3D Face Morphology Classification for Medical Applications

Thesis submitted to Cardiff University
in candidature for the degree of Doctor of Philosophy

Hawraa Hassan Abbas

School of Engineering
Cardiff University
2018

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.
Signed................................ (candidate) Date ....

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.
Signed .... (candidate) Date ....

STATEMENT 2

This thesis is the result of my own investigation, except where otherwise stated. Other sources are acknowledged by giving explicit reference.
Signed .... (candidate) Date ....

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organizations.
Signed .... (candidate) Date ....

STATEMENT 4

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, after expiry of a bar on access approved by the Graduate Development Committee.
Signed................................ (candidate) Date ....

ABSTRACT

Classification of facial morphology traits is an important problem for many medical applications, especially with regard to determining associations between facial morphological traits or facial abnormalities and genetic variants. A modern approach to the classification of facial characteristics (traits) is to use three-dimensional facial images. In clinical practice, classification is usually performed manually, which makes the process very tedious, time-consuming and prone to operator error. Also using simple landmark-to-landmark facial measurements may not accurately represent the underlying complex three-dimensional facial shape.

This thesis presents the first automatic approach for classification and categorisation of facial morphological traits with application to lips and nose traits. It also introduces new 3D geodesic curvature features obtained along the geodesic paths between 3D facial anthropometric landmarks. These geometric features were used for lips and nose traits classification and categorisation. Finally, the influence of the discovered categories on the facial physical appearance are analysed using a new visualisation method in order to gain insight into suitability of categories for description of the underlying facial traits.

The proposed approach was tested on the ALSPAC (Avon Longitudinal Study of Parents and Children) dataset consisting of 4747 3D full face meshes. The classification accuracy obtained using expert manual categories was not very high, in the region of 72%-79%, indicating that the manual categories may be unreliable. In an attempt to improve these accuracies, an automatic categorisation method was applied. In general, the classification accuracies based on the automatic lip categories were higher than those obtained using the manual categories by at least 8% and the automatic categories were found to be statistically more significant in the lip area than the manual categories.

The same approach was used to categorise the nose traits, the result indicating that the proposed categorisation approach was capable of categorising any face morphological trait without the ground truth about its traits categories.

Additionally, to test the robustness of the proposed features, they were used in a popular problem of gender classification and analysis. The re-

sults demonstrated superior classification accuracy to that of comparable methods.

Finally, a discovery phase of a genome wide association analysis (GWAS) was carried out for 11 automatic lip and nose traits categories. As a result, statistically significant associations were found between four traits and six single nucleotide polymorphisms (SNPs). This is a very good result considering that for the 27 manual lip traits categories provided by medical expert, the associations were found between two traits and two SNPs only. This result testifies that the method proposed in this thesis for automatic categorisation of 3D facial morphology has a considerable potential for application to GWAS.

I dedicate this thesis to my mother spirit,
to my late brother (Shalan), who is my role model, and
to my loving father, husband, children and other free-minded citisen of
Iraq.

ACKNOWLEDGEMENTS

LIST OF ABBREVIATIONS

2D	two-dimensional
3D	three-dimensional
3D-FER	3D Facial Expression Recognition
3D-FR	3D Face Recognition
ALSPAC	Avon Longitudinal Study of Parents and Children
ASM	Active Shape Model
BRIM	Bootstrapped Response-based Imputation Modeling
CH	Calinski–Harabasz index
CSL	Cost Sensitive Learning
DB	Davies Bouldin index
DI	Dunn index
DNA	deoxyribonucleic acid
DSM	Dense Surface Modeling
DZ	dizygotic
FOP	Fibrodysplasia Ossificans Progressiva
G	Gaussian curvature
GESSA	Geodesic Ensemble Surface Sampling Algorithm
GPA	Generalised Procrustes Analysis
GWAS	Genome Wide Association Study
H	mean curvature
HCI	Human-Computer Interaction
HK	Mean curvature and Gaussian curvature
k_1, k_2	principal curvatures
LDA	Linear Discriminant Analysis
MMP	Mitchell–Mount–Papadimitriou (method)
MZ	monozygotic
NMF	Non-negative Matrix Factorisation
NS	Noonan syndrome
PC	Principal Component
PCA	Principal Component Analysis
PLSR	Partial Least Squares Regression

R	curvedness
SHREC2008	3D Shape Retrieval Contest 2008 database
S_i	shape index
SI	silhouette index
SNP	Single Nucleotide Polymorphism
SVM	Support Vector Machine

List of Figures

List of Tables

CONTENTS

Chapter 1

INTRODUCTION

The development of facial morphology and its associated characteristics is the result of genetic and environmental interactions. Some characteristics are subject to greater influence from genetics; for example, face height and chin prominence are highly heritable [27, 144, 181, 182]. Other traits are subject to greater influence from environment; for example, urban pollution (e.g., vehicle exhaust fumes and cleaning products) is known to affect our facial development [40, 72, 233]. Twin studies provide evidence for discovering the relative contributions of genetics and environment on the individual's facial and holistic development. Monozygotic (MZ) twins share nearly 100% of their genes, which means that the differences between the twins are likely to be caused by the environment. Dizygotic (DZ) twins share about 50% of their genes [176, 215].

Generally, the shape of our face is mainly genetically determined. Therefore, to understand how the face morphology is defined genetically we need to determine the differences in facial appearance and link those patterns of variation to genetic influences. The most important purpose of understanding the genome is the ability to explore how genetic conditions related to certain diseases and syndromes; this enables us to develop new and functional ways to improve human health. Most previous studies in clinical practice tried to *manually* classify facial morphological characteristics (traits) which is time consuming and prone to operator error. In addition to that many studies used simple measurements such as Euclidean distances and angles between face landmarks for facial traits classification purpose (e.g., see [10, 158, 186, 201]). These features are easy to extract but poor in capturing facial 3D surface appearance.

To address these problems, this work proposes a new set of 3D features capable of representing complex 3D facial shapes. It also proposes

a new automatic approach for facial morphological traits classification and categorisation which may be used in conjunction with the analysis of relationships between facial traits and genetic variations based on these features.

1.1   Research objectives

The main objectives addressed in this thesis are:

- Proposing a new set of 3D geometric features more suitable for face morphology traits analysis and classification than the features used currently in clinical practice.

- Developing an effective automatic facial morphological traits classification and categorisation approaches.

- Applying the proposed automatic approaches to a number of face morphological traits.

- Proposing an approach to validate the discovered face morphological traits categories.

- Conducting a genome wide association study (GWAS) for automatic face morphological traits categories.

1.2   Thesis contributions

The contributions of the thesis can be summarised as follows:

- Establishing a set of novel 3D geometric features based on curvatures and geodesic paths between the facial anthropometric landmarks which improve the classification accuracy of 3D face morphology.

- Developing a new gender classification approach achieving higher prediction accuracy comparing with the state of art comparable methods using the novel 3D geometric features.

- Developing a new approach for automatic categorisation of morphological traits in a large population with application to lips and nose traits.

- A new method for visualising the influence of the discovered categories on the facial physical appearance to gain insight into the suitability of categories for the description of the underlying facial traits by using PLSR (Partial Least Squares Regression) approach.

- Performing a genome wide association study (GWAS) of the automatic lips and nose traits categories and analysis the results, to discover which geometric traits have a genetic basis.

1.3 Thesis overview

The remainder of this thesis is structured as follows:

- In Chapter 2 a review of the relevant literature relating to studying human face morphology, the existing methods of recording and analysis of facial morphology are presented. The literature pertinent to lips and nose morphological traits characterisation is also explored. It explores the research that addresses the gender classification and finds the relation between the gender and the face morphology characteristics. The applications of 3D face curvatures and geodesic paths in computer vision are explored in depth. Finally, the concept of genome wide association study (GWAS) is also investigated.

- The 3D dataset specification and its preprocessing (regularised and non-regularised) is presented in Chapter 3.

- A novel set of 3D geometric set of features is presented in Chapter 4. These features are a unique combination of geodesic path curvatures between the anthropometric landmarks.

- Chapter 5 focuses on using the 3D geometric features presented in Chapter 4 for lips trait classification and categorisation (which is also called clustering or unsupervised classification). Moreover, the clustering validation methods are explored and a new method is adopted to determine and visualise the effect and statistical significance of automatic lips categories on the lips area using the Partial least squares regression (PLSR) algorithm.

- In Chapter 6 a gender classification method is proposed using the geometric features explained in Chapter 4. The gender classification performance is compared with that using the state of art measurements (Euclidean and geodesic distances). Establishing the relation between face traits and gender for teenage Caucasian population is the important contribution in this chapter.

- Chapter 7 focuses on using the 3D geometric features presented in Chapter 4 for nose morphological traits categorisation. As in Chapter 5, the clustering validation methods are explored and Partial least squares regression (PLSR) method is used to determine the statistical significance of automatic nose categories on the appearance of the nose area. This chapter's main contribution is performing automatic nose trait categorisation without previous knowledge about the traits categories.

- Chapter 8 presents the results of the candidate gene association study. In this chapter, the automatic lips and nose traits categories proposed in Chapters 5 and 6 are used in GWAS. A simple method to visualise traits variation is proposed, this method can be used to visualise the morphological traits variation and the genes variation influence on the 3D face.

- Chapter 9 provides the conclusions and suggestions for future work.

1.4 Publications arising from this study

Below is a list of publications based on the novel contributions listed on the previous page.

1. Hawraa Abbas, Yulia Hicks, and David Marshall, Automatic classification of facial morphology for medical applications, *Procedia Computer Science* 60 (2015): 1649–1658.

2. Hawraa Abbas, Yulia Hicks, David Marshall, Alexei Zhurov, and Stephen Richmond, A 3D morphometric perspective for facial gender analysis and classification using geodesic path curvature features, *Computational Visual Media* (in press).

3. Hawraa Abbas, Yulia Hicks, Alexei Zhurov, David Marshall, Peter Claes, Caryl Wilson and Stephen Richmond, An automatic approach for classification and categorisation of lip morphology, *PLOS One* (submitted).

4. Hawraa Abbas, Yulia Hicks, David Marshall, Alexei Zhurov, Peter Claes and Stephen Richmond, Unsupervised nose morphological traits categorisation, *Journal of Anatomy* (to be submitted).

5. Hawraa Abbas, Yulia Hicks, Alexei Zhurov, David Marshall and Stephen Richmond, Genetic association for lips and nose traits phenotypes, *PLoS Genetics* (to be submitted).

Chapter 2

LITERATURE REVIEW

The scope of this thesis is fairly large and encompasses many different areas of research, from 3D facial morphology feature extraction to genes 3D face morphological traits association, via machine learning techniques. The main objective of this chapter is not to detail every technical field used in this thesis but rather focus on methods and strategies used for face morphology analysis and classification. Our primary aim is the analysis of the human facial morphology, from understanding of the face shape and its parts, to finding relationships between face morphology and gender, and finding the genetic association through automatic 3D facial morphology classification and categorisation.

This chapter follows the journey made throughout the Ph.D, finding the state of the art for each particular problem has been essential steps to determine the gaps in research that is aimed to fill: automatic 3D face morphology classification and analysis of their interaction with genotype. Here, different theoretical areas are linked to solve a particular problem and background information related to some of the tools or techniques used in this work will be presented separately throughout the thesis when appropriate.

First, an overview of the history of face morphology analysis is given. Then, the next sections are used to review some related topics like 3D face morphology analysis systems and brief description for 3D face morphology classification and its applications, a few pioneering papers related to thesis topics are presented. The local 3D geometric feature descriptors are highlighted in Section 2.4. The lips and nose traits classification literature review are presented in Sections 2.5 and 2.7. The relationship between gender variation and face morphology is closely looked in Section 2.6. A review on gene association and heritability literature is then given.

Finally, gaps in the current state of the art are analysed in Section 2.8.

2.1   Facial morphology overview

Face morphology analysis is a one of the oldest areas of research in art and science. It is hard to determine who the first person to measure a face. In fact, the study of facial morphology is believed to have originated in ancient Egypt more than 4500 years ago. Later artists and philosophers researched the ideal body and face proportions that society often uses to define beauty. The analysis of face proportion in art continued for a long period and still exists today [285]. However, the neoclassical canons of Leonardo da Vinci were one of the first attempts to define the proportions of the human body and head, see Fig. 2.1 [304].

Moreover, understanding of the face with the emergence of modern anatomy in the 18th century. Knowledge was gained regarding the mechanical interaction muscles and bones behind the skin that is fundamental in many face related field ofstudy, particularly the facial shapes and expressions studies [276]. In the Netherlands, Petrus Camper [284] was recognised for his theory of the facial angle. He determined that modern humans have facial angles between 70 and 80 degrees, with African and Asian angles closer to 70 degree, and European angles closer to 80 degrees. According to Camper's new portraiture technique, the facial angle is created by drawing two lines: one horizontally from the ear to the nostril; and the other perpendicularly from the advancing part of the upper jawbone to the most prominent part of the forehead. Blumenbach in 1776 [35] followed soon thereafter by establishing the formal system of analysing human skulls. On the basis of his craniometric measurements, Blumenbach divided the human species into five races: Caucasian or white race, Mongolian or yellow race, Malayan or brown race, Negroid or black race, and American or red race.

The fields of neurology, psychology and sociology, that flourished in the 20th century started to ask basic questions about how humans can see and identify faces and his capabilities limitation and how important is the face in human social interactions [304]. The ability to recognise faces from birth to adulthood has been studied, in normal individuals as well as in people with face syndromes [53, 255]. Furthermore, psychological

research has been performed to define our ability to judge gender, age, attractiveness and emotion from the facial shape and proportion [241].

Many research studies have found notable differences between the facial morphological proportions described in modern non-Caucasian ethnic populations [95–97, 308]. These investigations into the applicability of these proportions have generated considerable amounts of data on the facial dimensions of numerous ethnic groups from multiple regions around the world. Nevertheless, not all characteristics of the human face can be described by simple metrics, so the door is still open for categorical scales describing shapes of the entire face and its elements.
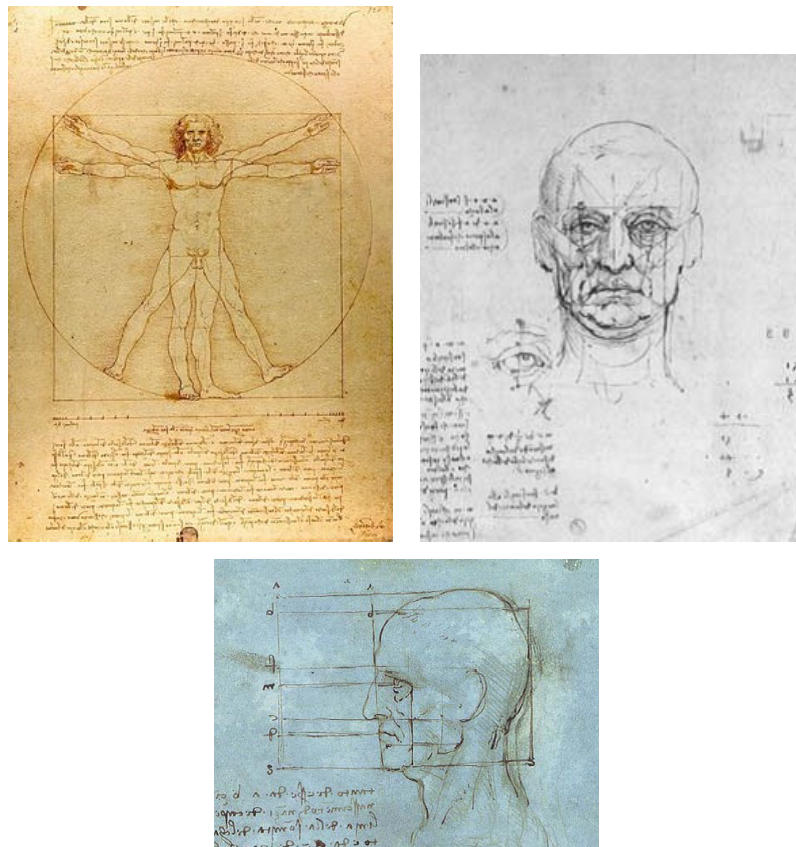
Figure 2.1: The oldest recorded body, face and head measurements record [304]

2.2 Face morphology recording techniques

Facial morphology is the study of facial form, shape and its structure. Analysis of the human face has a long tradition, as discussed earlier. Dif-

ferent techniques have been applied to analysing facial morphology and assessing growth development of the face for various medical purposes such as facial syndrome diagnosing, surgical treatment planning, facial asymmetry and facial anomalies genes association.

2.2.1   Visual assessment (Manual anthropometric)

Visual assessment is one of the oldest methods of examination that is still in use in medicine today. Visual assessment is the art of discovering or judging of a person's character from observing his visible features; it is a form of anthropology based on visual observation of the physical characteristics of the human body and exact measurements carried out in anthropometric measurements. The method has a high degree objective assessment of some characteristics for instances: skin color; hair color and eye color are among the more common characteristics assessed visually [95]. Also, this method can be used to characterise facial morphology, for example, Kataria at el. [146] studied facial index of North Indian population and to find out the distribution of their face type. Facial index is the facial height measured from Nasion to Menton with the help of sliding caliper and the facial width measured between the Zygion of each side with the help of spreading caliper, this measurement was conducted on a total number of 400 subiects (200 males and 200 females) of 18–25 years of age. Using the same measurement tool, Al-Sebaei [10] conducted a study on a group consisted of 168 healthy, Saudi Arabian dental students originating from the Arabian Peninsula (93 males and 75 females, age 20–24 years). By using a caliper, three neoclassical facial traits were measured to analyse different traits in this ethnic group, these traits are the vertical thirds of the face, the orbital trait, and the orbitonasal trait (intercanthal distance = nasal width). Visual assessment has also been used for research purposes in evaluating deformities such as cleft lip and palate deformity [158]. This method is accurate to some degree but requires special instruments and expert clinicians to make the measurements. The face curvature also can not be measured using this approach.

2.2.2   Two-dimensional (2D) photographs

The initial aim of photography is to provide a visual record of a particular object or condition at a particular time. A well taken photograph records the external appearance of health, disease or deformity. Moreover, there has always been the need for photographic records for purposes of research and publication, and for teaching and presentations. However, this method has faced some difficulties and limitations due to the varying degrees of resolution and accuracy of different photographic techniques. Limiting factors of 2D images are the effect of pose variations and it is sensitive to illumination changes, this effect on accurate manually or automatically landmarks placement.

