal curvatures at vertex v of the surface S.

It describes how gently curved a surface is. The dimension of curvedness is the reciprocal of length and ranges from negative infinity to positive infinity [Yoshida *et al.* 2002].

2.4.3 Support Vector Machine for the nose tip localization

The experiments were carried out on the FRGC v2.0 dataset [Phillips *et al.* 2005] using an SVM classifier with a standard Gaussian kernel function. We manually labelled all the face scans from FRGC v2.0 with 15 landmarks, including the nose tip and eye corners (described in section 2.3.4.1).

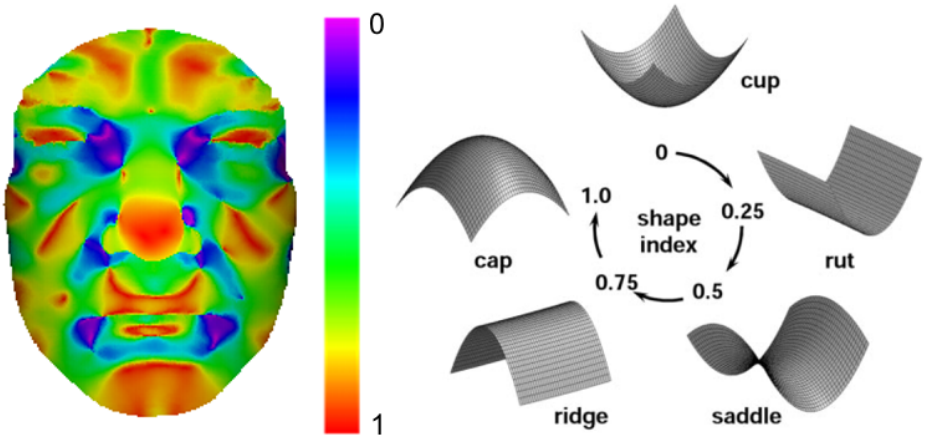

Figure 2.26: Shape Index decomposition on face model and the dictionary of its values [Yoshida *et al.* 2002].

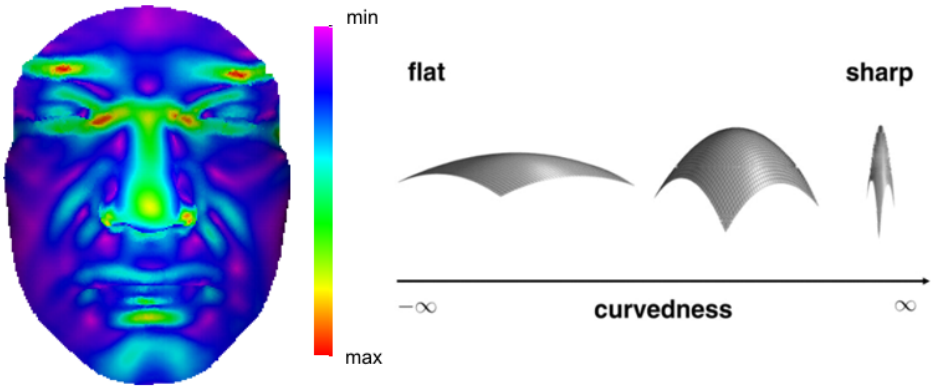

Figure 2.27: Curvedness, represents how gently curved a surface is, image source:[Yoshida *et al.* 2002].

2.4.3.1 SVM Training

Support Vector Machine (SVM) was used to identify the true nose tip from the set of nose tip candidate points. For this purpose, we made use of the publicly available Matlab LIBSVM[2] library.

For training, the manually marked nose tips were used as the positive samples while negative samples were randomly selected with proportion 1:1. For each training sample, the four previously described curvatures: Mean and the Gaussian curvature, Shape Index and Curvedness, were first calculated. They were fed alone or in combination to the classifier, giving birth to 15 configurations regarding the point signature (see table 2.4). To verify the localization rate by the SVM, we also varied the learning dataset size, ranging from 100 to 3000 points (figure 2.24).

Each configuration of the system (different point signature + training set size) was cross-validated 10 times, each time with different folds and reported results are the mean values of all tests. For each cross validation the training set was selected randomly while the remaining models formed the test set.

2.4.4 Experimental results

Once learned, the SVM is used for identifying the true nose tip among all the nose tip candidates. If the SVM classifies a nose tip candidate point as a correct one, this point is further checked with the manually labelled one. If the distance between the predicted nose tip and the manually marked one is less than 10mm, the True Acceptance Rate is increased.

Almost all configurations of the point signatures achieved very high True Acceptance Rate (TAR), close to 100%. The highest TAR with the lowest False Acceptance Rate (FAR) was achieved using together the Mean curvature, the Shape Index and the Curvedness where the TAR was 99.9% and the FAR was 6.7%. The lowest localization rate was achieved by single Gaussian Curvature where the TAR was 30.7% and the FAR 3.3%.

Based on the figure 2.28, one can observe that in case of increasing number of the training models the True Acceptance Rate as well as the False Rejection Rate are stable, respectively in a range of 99.9% and 0.1%. The change in the number (from 100 to 3000) of the training examples affects the True Rejection Rate and the False Acceptance Rate, respectively the TRR changed from 85.3% to 93.29% and the FAR changed from 14.7% to 6.71%.

2.4.5 Discussion

Looking at table 2.4 and figure 2.29, one can see that the True Acceptance Rate of the nose tip points is very high for almost all curvature based features, except the Gaussian Curvature which displays a 30.74% TAR for the nose tip. Meanwhile, this result is in line with the conclusion of the state of the art, for instance the one by

[2] http://www.csie.ntu.edu.tw/c̃jlin/libsvm/

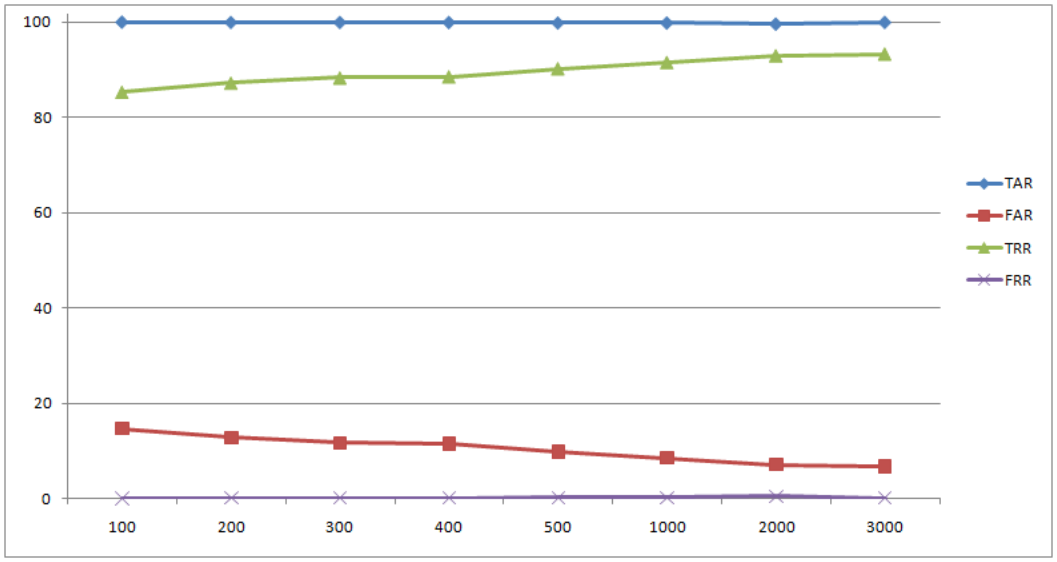

Figure 2.28: Results of the best verified combinations of the curvature properties (Mean curvature, Shape Index and Curvedness), horizontal axis represents number of models used for SVM learning, the vertical axis represents the percentage of TAR, FAR, TRR, FRR. (better seen in color)

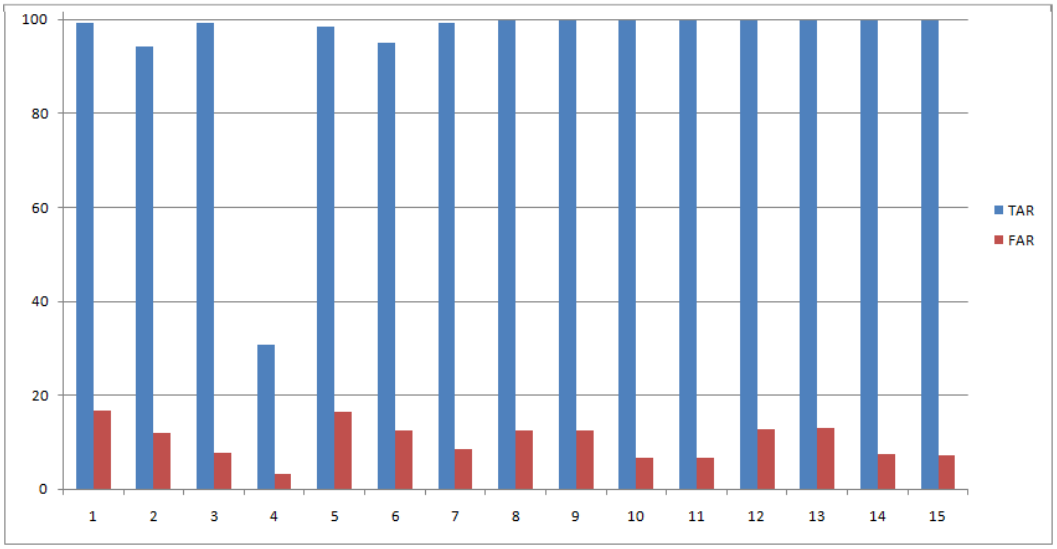

Figure 2.29: The TAR and FAR for each combination of the point properties for the shape signature, with 3000 models taken for the SVM learning. (C - Curvedness, SI - Shape Index, H - Mean Curvature, K - Gaussian Curvature) 1)C, 2)SI, 3)SI+C, 4)K, 5)K+C, 6)K+SI, 7)K+SI+C, 8)H, 9)H+C, 10)H+SI, 11)H+SI+C, 12)H+K, 13)H+K+C, 14)H+K+SI, 15)H+K+SI+C. (better seen in color)

Point Signature combination	TAR [%]	FAR [%]
C	99.39	16.72
SI	94.14	11.98
SI+C	99.19	7.87
K	30.74	3.35
K+C	98.58	16.45
K+SI	95.05	12.45
K+SI+C	99.19	8.52
H	99.90	12.49
H+C	99.90	12.42
H+SI	**99.70**	**6.70**
H+SI+C	**99.90**	**6.71**
H+K	99.70	12.88
H+K+C	99.70	12.94
H+K+SI	99.80	7.59
H+K+SI+C	99.90	7.15

Table 2.4: The obtained results (mean value of 10 cross-validations) for different point properties combinations with the fixed number of the training models (3000), (C - Curvedness, SI - Shape Index, H - Mean Curvature, K - Gaussian Curvature)

Ceron et al. [Ceron *et al.* 2010] who stated that the Gaussian Curvature has the worst descriptiveness for discriminating facial feature points.

When studying the False Acceptance Rates (FAR) of all these curvature features in table 2.4, one can see that the lowest FAR 6.7%, was achieved by the combination of the Shape Index, Curvedness and the Mean Curvature which also displays the highest 99.9% TAR. Very similar results were obtained from the configuration using both the Shape Index and Mean Curvature, displaying a 99.7% TAR and a 6.7% FAR respectively.

The False Acceptance Rate at the level of 6.7% means that the SVM classifier incorrectly classified non-nose tip points from the candidate set as the correct nose tips. Those mistakes generally happen when the point candidates have shapes similar to the nose, which occurs sometimes on cloths or hair regions. To exclude these incorrectly classified points one needs to consider further constraints or knowledge. The method guaranties very high True Acceptance Rate (99.9%), which gives reliability that the nose tip will not be missed.

2.4.6 Conclusion on landmarks localization by SVM algorithm

In this section, we have studied the relevance of four rotation invariant differential geometric properties, namely Mean, Gaussian curvatures, Shape Index and Curvedness, for the purpose of nose tip localization. For each 2.5D facial model, the set of nose tip candidate points was first automatically selected from a shape classification using a priori knowledge of nose shape. An SVM classifier trained on a subset of the

data set using the curvature features alone or with combinations was then invoked to select the true nose tip from the set of candidate points. Extensive experiments were carried out on the whole FRGC v2.0 dataset by varying the training dataset size. For each training size 10 times cross-validation was used with a random selection of the training set. The experimental results were evaluated in terms of the True Acceptance Rate (TAR) and the False Acceptance Rate (FAR) using manually labeled nose tip point as ground truth and expressed as a mean value of all cross-validation tests. 99.9% Nose Tip TAR with 6.7% FAR was achieved on the FRGCv2 dataset when the Mean curvature, the Shape Index along with the Curvedness was used as the input to the trained SVM classifier.

2.5 Curvatures Stability Across Different Model's Resolutions

Curvatures are used in many applications: face segmentation, facial landmarking, face recognition, point signatures calculation etc. While the curvatures are popular because of their rotation invariance, consistent repeatability over facial models and the simplicity to calculate, its most popular calculation method is sensitive to the noise and the model resolution changes.

In this section the stability of the modified curvature calculation method will be investigated, while changing the model's resolution and modifying the noise removal methods.

2.5.1 Invariance of the resolution

The most popular method for the Mean and Gaussian curvature calculation at a certain point is based on a bi-quadratic surface equation approximation [Colombo *et al.* 2006, Chang *et al.* 2006, Moreno *et al.* 2003, Sun & Yin 2008b, Trucco & Verri 1998, Toponogov & Rovenski 2006]. The method uses a local neighborhood of a point and the first and second derivatives of the approximated surface equation to estimate the Mean and Gaussian curvature. Local characteristic of the method exposes it to the surface noise and resolution changes. To overcome the problem of the surface noise authors usually use noise reduction techniques [Colombo *et al.* 2006, Chang *et al.* 2006, Moreno *et al.* 2003, Sun & Yin 2008b]. The noise is usually removed by smoothing filters, which also tend to remove fine details. Since filtering techniques relay on point's neighborhood, the strength of the filtering technique will be affected by model's resolution. Another source of errors in curvature calculation is the fact that the algorithm is also based on the point's neighborhood. While the scanner's resolution is constant, the model's resolution can change in case of the distance from the scanner, therefore leading to changes of the curvature values.

To achieve smooth curvatures decomposition, insensitive to resolution changes and noise, we propose to use geodesic distance expressed in [mm] for the definition

of a neighborhood instead of the first-neighbors. For each point on a surface the neighborhood will be taken using geodesic distance, which tends to preserve better surface topology than Euclidean one. The curvature calculation method is described in section 2.3.2.

We tested the proposed method for the computation of curvatures using models with different resolutions from the FRGC data set [Phillips *et al.* 2005]. To ascertain the stability of the curvature calculation, the curvatures were calculated and compared using an original model and several other models having different resolutions. We also tested the sensitivity of the proposed method with two different settings regarding noise: without any preprocessing or with preprocessing to remove noise.

2.5.1.1 Resolution change algorithm

To verify the stability of the curvature values in case of different resolution, the nose tip point was selected and its curvature values were examined. This point was selected as the one of the most salient points in the curvatures space, where its values of curvatures should be constant across different resolution. Instead of acquiring new models with lower resolution, the input model's resolution was changed, to assure that the position of the selected point will not change. Figure 2.30 shows four models with different resolutions, changed using a grid with different intervals between points.

The resolution change algorithm is based on a grid crossing 2.5D face model (fig. 2.30). The intersections of the grid lines create new points on the lower resolution face model, while the depth values are estimated based on the neighboring points from the model under change. The z value (depth), is calculated as a mean value of all z coordinates of points which belong to the neighborhood of the grid point.

By changing the intervals d between the grid lines, one can change the resolution of the model. If the neighborhood of the grid point contains the landmark point, the new created point becomes a new landmark on the lower resolution model and inherits the label.

2.5.2 Experimental results on curvature stability

Since the previous algorithms (based on first neighbors) for the curvature calculation need a preprocessing method, which will remove all spikes and smooth out the surface, additionally in our testing settings we have used noise rejection filters.