In spite of that photography still the cheapest technique for face recording. Photographs have been used by researchers to carry out facial morphology analysis via identifying certain landmarks on facial structures and extracting anthropometric measurements such as distances, angles, and ratios. Researchers [37, 176] worked on face measurements such nose width to identify genes influencing facial variation. Two-dimensional image techniques are very popular for after surgery assessment and treatment. For example, Mckearney et al. [186] used the lips images for forty-four participants aged 10 years, to access their upper lip symmetry after cleft lip surgery. In general, this method is a common but 2D photographs do not capture 3D face shape, which contains important morphological information.

2.2.3   Three-dimensional (3D) imaging technique

With the emergence of 3D imaging technique, it has become possible to obtain a life-like 3D image of the face with natural texture and color. Three dimensional imaging techniques which appropriate for capturing facial data can be broadly divided into surface imaging and volumetric images. Stereophotogrammetry and laser scanners are the two most commonly used surface imaging systems. Stereophotogrammetry projects a light pattern or mesh onto different regions of the face whereas laser scanners use laser rays to acquire depth information. Volumetric images may provide bony structure as well as soft-tissue facial information [137]. The understanding of facial morphology has been improved

dramatically with the development of precise and reliable 3D acquisition systems [278]. As previously discussed, face morphology analysis using 2D images is sensitive to illumination changes, but in 3D variations in illumination only affect the texture of the face, but the captured facial shape remains intact [7].

Many 3D imaging acquisition techniques and sophisticated software tools are currently available. The acquisition quality is mainly linked to the accuracy of recovering the z-coordinate (called depth information). Moreover, all systems suffer from a potential movement of the participant and facial expression changes between the multiple views that needed to construct a 3D model of the face [77]. Table 2.1 outlines the merits and demerits of popular systems.

Table 2.1: Comparison among 3D systems

Device	Advantages	Disadvantages
Traditional Stereo-Photogrammetry (3dMDface^{TM} System) [145, 147]	The images are captured quickly by two or more cameras take photos from different angles and stitch the images together into a 3D image	Low resolution 3D image quality
Laser [137]	Very accurate and high resolution 3D image produce by detecting the light source which is deflected onto a patient's face and a detector captures the distorted light	Expensive equipment and Technique-sensitive
Structured Light Technique [145, 147]	Very rapid 3D image is captured by projecting "structured light" onto the surface of the individual. Cameras capture the distorted light as it reflects from the surface	Technique-sensitive and Photorealistic varying resolution quality
Video-imaging [137, 145]	The 3D image is recorded by utilizing video-camera to capture number of frames per second, also the Speech and the animation can be captured	Low resolution, High memory capacity and time consuming for processing

Numerous studies has compared the performances of 3D imaging devices and the robustness of their data outputs. For example, Knoops at el. [156] conducted a study to compare the output quality of four 3D scanners. Head shapes of eight adults were captured with different 3D scanners: 3dMDface System, 3dMD Inc.; M4D Scan, Rodin4D; Structure Sensor, Occipital Inc.; and 1.5T Avanto MRI, Siemens; he discovered that: 3dMDface System and M4D Scan provide the high quality results. While

Schwenzer at el. [242] investigated accuracy and precision of two contacts free 3D imaging systems (white light vs laser). The laser system is superior concerning robustness but the complex white light system showed better accuracy at 0.2 s measuring time. In this work, laser based data is used.

2.3   Facial morphology analysis and classification methods

Facial morphology analysis and classification is important in many medical application. These applications include studying facial anomalies in comparison with normal facial morphology, recognising the facial key components of the particular syndrome, evaluating average facial growth in a cohort of subjects, studying the relation between face genes and face morphology, comparing facial morphologies for different ages, gender and ethnicity etc. Representing and extracting good quality facial features is an essential step in facial traits analysis and classification. For this purpose, two methods for facial morphology analysis have used:

- Dense surface model this method is computationally efficient as they reduce the computation space of the 3D faces by describing them with fewer dimensions. However, this method is not discriminative enough when the faces have small differences, classification of very similar facial traits or working on facial traits parts.

- Anthropometric measurements, this method is recommended for analysis of facial morphology by using measurements including distances, angles, ratios and proportions [95]. Anthropometric measurements remain a simple, inexpensive, efficient method for describing facial morphology. However, such measurements are unable to capture many face surface details and therefore discard or ignore many important facial shape information.

2.3.1   3D facial dense surface modeling

The 3D dense model can be used to analyse face morphology by establishing correspondence of thousands of points across each 3D face image. Prior to finding the dense correspondence among them, they need to be fitting them in the same virtual space by removing the translation,

rotation and size difference, as well as the registration process, to make all faces appropriately aligned [39].

Different approaches have been reported for constructing dense correspondences between 3D objects. For example, by extracting a set of landmarks on the 3D surfaces, then these landmarks are triangulated in a standardized manner and refined to construct the new mesh. Surface correspondence and triangulation are computed simultaneously in a hierarchical way, this method accuracy depends critically on the accuracy of the initially extracted landmarks [291]. Another approach was introduced by Kelemen et al. [150], where a parametric surface description for objects of spherical topology was developed. This method follows the pioneering work of Taylor and Cootes [71] on active shape models, but is based on a hierarchical parametric object description rather than a point distribution model.

The majority of the reported research in the field of facial morphology analysis establishes dense correspondence models through a deformation procedure [17, 65, 123, 188] in which a template or a generic model is deformed to the surface of an individual and dense correspondences are constructed by analysis the deformed template or generic model. Much research has been conducted using dense 3D models for analysis facial morphology variation in the wide medical application such as syndrome and face abnormality diagnosing, studying the heritability of the facial traits, and exploring the relationship between the face morphology variation and biological genes.

The dense surface model has been used successfully in studying facial abnormality syndrome and the caused heritable genes. For examples, Hammond et al. [123] utilised the dense surface model to visualise and recognise shape differences in a collection of 3D face images that includes 280 controls, 90 individuals with Noonan syndrome (NS), and 60 individuals with Velo-cardio-facial syndrome. In 2012, the same author analysed the facial traits for the Fibrodysplasia Ossificans Progressiva (FOP) syndrome, by applying dense surface modeling (DSM) technique [124].

Dense face models can be efficiently and effectively used to compare differences in facial morphologies for various populations and sexes. In general, various facial differences were spotted between the population groups, these differences could be seen in the lips, nasal, and lower facial

regions. Kau et al. [148] assessed the face variation in 473 subjects were recruited from 5 countries: Hungary, United Kingdom, Wales; United States, Slovenia, and Egypt. The linear differences between face surface models ranged from 0.37 to 1.00 mm in the male groups and from 0.28 to 0.87 mm in the women groups. Also Gor et al. [113] used the average faces as templates to compare the morphological differences at 200 subjects from 2 white population groups (Hungary and US), the mean facial differences were 0.55 ± 0.60 mm between the US and Hungarian women, and 0.44 ± 0.42 mm between the US and Hungarian men. The ranges of differences were -2.02 to 3.77 and -2.05 to 1.94 mm for the female and male pairings, respectively. Talbert et al. [269] used the dense surface face model for 200 subjects to compare facial morphologies of an adult African-American population to an adult Caucasian-American population using 3D surface imaging. He found the most distinct differences were in the forehead, alar base, and periocular regions and the average difference between African-American and Caucasian-American females was 1.18 ± 0.98 mm.

Strong evidence exists supporting the genetic influence on facial traits [157]. Much research has been conducted to find the relation between face morphology and biological genes and the heritability. Claes et al. [63] used dense quasi-landmarks to measure face shape in population subjects from three locations (United States, Brazil, and Cape Verde). Claes presented preliminary results and a validation strategy to create DNA-based facial composites. The effects of sex, genomic-ancestry, and 24 individual SNPs (Single Nucleotide Polymorphism) were modeled with respect to facial morphology using bootstrapped response-based imputation modeling (BRIM) [62].

The similarity of facial appearance within families, across many generations, establishes the idea that certain key genes exert particularly large effects on facial shape and appearance. Numerous studies have estimated facial shape heritability using anthropometric measurement [81, 153]. Some researchers have used the facial surface model to measure the heritability successfully. For example, a three-dimensional facial models have been acquired on a cohort of 952 twins recruited from the Twins UK data, then GESSA (Geodesic Ensemble Surface Sampling Algorithm) was used to place thousands of landmarks throughout the facial

surface and establishes point-wise correspondence across faces, yielding a heritability maps of the human face [281]. Koppen [160] produced measurement analysis for the 3D face model based on defining a set of coherent parts using non-negative matrix factorisation technique (NMF) whose shapes exhibit a large genetic influence [161], the shape variation in each part was used to study facial morphological features that are most likely heritable on twin data samples.

2.3.2   Anthropometric measurements

Facial anthropometric measurements based on identifying specific facial landmarks allow the quantification of changes in facial morphology as a result of growth or healthcare intervention. All faces share a set of landmarks all of which have a particular biological meaning. Facial landmarks can be divided into three broad categories [252]: biological or anthropometric landmarks, mathematical landmarks, and pseudo landmarks.

- Biological or Anthropometric landmarks, which are often used by scientists and physicians, are meaningful points that are defined as standard reference points on the face and head, such as the nasion, inner and outer canthi of the eyes, pronasale, subnasale, center of the upper lip, center of the lower lip, outer corners of the mouth (cheilions), and a chin point (pogonion) [92, 123]. Figure 2.2 illustrates major surface anthropometric landmarks.

- Mathematical landmarks which are defined according to certain mathematical or geometric properties of human faces, such as the middle point between two biological landmarks [190].

- Pseudo facial landmarks which are defined based on two or more mathematical or anatomical landmarks or hair contours. These landmarks do not have particular defined biological positions and they are relatively easy to acquire using computational methods [190], They are generally accurate enough for appearance based face recognition techniques.

Landmark localisation is often performed manually, but much research has been carried out on automatic accurate localisation. For example, Ruiz et al. [235] presented an algorithm for automatic localisation

of landmarks on 3D faces by combining the Active Shape Model (ASM) with the local shape model to jointly find the locations of a set of 14 facial landmarks. Similarly, Sang-Jun et al. [211] applied the Active Shape Models to extract the position of the eyes, the nose, and the mouth. the Principal Curvatures, Mean and Gaussian Curvature, the derivatives and the Shape and Curvedness Indexes used for 9 soft-tissue landmarks automatic identification [286]. Koppen [160] presented a pose-invariant method for the estimation of landmark positions in 3D images of faces.

The anthropometric evaluation of facial morphology begins with the identification of landmarks. A series of measurements between these landmarks are then taken using carefully specified procedures. As a result, repeated measurements of the same individual are very reliable, and measurements of different individuals can be successfully compared. The accuracy of these measurements depends dramatically on landmarks localisation and selection. Morecroft [197] conducted a study based on analysis predefined anthropometric facial landmarks to evaluate 3D shape analysis for facial identification. The results showed that 27 reproducible facial landmarks are important for facial comparison and identification. Among these landmarks are glabella, pogonion, endocanthion, exocanthion, cheilion, and stomion. Also, the influence of landmark labeling on the accuracy and precision of an indirect facial anthropometric was determined in [22]. Eighteen standard linear facial measurements were obtained from 19 anthropometric landmarks for 10 adults subjects, with landmarks labeled and without landmarks labeled before image acquisition, these were compared with anthropometric measurements. The recorded results showed no huge difference in accuracy and precision for the both labeling strategy.
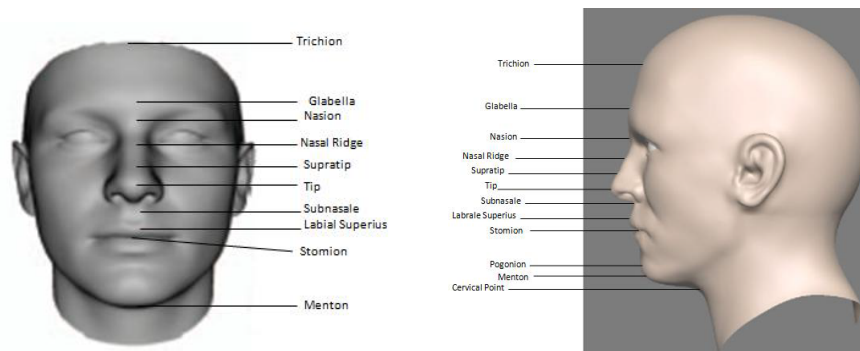
Figure 2.2: Facial anthropometric landmarks

Farkas et al. [95] described a total of 132 measurements on the face and head. Some measurements are paired, where there is a corresponding measurement on the left and right sides of the face. Such measurements are:

- The Euclidean distance between facial landmark pairs.

- Geodesic distance between two landmarks, the distance measured along a prescribed (shortest) path on the surface of the face.

- The angle of inclination between two landmarks with respect to one of the canonical axes. As an example is the inclination of the ear axis with respect to the vertical.

- The angle between locations, such as mento-cervical angle at the chin.

The systematic collection of anthropometric measurements has made possible a variety of statistical 3D face morphology investigations on groups of subjects for many computer vision and medical applications such as gender, race, or age classification, attractiveness and syndrome diagnosis. Seo et al. [244] characterized the facial dimensions of 144 Korean children three-dimensional imaging. Cluster analysis was performed on 16 facial dimensions to categorize them into different sizes, facial shapes were classified into three clusters: small, medium, and large. For genes association, Toma [276] used 250 measurements between 21 facial landmarks (Euclidean distance, angle, and ratio) using the ALSPAC dataset. A strong genetic association was identified between the common SNP rs7559271 in PAX3 gene on chromosome 2 and the 3D facial distance 'nasion to mid-endocanthion' (n-men). Wilamowska et al. [294] studied the genes that caused the 22q11.2DS facial syndrome on 53 subjects using 10 Euclidean measurements, these measurements were extracted from the lips and nose landmarks. To study the relationship between the anthropometric measurements of facial morphology and heritability, Djordjevic et al. [81] conducted a study at 1380 twin subjects, 1275 linear distances measured between 51 landmarks (37 manually identified and 14 automatically calculated), 1222 distances showed evidence of heritability control.

Some studies have attempted to compare between anthropometric measurements and face dense model for face morphology analysis. For example, Alqattan et al. [13] compared landmark and surface based three-dimensional analyses for facial asymmetry and they concluded facial asymmetry can be accurately quantified using landmark and surface based approaches. In contrast, Penget et al. [215] demonstrated that dense face model may substantially improve the detection and characterization of genetic association in common facial variation rather than the Euclidean anthropometric measurements between facial landmarks. Therefore, in this work, a new set of 3D face geometric features are developed, these features combine the advantages of anthropometric landmarks and surface shape information.

2.4   3D local feature descriptors

Local features have proven to be very successful in many vision tasks such as 3D object categorization and recognition, 3D modeling and scene reconstruction, 3D model retrieval and shape analysis and 3D biometrics [33, 42, 164, 171].

Local features have been widely investigated over the last few decades with the aim of designing descriptors which are distinctive and robust to to many compter vision applications [194]. A local feature based algorithm involves two major phases: keypoint detection and feature description. In the keypoint detection phase, keypoints with rich information content are first identified and their associated scales (spatial extents) are determined. In the feature description phase, the local geometric information around a keypoint is extracted and stored in a high-dimensional vector (i.e., feature descriptor). The feature descriptors of one surface are matched against the feature descriptors of other surfaces of interest to yield point-to-point feature correspondences [117].

A feature descriptor is descriptive if it is capable of encapsulating the predominant information of the underlying surface. A feature is robust if it is insensitive to a number of disturbances which can affect the data, e.g., noise and variations in the mesh resolution [279]. Although a large number of feature descriptors have been proposed such as tested Heat Kernel Signatures (HKS), spin image (SI), mesh histogram of gradient (Mesh-

HoG), Scale-Invariant Spin Image (SISI), local depth SIFT and generalized HKS (GHKS), they were exclusively designed for a specific application scenario (e.g., object recognition, and shape retrieval) [116].

Facial curvatures have been studied and used extensively. Gaussian curvature, mean curvature, principal curvatures, as well as shape index values are the most widely used ones. The major applications being in face recognition [101, 174, 198] or recognition a part of the face [219] and face expression recognition [290]. The mean and Gaussian curvature have been used as features for describing face regions [198]. Furthermore, the maximum, minimum, Gaussian, and mean curvature were used as features for 3D face recognition with different orientation on FRAV3D database [101]. In case of recognition parts of the face, Pflug et al. [219] employed the curvatures and semantic analysis of edge patterns to detect the ear in 3D profile face. Pflug's approach was robust against rotation and scaling, with the accuracy of detection reaching nearly 96%. Chang et al. [56] proposed a curvatures based approach for 3D face recognition in the presence of varied facial expressions, their approach based on combining the match scores from matching multiple overlapping regions around the nose to handle the problem of expression variation. In 2009, Chen and Biswas [58], focused on the global properties of ranking curvatures for small regions. Both Gaussian curvature and mean curvature of 3D shapes have been used to propose a multiscale method for 3D object analysis and similarity classification. They showed how the mean curvature can be used to find local features and extreme points in 3D facial data. Wang et al. [290] applied these features to 3D facial expression recognition. He used the SVM classification algorithm [115] on the 100 subjects of the Bosphorus database and a 76.56% recognition rate for six universal facial expressions was achieved. Junli et al. [311] used a geodesic network generated for each face with geodesics and iso-geodesics paths determined. Then, the mean curvature, Gaussian curvature, shape index, and curvedness were computed for each node of the network and combined together to produce an automated 3D facial similarity measurement. Tang et al. [270] presented a 3D face keypoint detection, description and matching framework based on principle curvature measures computed using *normal cycle theory*, their proposed method achieved recognition rates of 97.96% on the FRGCv2.0 database, and a

degradation of 3.9% between the neutral and the non-neutral expression subsets. Recently, an interesting study handled significant pose variations problems in 3D face recognition by using HK (H stands for mean curvature and K stands for Gaussian curvature) curvatures analysis for landmarks detection at half face then the matching step was performed with respect to frontal scans and side scans [175].

In this thesis, the curvatures are chosen but combined in a novel way, for face morphological traits analysis and classification.