In our work we have tested the filters with different kernel matrix sizes as well as different types of the neighborhood used for the surface equation approximation (including first neighbors neighborhood and the neighborhood based on [mm] metric).

First tests were conducted using different neighborhood types. To verify the curvature stability, one point (the nose tip) was selected over all facial models (the most prominent in curvatures space). The reported values are the differences in the

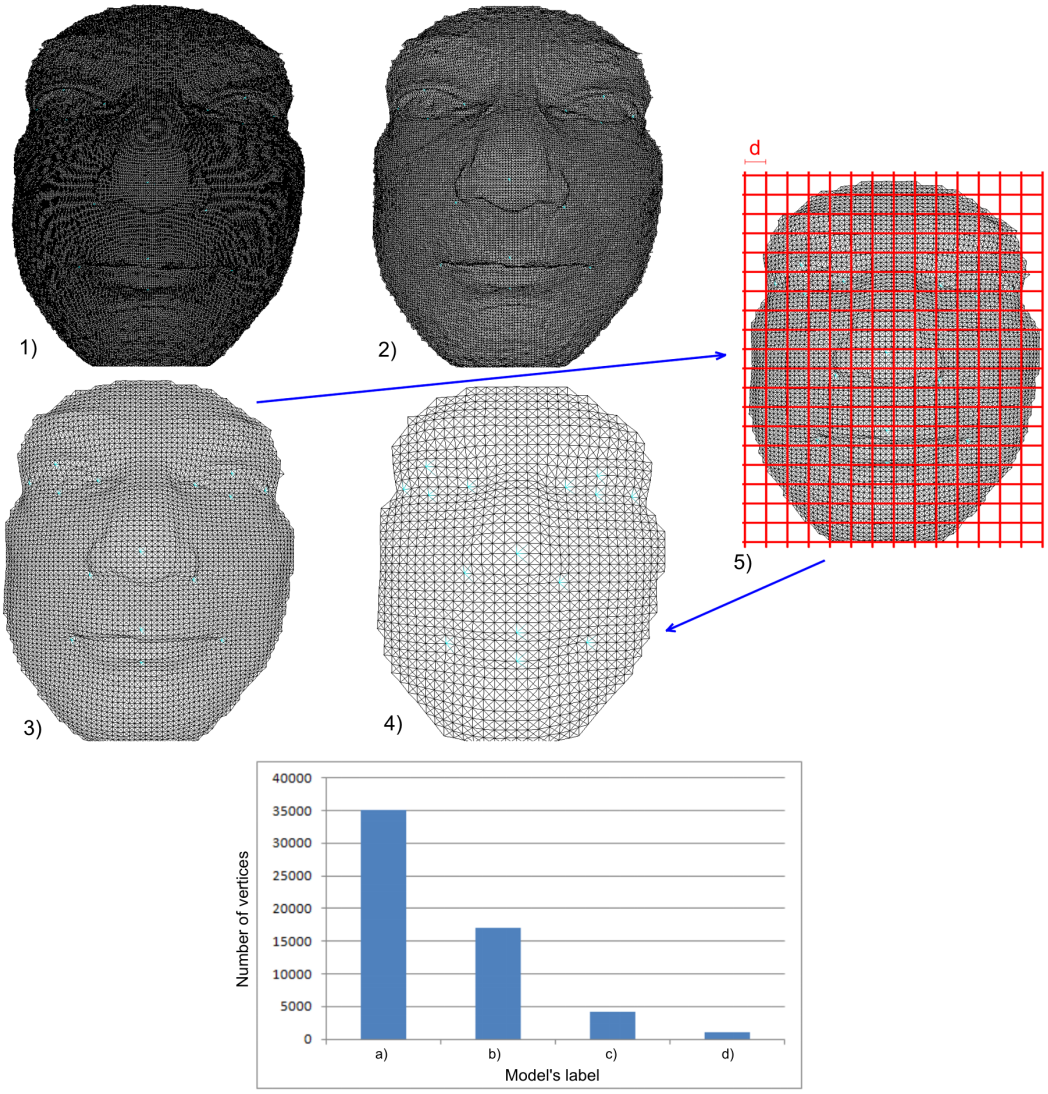

Figure 2.30: Example of different face model's resolution, 1) 35024 vertices, 2) 17015 vertices, 3) 4267 vertices, 4) 1068 vertices, 5) shows the grid used for resolution changes, by changing the intervals (d) between grid lines the resolution of the model can be changed.

values of the Mean and Gaussian curvature (please see fig. 2.31 and fig. 2.30) which correspond to the model's resolution.

Since the difference in the curvature values for the nose tip across different resolutions is the smallest using neighborhood within 15 mm of geodesic distance, for the next test we have compared it with different preprocessing algorithms. For the preprocessing algorithm, Gaussian filter was used with different sizes of the mask. Figure 2.32 shows the differences in curvatures between different resolutions and Gaussian filter sizes as well as different neighborhood types. In this test the 1-neighbor neighborhood as well as 15 mm neighborhood were compared.

The final, designed test was performed to verify the curvatures stability over the whole face. For this test, the high resolution models A were scaled to a lower resolution B by the scaling algorithm, presented in section 2.5.1.1. The scaling algorithm is able to obtain the dense correspondence between models A and B. To verify the curvature stability over the whole face, curvatures were calculated on both models (A and B). Using the correspondence, the difference over the whole model was calculated. Figure 2.33 shows the difference of the curvature values featured as a color decomposition on a few models obtained from the FRGC data set.

2.5.2.1 Discussion

Looking at figure 2.31 one can see that the most stable curvatures can be obtained using 15 mm in the neighborhood for the surface approximation. Still the biggest difference can be observed while comparing the highest resolution model with the lowest one. In case of higher resolutions the 10 mm neighborhood gives the best results, yet it has the biggest difference in case of a very low resolution. What should be also observed is the fact that the Mean curvature characterizes bigger difference than the Gaussian curvature. Looking further at the figure, one can note that using simple one-neighbor neighborhood for the surface approximation is affected by the resolution change. The difference between high and low resolution models is very high, which will affect other algorithms, for example thresholding HK-Classification. To summarize, the most stable curvatures can be obtained using 15 mm neighborhood for surface equation approximation.

Having chosen the best neighborhood for the surface approximation, later used to calculate the curvatures, the next test was to examine the influence of the filtering technique to the curvature stability. Looking closer at the figure 2.32 one can notice that the usage of filtering technique gives worse results in case of the curvatures stability using 15 mm neighborhood. This is due to the fact that the filter was working on range image and was smoothing the surface in different scale for different resolution. What should be noted is also the fact that while the filtering technique does not help in case of neighborhood expressed in mm, it helps in case of neighborhood expressed as a first-neighbor. Obtaining smoother surface helps to achieve more constant curvature values yet the stability is far away from the results obtained using [mm] metric.

The last test, figure 2.33, was carried out to verify the difference in curvatures

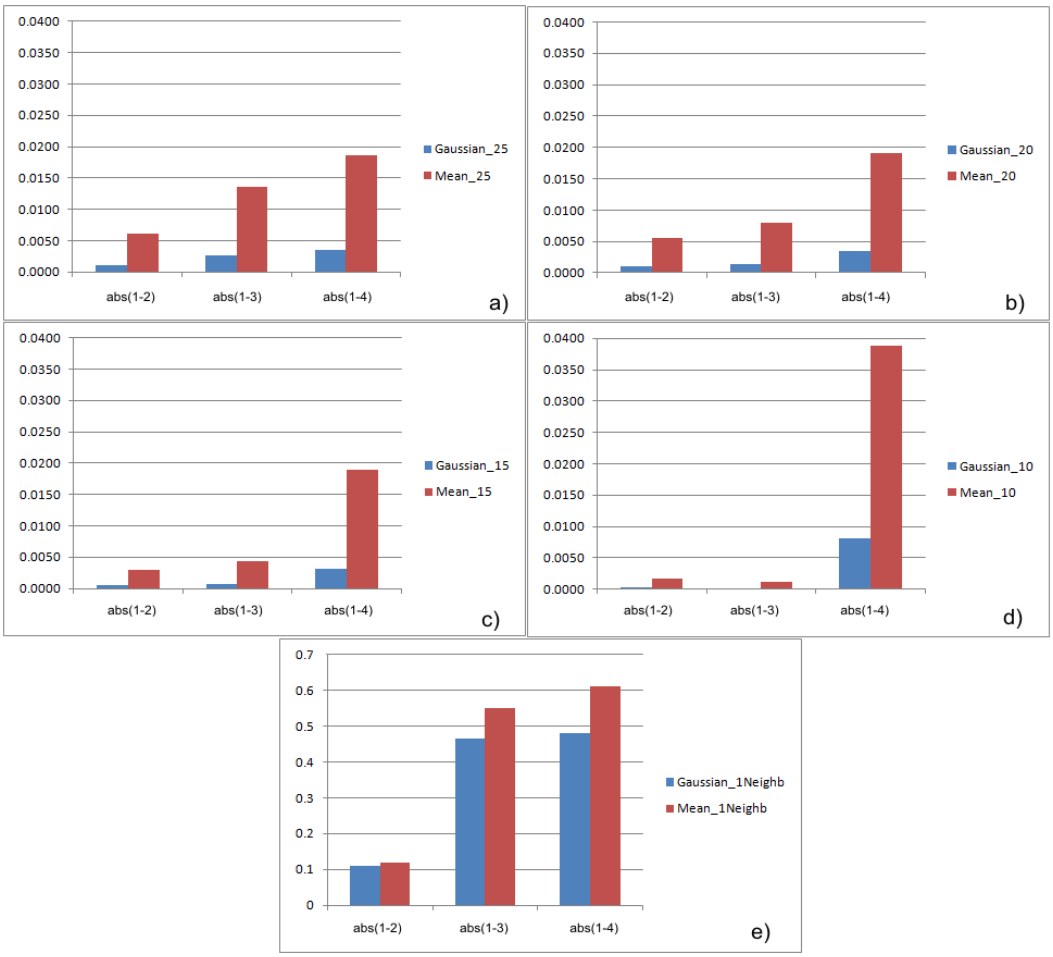

Figure 2.31: a, b, c, d - Show the absolute differences for Mean and Gaussian curvatures values using point neighborhood within 25, 20, 15, 10 mm of geodesic distance from the investigated point (the nose tip), e) shows the Mean and Gaussian curvatures calculated using "first neighbor" neighborhood for the surface approximation. Test performed without preprocessing.

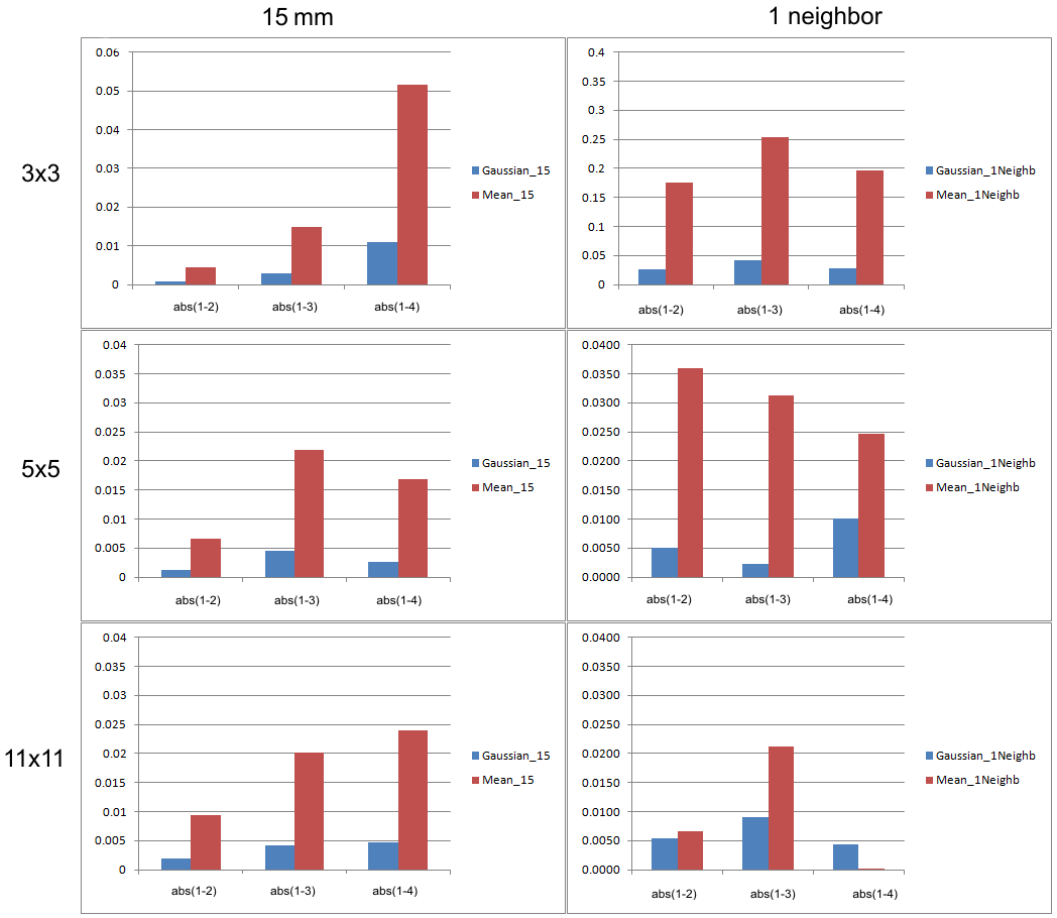

Figure 2.32: The absolute difference in the Mean and Gaussian curvature values using the reference model and smaller resolution models. The difference is conducted using preprocessed models by Gaussian filter with different kernel matrix size (indicated on left).

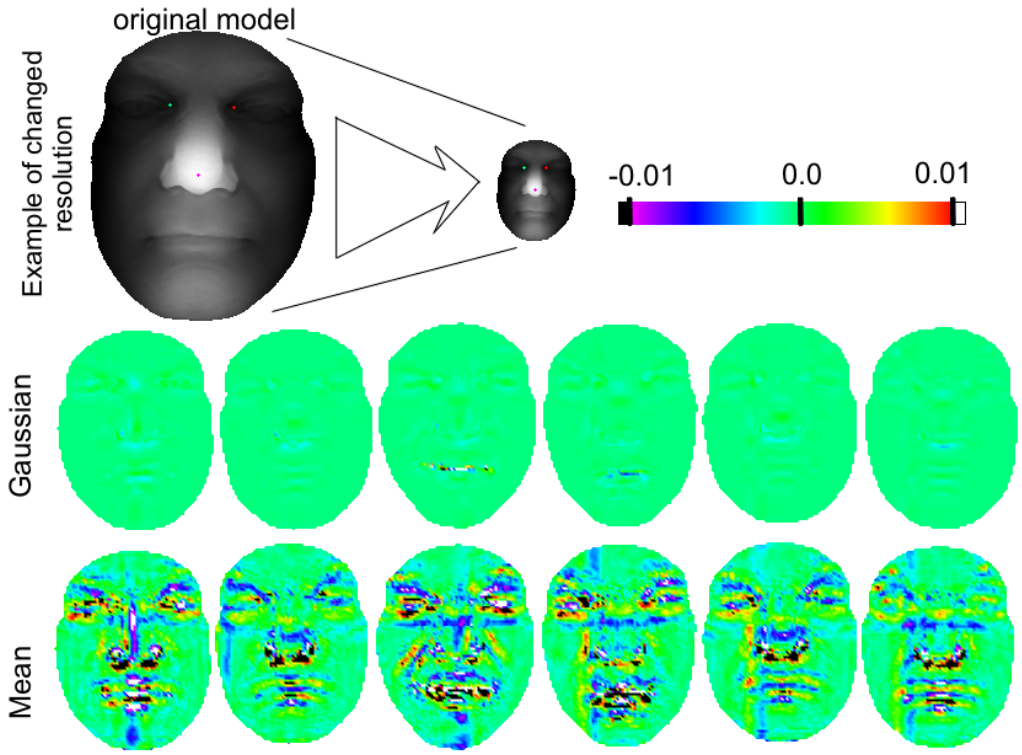

Figure 2.33: Upper row - example of the scaling to illustrate the resolution difference, two last rows - difference between the Mean and Gaussian curvature represented by color decomposition (computed using 15mm neighborhood for the surface approximation). Green color means no difference in the curvatures between different resolutions.

between models with different resolution. What should be noted is the fact that the Mean curvature is more pliant to resolution changes, where its decomposition introduces some areas with higher difference. The high difference regions are mainly the regions of fast surface changes, where the resolution change algorithm will modify their appearance by replacing a few pixels by their mean value.

2.5.3 Conclusion on curvatures stability

In this section we carried out extensive verification of the most popular curvatures calculation algorithm in case of model's resolution changes. We showed, that the most common method for the curvatures approximation, which is based on first-neighbors neighborhood, is influenced by model's resolution changes. To overcome the problem, we proposed a modification of the method, which relays on a different metric for selecting a neighborhood for a certain point. The modification allows us to achieve stable curvature values across different model's resolutions.

2.6 Conclusion on 2.5D Face Pre-processing, Analysis and Landmarking

In this chapter we have presented 2.5D face preprocessing method as well as the modification of the curvature calculation method, which was later used for the main facial points localization. Since the curvature calculation method is very sensitive to local noise, we proposed to change its locality by enlarging the neighborhood used for the curvature calculation. This change allows us to control the smoothness of the curvatures decomposition and allows us to design precise and robust facial main points localization algorithm. The algorithm achieved very high precision in localization of the nose tip and inner corners of eyes (based on the state-of-the-art presented in 2.2.2) . While the algorithm is rotation invariant, it can be used for the main points localization and face registration. We have also compared the algorithm to the other solution developed within our team, the SFAM method. Since the SFAM method is not rotation invariant, both methods are complementary.

Our modification of the curvature calculation method is based on a real geodesic distance expressed in mm, therefore the calculation method is invariant to the model resolution changes, which was proved by a set of tests on different models resolutions. To reduce the complexity of the main points validation, which is a bottleneck of the landmarking method based on curvature analysis, we proposed a learning-based technique, where a set of differential geometry descriptors was used. The test proved that the non-nose tip candidate points can be excluded with a high probability from the points candidates using SVM learning technique, thus reducing the time needed for the main points validation.

Chapter 3

Angle-preserving mapping-based face recognition

3.1 Introduction

Face recognition is potentially the best biometrics modality for people identification for it is: socially accepted, non-instructive and contactless. Unfortunately, face recognition in 2D proves to be a very challenging task as intra-class variations, due to factors as diverse as pose, lighting conditions, facial expressions, etc., are often much greater than inter-class variations [Zhao *et al.* 2003]. Recent years have witnessed 3D face models as a potential solution to deal with the two unsolved problems in 2D face recognition, namely lightning conditions and pose variations [Bowyer *et al.* 2006], thereby improving the effectiveness of face recognition systems. While 3D face models are theoretically insensible to lighting condition changes, they still require to be pose normalized before 3D facial shape-based matching.

A fundamental issue in 3D face recognition is 3D shape matching, which is a challenging task in noisy and cluttered scenes. Moreover, as 3D face models describe 3D facial shape, they are also more sensible to facial expression changes as compared to their 2D counterpart. Generally, the crux of surface matching is finding good shape representation, allowing to match two given free-form surfaces by comparing their shape representations [Wang *et al.* 2006]. Since facial expressions introduce non-ridig facial deformations, one of the shape representations, which has potential to deal with non-rigid deformations, is conformal mapping. According to conformal geometry theory, each 3D shape with disk topology can be mapped to a 2D domain through a global optimization. Having conformal parameterization of 3D shape, the problem of 3D shape matching can be simplified to a 2D image-matching problem of conformal geometric maps, thus allowing all the previously developed image matching techniques to work on 3D models. The conformal maps are insensitive to resolution changes, robust to noise as well as integrate geometric and appearance information onto 2D images. Therefore highly accurate and efficient 3D shape matching algorithms can be developped using conformal geometric maps [Wang *et al.* 2006].

3.2 Related Work on 3D Face Recognition Using Mapping Techniques

While 3D face models capture facial surface they are likely more sensitive to facial expression changes as compared to their 2D counterparts. Having higher dimensional data the whole 3D face processing algorithm is more complex and sophisticated. To deal with facial expressions, one of the first of 3D face recognition algorithms used surface matching based on ICP algorithm and cropping rigid facial parts [Amor *et al.* 2006, Chang *et al.* 2006] or using a combination of facial regions [Faltemier *et al.* 2008a]. More sophisticated algorithms used points signatures matching [Huang *et al.* 2010] or different face representations [Mpiperis *et al.* 2007, Samir *et al.* 2006].

The latest trend in 3D face recognition is based on removal of facial expressions. Expressions are removed by learning methods [Al-Osaimi *et al.* 2009, Pan *et al.* 2010] or by mapping [Bronstein *et al.* 2007b, Kakadiaris *et al.* 2007, Bronstein *et al.* 2005, Wang *et al.* 2006]. Additionally mapping techniques help to reduce complexity by moving 3D data to 2D domain. In [Wang *et al.* 2006], Wang et al. proposed conformal parameterization to reduce surface matching complexity. They studied a family of conformal geometric maps for recognition purpose. The authors projected Mean curvature on face images created by Least Square Conformal Maps. Those Least Square Conformal Shape Images were compared using normalized correlation coefficient, achieving 98.4% recognition rate at rank-one using 100 face models. Another, very interesting work is presented in [Kakadiaris *et al.* 2007], where the authors align each input data with Annotated Face Model (AFM) and create 2D geometry maps based on parametrization delivered by AFM. Reduction of facial expression is done by the Annotated Face Model fitting, which is also a bottleneck of the approach, making it more complex. Created 2D maps are then compared using two different distance metrics. For evaluation purposes the authors used the FRGCv2 data set and reported verification rate of 97.3% at 0.001 FAR. Another work, where the authors propose expression invariant face representation is presented in [Bronstein *et al.* 2007b]. In this work Bronstein et al. addressed the problem of constructing and analyzing expression-invariant representation of faces. In [Bronstein *et al.* 2007a] the authors demonstrate and experimentally justify a simple geometric model that allows to describe facial expressions as isometric deformations of the facial surface. This assumption is later used in [Bronstein *et al.* 2007b], where the authors isometrically embed face into a low-dimensional space, which is the core of their recognition system.

3.3 Discussion

Since facial expressions are known to be non-rigid, in our work we mainly focus on reduction of facial expressions by conformal mapping techniques. Since conformal mapping preserves angles, it can cope with non-rigid surface deformations. Like-

wise mapping introduced in [Bronstein *et al.* 2007b], conformal mapping requires the surface topology to be unchanged. Since surface topology is mainly affected by open mouth, first, in section 3.4 we propose an open mouth detection algorithm, which can be used later in conformal mapping technique. While Mean curvature can have ambiguous values, in contrast to [Wang *et al.* 2006], to create conformal face images we projected Shape Index, which is a normalized curvature representation. Unlike [Kakadiaris *et al.* 2007], where the authors used complex annotated face model fitting algorithm to deal with facial expressions in our work we used the Möbius transformation of UV conformal space to "compress" expression (section 3.5). We have also examined face recognition using only rigid facial part and conformal mapping which is presented in 3.6.

3.4 Dedicated Preprocessing for Conformal Maps Calculation

While facial surface deformation during expression is assumed to be near-isometric, open mouth significantly changes the surface topology and introduces anisometry. Removal of the open mouth part provides more reliable data for further 3D face processing pipeline, that requires consistently segmented faces and allows to treat face deformations as near-isometric. In this section, an automatic curvature-based open mouth detection algorithm is described.

Facial deformations can be modeled as near-isometric deformations of facial surface [Bronstein *et al.* 2007b, Bronstein *et al.* 2006, Bronstein *et al.* 2005]. Bronstein et al. [Bronstein *et al.* 2007b] presented an isometric model of facial expressions. The isometric model is based on an intuitive observation that facial skin stretches only slightly during expressions and all expressions of a face are assumed to be *intrinsically* equivalent. Nevertheless, this *intrinsicall* equivalence can be assumed only if the topology of the surface is preserved. Unfortunately open mouth, changes the topology of the surface by virtually creating a "hole". To handle expressions with open and closed mouth authors in [Bronstein *et al.* 2007b] proposed to fix the topology of the surface, by restricting the mouth to be open, disconnecting the lips in each case. A second method, to preserve the topology, enforces the mouth to be closed by "gluing" the lips when the mouth is open. In this section we: a) prove, using physically marked face, that the assumption to virtually close the open mouth, preserves better near-isometric deformations on the facial surface during expressions than the method used in [Bronstein *et al.* 2007b], b) present an automatic open mouth detection technique, c) propose modified geodesic distance to correctly measure near-isometric face deformations. To examine the assumption that facial deformations can be modeled as near-isometric, the distances between the nose tip and manually marked points on the facial surface were examined under various expressions. The open mouth detection is based on multi-scale curvature analysis. Correct metric preserving near-isometric deformations of the face is achieved by modified geodesic metric based on Dijkstra's algorithm, where the modification al-

lows us to create zero distance "bridges" between the upper and the lower lip in case of the open mouth detection.

3.4.1 The near-isometry hypothesis for open mouth expressions

An isometry is a distance-preserving map between metric spaces. Bronstein et al. in [Bronstein *et al.* 2007b] presented a facial expression invariant model based on the assumption that face expressions can be modeled as near-isometric. The near-isomtry assumption was derived from a test, where authors analyzed the changes of the geodesic distances between physically marked points during expressions. It should be noted, that the points on the used face model cover large areas, which undergo little or no motions during facial expressions, such as: forehead, nose ridge or points in the near neighborhood of those regions.

To reanalyze the facial expression near-isometric hypothesis, proposed by Bronstein, a face was physically marked using 10 points covering only the jaw part of the face, which moves the most during large expressions (fig. 3.1).

A subject was asked to introduce four expressions: neutral, small opening of the mouth part, larger opening of the mouth part, very large opening of the mouth part. For each 3D facial model the interior part of the mouth was removed to preserve surface topology. Each manually marked point was selected for all expressions giving nine traceable points (jaw part) and the nose tip (as a reference point) for each scanned model. Afterwards the geodesic, Euclidean and modified geodesic (see section 3.4.3) distances (d_g, d_e, d_{mg}) were calculated between the nose tip and all the traceable points, to verify if the isometric or near-isometric assumption holds. To verify the difference in the mentioned distances between facial expressions, the neutral model was taken as the reference face. Please refer to fig. 3.1 which shows differences in computed distances (d_g, d_e, d_{mg}) using reference model and models with other expressions.

It should be noted that the differences of the distances between traceable points and the nose tip are significantly reduced using modified geodesic distance. Smaller difference in the distances between points will help to achieve more consistent metric which will not be influenced by facial expressions.

3.4.2 Open mouth detection using high curvature edges

An analysis of Mean, Gaussian, principal curvatures, Shape Index and Curvedness leads to the conclusion, that the open mouth region characterizes very high k_1 curvature (fig. 3.2). In our research, a simple open mouth detection algorithm is proposed.

3.4.2.1 Selection of allowed curvatures

The algorithm starts from calculation of k_1 curvature using different neighborhood size (please refer to section 2.3.2 for more details about the curvature calculation method). To localize small and large open mouth regions, for each point p_i of a

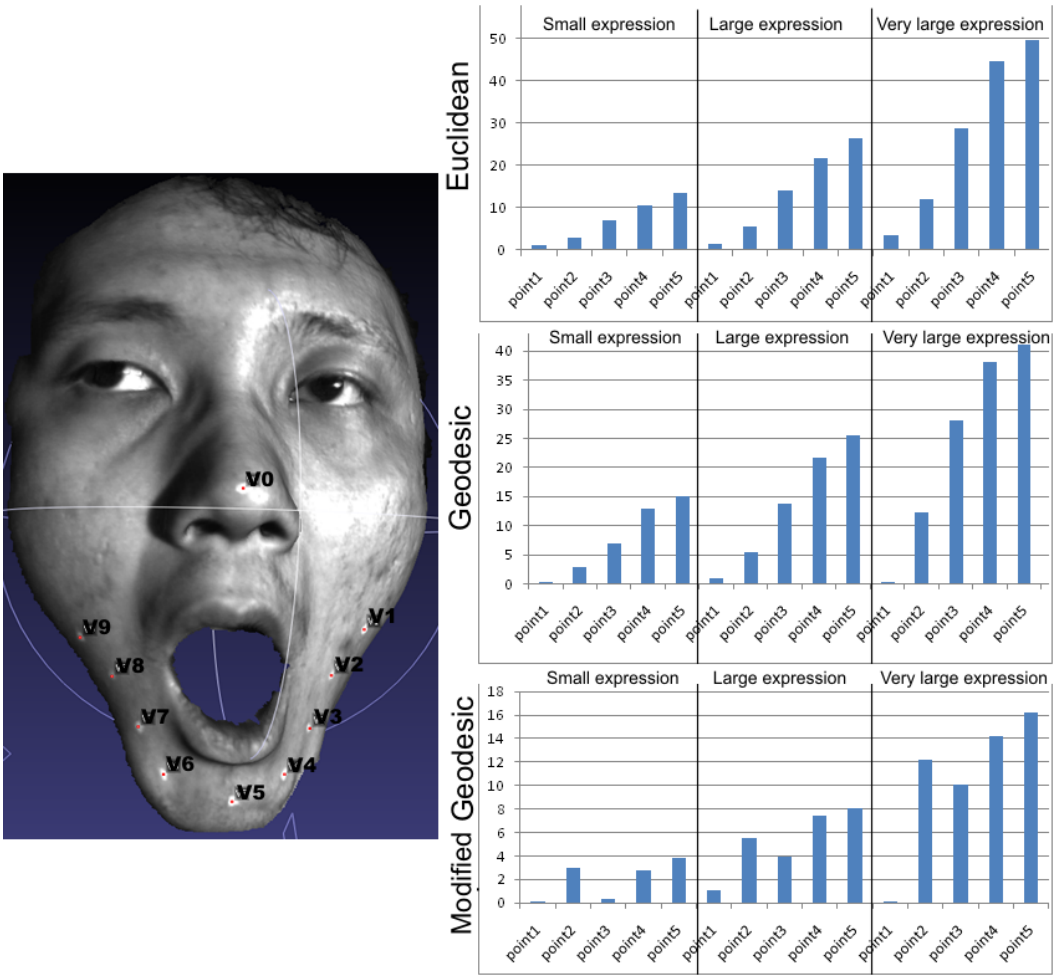

Figure 3.1: Marked face and differences in the distances measured between the nose tip and the traceable points using the neutral model and the models with other facial expressions.

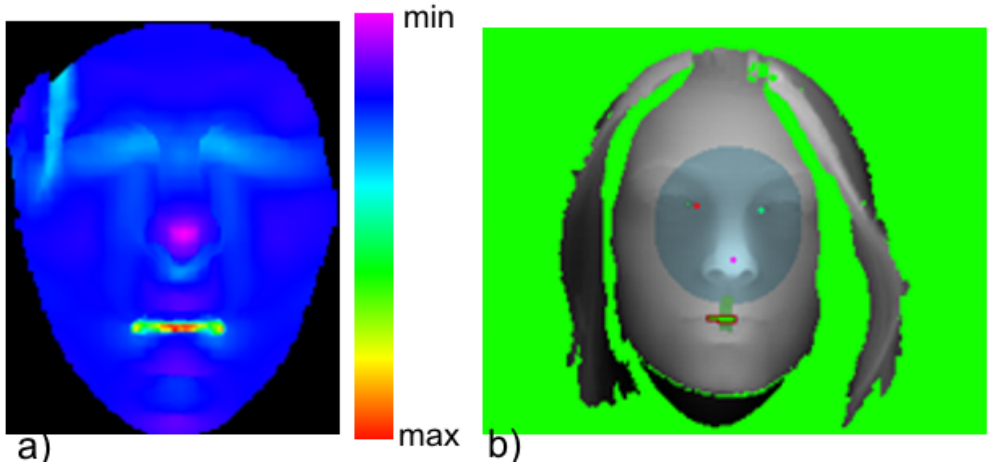

Figure 3.2: a) Decomposition of the principal curvature k_1 with 25mm neighborhood size in the surface equation approximation. Please note that the high k_1 curvature values correspond to a ridge between the lips. b) Static face region, determined based on three main points (light blue color), model with removed points which k_1 curvature exceeded the maximum allowed values.

surface, k_1 curvatures are calculated using different neighborhoods (10mm, 15mm, 20mm, 25mm, 40mm of geodesic distance from the investigated point p_i). Having the k_1 curvatures calculated (five values for each point p_i at different scale), localization of the open mouth part can be done by defining simple thresholds ($Tk1_{10}$, $Tk1_{15}$, $Tk1_{20}$, $Tk1_{25}$, $Tk1_{40}$). The thresholds are determined by the maximum k_1 curvature values on a rigid facial part.