2.5   Morphological lips traits classification

The lips are the key feature for the lower face [131]. Many studies have been undertaken to measure lip characteristics. Hwang et al. [135] have advocated the use of sliding calipers or transparent rulers to measure distances between clinically observable points (e.g. mouth width by measuring the distance from commissure to commissure). Heidari et al. [129] measured and classified lips morphology in 100 healthy Sistani and Baluch subjects, to analyse the lips characteristics difference between these ethnic groups. As previously explained, ordinary two-dimensional photography is the most common method of recording a human's facial morphology due to the ease of acquisition and low cost. For example, the studies [54, 251] used photographs of subjects to perform anthropometric measurements and classification of lips morphological traits, for accessing the lips shape and attractiveness in very young subjects.

In the last two decades, three-dimensional imaging has become quite common in various medical applications. Three-dimensional facial images are much more informative than two-dimensional photographs [47, 207] and are advantageous for studying lip morphology. For example, the study [199] used three-dimensional facial images of 109 subjects, aged 5–6 years, to produce four categories for philtrum shape: triangular, parallel, concave, and flat. Later on, Wilson et al. [295] produced 3D measurements of lip vermilion and Cupid's bow and described different morphological features of the vermilion of the lips and associated lip traits for 4747 subjects from the ALSPAC dataset [295]. Sforza et al. [245] assessed the nose and lip morphology using 3D scans of 64 North Sudanese subjects with Down's syndrome, age range 5–34 years. Although past research

indicates popularity of 3D imaging for research in facial morphology, the current methods in the area still usually rely on manual facial trait classification and categorisation, which is a very time consuming process. Consequently, such research would benefit significantly from an automatic method.

2.6   Morphometric perspective for facial gender analysis

The human face exhibits a set of significant social signals that give rise to traits such as gender, expression, identity, age and attractiveness. Traditional gender classification methods use a set of facial features in supervised classification algorithms; these methods differ according to the type of features and classification algorithms used.

It is logical to focus on biologically significant landmarks in order to extract features for facial gender classification since gender is a biological characteristic. Although the gender classification problem has been the subject of considerable research in recent years, current computer-based vision methods for facial gender recognition tend to ignore facial biological landmarks as the basis for gender classification despite their ability to efficiently classify gender with a minimum number of features when compared to the methods that use global 3D face geometric features. For example, Ballihi et al. [24] used a large set of geometric geodesic paths features (circular geodesic paths and radial geodesic paths) utilising the AdaBoost algorithm for feature selection to yield a gender classification rate of 86% at FRGCv2 dataset. Gender classification and face recognition using landmark-based and simple geometrical features were the subjects of much research in the past. Burton et al. [50] manually annotated 73 biological landmarks using a dataset of 91 male and 88 female faces, utilising a total of 2,628 Euclidean distance measurements. Due to limited computational capacity, the authors handpicked only 19 distances (and their corresponding ratios) and used these features to attain classification accuracy of 94%. Han et al. [126] utilised more intricate measures such as the volume and the area of face parts to classify gender in a small public dataset of only 61 subjects. They used a support vector machine (SVM) classifier to classify the areas and volumes of five local facial regions: the temple, eyes, nose, mouth, and cheeks. Using five-fold cross-validation,

they reported 82.6% gender classification accuracy. Gilani et al. [110] extracted the geodesic and Euclidean distances between 23 biological landmarks annotated manually for 64 3D facial meshes. Using these features, the authors proposed an approach that could determine a gender classification score for 3D faces with 89.6% classification accuracy. Toma et al. [276] derived 250 facial parameters 90 Euclidean distances between landmarks, 118 angles between 3D facial landmarks, and 42 Euclidean distance ratios in the large ALSPAC 3D face dataset to classify gender with approximately 80% accuracy.

Finding the relationship between facial morphology and gender variation also has received some attention. For example, Brown et al. [44] reported the results of such investigation in a database consisting of 32 photographs of male and female faces. His results showed the jaw, brows, eyes and chin (in descending order) effecting significant change in perceived gender. For 3D faces, Toma [276] worked on finding which parts of the 3D face are most effective in gender discrimination using the distance measurements between the anthropometric landmarks pairs. Gilani et al. [110] used the geodesic and Euclidean distances between the anthropometric landmarks pairs and reported that the distances between the eyes and forehead landmarks were the gender discriminative distances in 64 adult samples. Another study was conducted on 1555 subjects between 3 and 25 years [151]. The subjects were divided into six age groups to investigate facial sexual dimorphism between them, by using 29 traditional soft-tissue anthropometric measurements collected from 3D facial scans. Many differences between males and females faces relating to noses, jaw lines, protrusiveness of mandibles, prominence of cheeks and sizes of lips were reported at subjects from US and Hungary [113]. In general, these studies used either the global face features or the distances measurements between landmarks for gender classification. Therefore, focusing on gender classification and analysis relaying on biological landmarks is an interesting research topic.

2.7   Nose morphological traits classification

The nose is a characteristic component of the face, which is the best feature that distinguishes an individual [149]. Nose shapes and categorisa-

tion are, therefore, vital to forensic analysis, examinations of sexual differences and biological variability amongst ethnic groups and find the relation between the facial morphology traits variation and background genes and heritability. Shah et al. [248] and Adelaja et al. [3] found the relationship between the nose morphology variation such as nasal root width, nose height, nose width, nasal bridge length, nasal tip protrusion, nasal index, facial width and nasolabial proportion with ethnic groups. The objective of Tanikawa et al. [271] was to demonstrate sexual dimorphism in the entire three-dimensional facial surface form of adult humans depending on the nose shape variation. The sample consisted of female and male groups ($n = 200$; age range, 18–35 years). Barash et al. [26] investigated the face variation and especially the genes associated with nose morphology traits variation. This research studied the effect of certain genes on facial variation specifically, associations of nasal width with rs8035124 SNP, the cephalic index with rs16830498 SNP, nasal index with rs37369 SNP, transverse nasal prominence angle with rs59037879 and rs10512572 SNPs. Another interesting piece of research investigated the relationship between the nose shape and human characteristics [221]. Twelve shapes of noses according to human characteristics were described such as the straight nose presents intelligence, kindness, and faith, the Roman nose symbolizes bravery and understanding, while short nose indicates low self-confidence.

Recently, interesting research was conducted on 3D faces from a different population (Asia, Europe, and Africa) to prove the thought of differences in nose shape among populations are not the result of genetic effect only, but may be adaptations to climate. Linear distances were calculated from the 3D coordinates of nose landmarks to observe nares width, Alar base width, Nasal height, Nasal ridge length and Nasal tip protrusion. The spatial distribution of these nose traits measurements is compared with the global distribution of temperature, absolute humidity and relative humidity. The authors found that that some manifestations of nose shape may indeed have been driven by local adaptation to climate, as example width of the nares is correlated with absolute humidity and temperature [309]. Almost all these studies utilised basic geometric features such as the Euclidean distances and angles between landmarks or manual identification for nose traits categorisation.

2.8   The face morphology characteristics/genotype association

Cells are the main working units for all living systems. Chemical DNA (deoxyribonucleic acid) contains all the instruction needed to control the cell's activities. An organism's complete set of DNA is called *genome*. Genomes size varies widely, for example, the human genomes have more three billion DNA. Human cells have 23 pairs of chromosomes (22 pairs of autosomes and one pair of sex chromosomes), each chromosome contains many genes. That mean, DNA is in genes and genes are on chromosomes. The human genome is approximately contain 20,000–25,000 genes [276]. Single nucleotide polymorphisms, frequently called SNPs (pronounced "snips"), are the most common type of genetic variation among people. Each SNP represents a difference in a single DNA building block, called a nucleotide [99]. In spite of the numerous number of genes and SNPs that have been identified as pivotal for face morphology development. Nevertheless, little is know about the genes that are involved in facial variation, abnormalities or syndromes. Genome wide association study (GWAS) is the analytical method for mapping genes against human traits. It is powerful in detecting genes of small effect and disease genes [232]. Table 2.2 lists some of heritability and genetic association research. These studies have used the anthropometric measurements as facial features for heritability calculation and gene associations.

Table 2.2: Some studies related to heritability and genetic effects on facial morphology

Sample ethnicity	Sample number	Facial parameters	Effect	Reference
Twins from the East Flanders Prospective	79 pairs	By using direct measurements 5 facial proportions based on 4 vertical and 5 horizontal measurements were calculated	The highest heritability was for upper to lower facial height proportion	[240]
Caucasian	A set of 10 monozygotic twin pairs (20 individuals) and 11 same-sex dizygotic twin pairs	Thirteen 3D landmarks were collected from each facial surface. The resulting PC scores used to calculate heritability estimates	Moderate to high heritability were associated with variation in the breadth of orbital and nasal structures	[292]
Turkish Anatolian siblings	138 subjects (70 women) (68 men)	Six facial proportions and 6 soft tissue measurements for the 3D face	The genetic determination significantly higher in the soft-tissue measurements than in the facial proportions	[27]
Latin Americans	6,275 subject	Facial photographs of the individuals assessed 14 facial features on an ordered categorical scale reflecting the distinctiveness of each trait	Strong association was found between these facial traits: Columella inclination, Nose bridge breadth, Nose wing breadth, Nose wing breadth and Chin protrusion with these genes: DCHS2, RUNX2, GLI3, PAX1, EDAR	[4]
Children and their parents from Iceland	363 children (assessed at 6 and 16 years old)	Thirty three Euclidean and angular parameters direct measurement (Clinical examination)	Analysis of parental data can have predictive value for offspring	[144]
UK twin	952 twins	Twenty Euclidean and geodesic distance for the 3D face	Heritability estimates h^2 for 10 traits based on Euclidean Distances and 10 traits based on Geodesic Distances	[282]
Japanese, Hispanics, Chinese, Euro-Americans and African American	263 subjects	By direct linear measurement the following traits length calculated: Cranial base length, Maxillary length, total mandibular length, Mandibular corpus length and Mandibular ramus height	The relation between Mandibular ramus height and Growth Hormone polymorphisms genes	[280]

2.9   Summary: Gaps in the current state of art

As can be concluded from this literature review, one of the biggest challenges facing 3D facial morphology analysis is the automatic diagnosis and classification of facial traits. The process for determining the 3D geometric features that are most discriminative for facial morphology traits is still typically manual and uses simple anthropometric measurements. These measurements do not present the variation in face surface shape adequately. A lot of work is still needed in the study of the relationship between face morphology traits and gender (sex) variation in a certain age and ethnic groups. Solving these challenges will open the door for progress in many medical and survive applications such as:

- Measuring significant facial shape differences, normal and disordered facial asymmetry, and facial discordancy or the lack of harmony, useful for planning surgical interventions.

- Developing the criminal evidence science by predicting faces from DNA.

- Discovering the relationship between the medical syndrome and ethnic groups genes changes.

- Exploring traits variation/genotype associations in normal and abnormal subjects.

To address such issues, in this thesis the research effort was focused on exploring the advantages of face biological (Anthropometric) landmarks, by introducing a new set of 3D geometric features based around them. These 3D geometric features take the advantage of using the anthropometric landmarks but at the same time capture accurately the changes in face surface shape. The proposed geometric features are used for automatic morphological traits classification and categorisation and find the relationship between gender and facial traits in ALSPAC dataset. This work was validated by investigating genotypes associations with the automatic face morphological traits categories identified by our approach.

Chapter 3

DATASET AND
PREPROCESSING

Restricting analysis and classification of face morphology to 2D techniques makes measuring the trait differences between individuals difficult. Consequently, for analysis of complex 3D facial surfaces, acquisition techniques that provide a large increase in morphometric information are required, namely accurate 3D data scans.

However, the data from 3D scanners are often imperfect. Challenges are generally in the form of noise or holes due to the absorption of laser beam by the dark facial areas such as eyebrows and iris. Therefore, data preprocessing is a very important step in face morphology analysis systems. It eliminates noise and remove some undesired parts such as clothes, neck, ears, and hair from scan images. This step is common in all 3D face vision applications, including 3D face morphology characterisation and analysis.

In addition, the regularisation of a number of vertices by finding dense correspondences in the 3D meshes is often performed before analysis of faces. Surface regularisation is the process of finding the geometrical transformation between two or more surfaces that align the surfaces [60]. Many studies proved the use of dense sets of vertices is a ideal for analysing face morphology because the sets allow one to inspect and analyse facial images in detailed and holistic manner [138, 273].

In this work, two strategies for 3D face preprocessing are implemented, namely regularised and non-regularised preprocessing, to compare their effects on facial morphological traits classification. They are described in detail in Chapter 5.

The sections in this chapter are as follows. In Section 3.1, the 3D faces

dataset used in this work is described in detail. In Section 3.2, the prepro-
cessing steps and algorithms used in this work are explained. In Section
3.3, a general discussion of the preprocessing methods used in this work
is presented. Finally, in Section 3.4, this chapter is summarised.

3.1   ALSPAC Data Set

This study is based on data from the Avon Longitudinal Study of Parents
and Children (ALSPAC). ALSPAC is a long term health research project
involving the study of over 14,000 children as they grow up. This study
recruited pregnant women living in the former county of Avon in South-
West England who had an estimated delivery date of between April 1,
1991, and December 31, 1992. The cohort was made up of 14,541 preg-
nancies that resulted in 13,971 singletons/twins that were alive at one
year of age. Ethical approval for using this data in this work was ob-
tained from the ALSPAC Ethics and Law Committee and the Local Re-
search Ethics Committees [14]. When the children reached the age of
around 15 years, their faces were scanned in three dimensions using two
Konica Minolta Vivid 900 laser cameras [277]. The scanned images were
imported into Rapidform 2006 (a reverse engineering software package
for 3D images) [268] and processed by removing the color texture in or-
der to highlight morphological features and eliminate the influence of fa-
cial color tones. Then, 21 facial landmarks represented by a total of 63 (x,
y, and z) coordinate values were manually identified[1] and recorded for
each 3D facial image [277]. The biological landmark points, for this data
set, their location and meaning on the human face are shown in Fig. 3.1,
Table 3.1 contains their definitions.

---

[1]Manually obtained landmarks were already available for this dataset. However, they
can be obtained automatically using the following methods [78, 160, 235]

Table 3.1: Facial soft tissue landmarks definitions

Landmarks	Defnition
Glabella(g)	Most prominent midline point between eyebrows
Nasion(n)	Deepest point of nasal bridge
Endocanthion(en) L/R	Inner commissure of the left / right eye fissure
Exocanthion(ex)L/R	Outer commissure of the left / right eye fissure
Palpebrale superius(ps) L/R	Superior mid portion of the free margin of upper left / right eyelids
Palpebrale inferius(pi) L/R	Inferior mid portion of the free margin of upper left / right eyelids
Pronasale (prn)	Most protruded point of the apex nasi
Subnasale(sn)	Mid-point of angle at columella base
Alare(al)L/R	Most lateral point on left / right alar contour
Labiale superius(ls)	Mid-point of the upper vermilion line
Labiale inferius(li)	Mid-point of the lower vermilion line
Crista philtre(cph) L/R	Point on the left / right elevated margins of the philtrum just above VL
Cheilion(ch) L/R	Point located at left / right labial commissure
Pogonion(pg)	Most anterior midpoint of the chin
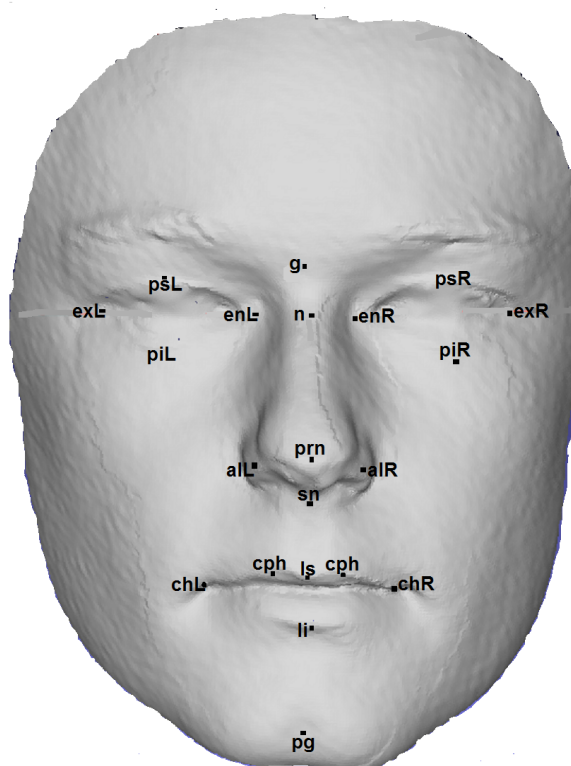
Figure 3.1: Landmarks used in this work.

Considerable amount of medical research has been conducted on the ALSPAC data. For example, Horwood et al. [133] investigated the prevalence, nature, and frequency of psychosis-like symptoms in 12-year-old

children and studied their relationship with IQ score. Another study was to examine the association between alcohol and cigarettes taking among adolescents and their early socioeconomic background [189]. The work presented in this thesis is the first study conducted on this data which is performed automatically.

3.2 Preprocessing

Preprocessing operations are aimed at improving the quality of data. In this work, two approaches for data preprocessing are used (Non-regularised and regularised). The motivation for using these methods is to compare the face morphology analysis and classification performance for raw images and registered images. The registration process is used here to regularise (downsample) the images to certain template resolution, any image from the data set can be used as a template and regularised all the images to its resolution [123]. However, in this work a template called AM is used, these template used successfully in many face morphology analysis research [60, 61, 185, 254]. Non-regularised (raw) images correspond to high resolution, while the regularised images correspond to low resolution. Consequently, the effect of high and low resolution mesh in face morphology classification is highlighted in Chapter 5^2.

- *Non-regularised preprocessing* in this case only the pose of each 3D face was corrected to a canonical form based on landmarks for each face. This step is required to reduce any errors due to the difference in poses before extracting 3D geometric features.

- *Regularised preprocessing* in this case the raw face meshes are marked with a small number of feature points and a 3D transform is used to align each face to a reference face, the choice of reference face will affect the final dense model. As the result, all the ALSPAC faces are represented as uniformly distributed dense meshes with 7150 verticies.

In the following sections, both methods are described in detail.

---

[2]Different template sizes can be used in future to test their effect in face morphology analysis and classification performance

3.2.1   Non-regularised preprocessing

Facial orientation is not a part of the facial shape, and alignment of their poses needs to be performed during or before features extraction. In a landmarks based facial representation, the correspondences between the faces are implicitly known via the landmark definitions. Thus, landmarks can be used to solve the alignment problem prior to further data extraction as in [123]. The ALSPAC data set faces do not have any holes. Therefore, only the rotation, translation, and scaling of the faces are carried out using a Generalised Procrustes Algorithm (GPA), followed by eliminating noises using well-defined filter.

Generalised Procrustes Algorithm (GPA)

Procrustes analysis (also called ordinary Procrustes analysis) is a form of statistical shape analysis that aligns faces by minimizing the Procrustes distance metric between their landmarks [114].

The algorithm steps are the following:

1. Choose a reference shape among the training set instances.

2. Align all other instances on the current reference by minimizing a measure of shape difference called the Procrustes distance between the instances.

3. Compute the mean shape of the current training set.

4. If the Procrustes distance between the mean shape and the reference is above a threshold, set the reference to mean shape and continue to step 2.

Suppose we have M faces each defined by 21 facial landmarks in three dimensions (x, y, z). Then each facial shape is represented by 63 coordinates.

The centroid of a shape is the point $(\overline{x}, \overline{y}, \overline{z})$ with the 21 landmarks mean coordinates. The centroid is taken to be the origin of coordinates and all shapes are now translated to the origin:

$(x_1 - \overline{x}, y_1 - \overline{y}, z_1 - \overline{z}), (x_2 - \overline{x}, y_2 - \overline{y}, z_2 - \overline{z})$ and so on.

The size of a face is defined as its centroid size, which is the root

square deviation of all landmarks from the centroid:

$$S = \sqrt{(x_1 - \overline{x})^2 + (y_1 - \overline{y})^2 + \cdots + (y_{21} - \overline{y})^2 + (z_{21} - \overline{z})^2} \qquad (3.1)$$

After the sizes of all faces $(S_1, S_2, ... S_M)$ have been calculated, the scale or size component is removed by scaling each object to the *average face size* $(\overline{S})$. Removing rotation is done by minimizing the Procrustes distance between the current shape and mean shape [276].

Denoising

In many 3D processes, it is important to handle noise properly to avoid distorted representation. The noise characterisation depends on the chosen scanner. Noise characterisation for depth sensors and for laser beams are different, most of the time, methods presume Gaussian and isotropic noise [104]. Many studies recommended Laplacian smoothing filter to smooth a polygonal mesh to remove the undesirable information (i.e. spikes) and clutter due to the 3D scanner (e.g. see [84,259]). In this stage, a Laplacian smoothing filter in Graph MATLAB Toolbox   [217] is used. Figure 3.2 illustrates the denoising process for 3D face mesh, the degree of 3D face smoothing depends on the number of iteration selected by the user. In this work the number of iterations was fixed to one and the filter parameters were fixed to the default parameters used in the toolbox examples $laplacian.type = distance$, $options.symmetrize = 0$ and $options.normalize = 1$.

(a) Original 3D image   (b) Smoothed 3D image
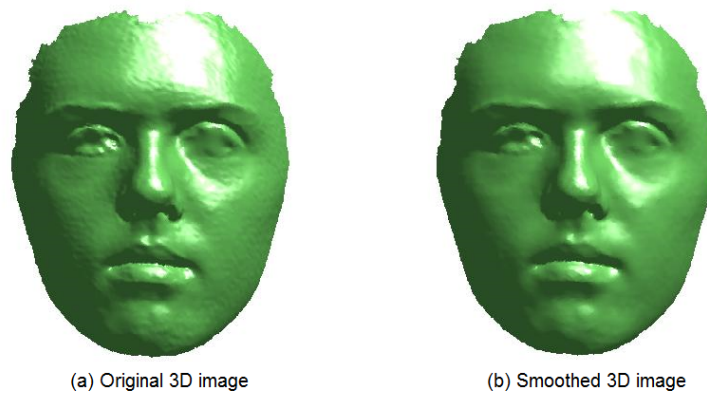
Figure 3.2: Denoising the 3D face. This result for one iteration only.

3.2.2   Regularised preprocessing

Scanners usually digitize 3D surfaces in the form of dense 3D point clouds. The point cloud representation is capable of completely representing 3D surfaces in face recognition systems [173]. However, 3D mesh representation allows for flexible and efficient manipulation of surfaces. For example, the deformation of 3D meshes is more flexible than point clouds [77]. This surface information when it is stored as a "mesh", is defined by the 3D locations of a large number of points and how these points are interconnected. It is something like a wire net, moulded onto the surface of the object being imaged Fig. 3.3.
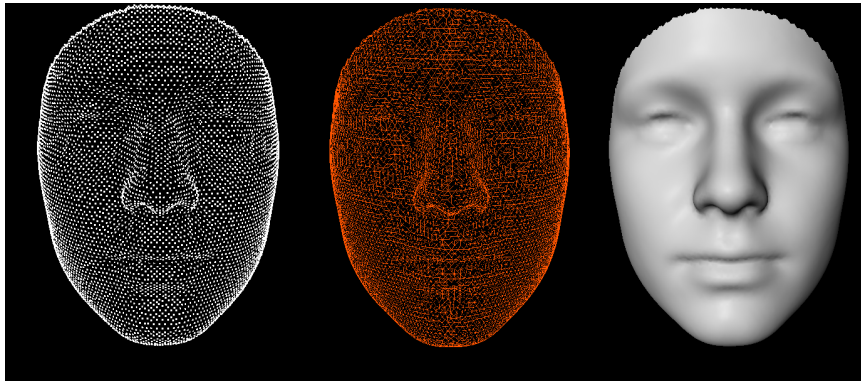
Figure 3.3: Components of a mesh object. Node, or vertex, locations (left), connectivity structure (middle) and the defined surface (right).

For all 3D representation types, the main aim of regularisation process is to establish a geometrical relationship between two or more surfaces (inter-subject) or between different morphologies of the same subject (intra-subject) [60]. In this work, the correspondence dense meshes are obtained by non-rigidly mapping of an anthropometric mask (AM) onto facial 3D images. An AM is essentially a predefined surface template covering the facial area of interest. This template defines the number of vertices used. By mapping the AM on the facial images, homologous spatially dense mesh configurations for all 3D images are obtained. In this context homologous means that vertex occupies the same position on the face relative to all other vertices for all individuals, Fig. 3.4 shows the corresponding nose tip points (i.e. points with the same indices) were colored in the mapped faces of the other individuals. The mapping of the (AM) onto the facial images is, therefore, standardises the facial 3D

images from different individuals so they can be analysed in a spatially dense way [60].

Figure 3.4: The same nose tip correspondence in 3 individuals.

Anthropometric Mask (AM)

The mask template used in this work is illustrated in Fig. 3.5, it represents the average face of 400 Western Australian healthy young individuals between the ages of 5–25 years captured with the 3dMD facial scanning system [65].
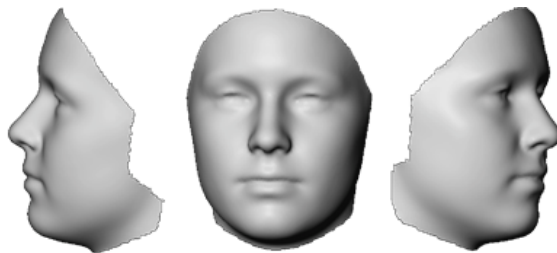
Figure 3.5: The anthropometric mask.

The steps for building the AM are as follows:

- A random face is taken from the database and used as an initial template.

- That initial face template is registered to all other faces in the database by using the Iterative closest point (ICP) method reported in [34]. The Iterative closest point (ICP) is an iterative registration algorithm alternating between correspondence update. The parameters update in every iteration step, until convergence occurs, to the nearest local minimum of the objective function.

- The resulting faces represent the shape of the target surfaces, but have corresponding points. This helps in easily calculating a geometric average face.

- The geometric average face is used as the template for another iteration. The iterative process is repeated until convergence and the template changes between iterations become negligible [65, 256].

Mapping of 3D Images

This mapping technique "fits" a template face (an anthropometric mask) to any number of faces. This process is, metaphorically, like stretching an elastic mask over a solid statue, as shown in Fig. 3.6. The result is that the face, after regularisation, has the same number of vertices and same connection structure as the template. The mask used for ALSPAC dataset faces regularisation is illustrated in Fig. 3.7, but without the neck part. The anthropometric mask consists of 7, 150 vertices only. This technique was developed by Peter Claes [60], which has been validated extensively on faces [63, 66, 67]. The technique steps as shown in Fig. 3.7 are as follows:

- Five marks need to be indicated manually on the 3D facial image (medial canthus of the right eye, medial canthus of the left eye, nasal tip, right labial commissure, left labial commissure).

- Prior to regularisation, the floating surface is to be brought into the vicinity of the target surface, by repositioning the floating surface, using a rigid or affine transformation, into the coordinate system of the target surface. This is performed by using the 5 initial starting points. The mask and the target face were roughly rotated and translated until the differences are harmonized. The Iterative closest point (ICP) is used for that purpose.

- Surfaces of different subjects are matched to each other using a non-rigid Thin Plate Spline (TPS) [82] based deformation model. The basic principle of TPS is the comparison of two different shapes is performed by distortion of the first one. It serves as a reference which is here the AM, while the compared specimen is termed the

target face. The deformation requires bending energy, which leads to exactly reproducible deformation patterns [234].

- In each iteration, each floating surface point is assigned to a corresponding point. A simple way to assign correspondences relies on using the Distance Weighted k-Nearest Neighbor Rule [87]. It is non-probabilistic classification procedure, which can be adapted to estimate the corresponding positions by weighing the contribution of k-Nearest target surface.
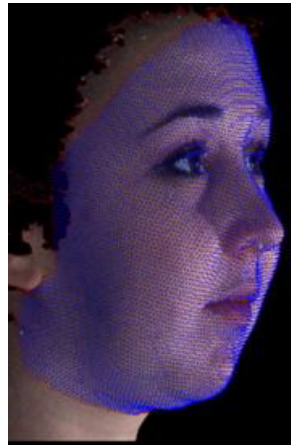
Figure 3.6: Antropometric mask mapped onto original 3D-image [67]
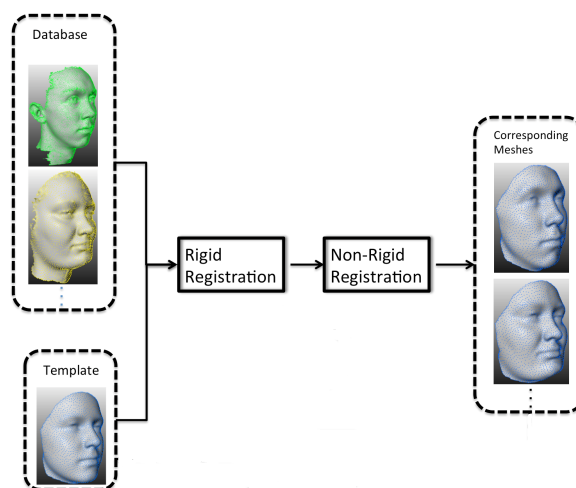
Figure 3.7: The regularisation general procedure steps.

3.3   Discussion

In this work, software written in MATLAB by Peter Claes (The University of Leuven) was used to regularise the facial surfaces to produce a dense correspondence for all the faces. The template used for faces regularisation contained only 7150 vertices in order to test the robustness of our 3D geometric features and classification and categorisation systems on very low resolution mesh comparing to the original mesh. Figure 3.8 illustrates the difference between the non-regularised and a regularised face. It is clear that the lips morphology is more clear for non-regularised faces. However, regularised images allow for faster extraction of facial features, which is a key advantage for testing the robustness of the proposed automatic classification method and if it is applicable for both high-resolution images (e.g., captured with laser scanners) and low-resolution images (e.g., obtained using stereo scanners).
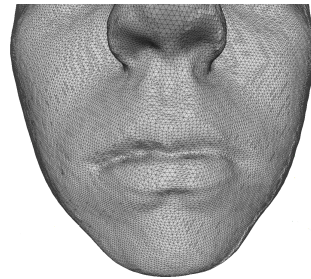
(a) Non-regularised mesh
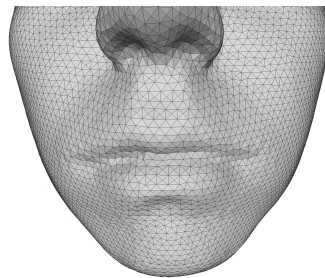
(b) Regularised mesh

Figure 3.8: The preprocessing methods. The difference between the methods of preprocessing (Non-regularised and regularised).

3.4   Summary

3D face preprocessing is a necessary step for all computer vision applications, specifically for face morphology classification and categorisation. In this chapter two methods for preprocessing are described (regularised and non-regularised), the effect of using these methods on face morphology classification performance in-depth for lips traits classification will be tested in Chapter 5.

Chapter 4

NOVEL 3D GEOMETRIC FEATURES FOR FACE MORPHOLOGY ANALYSIS

Once landmarks are annotated manually or automatically in zones of the faces, it becomes possible to extrapolate important geometric information their particular position gives them. For face morphology analysis applications, the computation of the Euclidean distances or the angles between landmarks is a widely used method. They are considered measures, rather than surface features. These measures are called *anthropometric*, and one measure may involve more than one landmark. Nonetheless, these facial measurements do not provide explicit information about facial shape. Instead, feature descriptors based on 3D geometry give immediate information about the shape of eyes, mouth, nose, and global traits variation. Different from all the existing 3D face morphology analysis and classification approaches, this work is the first to study the effectiveness of the curvature for the practical automatic 3D face morphology classification and categorisation application. The curvatures are extracted for face parts and for the geodesic paths between the biological landmarks.

This chapter is organised as follows. The face region features concept are explained in Section 4.1. The basic curvature estimation methods are described in Section 4.2. Curvatures robustness in facial morphology classification is described in Section 4.3. Section 4.4 describes the novel 3D geometric features based on geodesic paths. Section 4.5 provides a discussion of the results. Finally, Section 4.6 records this chapter summary.

4.1   Parts vs whole

The existence of part based face perception has been demonstrated in computer vision applications. For example in 3D face recognition it can provide robustness to facial expression by excluding affected parts [212]. In facial morphological traits studies, the face representation of choice has always been holistic [63,73]. However, in this work, we argue that part part-based facial morphology analysis can produce better results at least for some of the traits. For example, for some detailed local facial traits, such as chin dimple, part based morphology analysis may be a more successful approach.

In this chapter we propose novel 3D geometric features based on face parts, which are applied to classification and analysis of face morphology in this and next chapters. These features can be divided into two types according to the properties of the traits we analyse:

- Face region based geometric features. This type of 3D geometric features are used to analyse face area based traits, such as philtrum shape and chin shape.

- Geometric features based on the paths between anthropometric landmarks.These features are used to analyse very specific morphological traits such as lips contours and tip nose shape.

For the both types, on the basis of our literature review, we decided to use the curvature descriptors as the most effective for description of face shapes.

4.2   Curvature estimation methods

In computer vision and geometric design applications, the surfaces of discrete form such as meshes, and point surfaces become more and more important. At the same time, various curvature estimation techniques have been developed and implemented. From a mathematical standpoint, the curvature information can be represented by the first and second partial derivatives of the local surface [77]. Curvature estimation techniques on mesh surfaces can be broadly divided into three groups; continuous, discrete and estimation of a curvature tensor [103].

The continuous estimation includes fitting a surface locally (at least second order) then computing the curvatures by interrogating the fitted surface. This method produces the best approximation of the underlying surface. It has been one of the more common and robust approaches for curvatures estimation. There are numerous curvature estimation methods depending on surface fitting principle [227], such as the quadric fitting method suggested by Hamann [121] and a cubic-order algorithm for approximating principal direction vectors by Goldfeather et al. [111].

The discrete estimation method utilises a direct approximation formula for the curvature. These methods do not only involve solving a least squares problem but also appear very fast and flexible. Research using such discrete estimations for curvature includes Taubin [272] where the tensor of curvature of the surface at the vertices in a polyhedral approximation was estimated. A key advantage of Taubin's method is its simplicity, with the complexity being linear in both time and space. Meek and Walton [187] and Meyer et al. [193] estimated curvatures directly on the discrete triangle meshes on the basis of the Gauss-Bonnet theorem, known as the angle deficit method, which approximates Gaussian curvature as 2π minus the sum of the angles for the mesh faces at a vertex, divided by an area associated with the vertex.

The curvature tensor estimation is similar to the discrete methods, except that instead of estimating the curvature directly, a discrete estimation of the curvature tensor is created, and the curvatures and principal directions are calculated from the curvature tensor. These methods tend to have computational complexity lower than the fitting methods, but slightly higher than the discrete methods but more estimated accuracy [103]. In this work the *Normal Cycle* curvature tensor method is used to calculate the curvature features for the 3D images. This method was implemented, in particular, by [100] to give a general method to define curvatures of a large class of objects, without any assumptions of smoothness or convexity and it has been successfully developed by the author of [69].