Having the three main points, automatically selected by the algorithm previously described in section 2.3.3, the rigid region can be localized by selecting points where the sum of distances of which to the three main points is lower than the sum of distances between the three points increased by 15% (value chosen empirically). Fig. 3.2 shows the rigid region marked by a light blue color. This type of calculation leads to the selection of the static, rigid facial region, which is parametrized by the distances between the three points, consequently invariant to the face size.

All points on the facial surface, which k_1 curvature, at a certain scale, exceeds the threshold ($Tk1_{10}$, $Tk1_{15}$, $Tk1_{20}$, $Tk1_{25}$, $Tk1_{40}$) can be considered as points of open mouth part. In reality, those points also correspond to ridges between face and hair parts (can be seen on fig. 3.2b).

3.4.2.2 Detection of the curvatures belonging to the open mouth

The high curvature regions (exceeding the allowed curvature thresholds), belong to the open mouth part as well as appear on the side of facial surface, on cloths or hair parts (fig. 3.2b). To detect those corresponding to the mouth part, the mouth position needs to be known. Since the most precise point, which can be automatically

Figure 3.3: Virtual bridges between the upper and the lower lip (the colors from the upper lip correspond to the colors on the lower lip).

detected, is the nose tip, the average distance (dl) and standard deviation (σ) from it to the upper lip was calculated using manual landmarks on the FRGC dataset.

The open mouth detection is based on the assumption that a stripe followed between the inner corners of the eyes and covering the upper lip position ($dl + -\sigma$), will cross the high curvature region; we assume that this region belongs to the mouth part. If the stripe will cross more than one region, the biggest one is selected as the correct open mouth part.

3.4.3 Dealing with face near-isometric expression through a modified geodesic distance

To deal with face near-isometric expressions the open mouth part needs to be localized. As showed in the section 3.4.1, surface topology can be better preserved while moving the geodesic distance from upper lip to lower one. For this purpose, the Dijkstra's algorithm can be used. For a given source vertex in the graph (mesh), the algorithm finds the path with the lowest cost between that vertex and every other vertex. The method can be extended to calculate the geodesic distances over the whole facial surface, treating the euclidean distance between neighboring vertices as the weights.

To be able to "glue" the upper and the lower lip in the case of open mouth, the main modification of the Dijkstra's algorithm was to add zero weight "bridges" between the upper and the lower lip (fig. 3.3). The upper and lower lip were localized analyzing the open mouth region. Those bridges "continue" the calculated geodesic distance across the open lips and virtually glue them making the lower face boundary consistent between expressions. Figure 3.4 shows the difference between face cropping using geodesic distances calculated using virtual bridges between lips and without them while figure 3.5 shows modified geodesic distance decomposition on faces with neutral as well as non-neutral expressions.

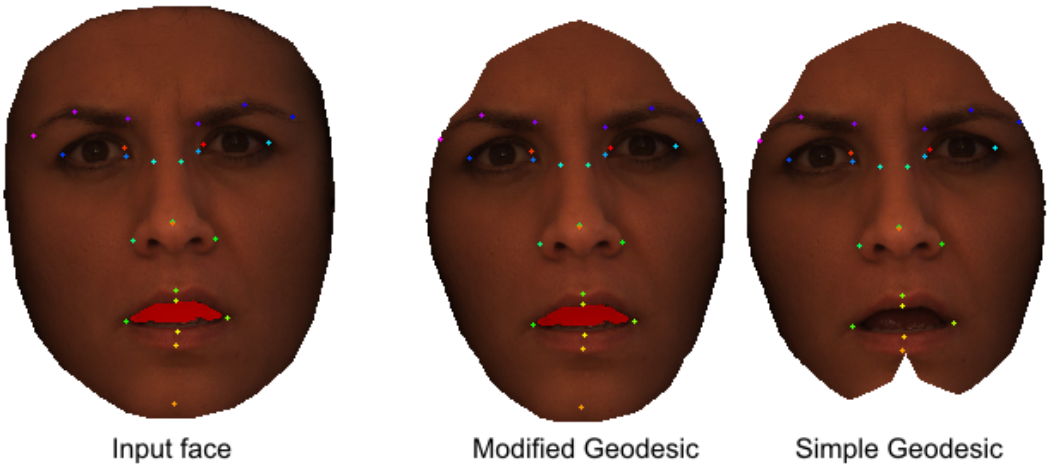

Figure 3.4: Difference between the modified geodesic and a standard geodesic face cropping (the boundary of the models is based on 100mm geodesic distance from the nose tip, the model from the Bosphorus data set [Savran *et al.* 2008]).

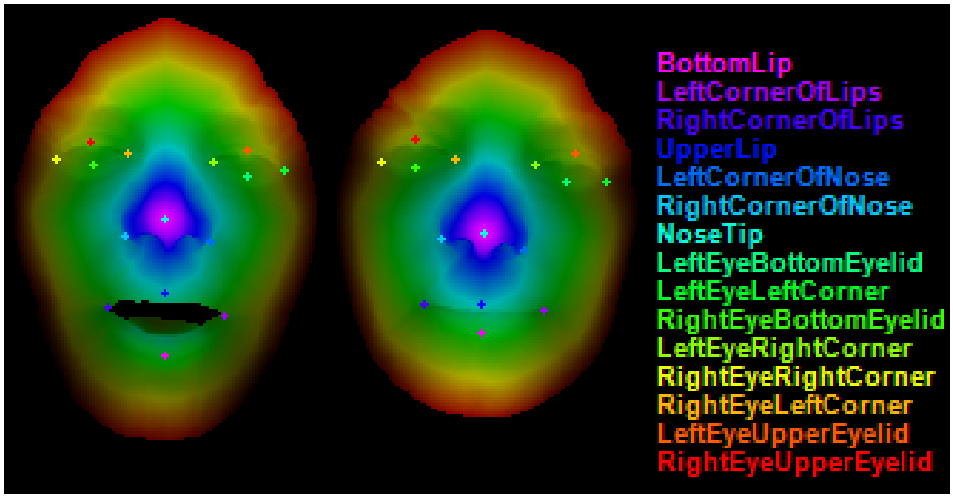

Figure 3.5: Modified geodesic distance on neutral and expression faces (marked as a color decomposition). Please note how, the modified geodesic distance is moved from the upper lip to the lower one (the models are from the FRGC data set [Phillips *et al.* 2005]).

3.4.4 Experimental results on open mouth detection

This research focuses on evaluation of the open mouth part localization. Since there is no ground truth data on open mouth boundary, the localization will be expressed in terms of the size as well as the frequency and position of the localized open mouth parts. The size of the localized part will be expressed in case of marked expressions on two publicly available 3D face datasets.

3.4.4.1 Test datasets

In our work we used two datasets: the FRGCv2 [Phillips *et al.* 2005] and Bosphorus [Savran *et al.* 2008] dataset. The FRGCv2 database contains 4007 3D images from 466 subjects, acquired between 2003 and 2004. The Bosphorus Database is a database of 3D faces which includes a rich set of expressions, systematic variations of poses and different types of occlusions. The dataset contains 105 subjects in various poses, expressions and occlusion conditions with 24 labeled facial landmarks each.

3.4.4.2 Mouth part localization

For the two data sets: we calculated the size of the localized mouth region in the case of expressions labeled in the data sets. In addition the mouth position was also examined using Bosphorus data set.

Analyzing the results (fig. 3.7 and fig. 3.6) one can notice that expressions influence the size of the localized region, as can be especially seen on the Bosphorus data set, designed for facial expression recognition. At a closer look at the facial action units and facial expressions from the Bosphorus data set, facial lower action unit 27 has the biggest detected facial opening region. The neutral expression during this experiment was placed near zero. It is worth noting that the "nonExpression" models from FRGC data set contain some open mouth regions, possibly due to errors in labeling.

The final test we have performed, was to investigate the detected open mouth region position. In this case, each expression from the Bosphorus data set was processed separately. Figure 3.8 shows the computed open mouth position and its frequency over all expression models. The white area on the black square is the frequency of appearance of the face shape. All faces were previously normalized by translating the nose tip in to the center (the black dot in the center of each image, there is no need to normalize rotations since the models are all frontal). One can notice that specific expressions introduce the mouth opening region in a certain position and shape on the facial model. Our method successfully computes these positions and shapes.

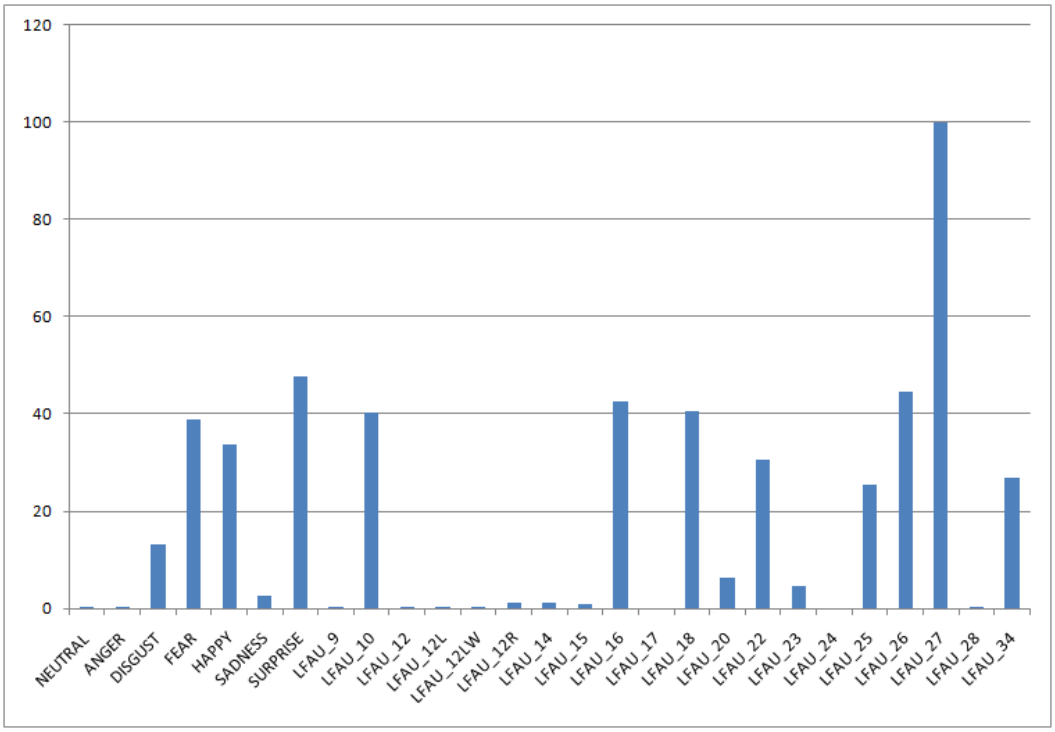

Figure 3.6: Size of the open mouth part (number of points) in case of expressions from the Bosphorus data set.

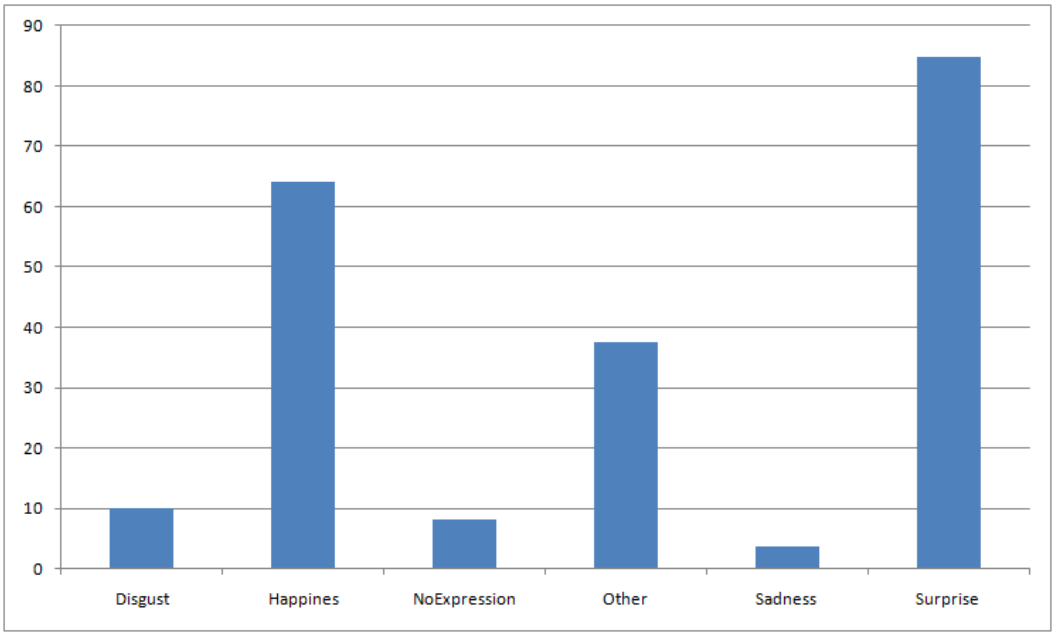

Figure 3.7: Size of the open mouth part (number of points) in case of expressions from the FRGCv2 data set.

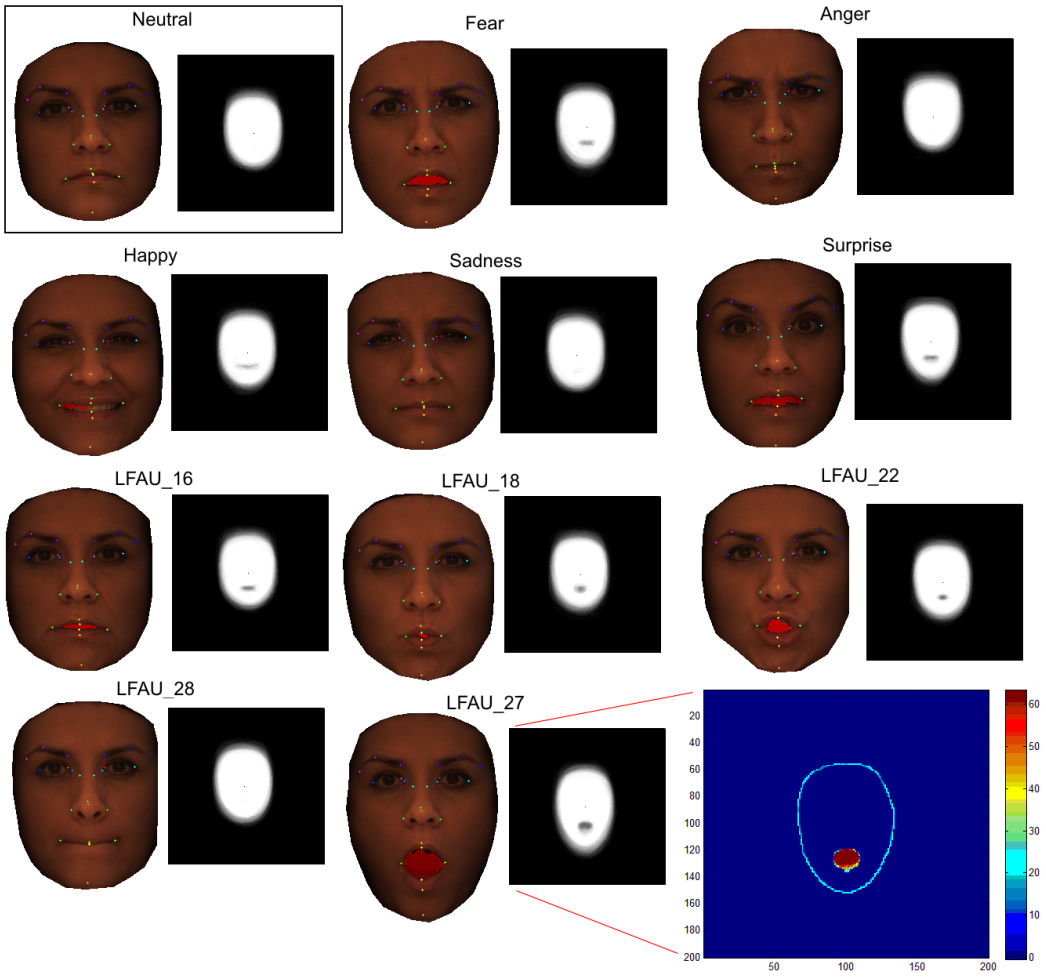

Figure 3.8: Bosphorus Expression Maps, created based on first 46 subjects. Left column introduces the model expression while right column (black square) introduces the expression map. The map shows mean face shape (white) with inner black-white gradient showing the frequency of localized mouth opening part (darker - more frequent)

3.4.5 Conclusion

Open mouth detection and usage of the modified geodesic distance are able to preserve near-isometric facial deformations and thus to provide more reliable data for further processing. The algorithm is based on high curvature edges detection and analysis of the rigid facial region. Such an analysis and removal of vertices, the curvatures of which exceed the allowed values, permits the removal of ridges between the facial surface and hair parts as well as the detection and removal of the open mouth part. Detection of the open mouth part is a crucial point in near-isometric facial cropping which later can be used for conformal face mapping. Based on the performed tests, the designed algorithm is simple yet effective in open mouth detection and can be ameliorated with additional verification step by mouth shape learning techniques and analysis.

3.5 Face Recognition Based on Mapping Preserving Non-rigid Surface Deformations

In this section, we propose to deepen conformal mapping-based approach for 3D face recognition. The proposed approach makes use of conformal UV parameterization for mapping purpose and Shape Index decomposition for similarity measurement. The 3D facial surface matching problem is reduced to 2D image matching thanks to the resulted 2D conformal geometric maps. To deal with facial expressions, the Möbius transformation of UV conformal space is also used to "compress" face mimic region. Rasterized images are used as an input for $(2D)^2PCA$ recognition algorithm [Kukharev & Forczmanski 2004].

3.5.1 Conformal UV parameterization for face normalization

It can be proven that there exists a mapping from any surface with a disk topology to a 2D unit disk [Haker *et al.* 2000], which is one-to-one, onto, and angle preserving. This mapping is called conformal mapping and keeps the line element unchanged, except for a local scaling factor [Floater & Hormann 2005]. Conformal maps have many appealing properties: (1) If the parameterization is conformal, then the surface is uniquely determined (up to a rigid motion) by the mean curvature with area stretching factor defined on the parameter domain. (2) Conformal parameterization depends on the geometry itself, not the triangulation of the surfaces. From a practical point of view, conformal parameterization is easy to control. Hence conformal parameterization is crucial for 3D shape matching and recognition. Consider the case of mapping a planar region S to the plane D.

Suppose S is a topological annulus, with boundaries $\partial S = \gamma_0 \ . \ \gamma_1$ as shown in Figure 3.9. First, we compute a path γ_2 connecting γ_0 and γ_1. Then we compute a harmonic function $f : S \rightarrow R$, such that:

$$\begin{cases} f_{\gamma_0} &= 0 \\ f_{\gamma_1} &= 1. \\ \Delta f &= 0 \end{cases} \tag{3.1}$$

The level set of f is shown in Figure 3.9. Then ∇f is a harmonic 1-form.

We slice the surface along γ_2 to get a new surface $\tilde{S}$ with a single boundary. γ_2 become two boundary segments γ_2^+ and γ_2^- on $\tilde{S}$. Then we compute a function $g_0 : \tilde{S} \to R$, such that:

$$\begin{cases} g_0|_{\gamma_2^+} &= 1 \\ g_0|_{\gamma_2^-} &= 0 \end{cases} \tag{3.2}$$

g_0 takes arbitrary value on other vertices. Therefore ∇g_0 is a closed 1-form defined on S. Then we find another function $g_1 : S \to R$, such that $\nabla g_0 + \nabla g_1$ is a harmonic 1-form $\nabla.(\nabla g_0 + \nabla g_1) = 0$.

Then we need to find a scalar λ, such that $*\nabla f = \lambda(\nabla g_0 + \nabla g_1)$, where $*$ is a the Hodge star operator ($*(fdx + gdy) = fdy - gdx$, where $fdx + gdy$ is a differential one form). The holomorphic 1-form is given by:

$$\omega = \nabla f + i\lambda(\nabla g_0 + \nabla g_1). \tag{3.3}$$

Let $Img(\int_{\gamma 0} \omega) = k$, the conformal mapping form S to a canonical annulus given by:

$$\Phi(p) = exp^{\frac{2\Pi}{k} \int_q^p \omega}, \tag{3.4}$$

where q is the base point, the path form q to p is arbitrary chosen.

For more details about conformal parameterization please refer to [Wang *et al.* 2006].

The result of conformal UV parameterization can be seen on Figure 3.10, where inner face hole created in lips part has been mapped to inner circle and outer 3D face boundary has been mapped to unit circle.

3.5.2 Process overview

The main idea underlying this approach is to transform a 3D facial shape matching problem to a 2D one using conformal parameterization. Furthermore, to deal with facial expression variations, we make use of Möbius transformation to "compress" the elastic facial region, leading to a 2D conformal map less sensitive to facial expressions.

The overview of the whole pipeline to create conformal maps from 3D face models can be seen on Figure 3.11. Following sections will describe each step in details.

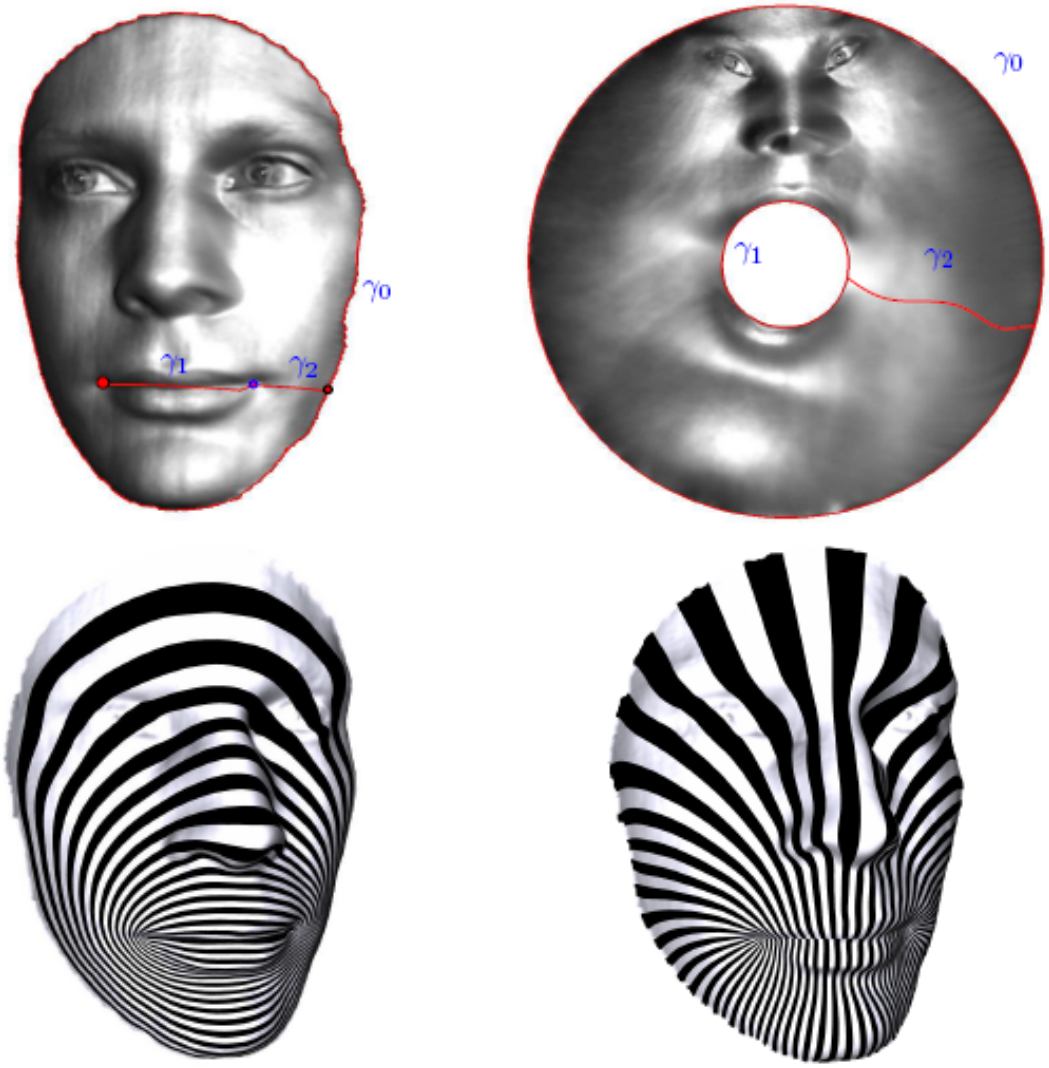

Figure 3.9: Harmonic 1-forms. Top row, the cut on the surface. Bottom row, the level sets of the harmonic 1-form ∇f and its conjugate harmonic 1-form $\lambda(\nabla g0 + \nabla g1)$.

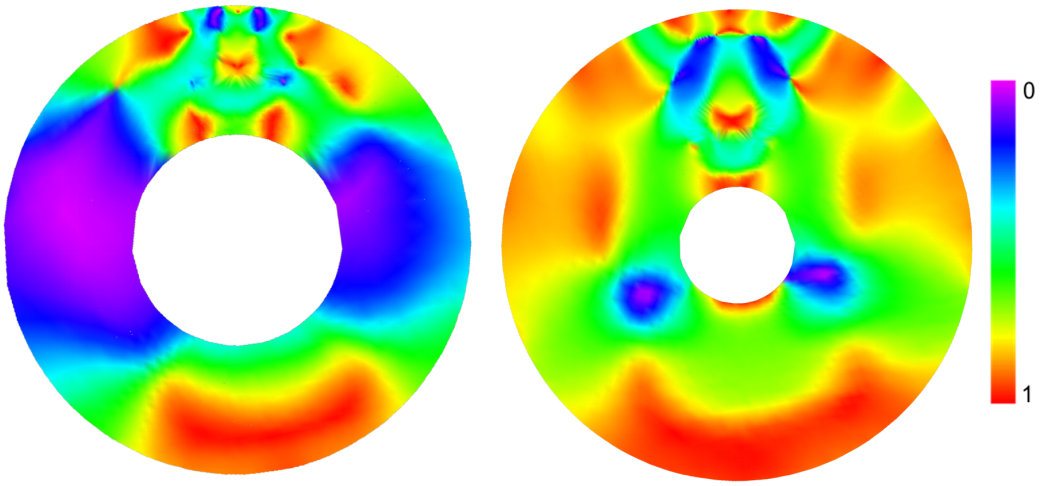

Figure 3.10: Conformal Map created from parametrized UV coordinates, colors represent Shape Index value. The mouth part has been mapped to the inner hole while external face boundary to the boundary of circle (genus 0 surface, with a single boundary).

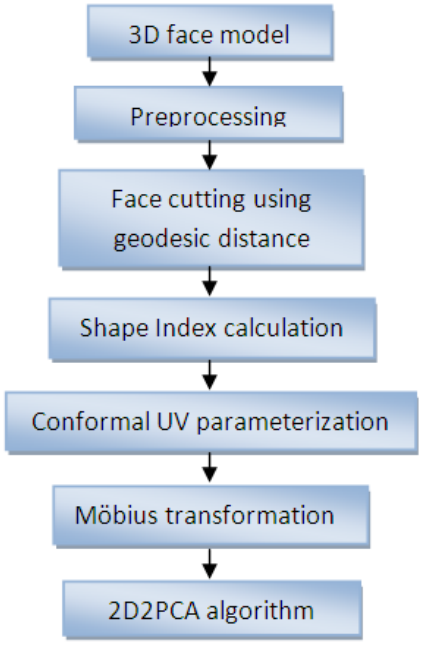

Figure 3.11: Flowchart of the face recognition algorithm.

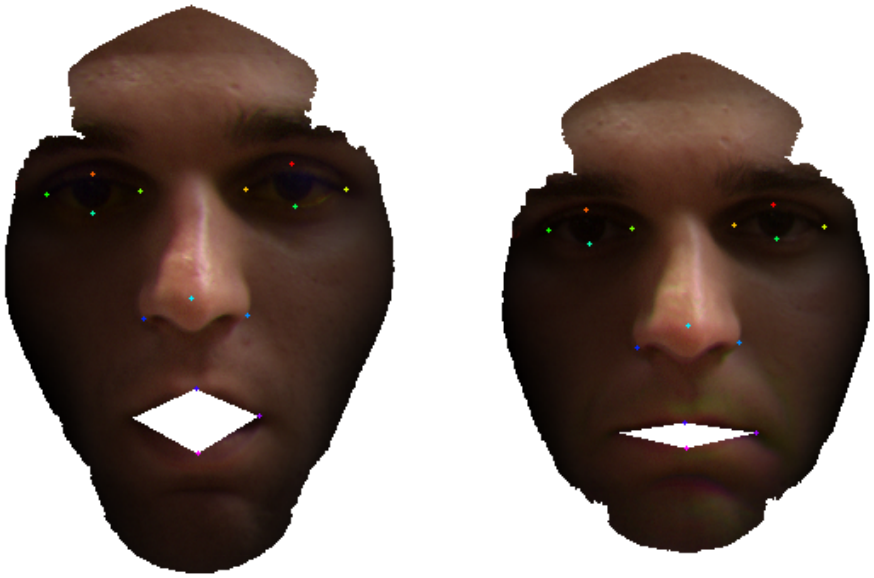

Figure 3.12: Geodesic face cutting result with zero distance between lips in presence of expression.

3.5.3 Generation of face conformal maps

3.5.3.1 3D Face preprocessing

Direct application of conformal mapping introduced in [Wang *et al.* 2006] is not feasible on 3D face models as it requires surface with disk topology (genus 0 surface, with a single boundary). For this purpose, we have "closed" the mouth region based on manual landmarks, setting to zero the distance between the upper and the lower lips.

Conformal mapping is also sensitive to outer boundary [Wang *et al.* 2006]. To deal with this problem, faces are cropped using a fixed geodesic distance, 100 mm in this work, from the nose tip.

Once this preprocessing is carried out, the 2D UV conformal parameterization (section 3.5.1) of a 3D face model can be calculated according to [Wang *et al.* 2006]. Figure 3.12 shows the result of this preprocessing step. As we can see, "closing the mouth" while using geodesic distance for cropping 3D faces leads to a more consistent outer boundary, especially in the chin region, displaying roughly the same border the mouth being opened or closed.

3.5.3.2 Shape Index

Since UV conformal parameterization transfers 3D model to 2D map, some 3D property has to be moved over 2D face map. To deal with variations due to lighting

conditions on texture images, we chose to project Shape Index values. Alternatively, other geometric measures such as normal vectors, curvatures, etc. can be projected.

Shape Index (Figure 2.26) is a normalized curvature representation in a certain point of a surface within 2.5D image (for more information please see section 2.3.2).

3.5.3.3 Conformal Map Normalization

Facial conformal maps generated by the harmonic energy minimization from 3D face models can have different size and 2D rotation. To facilitate matching of 2D facial conformal maps, they need to be size and rotation normalized.

For rotation normalization we make use of the two inner eyes corners mapped on the conformal map then rotate the underlying conformal map so that both the two inner eye corners lie at the horizontal line. Once the pose is corrected, the size of the underlying conformal map is also normalized, using the radius min-max rule, setting radius of conformal map to 50 units.