4.2.1   Estimating curvature using normal cycle theory based method

This theory provides a unified, simple, and accurate way to define curvature for both smooth and polyhedral surfaces [69, 262]. The main idea of *normal cycle theory* is that in order to acquire a continuous tensor field over an entire surface, a piecewise linear curvature tensor field should be calculated by estimating the curvature tensor at each vertex and then adding those values linearly across triangles. Fig. 4.1 shows the principal method that is used to calculate the curvature tensor for each vertex along the edge e, where for every edge e of the mesh there is a minimum curvature (along the edge) and maximum curvature (across the edge). These line dense tensors can be averaged over an arbitrary mesh region B according to the following equation:

$$\mathcal{T}(v) = \frac{1}{|B|} \sum\nolimits_{\text{edges}} \beta(e)|e \cap B| \, ee^T \tag{4.1}$$

where v represents the vertex position on the mesh, $|B|$ is the surface area around v over which the curvature tensor is estimated, $\beta(e)$ is the signed angle between the normal vector to the two oriented triangles incident to edge e, $|e \cap B|$ is the length of $e \cap B$, and e is a unit vector in the same direction as e [12]. In practical terms, the normal cycle method is fast, and it provides excellent results, although the important issue of how the user should choose the neighbourhood B that approximates a geodesic disk around the vertex v still remains. The selection of the neighbourhood size can significantly affect the results: small neighbourhoods provide better estimates, while an increase in the neighbourhood (ring) size smooths the estimates, which leads to less sensitivity to noise [69, 236]. Generally, the eigenvectors of $\mathcal{T}(v)$ associated with the eigenvalue magnitude are used to estimate curvatures at each vertex. The principal curvatures k_1 and k_2 at v are estimated by the eigenvalues, while the eigenvectors represent the curvature directions [12].
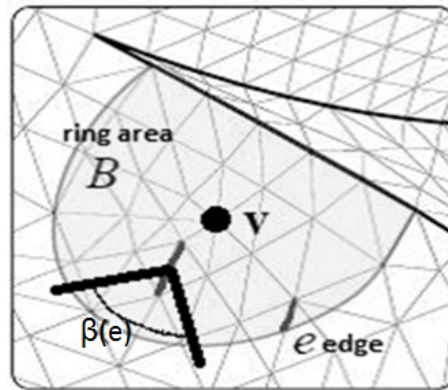
Figure 4.1: Curvature estimation according to normal cycle theory: Eq. (4.1)

Using principal curvatures (k_1 and k_2), Mean (H), and Gaussian (G) curvature are calculated as:

$$H = \frac{k_1 + k_2}{2} \tag{4.2}$$

$$G = k_1.k_2 \tag{4.3}$$

H is the average of the maximum and minimum curvature at a vertex v. G is the multiplication of the principal curvatures; its sign denotes if the surface is locally elliptic or hyperbolic [77]. Figure 4.2 illustrates the four curvature features.
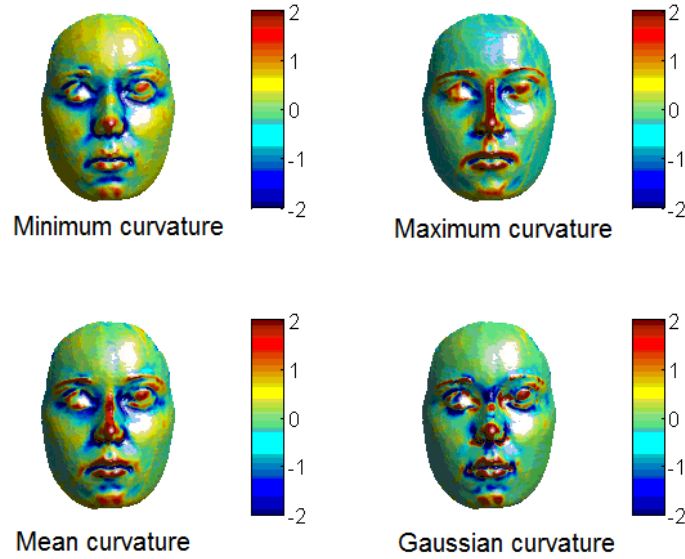
Figure 4.2: The minimum, maximum, mean, and Gaussian curvature information

Shape index (S_i) quantitatively measures the shape of a surface at a vertex v and captures the intuitive notion of local shape of a surface. Every distinct surface shape corresponds to a unique value of S_i (except for the planar shape). The shape index for any surface point can be calculated from the principal curvature (k_1 and k_2) at that point.

$$S_i = 0.5 - \frac{1}{\pi}.tan^{-1}(\frac{k_1 + k_2}{k_1 - k_2}) \tag{4.4}$$

Another surface feature called, curvedness (R), measures how highly or softly bent a surface is. Curvedness can define the scale difference between objects: for example the difference between a soccer ball and a cricket ball. This surface point feature can be also calculated from the principal curvatures (k_1 and k_2) as follows [83].

$$R = \sqrt{\frac{(k_1{}^2 + k_2{}^2)}{2}} \tag{4.5}$$

In general, the shape index (S_i) and the curvedness (R) are robust surface information from the 3D image and they are invariant to change in 3D image orientation. Figure 4.3 shown that the variations in shape and

curvedness, S_i and R, can be represented as polar coordinates within a Cartesian coordinate frame given by the two principal curvatures (k_1 and k_2).
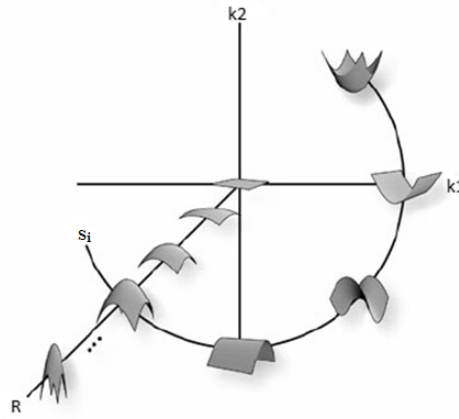
Figure 4.3: The shape index and curvedness described in 2D space. Indices (S_i,R) are viewed as coordinates in the (k_1 and k_2) plane, with planar points mapped to the origin. The effects on surface structure from variations in the curvedness (radial coordinate) and Shape Index [166]

4.3   Facial morphology classification based on region features

In this section, the curvatures capability in classifying facial morphological traits are evaluated. Curvatures are used to classify the philtrum area width into three classes predetermined manually by Wilson [295]. Multiclass SVM (supper vector machine) machine learning method is used for classification purpose. The philtrum is the vertical groove on the surface of the upper lip, below the septum of the nose Fig. 4.4. For the assessment, 1000 3D face meshes are used from ALSPAC dataset.
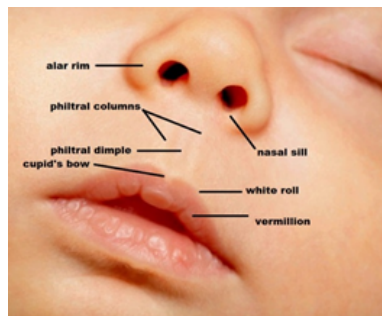
Figure 4.4: The philtrum area

4.3.1   Curvatures extraction

The first step in creating any 3D face classification system is building the features descriptors or vectors for the 3D faces. The features vector is extracted from the 3D face as a whole or from certain face region as our case. In other words, for data set with n 3D faces m features vectors can be produced to construct $n * m$ training and testing matrix. In this work, after the non-regularised preprocessing method implementation, the maximum, minimum, mean, Gaussian, maximum direction, minimum Direction curvature information was extracted for each face philtrum area, each curvature is used as a separate entity to form the feature vector in the training and testing classification set. After that, the feature vectors are equally quantised to a certain number of features. In general, the procedure to extract training and testing sets are:

- Firstly, extracting out the philtrum area of each face according to the landmarks (sn, ls, cphL and cphR) Fig. 3.1.

- Next, calculating the curvatures for the philtrum area using the Normal Cycle method as previously described. Different sizes of the ring (the number of vertices surrounding the vertex that the curvature is calculated for it) were tested. Algorithm 1 below summarises the process of the computation of curvatures using the *Normal Cycle* approach.

- The number of features in the curvature vectors are not equal for each subject's philtrum, for example, if the philtrum 1 has $3 * 2314$ vertices then the curvature vectors will be $1 * 2314$ and philtrum 2 has $3 * 1113$ vertex, the curvature vectors will be $1 * 1113$. Therefore, it became important to quantise them equally, therefore, each curvature vector is quantised to 5, 10, 20,30,40, or 50 bins using histogram normalisation technique. The histogram is obtained by splitting the range of the data into equal-sized bins (called classes). Then for each bin, the number of points from the data set that fall into each bin are counted. The normalized count is the count in the class divided by the number of observations times the class width. For this normalisation, the area under the histogram is equal to one. From a probabilistic point of view, this normalisation results in a

relative histogram that is most akin to the probability density function and a relative cumulative histogram that is most akin to the cumulative distribution function [15, 196].

Algorithm 1: Curvatures estimation using Normal Cycle method

Data: Set vertex V of philtrum area; F mesh faces for philtrum area; R: the ring size

Result: The maximum and minimum curvature vectors $Cmax$, $Cmin$; The maximum and minimum curvature direction vectors $Umin$, $Umax$; The mean curvature vector $Cmean$; The Gaussian curvature vector $Cgauss$.

initialization n is the number of V;

For i:=1 TO n

Specified surface area around V_i according to R which the curvature tensor is estimated

Compute the tensor using Eq. 5

Calculate the eigenvector u and eigenvalue d of $\mathcal{T}(v)$

$[temp, I] = sort(d)$

$D(:, i) = d(I)$

$U(:, :, i) = u(:, I)$

END

$Umin = U(:2,:)$ after singleton removing from U

$Umax = U(:1,:)$ after singleton removing from U

$Cmin = D(1,:)$

$Cmax = D(2,:)$

$Cmean = (Cmin + Cmax)/2$

$Cgauss = Cmin * Cmax$

4.3.2   Classification methods

The aim of supervised learning is to build a concise model of the distribution of class labels in terms of predictor features. The resulting classifier is then used to assign class labels to the testing instances where the values of the predictor features are known, but the value of the class label is unknown [163]. Their are numerous methods for supervised learning such as:

- **Decision Trees** are a classifier that characterises the observations according to the value of the features. Each feature is represented

by a node in the tree; with each branch from the node representing a value that the feature can assume. The classification starts at the root node. The features that best divide the data are selected as nodes. Different measures are used to find the best features such as: information gain, gini index, and refliefF algorithms. The data division continues until the data is split into two subsets belonging to the same class [86].

- **Neural Networks** the original idea behind the neural network is inspired by the mechanism of patterns recognition in the brain. A neural network can approximate any relation between the class label and the features, and they can also deal with multi-class data [127].

- **K-Nearest Neighbor Classifier** this classifier is based on the principle that the observations in the data are generally close to the other similar observations which belong to the same class. KNN assigns to an unlabeled observation the dominant class among the K nearest neighbors. Many different metrics have been used to calculate the distance between the observations within n-dimensional space, where n is the number of features in the dataset [163].

- **Random Forest** RF is an ensemble classifier that consists of many decision trees. RF seeks to address the problem of instability associated with single trees and their sensitivity to the training data. The output class of RF is the statistical model of the output of individual trees. RFs combine the "Bagging" concept, with a random subset of features. Each tree is built on a separate bootstrapped sample, and only a randomly selected feature subset is used at each node. In this case a variation among the trees is obtained. The performance of the RF is not sensitive to the values of their parameters [302].

In this work, Support Vector Machine (SVM) with linear kernal is used for supervised learning because it is robust, accurate and is effective even when using a small training sample [108]. This is achieved by mapping the data into a higher dimensional space by using Kernel functions. A linear kernel has been reported to provide the best performance in many applications, and requires only one parameter to be tuned [32]. SVM have

been applied in many domains such as image classification and biomedical problems, it have reported a high classification accuracies [205, 222].

Classification using SVM

After extracting the feature descriptors from 1000 faces, they were classified using Support Vector Machine (SVM) [142] into three classes (narrow, Average and wide) available to us in this study [295]. For evaluating the classification performance 5-fold cross-validation method was used. In this method, the dataset is divided into k subsets. Each time, one of the k subsets is used as the test set and the other $k-1$ subsets are put together to form a training set. Then the average error across all k trials is computed [118].

The SVM is a popular supervised learning method. SVM is inherently a two class classifier. Given a set of training examples, each data is marked as belonging to one of the two classes and the SVM builds a model that categorises the new example data to one class or another. SVM maps an input sample to a high dimensional feature space and tries to find an optimal hyperplane that minimises the classification error for the training data using the non-linear transformation function. The boundaries between classes are hyperplanes (a line in Fig. 4.5). The largest margin between the two classes means the best hyperplane for the SVM (two dash lines in Fig. 4.5). The margin means the maximal width of the slab parallel to the hyperplanes that has no interior data points. The support vectors are the width constrain of that margin [142].
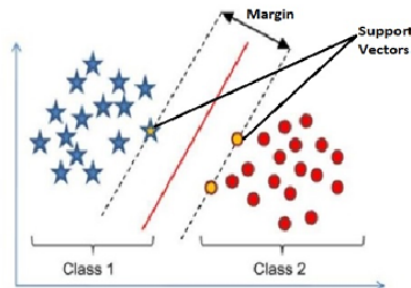
Figure 4.5: Binary SVM
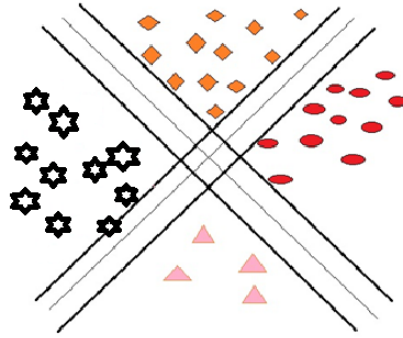
Figure 4.6: Multiclass SVM

However, facial Morphology classification is not a binary classification. By their nature SVM is essentially binary classifier, but, it can be adapted to handle the multiple classification tasks. Figure 4.6 shows a multiclass SVM problem where we have four classes which are separated by gaps. The two methods commonly used are the One-Against-One (1A1) and One-Against-All (1AA) techniques. In 1AA approach an SVM is determined for each class by discriminating that class against the remaining (N-1) classes. On the other hand, the 1A1 approach involves constructing a classification results for each pair of classes resulting in N(N-1)/2 classifiers.

The acknowledged drawback for 1A1 that it is more computationally intensive since it requires many SVM classifiers to be built. However, both approaches have approximately the same accuracy according to many studies [18, 90, 293]. In this work, 1AA is chosen for its simplicity, practicality and to avoid the intensive computation of 1A1. In this work, publicly available software $SVM^{multiclass}$ is used [143] with the parameters c =100, g= 0.005 and t= 0.

4.3.3 Experimental results

The objective of this experiment is to classify the philtrum area width automatically to three classes, baseline compared to human expert manual annotation [295] and to investigate for the best curvatures combination for this purpose.

The extraction of curvatures depends on the selection of the ring size, the size of the neighborhood (number of mesh layers) around the vertex

used to calculate the curvature tensor, there is no standard number of the ring size. Therefore, the range between 1 and 10 was tested for the best ring size selection.

Figure 4.7 illustrates the affect of ring size on mean curvature values, it is clear the selection of this factor affect the surface smoothing which it is required to be high value if noise is present, but smoothing can also mask surface detail.
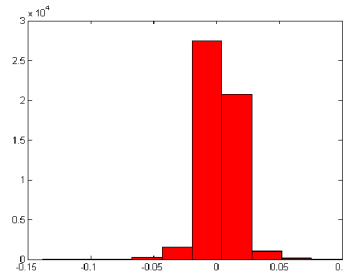
The mean curvature with ring size =1
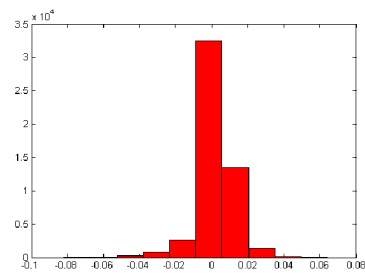
The mean curvature with ring size =3
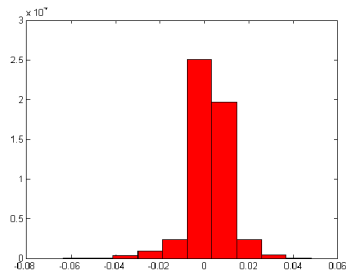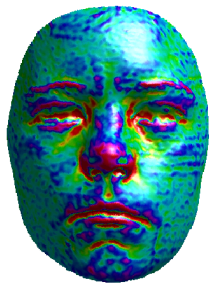
The mean curvature with ring size =5
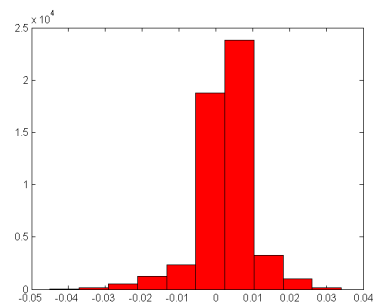
The mean curvature with ring size = 10

Figure 4.7: The ring size

After comparing the classification accuracy for each case, the best value was found to be 5 with a classification accuracy reaching 92% when using the mean curvature feature, and 89% and 87% when using the max,

min curvature respectively. The same was done when specifying the feature descriptor size as they were quantised into to 5, 10, 20, 30, 40, and 50 bins. The best classification accuracies were acquired when the descriptor size was between 20–50 bins, the accuracies did not increase dramatically when the descriptor size was larger than 50 bins.

The bar chart in Fig. 4.8 illustrates the classification accuracy for the curvatures as described above. It shows that the curvature mean, max, and min respectively offered the best classification accuracy. In contrast, the worst accuracy was achieved by using the Gauss curvature and the max direction. Also, it illustrates the dependence of the classification accuracy on the number of bins used in the classification descriptor. Additionally, Fig. 4.9. illustrates the classification accuracy when different curvatures were combined into a single vector to acquire better classification outcome. The results for feature combinations did not produce better results in all cases. For example, when the mean and Gauss curvatures were combined, the accuracy of 65% was achieved. While, the best combination was the max and min with the classification accuracy reaching 92% and exceeding it, but when the Mean, max and min were combined, better performance accuracy of 97% is achieved.
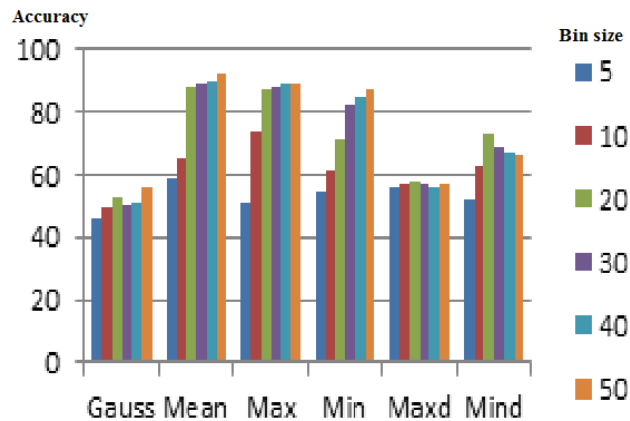
Figure 4.8: Classification accuracy for different descriptor size
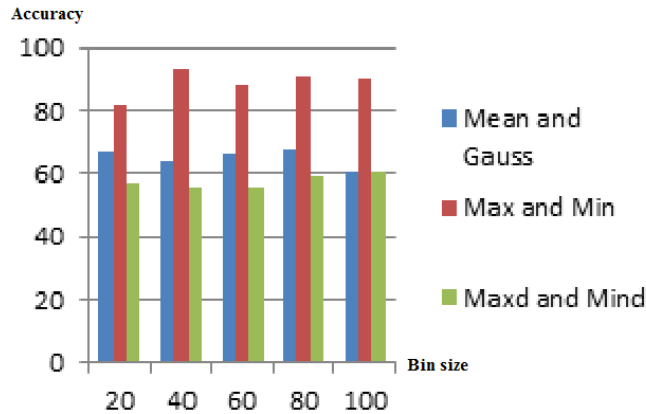
Figure 4.9: Classification rate for compound features

4.4 Novel geodesic path curvature feature descriptor

Geodesic paths have been used extensively in face recognition systems (e.g., see [5, 24, 139]). These studies employed the radial geodesic paths or iso-geodesic paths of the whole face as features for FR purposes.