3.5.3.4 Compressing facial expression sensitive regions by Möbius transformation

Variations by facial expressions are a major challenge in 3D face recognition. Facial conformal maps so far generated have reduced a 3D shape matching problem to a 2D one while preserving facial topology. However, they are still facial expression sensitive. In order to decrease such a sensibility, we propose to make use of Möbius transformation to "compress" facial elastic regions, i.e. the lower part of a face model. For this purpose, the center of a conformal map is moved to the nose tip of the face. Then Möbius transformation is carried out on UV conformal coordinates, using the following formula:

$$f(\theta, z_0, z) = e^{i\theta} \frac{z - z_0}{1 - \bar{z_0} z}, \tag{3.5}$$

where $z = (u + iv)$ is a complex number within the unit disk (UV coordinates). θ, z_0 and z are parameters. The mapping will move z_0 to the origin.

Figure 3.13 shows some results of this transformation on two facial conformal maps. Finally conformal maps resulting from the Möbius mapping are rasterized.

3.5.4 (2D)^{2}PCA recognition algorithm

In this work (2D)^{2}PCA [Kukharev & Forczmanski 2004, Yang *et al.* 2004, Chen *et al.* 2005], a variant of PCA with better performance, is used for feature dimension reduction and similarity computation.

Principal Component Analysis (PCA) is a well-known feature extraction and data representation technique. However it has one serious drawback for 2D images, because 2D image matrix has to be previously transformed to 1D vector by columns or rows concatenation. This type of concatenation into 1D vector often leads to a high-dimensional vector space, where it is difficult to evaluate covariance matrix

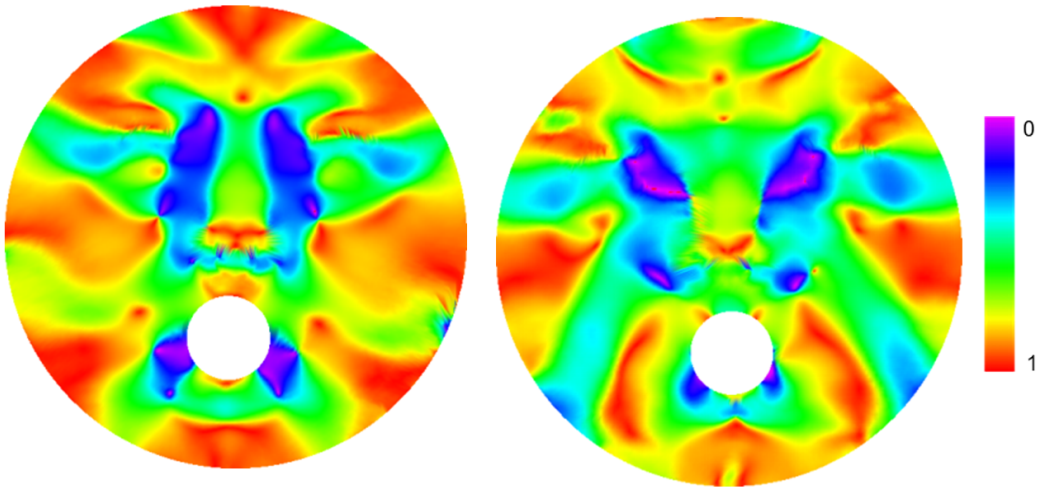

Figure 3.13: Conformal maps (Figure 3.10) transformed by Möbius transformation with center point in the nose tip.

accurately due to its large size and relatively small number of training samples [Kukharev & Forczmanski 2004, Yang *et al.* 2004, Zhang & Zhou 2005]. Also eigen decomposition of large covariance matrix is also very time-consuming.

To overcome those problems 2DPCA was proposed [Kukharev & Forczmanski 2004, Yang *et al.* 2004]. 2DPCA technique computes eigenvectors directly from so-called image covariance matrix, without conversion to 1D vector. 2DPCA is more efficient method than standard PCA what was reported in [Kukharev & Forczmanski 2004, Yang *et al.* 2004].

As a standard 2DPCA method works in row directions, the alternative 2DPCA works in columns directions of images, $(2D)^2$PCA algorithm combines both of them [Chen *et al.* 2005]. In more details 2DPCA optimal matrix X reflecting information between rows of images, alternative 2DPCA learns optimal matrix Z reflecting informations between columns of images. $(2D)^2$PCA uses both matrixes X and Z to create coefficient (feature) matrix C:

$$C = Z^T A X, \tag{3.6}$$

For classification the nearest neighbor classifier can be used [Yang *et al.* 2004]:

$$d(C, C_k) = \| C - C_k \| . \tag{3.7}$$

The method was tested on standard face image database along with PCA and 2DPCA and gained higher accuracy with lower feature dimensionality. For more details of the whole method please refer to [Kukharev & Forczmanski 2004, Chen *et al.* 2005].

	I	II	III
ShapeIndex	86.43%	97.65%	69.38%
Mean Curv.	86.84%	94.29%	75.51%
Curvadness	86.23%	96.30%	70.91%
Kakadiaris 2007 PAMI[Kakadiaris *et al.* 2007]	97%	-	-
Wang 2006 CVPR[Wang *et al.* 2006]	95.7%	-	-

I - Neutral vs. All
II - Neutral vs. Neutral
III - Neutral vs. Expression

Table 3.1: Rank-1 recognition rate on 62 subjects of FRGCv2.0 data set.

3.5.5 Experimental results

3.5.5.1 Testing Data Set

To experiment our approach, 62 subjects were randomly selected from FRGCv2 database [Phillips *et al.* 2005]. FRGCv2 dataset contains 4007 3D scans of 466 persons. The data set contains also labeled expression variations like: NoExpression, Disgust, Happiness, Sadness, Surprise, Other, used in different scenarios for face recognition.

3.5.5.2 Experimental Settings

For each 3D face model, the corresponding facial conformal map is generated using the whole process described above, including 3D face cropping, UV paramterization calculation, normalization, Möbius transformation and rasterization. The resulting 2D conformal maps were used as input for $(2D)^2$PCA algorithm for recognition, keeping 99% of eigenvalues variation.

One model with neutral expression from each selected subject is taken for the gallery and the remaining models according to labeled expression are used as a probe in different scenarios: 1) Neutral vs. Neutral, 2) Neutral vs. Expression, 3) Neutral vs. All. In case of first scenario probe models are selected within "NoExpression" labels, in the second scenario probe models come from all models except those marked "NoExpression", and finally in the last scenario we take all expression and no-expression models as a testing probe.

3.5.5.3 Results and Analysis

Using this experimental setting, our approach has achieved 97.65% rank-one recognition rate for scenario where models labeled as "NoExpression" are presented to the system. All test scenarios and different combinations of curvatures are presented in Table 2.4.

For comparison reasons we have also projected different vertex features as color maps. While the Shape Index is a normalized curvatures representation with value in range $< 0, 1 >$, the mean curvature or curvedness have no range. To create images with the same values range, average maximum and minimum values were calculated using models form the Gallery.

As we can see in Table 3.1 Shape Index Maps achieved best performance in the test Neutral vs. All at 97.65%, but Cuvardness Index is not far away with result at 96.3%. While in case of Expressions Mean curvature maps outperforms Shape Index maps with difference of 5%.

Comparing our approach to article [Wang *et al.* 2006] we have achieved around 9% lower performance in the scenario of Neutral-All, but our algorithm has been tested on much bigger data set containing large expressions, while the tests made in the previous article were made only on 100 models without any mention about expressions. In [Kakadiaris *et al.* 2007] authors did not mention about recognition results in different scenarios and no results are provided to evaluate the sensibility of their algorithm with respect to expression variations. Their technique requires first an ICP based accurate registration of 3D face scans which is exposed to facial expressions and correct models cropping.

3.5.6 Conclusion

In this work we deepen the conformal geometry-based approach for face recognition proposed in [Wang *et al.* 2006], using mouth as the inner boundary and the Möbius transformation to "compress" facial expression sensitive regions. The major advantage of such an approach is to convert an initially 3D facial shape matching problem to a 2D one, thus making available all the techniques so far developed in 2D for 3D face recognition. The algorithm has been tested on a subset of FRGCv2 database with different scenarios and achieved 86.43%, 97.65% and 69.38 rank-one recognition rate in respectively Neutral vs. All, Neutral vs. Neutral and Neutral vs. Expression scenarios.

3.6 Partial Face Biometry Using Shape Decomposition on 2D Conformal Maps of Faces

While pose changes and lighting variance can be resolved by the use of third dimension, by pose normalization and lighting insensitive scanning techniques, the expression changes is still a problem. In this work we avoid this problem by recognition based only on the rigid part of the face, which cannot be severely affected by muscles. To deal with the computational cost of 3D face recognition we utilize conformal maps of 3D surface to a 2D domain, thus simplifying the 3D mapping to a 2D one (for which all standard image classification methods are applicable). While conformal maps requires a certain amount of computation, they have global optimal solution unlike ICP which suffers from local minimums. Since face appear-

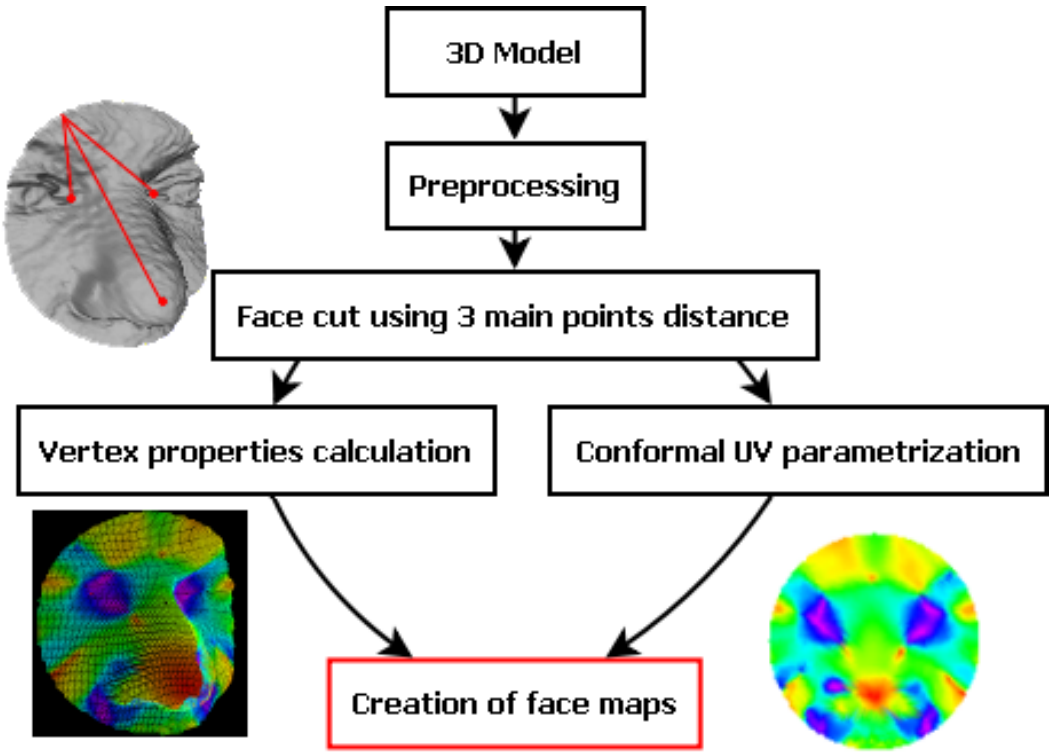

Figure 3.14: Face maps creation flow chart.

ance is influenced by expressions in this research we explore the descriptiveness of shape dissimilarity. We transform a 3D shape to a 2D domain using conformal mapping and use shape decomposition as a similarity measurement. In this work we investigate several classifiers as well as several shape descriptors for recognition purposes.

3.6.1 Process overview

The principal issue addressed in this chapter is to create facial feature maps which can be used for recognition by applying previously developed 2D recognition techniques. Creation of 2D maps from 3D face surfaces can handle model rotation and translation, and allows to use well known 2D recognition techniques.

To create face maps which are later used for recognition, we started with models preprocessing (hole, spike removal). Next step was to segment the rigid part of a face, which has less potential to change during expression [Amor *et al.* 2006]. Finally we performed UV conformal parameterization as well as shape index calculation for each vertex. To create a face map we combined UV parameters and Shape Index values to construct a shape value distribution over the conformal parameterization (the process can be seen on figure 3.14).

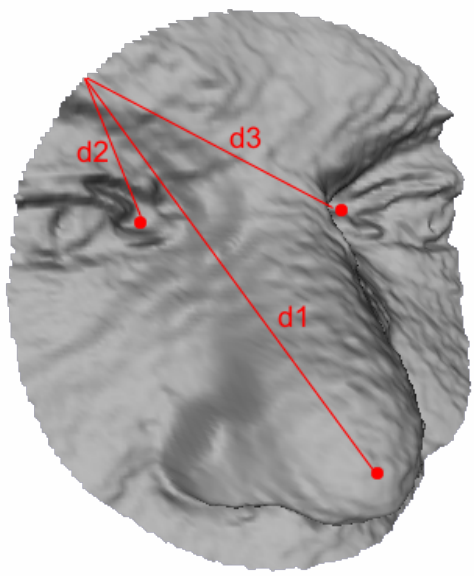

Figure 3.15: Discriminative part of a face used for recognition. Cropped part, where the sum of the distances from the nose tip and the left and right eye inner corners is less than 50 mm (better seen in color).

3.6.1.1 Face discriminative part segmentation using 3 landmarks

To deal with expressions and to explore the field of partial face recognition [Drira *et al.* 2009]. We use the upper rigid part of a face defined by [Amor *et al.* 2006] which is less influenced by expression than the lower part.

To crop this region of the face we used three main face landmarks, the nose tip and the inner corners of eyes; in our experiments we used the manual points. We crop the discriminative face part using the sum of Euclidean distances of each face point to 3 landmarks and setting a threshold to 50 mm. The result of cropping can be seen on figure 3.15.

3.6.1.2 3D face mapping to 2D space

Direct application of conformal mapping as introduced in [Wang *et al.* 2006] is not feasible on 3D face models, as it requires a surface with disk topology (genus 0 surface, with a single boundary). For this purpose, we have made a small hole in the face surface (by removing one vertex) in the center of the cropping region; later this inner boundary will become the inner boundary of the conformal map and the border of the face will become the outer boundary of the conformal map, as can be seen on figure 3.16.

To create such a conformal face map, the UV conformal parameterization of a 3D face is used (section 3.5.1); each 3D vertex after parameterization has 2D UV coordinates through which the whole model can be projected to 2D space.

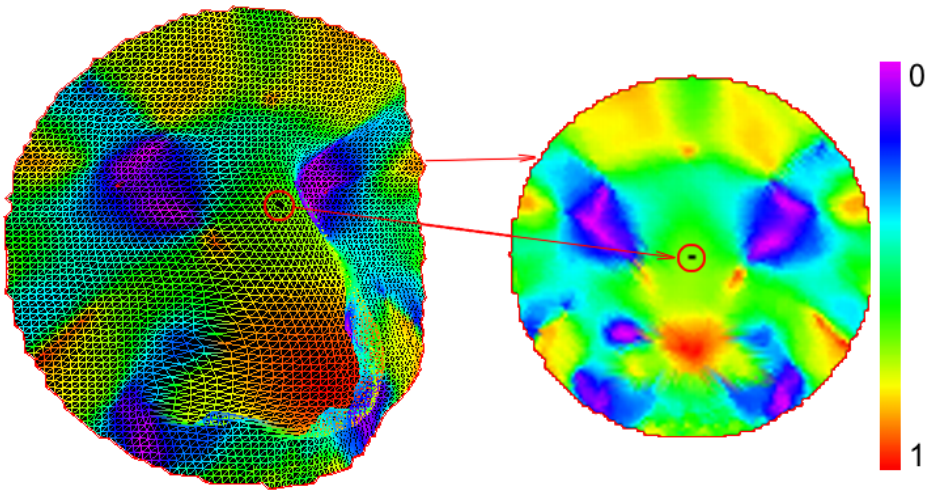

Figure 3.16: Projected 3D face model to a 2D space conformal mapped to a disc, (better seen in color).

All conformal maps were also normalized for size and rotation using above mentioned manual landmarks, projected from the 3D space.

3.6.2 Face maps similarity computation

The previous step results in face map images which are the input for the recognition algorithm. In this work we used $(2D)^2PCA$ [1][Kukharev & Forczmanski 2004, Yang *et al.* 2004, Chen *et al.* 2005], as well as nearest neighbor classifier with L_1 and L_2 norms.

For test purposes we randomly chose 231 subjects from The FRGC2.0 data set [Phillips *et al.* 2005]. Table 3.2 shows all results for different neighborhood sizes for Shape Index and Mean curvature, as well as different similarity measurement methods.

3.6.3 Conclusion

In this work we proposed to combine the conformal geometry, partial face biometry and differential geometry tools for recognition. 3D face recognition has many advantages, but a large computation cost. Using 3D data as an input and projecting 3D features to 2D maps has advantages of 3D invariance to lighting and pose like also 2D similarity efficiency.

In our work we used the rigid face part from a subset of 231 subjects and 1249 models as a query, we achieved approximately 80% rank-one recognition rate; compared to [Drira *et al.* 2009] where authors used nasal curves and achieved 76.1%

[1] $(2D)^2PCA$ - a variant of PCA with better performance, is used for feature dimension reduction and similarity computation.

	I	II	III
$(2D)^2PCA$			
ShapeIndex 25mm	72.85%	81.65%	65.71%
ShapeIndex 20mm	75.34%	82.1%	69.46%
ShapeIndex 15mm	77.1%	82.78%	72.15%
ShapeIndex 10mm	76.14%	84.5%	68.86%
Mean Curv. 15mm	67.09%	72.8%	62.12%
Nearest Neighbor			
L_1			
ShapeIndex 25mm	74.77%	82.27%	68.26%
ShapeIndex 20mm	75.26%	82.09%	69.31%
ShapeIndex 15mm	**79.18**%	84.5%	74.55%
ShapeIndex 10mm	77.42%	85.19%	70.65%
L_2			
ShapeIndex 15mm	75.74%	82.96%	69.46%
1 Loop ICP	70.21%	-	-

I - Neutral vs. All
II - Neutral vs. Neutral
III - Neutral vs. Expression

Table 3.2: Rank-1 recognition rate on 231 subjects from FRGCv2.0 data set.

rank-one recognition rate and [Gokberk *et al.* 2008] where authors achieved 75.3% rank-one recognition rate using the shape index decomposition.

3.7 Conclusion on Conformal Mapping Based Face Recognition

Face recognition in 3D has been proposed to overcome shortcomings of 2D face recognition techniques. Having many advantages, 3D face recognition is more complex thus making difficult real time applications. Furthermore, since 3D facial surface describe facial geometry, those are more sensitive to facial expressions than 2D counterparts. While facial expressions are know to be non-rigid, one of the solutions to deal with non-rigid surface deformations is conformal mapping. Conformal mapping is a non-rigid mapping technique, which preserves angles, moving 3D surface into 2D domain yet reducing the complexity of 3D surface matching to 2D image matching. Conformal mapping-based 3D face recognition, while highly effective when used with face models without prominent facial expression (97% recognition rate rank-one) [Wang *et al.* 2006], suffers form drastical performance drops drastically when facial expressions are present (86.7% recognition rate rank-one). The reason why the conformal mapping performs worse in case of expressions is the boundary con-

sistency as well as topology preservation condition. Large facial expressions cause changes in the facial surface topology, where open mouth generates additional part between open lips. Facial surface boundary is also a difficult problem, where besides facial expressions additional factors like: hair style changes or partial occlusions cause inconsistent boundary between models thus causing incorrect mappings. One of the solutions to create illumination insensitive recognition system is to project some facial surface properties on 2D conformal maps. Meanwhile those properties are highly affected by facial expressions which change facial surface geometry thus affecting recognition scores.

One of the solutions for the surface topology and boundary consistency preservation can be the open mouth detection algorithm, presented in section 3.4. Accurate and reliable open mouth detection and removal can theoretically handle surface topology changes but facial boundary preservation during expressions is still an open problem. Having the two problems resolved, one can think about conformal mapping as a tool which can deal with 3D facial expressions and complexity of 3D surface matching.

Chapter 4

Conclusion and Future Work

This research work mainly addresses the problems of 3D face analysis, including facial feature points and open mouth localization as well as non-rigid mapping based face recognition using differential geometry tools. The whole work can be combined to create a full 3D face recognition system, starting from 3D face models preprocessing, main points localization and mapping-based recognition.

In the following part of the thesis contributions, conclusions and perspectives will be presented.

4.1 Contributions

- The new face recognition modality in 3D has been proposed to overcome shortcomings of 2D counterpart. Since nowadays 3D scanning techniques suffer from scanning artefacts, one of our contributions relates to spikes removal technique. Models suffer from holes and spikes which are the main quality issues. Our first contribution refers to the preprocessing method, which increases models quality without removing fine details on facial surface. The method consists of adding simple threshold to the median filtering technique, which helps to remove shot noise from processed models.

- The most popular method for 3D face models analysis and processing relies on curvatures analysis. While the processing of the curvatures allows facial expressions analysis, landmarks localization, face recognition, etc., their calculation method is very sensitive to the surface noise. Our second contribution relies on modification of the curvatures calculation method, which incorporates noise reduction and helps to control smoothness of the curvatures decomposition. Additionally, the modification helps to achieve consistent curvatures values across different models resolutions.

- Subsequently, the developed curvatures calculation method was used for precise facial main points localization. Our next contribution relies on analysis of very smooth curvatures, where the main patches on facial surface can be distinguished thus excluding noise and small facial irregularities. The developed method achieved the highest precision in the nose tip localization in adverse conditions including face rotations and noise.

- While curvatures analysis generally leads to many facial anthropometric candidate points, our next contribution relies on fast incorrect candidate points

rejection, based on differential geometry point signatures and a learning technique. The nose tip can be alternatively localized using SVM classifier and geometry based point properties which were carefully examined by numerous experiments.

- Face recognition, especially 2D modality, has been studied for many years. Since face recognition in 3D is fairly new, we propose to map 3D facial models to 2D domain, where all the previously developed face recognition techniques can be used. Additionally, since the facial expressions are known to be non-rigid, we propose to use non-rigid conformal mapping which has potential to deal with facial expressions.

- The final contribution of this thesis relies on the specific preprocessing algorithm needed to preserve surface topology as well as surface boundary, which are the main conditions for conformal mapping to achieve comparable face maps. Since the surface topology mainly changes due to large facial expressions, where open mouth causes appearance of additional vertices between lips, the open mouth part needs to be removed. The specific preprocessing consists of face curvatures analysis and removal of incorrect vertices with curvatures exceeding the maximum thresholds.

4.2 Perspectives for Future Work

Extensions of this work, that we envisage are presented in the following.

In this thesis curvatures analysis was used for the main facial points localization, open mouth detection as well as for face recognition purposes. One of the bottlenecks of the presented landmarks localization technique concerns the step of main facial points validation. Numerous of the main points candidates constrain a long and exhaustive verification process, which aim is to localize the true main facial points. To reduce the complexity, we have proposed a learning technique which aim was a fast rejection of incorrect points. Since many of the main points candidates were localized on facial hair or cloths, face curvatures model could be helpful. The face curvatures model can be used in a similar way as a face skin color model (forming a cascade filtering), rejecting non-face vertices thus simplifying further calculations.

Another improvement may concern the time consumption of curvatures calculation method. A larger neighborhood used in the curvatures calculation method causes significant increase in the time needed to estimate curvatures values. Since proposed modification in the curvatures calculation method is invariant to resolution changes, the facial models' resolution can be decreased. A large decrease of facial models' resolution has a small influence on curvatures values. Since we are targeting only the main facial points (the nose tip and inner eyes corners), which are still present even in a very small resolution, thus the resolution can be changed probably without influence on the main points localization.

Since open mouth significantly changes surface topology, its detection and removal is a crucial step for face topology preservation. In our work we have used a simple method for an open mouth detection relying on curvatures analysis. In this point the simple method can be exchanged to a more sophisticated one relying for example on open mouth shape learning-based and fitting techniques.

Since facial expressions are known to be non-rigid and significantly influence face recognition based on 3D models, in our work we propose conformal mapping-based 3D face recognition. Conformal mapping algorithms map 3D surface to 2D domain preserving angles thus have potential to deal with non-rigid surface deformations. To create correct 2D conformal maps the surface topology as well as a surface boundary needs to be preserved. For our future fork we plan to use the designed preprocessing which aims to remove open mouth part (which changes significantly surface topology), to create comparable 2D conformal maps. After removing open mouth part, the upper and lower lip should be connected through correct triangulation.

Finally, an interesting research direction could be towards Active Curvatures Face Models, which can be used for many purposes like: face localization, facial landmarks extraction, open mouth detection, face analysis, face expressions recognition, occlusions detection etc. Those models will be invariant to illumination changes and since curvatures are rotation-invariant properties of 3D surfaces, the model will also inherit rotation-invariance.

Publications

The results obtained during my Ph.D study have been the subject of four publications at international conferences and two at a national conference. Moreover two conference papers have been submitted to international conferences.

International Conferences

1. P. Szeptycki, M. Ardabilian, L. Chen, W. Zeng, D. Gu, D. Samaras, "Partial face biometry using shape decomposition on 2D conformal maps of faces", International Conference on Pattern Recognition, Istanbul, Turkey, pp. 1-4, 2010;

2. P. Szeptycki, M. Ardabilian, L. Chen, W. Zeng, D. Gu, D. Samaras, "Conformal mapping-based 3D face recognition", Fifth International Symposium on 3D Data Processing, Visualization and Transmission, Paris, France, pp. 1-8, 2010;

3. X. Zhao, P. Szeptycki, E Dellandréa, L. Chen, "Precise 2.5D Facial Landmarking via an Analysis by Synthesis approach", IEEE Workshop on Applications of Computer Vision (WACV 2009) , Snowbird, Utah, pp. 1-7, ISBN 978-1-4244-5497-6, ISSN 1550-5790, 2009;

4. P. Szeptycki, M. Ardabilian, L. Chen, "A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking", International Conference on Biometrics: Theory, Applications and Systems, Washington, USA, 2009.

National Conferences

1. P. Szeptycki, M. Ardabilian, L. Chen, "Is it a face? How to find and validate a face on 3D scans?", COmpression et REprésentation des Signaux Audiovisuels, Lyon, France, pp. 1-6, 2010;

2. P. Lemaire, P. Szeptycki, M. Ardabilian, L. Chen, "Reconnaissance de visages en 3D orientée région", COmpression et REprésentation des Signaux Audiovisuels, Lyon, France, pp. 1-6, 2010.

Submissions to International Conferences

1. P. Szeptycki, M. Ardabilian, L. Chen, "Nose Tip localization on 2.5D facial models using differential geometry based point signatures and SVM classifier", International Joint Conference on Biometrics, Washington, USA, 2011;

2. P. Szeptycki, D. Samaras, M. Ardabilian, L. Chen, "Automatic 3D face open mouth detection for better near-isometric facial deformations preservation", International Joint Conference on Biometrics, Washington, USA, 2011.

Bibliography

[Al-Osaimi *et al.* 2009] F. Al-Osaimi, M. Bennamoun and A. Mian. *An Expression Deformation Approach to Non-rigid 3D Face Recognition.* Int. J. Comput. Vision, vol. 81, pages 302–316, 2009. 54

[Amor *et al.* 2006] Boulbaba Ben Amor, Mohsen Ardabilian and Liming Chen. *Enhancing 3D Face Recognition By Mimics Segmentation.* Sixth International Conference on Intelligent Systems Design and Applications, vol. 3, pages 150–155, 2006. 54, 73, 74

[Besl & Jain 1986] P.J. Besl and R.C. Jain. *Invariant surface characteristics for 3-d object recognition in range images.* Comput. Vision, Graphics Image Proc., vol. 33, pages 33–80, 1986. 23

[Besl & McKay 1992] Paul J. Besl and Neil D. McKay. *A Method for Registration of 3-D Shapes.* PAMI, vol. 14, no. 2, pages 239–256, 1992. 8

[Bishop 2006] Christopher M. Bishop. Pattern recognition and machine learning (information science and statistics). Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006. 20

[Boehnen & Flynn 2005] Chris Boehnen and Patrick Flynn. *Accuracy of 3D Scanning Technologies in a Face Scanning Scenario.* Proceedings of the Fifth International Conference on 3-D Digital Imaging and Modeling, pages 310–317, 2005. 9

[Bowyer *et al.* 2006] Kevin W. Bowyer, Kyong Chang and Patrick Flynn. *A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition.* Computer Vision and Image Understanding, vol. 101, no. 1, pages 1–15, 2006. ix, 7, 9, 10, 53

[Breitenstein *et al.* 2008] Michael D. Breitenstein, Daniel Küttel, Thibaut Weise, Luc J. Van Gool and Hanspeter Pfister. *Real-time face pose estimation from single range images.* CVPR, pages 1–8, 2008. 17, 37

[Bronstein *et al.* 2005] A. M. Bronstein, M. M. Bronstein and R. Kimmel. *Three-dimensional face recognition.* Int. JOURNAL of Computer Vision, vol. 1, pages 5–30, 2005. 8, 10, 12, 17, 54, 55

[Bronstein *et al.* 2006] Alexander M. Bronstein, Michael M. Bronstein and Ron Kimmel. *Efficient Computation of Isometry-Invariant Distances between Surfaces.* SIAM Journal on Scientific Computing, vol. 26, no. 5, pages 1–28, 2006. 55

[Bronstein *et al.* 2007a] Alexander M. Bronstein, Michael M. Bronstein and Ron Kimmel. *Expression-invariant representations of faces.* IEEE Trans. PAMI, vol. 2004, pages 1042–1053, 2007. 54

[Bronstein *et al.* 2007b] Er M. Bronstein, Michael M. Bronstein, and Ron Kimmel. *Expression-invariant representations of faces.* PAMI, vol. 2004, pages 1042–1053, 2007. 54, 55, 56

[Ceron *et al.* 2010] Alexander Ceron, Augusto Salazar and Flavio Prieto. *Relevance analysis of 3D curvature-based shape descriptors on interest points of the face.* International Conference on Image Processing Theory, Tools and Applications, pages 1–6, 2010. 17, 37, 38, 44

[Chang *et al.* 2006] K. I. Chang, Kevin W. Bowyer and Patrick J. Flynn. *Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 10, pages 1695–1700, 2006. 8, 10, 12, 17, 18, 19, 37, 38, 45, 54

[Chen *et al.* 2005] Liming Chen, Georgy Kukharev and Tomasz Ponikowski. *The PCA reconstruction based approach for extending facial image databases for face recognition systems.* Enhanced methods in computer security, biometric and artificial intelligence systems, 2005. 69, 70, 75

[Cohen-Steiner & Morvan 2003] David Cohen-Steiner and Jean-Marie Morvan. *Restricted delaunay triangulations and normal cycle.* Proceedings of the nineteenth annual symposium on Computational geometry, no. 10, pages 312–321, 2003. 18

[Colbry *et al.* 2005] D. Colbry, G. Stockman and Jain Anil. *Detection of Anchor Points for 3D Face Verification.* Computer Vision and Pattern Recognition, pages 118–118, 2005. 13, 36

[Colombo *et al.* 2006] Alessandro Colombo, Claudio Cusano and Raimondo Schettini. *3D face detection using curvature analysis.* Pattern Recognition, vol. 39, no. 3, pages 444–455, 2006. 8, 9, 12, 17, 18, 19, 21, 36, 37, 38, 45

[Conde *et al.* 2005] Cristina Conde, Roberto Cipolla, Licesio J. Rodríguez-Aragón, Ángel Serrano and Enrique Cabello. *3D Facial Feature Location with Spin Images.* Machine Vision Applications, pages 418–421, 2005. 15, 37