The previous section demonstrated the robustness of the curvatures in the classification of the philtrum area width. This leads us to use the curvatures but in a novel strategy to classify and cluster the face morphological traits. The *geodesic path* between anthropometric landmarks is used to define key points, from which the curvature features descriptor can be extracted. These geometric features are introduced because we would like to classify and categorise specific anatomical facial traits, such as nose tip shape, nose ridge shape, Cupid pow shape, etc., and to investigate the effect of anthropometric landmarks and the geodesic paths between them on facial traits analysis. In the following chapters, these novel geometric features are used to classify the lips and nose morphological traits and find the relation between the gender and facial morphology.

Many previous studies have used geodesic and Euclidean distances as features for 3D facial morphology analysis (e.g., see [81, 109, 276, 287]). These quantities are also calculated and their effectiveness is assessed for lips trait classification and categorisation, as well as for gender classification for the purpose of comparison.

4.4.1   Geodesic path and distance

The geometric features used in this present work were curvatures of the *geodesic path* between two anthropometric landmarks. The geodesic path is the shortest curve or route between two points on a surface and the *geodesic distance* is the length of this curve [200]. There are a number of algorithms that may be used to compute geodesic paths and distances on triangular meshes; some are approximate, such as the fast marching method [218], while others are exact (and relatively slow). The exact algorithms include the Mitchell–Mount–Papadimitriou (MMP) method [264], the Chen and Han (CH) method [57], Xin and Wang method [299], and heat method [74]. Figure 4.10 illustrates the difference between the geodesic distance or path and the Euclidean distance. The Euclidean distance $\mathscr{D}_{BA}$ between points A and B in three dimensions defined by their coordinates (X_A, Y_A, Z_A) and (X_B, Y_B, Z_B) is calculated using the classic formula:

$$\mathscr{D}_{BA} = \sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2 + (Z_B - Z_A)^2} \qquad (4.6)$$
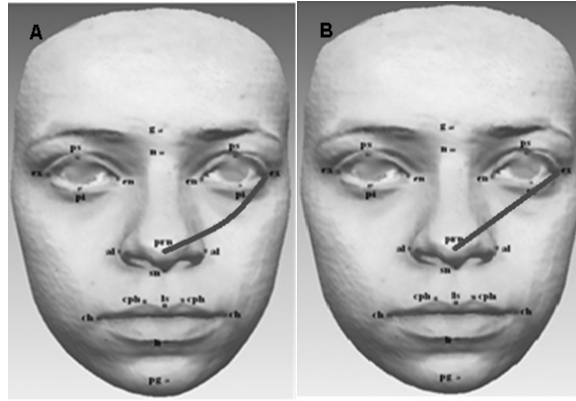
Figure 4.10: Geodesic distance (A) is the shortest surface distance between any two landmarks on the mesh, while Euclidean distance (B) is the straight-line distance distance between the two landmark

Fast marching method

Fast marching algorithm is a widely used algorithm in computer vision and computer graphics [23, 68]. This algorithm can be described as follows. Suppose we are given a metric $M(s)\,ds$ on some manifold $\mathscr{S}$ such that $M > 0$. If we have two points, $n_0 \in \mathscr{S}$ and $n_1 \in \mathscr{S}$, the weighted

geodesic distance between n_0 and n_1 is defined as

$$d(\mathscr{S}, n_0, n_1) \stackrel{\text{def}}{=} \min_{y}\left(\int_0^1 ||y'(t)|| M(y(t))\, dt\right) \tag{4.7}$$

where the y's are all possible piecewise regular curves on $\mathscr{S}$ such that $y(0) = n_0$ and $y(1) = n_1$. Fixing the point n_0 as the starting point, the distance $U(n) = d(\mathscr{S}, n_0, n)$ to all other points, n, can be computed by propagating the level set curve $C_t = \{n : U(n) = t\}$ using the evolution equation $\partial C_t / \partial t(n) = N_n / M(n)$, where $\{N_n$ is the exterior unit normal to C_t at the point n and $U(n)$ satisfies the nonlinear *Eikonal* equation [218]:

$$\|\nabla U(n)\| = M(n) \tag{4.8}$$

Exact geodesic algorithm

The basic idea of the exact algorithm is to track together groups of shortest paths that can be parameterised automatically. This is achieved by partitioning each mesh edge into a set of intervals that are called windows. The windows are then propagated across faces of the mesh in a Dijkstra-like sweep. Therefore, for a triangle mesh surface S and a source vertex $v \in S$, the exact algorithm computes an explicit representation of the geodesic distance function $D : S \rightarrow R$. For any point $p \in S$, this function $D(p)$ returns the length of the geodesic path from p back to the source v. Once a complete representation of D has been computed, one can quickly apply a "backtracking" algorithm to compute the shortest path from any query point to the source. These shortest paths are ruled by the following three attributes. Interior to a triangle, the shortest path must be a straight line. When crossing over an edge, the shortest path must correspond to a straight line [204, 264]. Exact methods are more accurate in tracking geodesic paths than approximate methods (Fast marching) and are especially advantageous for low-resolution meshes, where processing overhead is not so critical. Figure 4.11 highlights the difference between an exact and a fast geodesic algorithm in determining geodesics in a synthetic low-resolution mesh, where the black exact trajectory is clearly following the mesh edges more regularly than the red fast trajectory.
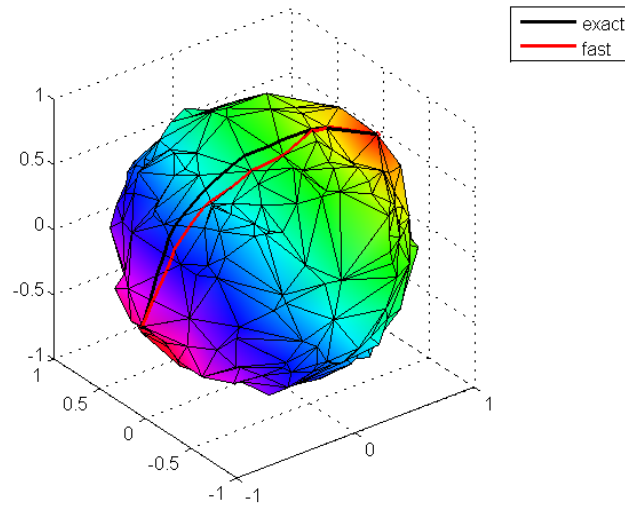
Figure 4.11: The exact and fast marching algorithms. Two paths extracted for a synthetic mesh.

4.5   Discussion

The new 3D geometric features proposed in this chapter combined the advantages of the curvatures with a geodesic path exist between the anthropometric landmarks for efficient face shape analysis. Specifically, these features exploit the mean and Gaussian curvatures, shape indices, and curvedness measures obtained from the *geodesic paths* between anthropometric landmarks of facial morphological traits.

These curvatures can be computed directly on the triangle meshes, making the curvature based information more sensible and providing a credibility in representing the triangle mesh shape changes. In this chapter, the classification accuracy proved that the curvatures combination can be used to enhance the 3D face morphology classification performance. Also, the curvatures are robust geometric features, easy to extract, the time consumption for processing is very little.

Figure 4.12 illustrate, as example, the geodesic path for the lip area, curvatures are calculated for these paths vertices to produce the feature descriptors, which are utilised for supervised or unsupervised classification. The main problem is normalising the feature descriptors. In Chapters 5, 6 and 7 the procedure of extracting the geodesic curvatures for each

facial traits is explained in detail, and the method of normalising the feature descriptors is described minutely.
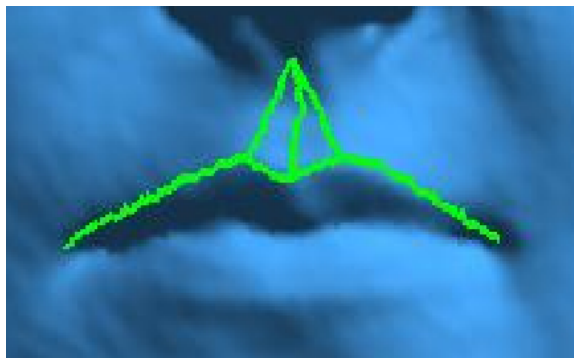
Figure 4.12: The lip geodesic paths

4.6   Summary

A new set of geometric features based on curvature measures computed from the geodesic path between 3D facial landmarks is proposed. These features are general features and can be tested in many 3D face applications, such as face recognition (FR) and Facial Expression Recognition (FER). In the following chapters they will be used for face morphology classification and categorisation, also for gender analysis and discrimination. Lips traits classification and categorisation method using these features is presented in the next chapter.

Chapter 5

AUTOMATIC LIPS MORPHOLOGY TRAITS CLASSIFICATION AND CATEGORISATION

The face is the most expressive part of the human body and is essential in everyday social interaction. The lips are one of the key components of the face (Fig. 5.1); the lips area runs from the base of the nose to the tip of the chin and therefore constitutes most of the lower third of the face. The lips include the philtrum and Cupid's bow. The lip vermillion is the thin layer of skin, red in colour, overlying a highly vascularized region. The appearance of the lips varies with facial movement; therefore, the lips should be assessed when the subject is relaxed and has a natural head position [257]. The morphological features of the lips vary greatly between individuals and are particularly dependent on age, sex and ethnicity [79, 246].
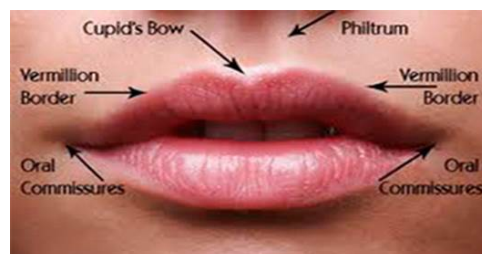
Figure 5.1: Lip traits: basic morphological lip features, which vary greatly between individuals.

Studying lips morphology is important for a number of diverse ap-

plications, including face recognition and gender classification [159, 237, 249, 260]. The lips have also been reported to contribute to facial attractiveness and a number of studies have attempted to evaluate lip aesthetics by creating norms and standards of ideal lip position and shape based on cephalometric analysis and facial measurements [49, 131, 192, 231]. Furthermore, lip morphology plays an important role in the diagnosis and analysis of many medical syndromes and face shape abnormalities, as well as in finding genetic variants associated with these syndromes.

The cleft lip (CL) is one of the most recognisable facial anomalies which has been the focus of clinical research for many decades. For example, an in-depth review of 20 CL classification schemes was provided in [289], while the association between genetic variants and environmental factors associated with cleft lip was analysed in [29]. Fetal alcohol syndrome (FAS) is another medical condition where lip morphology is used in a diagnostic role. FAS is reported to be associated with certain characteristic facial features, including smooth philtrum, thin upper lip vermilion and short palpebral fissure length [265].

Thus, the identification and classification of lip shape characteristics is important in many applications. So far, the categorisation (clustering) of lip morphological characteristics (traits) has mostly been performed manually in clinical practice which is time consuming and prone to operator error.

The aim of the present work is to devise a method for automatic classification and categorisation of lip morphological traits. The method relies on a combination of 3D geometric features. Six lips area shape traits (Philtrum, Cupid's Bow, lips contours, chin, and lower lip tone) Fig. 5.2 are investigated using our approach. These traits were categorised manually by medical experts in previous research and the results of manual categorisation and classification are available to us in this study [295] to benchmark our approach results against a clinical expert results.
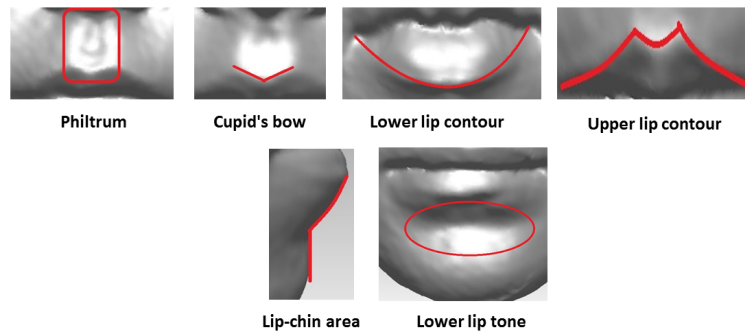
Figure 5.2: Lip traits

Other research contributions is a new method for visualisation and analysis of the effect of the found categories on the model of the lip area. To this end, partial least squares (PLS) regression [136] is used. Many studies proved the PLSR effectiveness in analysing the effects of biological and environmental variables on facial shape. Shrimpton et al. [254] used a spatially-dense multivariate regression to investigate the effect of metadata (age, gender, and BMI) on facial form and tissue depth in 67 males and 89 females. Matthews et al. [185] showed the effect of gender and head size variation in 473 one-year-old infants. Claes et al. [61] explored sex and ethnic variations in 592 Africans and Europeans and studied the effect size and statistical significance of sex in terms of local shape characteristic changes (area, curvature and normal displacement). Consequently, PLSR is adopted with dummy variables for visualising the influence of the discovered categories on the facial physical appearance to gain insight into suitability of categories for description of the underlying facial traits.

The remainder of this chapter is organised as follows: the next section outlines the proposed approach, while Section 5.2 describes six computational experiments performed to evaluate the proposed automatic classification and categorisation approach. Section 5.3 provides a discussion of the results. Finally, Section 5.4 presents summary remarks.

5.1   The proposed approach overview

In this section, an automatic lip morphology classification and categorisation approaches are described in detail. This is the first time that such an automatic system has been developed, with past approaches being de-

pendent on manual work of highly trained clinicians. An automated approach has clear benefits in bringing such technology to a wider general use, as well as opening up potential for the analysis of large datasets with direct benefit to such research areas as finding genetic associations and automatic facial syndromes diagnosis.

Figures 5.3 and 5.4 shows the block diagrams of the proposed approaches which are then further explained below.
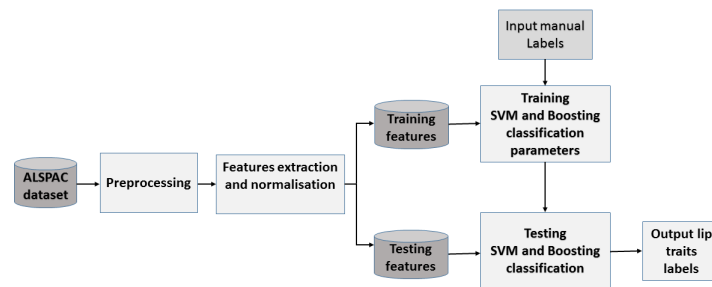
Figure 5.3: Block diagram of the proposed automatic lip traits classification approach
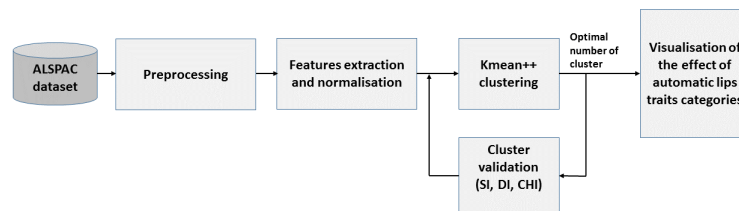
Figure 5.4: Block diagram of the proposed automatic lip traits categorisation approach

5.1.1   The ALSPAC dataset and lips traits annotation

The ALSPAC dataset includes 4747 three-dimensional images of normal young subjects, each with 21 facial landmarks. Eight anthropometric

landmarks (sn, ls, li, cphL, cphR, chL, chR and pg) were localized in the lips region Fig. 5.5.
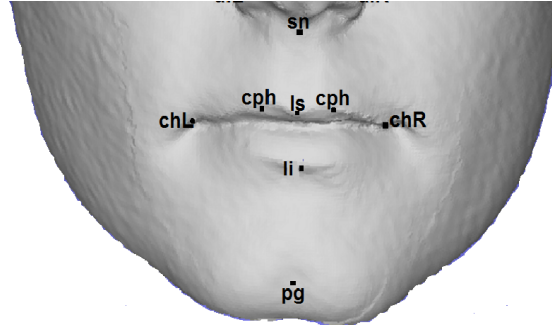
Figure 5.5: Lip landmarks

In this work, six lips traits previously identified in [295] are considered namely, philtrum shape, Cupid's Bow shape, lips contours shape, chin shape, and lower lip tone shape (Fig. 5.2). Depending on its appearance each trait was manually categorised into between three and seven classes, with the corresponding labels assigned to all ALSPAC images [295]. Some of the categories in the ALSPAC dataset were unbalanced, i.e, only a few images had the corresponding labels. For example, the philtrum shape was categorised into seven classes, two of which were found to be minority categories (contain only 150 and 180 subjects). Therefore, choosing a suitable method to resolve this problem was essential.

5.1.2 Data preprocessing

The subjects were scanned with a laser camera at high resolution resulting in high density meshes. In this work, the proposed approach is tested for both the non-regularised and regularised meshes.

5.1.3 Feature extraction and normalisation

Two methods to extract geodesic curvature features are employed. The Euclidean and geodesic distances are also calculated to assess their effectiveness for lip traits classification and categorisation as compared to geodesic curvature features. Because, many previous studies have frequently used these quantities as features for 3D facial morphology analysis (e.g., see [81, 109, 276]).

Extracting geodesic paths

Gabriel Peyre's MATLAB fast-marching toolbox [216] and the exact geodesic toolbox [154] were used to find geodesic paths and calculate the geodesic distances between two landmarks. The former toolbox was used for high-resolution data (non-regularised meshes), while the latter one was used for low-resolution data (regularised meshes) because as we explained in Section 4.3.1 the exact method is more suitable for tracking low resolution mesh and this compatible with [38] finding.