[D'Hose *et al.* 2007] J. D'Hose, J. Colineau, C. Bichon and B. Dorizzi. *Precise Localization of Landmarks on 3D Faces using Gabor Wavelets.* Biometrics: Theory, Applications, and Systems, pages 1–6, 2007. 16, 32

[Dibeklioglu *et al.* 2008] Hamdi Dibeklioglu, Albert Al Salah and Lake Akarun. *3D Facial Landmarking under Expression, Pose, and Occlusion Variations.* BTAS, pages 1–6, 2008. 16

[Dorai & Jain 1995] Chitra Dorai and Anil K. Jain. *COSMOS - A Representation Scheme for 3D Free-Form Objects.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pages 1115–1130, 1995. 39

[Drira *et al.* 2009] Hassen Drira, Boulbaba Ben Amor, Mohamed Daoudi and Anuj Srivastava. *A Riemannian Analysis of 3D Nose Shapes For Partial Human Biometrics.* ICCV, vol. 1, no. 1, pages 1–8, 2009. 74, 75

[Eggert *et al.* 1997] D.W. Eggert, A. Lorusso and R.B. Fisher. *Estimating 3-D rigid body transformations: a comparison of four major algorithms.* Machine Vision and Applications, vol. 9, pages 272–290, 1997. 26

[Faltemier *et al.* 2007] T.C. Faltemier, K.W. Bowyer and P.J. Flynn. *Using a Multi-Instance Enrollment Representation to Improve 3D Face Recognition.* Biometrics: Theory, Applications and Systems, pages 1–6, 2007. 9, 11, 16

[Faltemier *et al.* 2008a] T. Faltemier, W. Bowyer and P.J. Flynn. *A Region Ensemble for 3-D Face Recognition.* IEEE Transactions on Information Forensics and Security, vol. 3, no. 1, pages 62–73, 2008. ix, 8, 9, 11, 16, 54

[Faltemier *et al.* 2008b] Timothy C. Faltemier, Kevin W. Bowyer and Patrick J. Flynn. *Rotated Profile Signatures for Robust 3D Feature Detection.* IEEE Int. Conf. on Automatic Face and Gesture Recognition, 2008. 9, 11, 14, 17, 32

[Floater & Hormann 2005] Michael S. Floater and Kai Hormann. *Surface Parameterization: a Tutorial and Survey.* Advances in multiresolution for geometric modelling, pages 157–186, 2005. 64

[Gatzke & Grimm 2006] Timothy D. Gatzke and Cindy M. Grimm. *Estimating curvature on triangular meshes.* International Journal of Shape Modeling, vol. 12, no. 1, pages 1–28, 2006. 18

[Gokberk *et al.* 2008] Berk Gokberk, Helin Dutagaci, Aydin Ulas, Lale Akarun and Bulent Sankur. *Representation plurality and fusion for 3D face recognition.* IEEE Transactions on Systems Man and Cybernetics Part B, vol. 38, no. 1, pages 155–173, 2008. 76

[Gupta *et al.* 2010] Shalini Gupta, Mia K. Markey and Alan C. Bovik. *Anthropometric 3D Face Recognition.* International Journal of Computer Vision, 2010. ix, 8

[Haker *et al.* 2000] Steven Haker, Sigurd Angenent, Allen Tannenbaum, Ron Kikinis, Guillermo Sapiro and Michael Halle. *Conformal Surface Parameterization for Texture Mapping.* IEEE Transactions on Visualization and Computer Graphics, vol. 6, no. 2, pages 181–189, 2000. 64

[Huang *et al.* 2010] Di Huang, Guangpeng Zhang, Mohsen Ardabilian, Yunhong Wang and Liming Chen. *3D Face Recognition using Distinctiveness Enhanced Facial Representations and Local Feature Hybrid Matching.* International Conference on Biometrics: Theory, Applications and Systems (BTAS), pages 1–6, 2010. 54

[Kakadiaris *et al.* 2007] Ioannis A. Kakadiaris, Georgios Passalis, George Toderici, Mohammed N. Murtuza, Yunliang Lu, Nikos Karampatziakis and Theoharis Theoharis. *Three-Dimensional Face Recognition in the Presence of Facial Expressions: An Annotated Deformable Model Approach.* PAMI, vol. 29, no. 4, pages 640–649, 2007. 17, 54, 55, 71, 72

[Kalogerakis *et al.* 2007] Evangelos Kalogerakis, Patricio Simari, Derek Nowrouzezahrai and Karan Singh. *Robust statistical estimation of curvature on discretized surfaces.* Proceedings of the fifth Eurographics symposium on Geometry processing, pages 13–22, 2007. 18

[Kukharev & Forczmanski 2004] Georgy Kukharev and Pawel Forczmanski. *Data Dimensionality Reduction for Face Recognition.* Machine GRAPHICS and VISION, vol. 13, no. 1, pages 99–121, 2004. 64, 69, 70, 75

[Lu *et al.* 2006] X. Lu, A. K. Jain and D. Colbry. *Matching 2.5D Face Scans to 3D Models.* PAMI, vol. 28, no. 1, pages 31–43, 2006. 8, 13

[Malassiotis & Strintzis 2005] Sotiris Malassiotis and Michael G. Strintzis. *Robust real-time 3D head pose estimation from range data.* Pattern Recognition, vol. 38, no. 8, pages 1153–1165, 2005. 8, 13

[Matthews & Baker 2004] Iain Matthews and Simon Baker. *Active Appearance Models Revisited.* International Journal of Computer Vision, vol. 60, no. 1, pages 135–164, 2004. 15

[Mian *et al.* 2005] A. S. Mian, M. Bennamoun and R. A. Owens. *Region-based Matching for Robust 3D Face Recognition.* British Machine Vision Conf., vol. 1, pages 199–208, 2005. 14

[Mian *et al.* 2006] Ajmal Mian, M. Bennamoun and R. Owens. *3D Face Recognition by Matching Shape Descriptors.* European Conf. on Computer Vision, pages 344–355, 2006. 10, 14

[Mian *et al.* 2007] Ajmal Mian, Mohammed Bennamoun and Robyn Owens. *An Efficient Multimodal 2D-3D Hybrid Approach to Automatic Face Recognition.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 11, pages 1927–1943, 2007. ix, 8, 9, 10, 11, 14

[Moreno *et al.* 2003] A.B. Moreno, A. Sanchez, J.F. Velez and F.J. Diaz. *Face recognition using 3D surface-extracted descriptors.* Irish Machine Vision and Image, 2003. vii, 12, 18, 19, 45

[Morvan 2008] Jean-Marie Morvan. Generalized curvatures. Springer-Verlag Berlin Heidelberg, Berlin, 2008. 18

[Mpiperis *et al.* 2007] Iordanis Mpiperis, Sotiris Malassiotis and Michael G. Strintzis. *3-D Face Recognition With the Geodesic Polar Representation.* IEEE Transactions on Information Forensics and Security, vol. 2, no. 3, pages 537–547, 2007. 54

[Nair & Cavallaro 2007] P. Nair and A Cavallaro. *Region Segmentation and Feature Point Extraction on 3D Faces using a Point Distribution Model.* ICIP, pages 85–88, 2007. 17

[Nair & Cavallaro 2009] Prathap Nair and Andrea Cavallaro. *3-D face detection, landmark localization, and registration using a point distribution model.* IEEE Transactions on Multimedia, vol. 1, no. 4, pages 611–623, 2009. 8, 13, 15, 17, 36, 40

[Niese *et al.* 2008] R. Niese, A. A. Hamadi, F. Aziz and B. Michaelis. *Robust facial expression recognition based on 3-d supported feature extraction and SVM classification.* Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition, pages 1–7, 2008. 8

[Pan *et al.* 2010] Gang Pan, Song Han, Zhaohui Wu and Yuting Zhang. *Removal of 3D facial expressions: A learning-based approach.* CVPR'10, pages 2614–2621, 2010. 54

[Pears 2008] Nick Edwin Pears. *RBF shape histograms and their application to 3D face processing.* IEEE Int. Conf. on Automatic Face and Gesture Recognition, 2008. 15, 37

[Petrovska-Delacrétaz *et al.* 2008] Dijana Petrovska-Delacrétaz, Sylvie Lelandais, Joseph Colineau, Liming Chen, Bernadette Dorizzi, Emine Krichen, Mohamed Anouar-Mellakh, Anis Chaari, Souhila Guerfi, Mohsen Ardabilian, Johan D Hose and Boulbaba Ben Amor. *The IV2 Multimodal Biometric Database (Including Iris, 2D, 3D, Stereoscopic and Talking Face Data) and the IV2-2007 Evaluation Campaign.* BTAS 08, pages 1–7, 2008. 25

[Phillips *et al.* 2005] P. Jonathon Phillips, Patrick J. Flynn, Todd Scruggs, Kevin W. Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min and William Worek. *Overview of the face recognition grand challenge.* Computer Vision and Pattern Recognition, vol. 1, pages 947–954, 2005. xii, 29, 40, 46, 60, 61, 71, 75

[Pottmann *et al.* 2007] Helmut Pottmann, Johannes Wallner, Yong-Liang Yang, Yu-Kun Lai and Shi-Min Hu. *Principal curvatures from the integral invariant viewpoint.* Comput. Aided Geom. Des., vol. 24, no. 1, pages 428–442, 2007. 18

[Queirolo *et al.* 2010] Chaua C. Queirolo, Luciano Silva, Olga R.P. Bellon and Mauricio Pamplona Segundo. *3D Face Recognition Using Simulated Annealing and the Surface Interpenetration Measure*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 1, pages 206–219, 2010. 10

[Samir *et al.* 2006] Chafik Samir, Anuj Srivastava and Mohamed Daoudi. *hree-Dimensional Face Recognition Using Shapes of Facial Curves*. IEEE Trans. Pattern Anal. Mach. Intell., pages 1858–1863, 2006. 54

[Savran *et al.* 2008] Arman Savran, Nese Alyuz, Hamdi Dibeklioglu, Oya Celiktutan, Berk Gokberk, Bulent Sankur and Lale Akarun. *Bosphorus Database for 3D Face Analysis*. Biometrics and Identity Management: First European Workshop, pages 47–56, 2008. xii, 16, 29, 30, 60, 61

[Shi *et al.* 2006] J. Shi, A. Samal and D. Marx. *How effective are landmarks and their geometry for face recognition?* Computer Vision and Image Understanding, vol. 102, no. 2, pages 117–133, 2006. 8

[Soltana *et al.* 2010] Wael Ben Soltana, Di Huang, Mohsen Ardabilian and Liming Chen. *Comparison of 2D/3D Features and Their Adaptive Score Level Fusion for 3D Face Recognition*. 3D PVT, pages 1–8, 2010. 8

[Sun & Yin 2008a] Y. Sun and L. Yin. *Facial Expression Recognition Based on 3D Dynamic Range Model Sequences*. ECCV, 2008. 8

[Sun & Yin 2008b] Yi Sun and Lijun Yin. *Automatic Pose Estimation of 3D Facial Models*. Int. Conf. Pattern Recognition, 2008. 8, 12, 17, 18, 19, 45

[Szeptycki *et al.* 2009] Przemyslaw Szeptycki, Mohsen Ardabilian and Liming Chen. *A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking*. BTAS, pages 1–6, 2009.

[Szeptycki *et al.* 2010] Przemyslaw Szeptycki, Mohsen Ardabilian, Liming Chen, Wei Zeng, David Gu and Dimitris Samaras. *Conformal mapping-based 3D face recognition*. 3DPVT, pages 1–8, 2010.

[Tang *et al.* 2008] X. M. Tang, J. S. Chen and Y. S. Moon. *Accurate 3D face registration based on the symmetry plane analysis on nose regions*. European Signal Processing Conference, 2008. 14

[Tong & Tang 2005] Wai-Shun Tong and Chi-Keung Tang. *Robust Estimation of Adaptive Tensors of Curvature by Tensor Voting*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, pages 434–449, 2005. 18

[Toponogov & Rovenski 2006] Victor Andreevich Toponogov and Vladimir Y. Rovenski. Differential geometry of curves and surfaces: A concise guide. Birkhauser, 2006. 19, 20, 45

[Trucco & Verri 1998] Emanuele Trucco and Alessandro Verri. Introductory techniques for 3-d computer vision. Prentice Hall PTR, Upper Saddle River, NJ, USA, March 1998. vii, ix, 19, 23, 24, 27, 37, 38, 39, 45

[Umeyama 1991] S. Umeyama. *Least-squares estimation of transformation parameters between two point patterns.* Transactions on Pattern Analysis and Machine Intelligence, vol. 13, pages 376–380, 1991. 26

[Wang *et al.* 2006] Sen Wang, Yang Wang, Miao Jin, Xianfeng Gu and Dimitris Samaras. *3D Surface Matching and Recognition Using Conformal Geometry.* IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 2453–2460, 2006. 53, 54, 55, 65, 68, 71, 72, 74, 76

[Wang *et al.* 2008] Yueming Wang, Xiaoou Tang, Jianzhuang Liu, Gang Pan and Rong Xiao. *3D Face Recognition by Local Shape Difference Boosting.* ECCV, pages 603–616, 2008. 14

[Wang *et al.* 2009] Yueming Wang, Jianzhuang Liu and Xiaoou Tang. *Robust 3D Face Recognition by Local Shape Difference Boosting.* PAMI, vol. 99, pages 1–15, 2009. 9, 14

[Wen *et al.* 2006] Gaojin Wen, Zhaoqi Wang, Shihong Xia and Dengming Zhu. *Least-squares fitting of multiple M-dimensional point sets.* The Visual Computer, vol. 22, no. 6, pages 387–398, 2006. 26

[Xu *et al.* 2006] Chenghua Xu, Tieniu Tan, Yunhong Wang and Long Quan. *Combining local features for robust nose location in 3d facial data.* Pattern Recognition Letters, vol. 27, no. 13, pages 1487–1494, 2006. 15, 17, 37

[Yang *et al.* 2004] Jian Yang, David Zhang, Alejandro F. Frangi and Jing yu Yang. *Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition.* PAMI, vol. 26, no. 1‘, pages 131–137, 2004. 69, 70, 75

[Yin *et al.* 2006] Lijun Yin, Xiaozhou Wei, Jun Wang and Matthew J. Rosato. *A 3D Facial Expression Database For Facial Behavior Research.* Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, pages 211–216, 2006. 12

[Yoshida *et al.* 2002] Hiroyuki Yoshida, Janne Näppi, Peter MacEneaney, David T. Rubin and Abraham H. Dachman. *Computer-aided Diagnosis Scheme for Detection of Polyps at CT Colonography.* Imaging and Therapeutic Technology, vol. 22, no. 4, pages 963–979, 2002. x, 40, 41

[Zhang & Zhou 2005] Daoqiang Zhang and Zhi-Hua Zhou. *(2D)2PCA: 2-Directional 2-Dimensional PCA for Efficient Face Representation and Recognition.* Neural Networks in Signal Processing, vol. 69, pages 224–231, 2005. 70

[Zhao *et al.* 2003] W. Zhao, R. Chellappa, P. J. Phillips and A. Rosenfeld. *Face recognition: A literature survey.* ACM Comput. Surv., vol. 35, no. 4, pages 399–458, 2003. 53

[Zhao *et al.* 2009a] Xi Zhao, Emmanuel Dellandréa and Liming Chen. *A 3D statistical facial feature model and its application on locating facial landmarks.* Advanced Concepts for Intelligent Vision Systems (ACIVS 2009), pages 686–697, 2009. 9, 11, 17

[Zhao *et al.* 2009b] Xi Zhao, Przemyslaw Szeptycki, Emmanuel Dellandréa and Liming Chen. *Precise 2.5D Facial Landmarking via an Analysis by Synthesis approach.* IEEE Workshop on Applications of Computer Vision (WACV 2009), pages 1–7, 2009. x, 9, 11, 15, 16, 17, 31, 32, 33, 34, 36

[Zhao *et al.* 2010] Xi Zhao, Di Huang, Emmanuel Dellandréa and Liming Chen. *Automatic 3D facial expression recognition based on a Bayesian Belief Net and a Statistical Facial Feature Model.* ICPR, pages 1–4, 2010. 8