Figure 5.6(a) illustrates the paths used for all lip traits, apart from the lower lip tone trait Table 5.1. For the lower lip tone trait we used the geodesic path between Lower lip contour landmarks (chL, li and chR) and four extra geodesic paths connecting the above landmarks shifted down by a certain distance Figure 5.6(b).
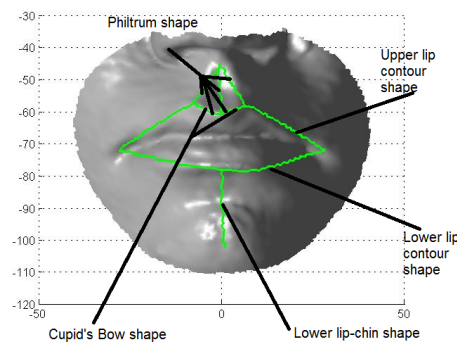
((a)) Paths for the lips traits
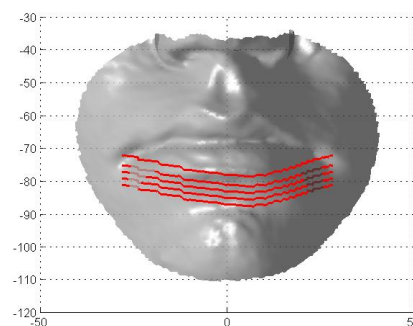
((b)) Paths for the lower lip tone

Figure 5.6: The geodesic paths used for lips trait classification and categorization

Table 5.1: List of geodesic paths defining morphological lip traits

Trait name	List of related geodesic paths
Philtrum shape	sn-cphL, sn-ls, sn-chpR, cphL-ls, cphR-ls
Upper lip contour	chL-cphL, cphL-ls, ls-cphR, chpR-chR
Cupid's bow	cphL-ls, ls-cphR
Lower lip contour	chL-li, li-chR
Lip-chin area	li-pg

Curvature features

The local surface principal curvatures are first calculated for the points along the geodesic path; the other features (mean curvature, Gaussian curvature, shape index, and curvedness) are then calculated from the principal curvatures. The extraction of these features depends on the selection of ring size constraints which is the neighborhood around a vertex from which estimate of the curvature tensor is calculated. In this work, the ring size was selected experimentally to acquired better classification results, thus the ring size 2 is recommended. See Chapter 4 for more information on the curvatures calculation algorithm.

Normalisation of curvature features

Histogram normalisation is used to normalize the features descriptors. It is an estimate of the probability distribution of a continuous variable. It is a kind of bar graph, to construct a histogram, the first step is to "bin" the range of values—that is, divide the entire range of values into a series of intervals—and then count how many values fall into each interval. The bins are usually specified as consecutive, non-overlapping intervals of a variable. If the bins are of equal size, a rectangle is erected over the bin with height proportional to the frequency—the number of cases in each bin. A histogram may also be normalized to display "relative" frequencies. It then shows the proportion of cases that fall into each of several categories, with the sum of the heights equaling one [134].

In this work, each geodesic path has a different number of nodes (vertices) for curvature calculation. To deal with this, a normalised histogram distribution was calculated for each path feature; for this purpose, the number of bins selected was 5, 10, 15, 20, or 25, depending on the minimum number of nodes in a path across the entire sample, the node num-

bers for the longest path (li-pg) were around 100 node. The histogram representation has used for features normalisation and representation in many image and pattern recognition such as Bag-of-Words (BoW) image representation [168], this motivate us to use the histogram representation to normalise the geodesic path curvature features.

Let P_1^k, ..., P_n^k denote the vertices of a path P^k on facial mesh k and let m_i^k, g_i^k, c_i^k, and s_i^k denote, respectively, the mean curvature, Gaussian curvature, curvedness value, and shape index value evaluated at vertex P_i^k ($i = 1, \ldots, n$). For each path, we choose a number $b = 5, 10, 15, 20$, or 25 such that $b \leq \min n$, where $\min n$ is the minimum number of vertices in all paths P^k across the sample. After the histogram normalisation (using the MATLAB function `histnorm`) with b bins, we get exactly $4b$ characteristic curvature features for path P^k:

$$\mathbf{m}^k = \left[\hat{m}_1^k, \ldots, \hat{m}_b^k \right] \tag{5.1}$$

$$\mathbf{g}^k = \left[\hat{g}_1^k, \ldots, \hat{g}_b^k \right] \tag{5.2}$$

$$\mathbf{c}^k = \left[\hat{c}_1^k, \ldots, \hat{c}_b^k \right] \tag{5.3}$$

$$\mathbf{s}^k = \left[\hat{s}_1^k, \ldots, \hat{s}_b^k \right] \tag{5.4}$$

where ˆ denotes the respective values resulting from the histogram normalisation. Then a features descriptor is composed $\mathbf{D}^k = \left[\mathbf{m}^k, \mathbf{g}^k, \mathbf{c}^k, \mathbf{s}^k \right]$ consisting of $4b$ components. Because a different number of geodesic paths are extracted for each traits, the features descriptor for each trait can be calculated by concatenating its paths' descriptors see Fig. 5.6 and Table 5.1.

5.1.4   Data balancing and classification

The manual lip traits categories obtained in [295] are imbalanced (different classes have different number of subjects), and hence, balancing data method is essential to handle this problem. The problem of imbalanced datasets can arise in classification when the number of elements in one class is much lower than that in other classes. The main challenge in the imbalance problem is that such small classes are often equally or even more important than other classes. Standard classifiers naturally tend to overestimate the importance of the larger classes and underesti-

mate the importance of the smaller classes. To cope with this problem, several methods have been suggested [48, 223, 239] which can be divided into two groups: (i) data dependent methods and (ii) classifier dependent methods.

Data dependent methods use data resampling to reduce the effect caused by class imbalance. One approach is oversampling, in which the dataset is balanced by replicating elements of a minority class. The advantage of this approach is that there is no data loss. The drawback is that it may cause over fitting and additional computational overheads. Another approach uses undersampling, suggesting that the dataset is balanced by deleting some elements of a majority class. The main drawback of this approach is that potentially useful information is neglected [223].

Classifier dependent methods are based on creating innovative algorithms or modifying existing ones to tackle the class imbalance problem by constructing an efficient classifier. For example, Cost Sensitive Learning (CSL) is associated with misclassifying patterns. A cost matrix is utilised to decrease the relative weight of majority classes and increase the relative weight of minority classes to ensure they have similar significance. The advantage of the CSL method is that no data is replicated or eliminated; however, weighing data to specify relative costs for misclassification can be a difficulty [179]. A common approach to balance datasets is to use the so-called boosting algorithm. Many studies (e.g., [263, 307]) have shown that the boosting method is quite effective in handling different types of unbalanced data. This method is an iterative technique to enhance the performance of weak classifiers. This method is much more helpful in tackling class imbalance problems, because it is the minority class elements that are mainly expected to be misclassified and, hence, should be assigned higher weights in subsequent iterations [162, 239].

Multiclass boosting

The present work uses the multiclass boosting method to classify the unbalanced ALSPAC dataset [307] and compares the results with those obtained using the multiclass SVM (supper vector machine) method which does not do data balancing. The basic idea of boosting is instead of learning a single complex classifier, learn several simple classifiers, then combine the output of the simple classifiers to produce the classifica-

tion decision. In this case, any weak learner can be potentially iteratively boosted to become also a strong learner [70]. In this work, public MATLAB software called *Multiclass Gentle Adaboosting* is used for classification purpose [210] with the parameters $options.weaklearner = 0$, $options.epsi = 0.1$, $options.lambda = 1e-2$, $options.max-iteration = 2000$ and $options.T = 10$. This Software uses Gentle Adaboost classifier and the multiclass problem is performed with the 1AA strategy. Gentle AdaBoost (GA) is one of the AdaBoost (Adaptive Boosting) variants [313]. The AdaBoost algorithm's general steps are:

- Individual weights are given for data samples.

- Successive classifiers are learned on weighted versions of the data.

- The entire training set is considered in order to learn each classifier.

- The weight is updated by the misclassification of the previous classifiers.

- The next classifier focus on the most difficult patterns

- The algorithm stops when the error rate of a classifier is >0.5 or it reaches the maximum number of iteration which is initialised by the user.

Gentle AdaBoost (GA) improves it by using weighted least-squares regression. Commonly, GA produces a more stable and reliable classification ensemble [70, 313].

5.1.5   Automatic categorisation

The aim of data clustering (categorisation), or cluster analysis, is to discover the natural grouping of a set of points, objects, or patterns. Webster [191] defines cluster analysis as "a statistical classification technique for discovering whether the individuals of a population fall into different groups by making quantitative comparisons of multiple characteristics".

This work investigates a way for automatically discovering lip traits categorisations by using 3D geometric features. Unsupervised learning (partitioning) has been investigated before in many applications. For example, in character recognition [30]. It has also been explored in biometrics, such as in fingerprint recognition where the fingerprints divided

into classes before the classification [184]. For example, Liu et al. [177] applied Kmeans in PCA space to partition the gallery set into a number of clusters, each containing a subset of face images having similar characteristics. Yank et al. [301] investigated the problem of automatically discovering human faces categorization from a collection of images, using also Kmeans for clustering the SIFT images feature.

Since the task of clustering is subjective, the means that can be used for achieving this goal are plenty. Every methodology follows a different set of rules for defining the 'similarity' among data points. In fact, there are more than 100 clustering algorithms known. But few of the algorithms are used popularly. In general, the clustering methods can be divided according to similarity measurement as follow:

- **Connectivity models**: These models are based on the notion that the data points closer in data space exhibit more similarity to each other than the data points lying farther away. These models can follow two approaches. In the first approach, they start with classifying all data points into separate clusters and then aggregating them as the distance decreases. In the second approach, all data points are classified as a single cluster and then partitioned as the distance increases. Examples of these models are hierarchical clustering algorithm and its variants.

- **Centroid models**: These are iterative clustering algorithms in which the notion of similarity is derived by the closeness of a data point to the centroid of the clusters. K-Means clustering algorithm is a popular algorithm that falls into this category. In these models, the no. of clusters required at the end have to be mentioned beforehand, which makes it important to have prior knowledge of the dataset. These models run iteratively to find the local optima.

- **Distribution models**: These clustering models are based on the notion of how probable is it that all data points in the cluster belong to the same distribution (For example: Normal, Gaussian). These models often suffer from overfitting. A popular example of these models is Expectation-maximization algorithm which uses multivariate normal distributions.

- **Density Models**: These models search the data space for areas of varied density of data points in the data space. It isolates various different density regions and assign the data points within these regions in the same cluster. Popular examples of density models are DBSCAN and OPTICS [238].

Many studies have compared the computing performance and clustering accuracy of different clustering algorithms [94, 214, 275]. In spite of that, they concluded there was no superior clustering algorithm, but the performance of the clustering algorithm depends heavily on the nature of dataset. The present study carries out automatic categorisation of lip traits by using Kmeans++ clustering technique. The standard Kmeans has been successfully utilised in many computer vision application especially in medical image processing (e.g., see [167, 209]. The Kmeans++ clustering algorithm has been shown to provide better categorisation (clustering) results comparing to standard Kmeans [19, 152] by updating the initialization mechanism to select cluster centroids.

Kmeans++

This method work on the same concept of kmeans for data clustering by choosing C initial centroids, where C is a user specified parameter. The most common form of standard kmeans is deceptively simple: the collection of points are assigned to the closest centroid to form the cluster, then compute new centroids based on new clusters, these steps repeated until the centroids remain the same (no change). kmeans++ outperformed kmeans, both by achieving a lower potential value and also by completing faster because Kmeans++ addressed the problem of choosing the cluster center by specifying a procedure to initialize the cluster centers before proceeding with the standard kmeans optimization iterations. With the kmeans++ initialization, the algorithm is guaranteed to find a solution that is $O(logC)$ competitive to the optimal kmeans solution [19, 152].

Cluster validation approaches

Determining the number of clusters relies on the cluster validity indexes. In order to determine the optimal number of clusters C^*, other parameters are fixed and parameter C is optimised by the validity indexes. The

steps for determining the optimal number of clusters is shown in Fig. 5.7. Given the dataset M, a specific clustering algorithm and a fixed range of number of clusters $[Cmin, Cmax]$, thus basic procedure involves the following:

- Repeat a clustering algorithm successively for the number of clusters C from the predefined range $[Cmin, Cmax]$.

- Select the C^* for which the grouping provides the best result according to the validity index.

- Compare the optimal C^* value with external information if available.
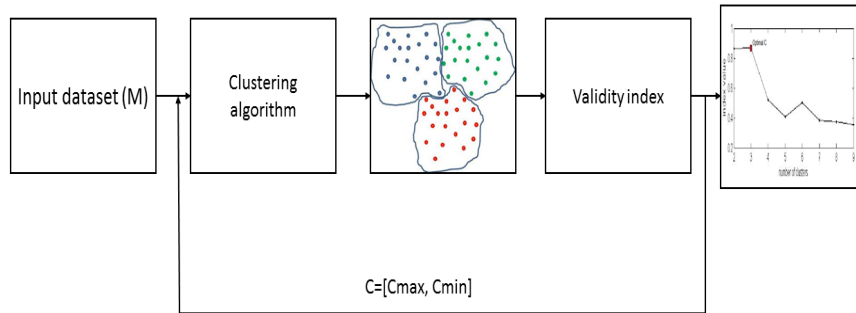
Figure 5.7: Validating the quality of clustering

In this work, three cluster validation methods were used to specify the number of clusters for each lip traits.

Silhouette index(SI)   The silhouette index (SI) validates the clustering performance based on the pairwise difference of between and within-cluster distances. Moreover, the optimal cluster number is specified by maximising the value of this index. This index is computed by:

$$s(i) = \frac{(b(i) - a(i))}{Max\{a(i), b(i)\}} \tag{5.5}$$

where $a(i)$ is the average distance between the ith element and all other elements within the same cluster while $b(i)$ is the minimum average distance between the element ith to any other cluster, of which the element ith is not a member [178, 229].

Dunn index   the Dunn index (DI) measures the minimum pairwise distance between objects in different clusters as the inter-cluster separation and the maximum diameter among all clusters as the intra-cluster compactness[1]. The optimal cluster number is specified by maximizing the value of this index. To calculate the Dunn index we firstly have to compute the distances between all the data points as follow:

$$Dunn = min_{1 \leq i \leq c} \left\{ min \left\{ \frac{d(c_i, c_j)}{max_{1 \leq k \leq c}(d(X_k))} \right\} \right\} \tag{5.6}$$

where $d(c_i, c_j)$ defines the intercluster distance between cluster X_i and X_j and $d(X_k)$ is the intracluster of cluster (X_k) and c is the number of cluster [178, 229, 230].

Calinski-Harabasz index(CH)   The Calinski–Harabasz index (CH) evaluates the cluster validity relying on the average between- and within cluster sum of squares. CH index calculates separation based on the maximum distance between cluster centers, and measures compactness depending on the sum of distances between objects and their cluster center. In addition, the optimal cluster number is specified by maximizing the value of this index. This index is computed by:

$$CH = \frac{trace(S_B)}{trace(S_W)} \cdot \frac{n_p - 1}{n_p - k} \tag{5.7}$$

where (S_B) is the between cluster scatter matrix, (S_W) the internal scatter matrix, n_p the number of clustered samples, and k the number of clusters [229, 230].

5.1.6   Visualisation using partial least squares regression

Finally, we would like to analyse the relationship between automatically or manually determined traits categories and the geometric characteristics of the corresponding facial region. A common approach to discover the relationships in data is regression. The challenge in using dense data is the large number of correlated dependent variables in comparison to the number of observations, leading to model instability when using the linear least squares regression. We have addressed this problem

---

[1]Compactness measures how closely related the objects in a cluster are [178]

by using the more advanced technique of partial least squares regression (PLSR) [136].

The aim of partial least squares regression is to establish a linear relationship between two sets of variables, X and Y, where X is the set of dependent variables and Y is the set of independent variables. In our case, Y is the set of manual or automatic categories dummy variables while X is the set of the x, y, and z coordinates of all vertices in a regularised facial mesh.

The main difference between linear regression and Partial Least Squares regression is that the latter treats both sets of variables symmetrically since they are both assumed to be subject to a common underlying cause, whereas linear regression considers one set to determine the other.

In addition, linear regression is based on a statistical model where all error is attributed to the dependent variable since it is assumed that the independent variable is measured with zero error. This stands in sharp contrast to PLS where no underlying model is used so no errors are attributed to either set of variables.

A third significant difference in the case of more than one measurement lies in the fact that PLS assesses the covariance between two complete sets of variables, whereas multiple linear regression calculates the regression coefficient for a dependency on each parameter separately keeping all other parameters constant [310]. However, also with PLS it is possible to keep every parameter but one approximately constant since it depends on how X in equation (5.9) is constructed.

Taking the above properties of the different kinds of regression into account it becomes clear that PLS regression is the more suitable approach when different effects in the lips need to be investigated where it is not clear yet what their mutual effects will be. Indeed, if the first of two highly correlated parameters is changed without constraining the second one to remain constant, the effect on the second parameter caused by the first one can be visualized. In equation (5.8) X and Y are both matrices where each column represents the values for a different 3D face or the lips area of the face of the ALSPAC database. Each row of X represents the values for a different parameter and each row of Y represents an x, y or z coordinate of a point on the 3D face or its lips area. X and Y in equation (5.9) on the other hand are column vectors where X contains

values that indicate the degree of change for every parameter in X, which means that for every parameter of X that needs to be kept constant, the corresponding value in X is equal to zero. The column vector Y is the calculated effect on the coordinates of the 3D face lips area or the whole face due to X. Equation(5.10) explains how a new shape $\hat{Y}$ can be obtained from the old one $\tilde{Y}$ [224].

$$Y = \beta X \tag{5.8}$$

$$\Delta Y = \beta \Delta X \tag{5.9}$$

$$\hat{Y} = \tilde{Y} + \Delta Y \tag{5.10}$$

The calculation of is preceded by an orthogonal projection of the matrices X and Y similar to the projections in PCA. However this time the projections are not calculated based on eigenvalue decomposition of the variance-covariance matrix, as is the case in PCA, but on singular value decomposition (SVD) of the covariance matrix describing the covariance between the two sets. The reason for choosing SVD is that this covariance matrix is not necessarily a square symmetrical matrix in contrast to ordinary covariance matrices, because it is the covariance between two sets of variables that are in general not of the same size [310]. The orthogonal space is a simplified space where also the search for the direction of main covariance between the two sets to calculate beta is greatly simplified, just like in PCA where the simplified orthogonal space made it obvious to find the direction of main variance in a dataset.

However, whilst categorical variables with two values may be directly entered as predictor or predicted variables in a multiple regression model, the categorical variables with more than two values cannot be entered directly into a regression model, therefore, *dummy variables* will be used [261].

A dummy variable is an artificial variable created to represent an attribute with two or more discrete values rather than continuous values as in standard regression. Therefore, dummy variables are created in such situations to force the regression algorithm to analyse attribute variables correctly. The categorical variables (labels) were converted into dummy

variables [261]. For C categories, we need to $C-1$ dummy variables before starting the regression process to determine their multiple and partial effects on the lips region.

The effects of traits categories on lip morphology can be illustrated using color maps. Regression coefficients define a set of weights on each 3D face vertex. Together they define the magnitude and the direction of the vertex displacement per unit of the predictor (the predictors here are the label dummy variables). The values of interest represented in the color maps are: (1) the 'partial coefficients' (magnitude); (2) the proportion of the variance at each vertex explained (partial R^2) by the predictor; and (3) the significance of the effect at each vertex [61, 185].

5.2   Results

Six computational experiments were designed in order to assess the performance of proposed approaches for lips traits classification and categorisation. Experiments 1, 2, 3 and 4 were designed to investigate the best features to classify lip traits using the manual labels provided in [295]. In these experiments, the classification performance is measured using classification accuracy and AUC (Area Under ROC Curve) values. These values are the average results for 5-folds[2] cross validation runs. In Experiment 5, the proposed automatic approach for lips traits catogrisation is assessed. Finally, in Experiment 6, the effectiveness of manual and automatic categorisation are compared using visualisation method. A detailed explanation of these experiments is provided in the subsequent subsections.

5.2.1   Experiment 1: Classification using 3D Euclidean distances

The Euclidean distances between the landmarks defining the lip traits, as shown in Fig. 5.6(a), were used to classify the lip traits in this experiment because this distance measurement has used in many face morphology analysis studies (Section 2.3.2). The Euclidean distances between the landmarks pairs related to certain lip traits are combined together to form the classification descriptor, for example, five distances combined to classify the philtrum shape. Table 5.2 lists the accuracies as well as the AUC

---

[2]4745 ALSPAC face meshes were used in the classification task

values for the classification performed using SVM and boosting classifiers for all lip traits except for the lower lip tone. These accuracies indicate the Euclidean distance is shown to be unsuitable for lips traits classification, providing poor classification accuracies. However, the results are approximately the same for both preprocessing methods, which indicates the Euclidean distance measurement is not very sensitive to the mesh resolution. The results are shown for both regularised and non-regularised facial meshes.

Table 5.2: Classification results based on Euclidean distance: accuracies and AUC values. The classification is performed using the SVM and boosting methods for the regularised and non-regularised meshes.

Lip traits	non-regularised mesh		Regularised mesh	
	SVM	boosting	SVM	boosting
Accuracy				
Philtrum shape	56.7	**62**	56	60.9
Cupid's bow	56	**62.7**	55.4	62.5
Upper lip vermilion contour	56.6	**60**	57	58.9
Lower lip vermilion contour	57	**61.5**	56	60.7
Lower lip-chin shape	55	**60.8**	55.5	60.2
AUC values				
Philtrum shape	0.558	0.628	0.552	0.615
Cupid's bow	0.566	0.633	0.560	0.622
Upper lip vermilion contour	0.570	0.620	0.577	0.610
Lower lip vermilion contour	0.567	0.605	0.548	0.610
Lower lip-chin shape	0.553	0.605	0.549	0.600

5.2.2   Experiment 2: Classification using 3D geodesic distance

The second experiment used geodesic distances for the lip traits classification. These were used because many studies (e.g., see [119, 125]) suggest that geodesic distances could describe 3D models better than Euclidean distances. The geodesic distances were calculated using the fast marching and exact geodesic algorithms; these distances were between landmarks in the lips region, as shown in Fig. 5.6(a). Table 5.3 lists the classification results obtained using these features with both classification methods (SVM and boosting), for the regularised and non-regularised meshes. The accuracies were better when they compared with the results obtained using an Euclidean distance measure. The results also show little sensitivity to the preprocessing method, since the exact method is utlised to calculate the geodesic distances for regularised mesh and this provides better performance comparing to fast marching

method (Section 4.3.3).

Table 5.3: Classification results based on geodesic distance: accuracies and AUC values. The classification is performed using the SVM and boosting methods for the regularised and non-regularised meshes.

| | non-regularised mesh | | Regularised mesh | |
Lip traits	SVM	boosting	SVM	boosting
	Accuracy			
Philtrum shape	61.8	**66.9**	61	65.7
Cupid's bow	60.9	**65.6**	60.2	64.5
Upper lip vermilion contour	59.6	**65**	58.7	63
Lower lip vermilion contour	59.9	**65.6**	58.6	64
Lower lip-chin shape	62	**67.3**	61	65.6
	AUC values			
Philtrum shape	0.614	0.668	0.61	0.653
Cupid's bow	0.595	0.664	0.595	0.643
Upper lip vermilion contour	0.604	0.667	0.608	0.650
Lower lip vermilion contour	0.605	0.672	0.609	0.630
Lower lip-chin shape	0.624	0.659	0.618	0.654

5.2.3   Experiment 3: Classification using 3D geodesic path curvatures

The previous two experiments utilised the Euclidean and geodesic distances as morphological lip features, which are traditionally used for classification. In contrast, the third experiment is based geodesic path curvature descriptor. As explained in Section 4.3 this descriptor combines the mean curvature, Gaussian curvature, shape index, and curvedness calculated for the points of the geodesic path between landmarks (Fig. 5.6(a)). The combination of paths' curvatures for each lip trait represent the classification feature descriptor for this trait.

The lower lip tone was the most difficult morphological trait to classify, because its geometric features have to be determined in a narrow area very close to and below the lower lip contour. To this end, we used several geodesic paths approximately parallel to the lower lip contour to cover the desired area, as shown in Fig. 5.6(b). To find the optimal number of such paths and the distance between them, tuning tests were carried out that compared the classification accuracies. The best accuracy was obtained for five paths separated in the vertical direction approximately by 2 mm. Table 5.4 lists the classification accuracies and AUC values for both classification methods (SVM and boosting) for the regularised and non-regularised meshes. The classification accuracies increase markedly,

for non-regularised meshes the boosting classification accuracies are approximately between $71\% - 77\%$, while the classification accuracies for regularised meshes are approximately between $67\% - 72\%$. The classification accuracies depend on mesh resolution because the curvature features are sensitive to mesh resolution, which confirms the inference made in [102].

Table 5.4: Classification results based on geodesic path curvatures: accuracies and AUC values. The classification is performed using the SVM and boosting methods for the regularised and non-regularised meshes.

Lip traits	non-regularised mesh		Regularised mesh	
	SVM	boosting	SVM	boosting
Accuracy				
Philtrum shape	68	**74.8**	64.9	70.6
Cupid's bow	65	**70.7**	62	67.4
Upper lip vermilion contour	66	**75.7**	62.8	69.7
Lower lip vermilion contour	66	**74**	61.8	70.6
Lower lip-chin shape	66.8	**75.5**	63.7	71.8
Lower lip tone	67	**76.8**	63.7	72
AUC values				
Philtrum shape	0.665	0.759	0.640	0.712
Cupid's bow	0.637	0.694	0.615	0.662
Upper lip vermilion contour	0.654	0.738	0.618	0.690
Lower lip vermilion contour	0.649	0.732	0.625	0.689
Lower lip-chin shape	0.653	0.750	0.655	0.705
Lower lip tone	0.675	0.756	0.647	0.714

5.2.4   Experiment 4: Classification using combination of features

In this experiment, the classification performance was explored by using a combination of the Euclidean distance, geodesic distances and geodesic curvature features as classification descriptor since many studies used different geometric features combination in many 3D face applications (e.g. see [89, 170]. The Min-Max scaling approach was used to normalise the data to a fixed range zero to one [225].

The best performance was achieved when the geodesic distance and curvature features were combined in classification experiment. Table 5.5 displays the classification results using different combinations of features for the non-regularised and regularised faces. Using a combination descriptor from the geodesic curvature features and the geodesic distances increase the classification accuracy. For non-regularised meshes boosting classification accuracies are approximately between $72\% - 79\%$. While the classification accuracies for regularised meshes are between approxi-

mately $70\% - 74\%$.

Table 5.5: Classification results based on different combinations of features: error rates and AUC values. The classification is performed using the SVM and boosting methods for the non-regularised and regularised meshes. GD stands for geodesic distances, ED for Euclidean distances, and GC for geodesic path curvature features.

| Lip traits | Combination of features | | | | | |
| | GC + ED | | GC + GD | | GC + ED + GD | |
	SVM	boosting	SVM	boosting	SVM	boosting
	Accuracy (non-regularised faces)					
Philtrum shape	66.8	72	69	**76.8**	69	76.5
Cupid's bow	65	70	67	**72**	65.6	71.4
Upper lip contour	65	74.5	67	**78.5**	64.8	78
Lower lip contour	64	71.5	66	**76.4**	65.8	75.5
Lip-chin shape	64.8	73.2	69	**78**	67	76.8
	AUC values (non-regularised faces)					
Philtrum shape	0.655	0.725	0.664	0.770	0.680	0.772
Cupid bow	0.635	0.683	0.656	0.705	0.648	0.690
Upper lip contour	0.65	0.747	0.651	0.778	0.651	0.761
Lower lip contour	0.656	0.699	0.645	0.752	0.657	0.743
Lip-chin shape	0.640	0.723	0.690	0.759	0.665	0.757
	Accuracy (regularised faces)					
Philtrum shape	64.6	79.8	65.9	**72.6**	65.4	71.8
Cupid bow	61.5	65.3	63.7	**69.7**	61.9	68.8
Upper lip contour	60	69.5	63.8	**73.2**	62.8	71.7
Lower lip contour	62	70	64.8	**72.6**	64.9	72.6
Lip-chin shape	62	69.5	64.7	**74.4**	62.9	73.6
	AUC values (regularised faces)					
Philtrum shape	0.637	0.700	0.648	0.723	0.646	0.729
Cupid bow	0.605	0.645	0.643	0.687	0.615	0.674
Upper lip contour	0.604	0.685	0.629	0.728	0.630	0.697
Lower lip contour	0.600	0.680	0.640	0.698	0.635	0.692
Lip-chin shape	0.61	0.686	0.638	0.737	0.630	0.709

5.2.5   Experiment 5: Unsupervised categorisation of lips traits

In this experiment, the combination of geodesic curvature features and geodesic distances were used to categorise (cluster) the lip morphological traits without relying on manual categories to propose an automatic categorisation system. An unsupervised clustering scheme was utilised to partition these features into multiple clusters, each defined by a centroid, the Kmeans++ algorithm was used to perform the clustering.

Kmeans++ starts with allocation one cluster center randomly and then searches for other centers given the first one. So this algorithm uses random initialization as a starting point, this can give different results on different runs. Therefore, the Kmeans++ clustering algorithm was run 100

times. Out of 100 results, the clustering results that produced the minimal sum of squared distances score was chosen.

An analysis for the optimum number of clusters was carried out using internal cluster validation techniques. We computed the validity indexes CH, DI, and SI for all traits with the number of clusters C, ranging from 2 to 9, with the tests repeated 50 times to find a stable number of clusters. For example, Fig. 5.8 plots the validation results for lips traits, for the philtrum shape; the optimum number of clusters is 5 (based on DI and SI) or 7 (based on CH). For the other traits, the optimum number of clusters was obtained in the same manner : 3 for Cupid's bow, 3 or 4 for the upper lip contour, 3 or 4 for the lower lip contour, 5 for the lower lip-chin shape, and 5 for the lower lip tone shape. These results are almost the same as those produced by a medical expert in [295]. There are slight differences: for example, the philtrum shape was categorised manually into 7 clusters, while automatic categorisation sometimes produced five clusters and sometimes seven clusters. To be able to compare to the manual results in the next experiment, the philtrum shape with 7 automatic clusters are used. The numbers in Table 5.6 shows the percentage of the number of times the validation methods chose a certain number of clusters.

Philtrum shape
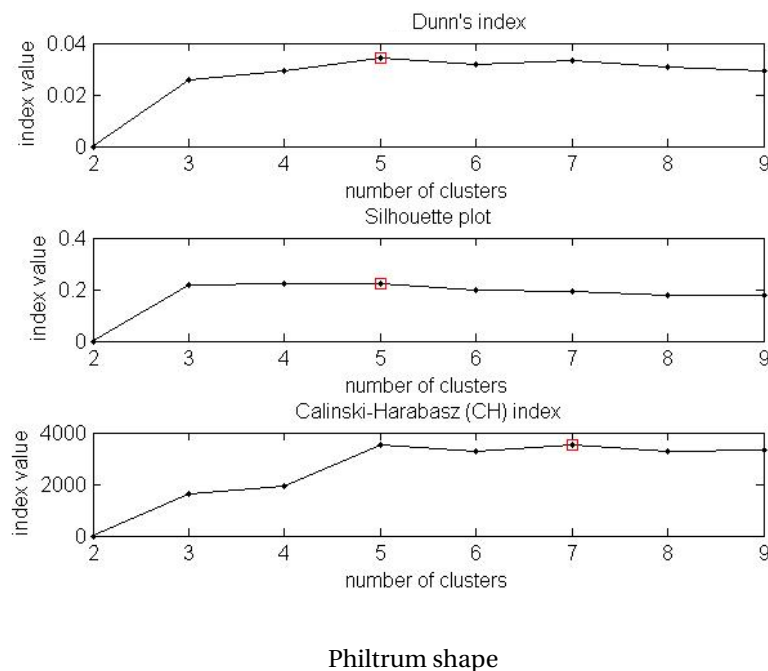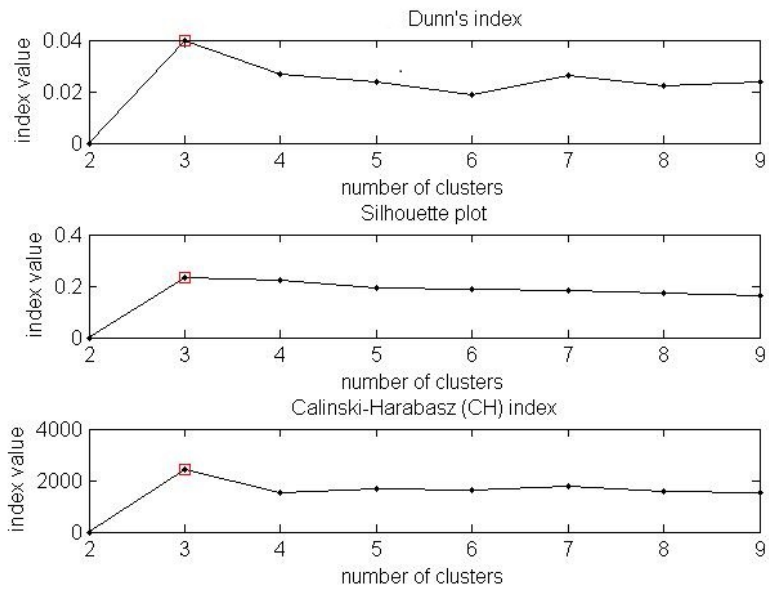
Cupid's Bow shape
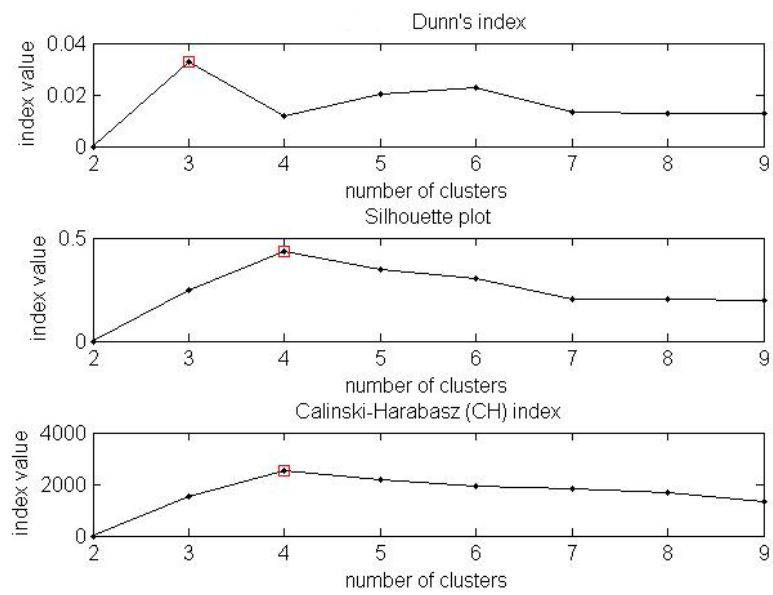
Upper lip contour shape
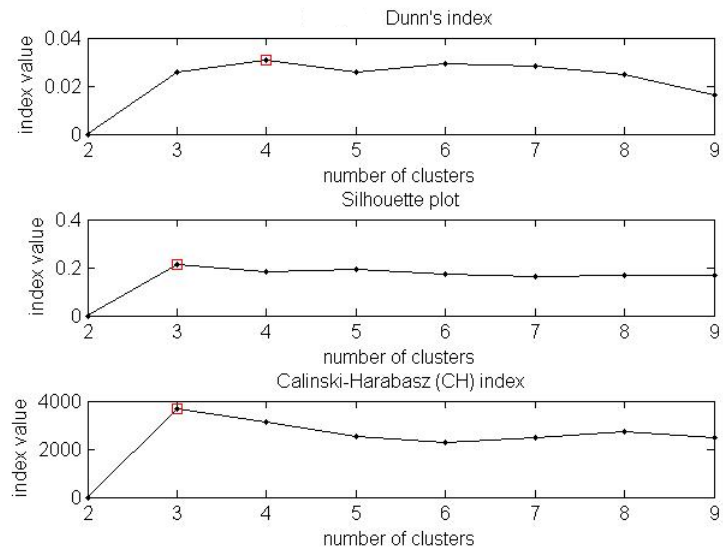
Lower lip contour shape
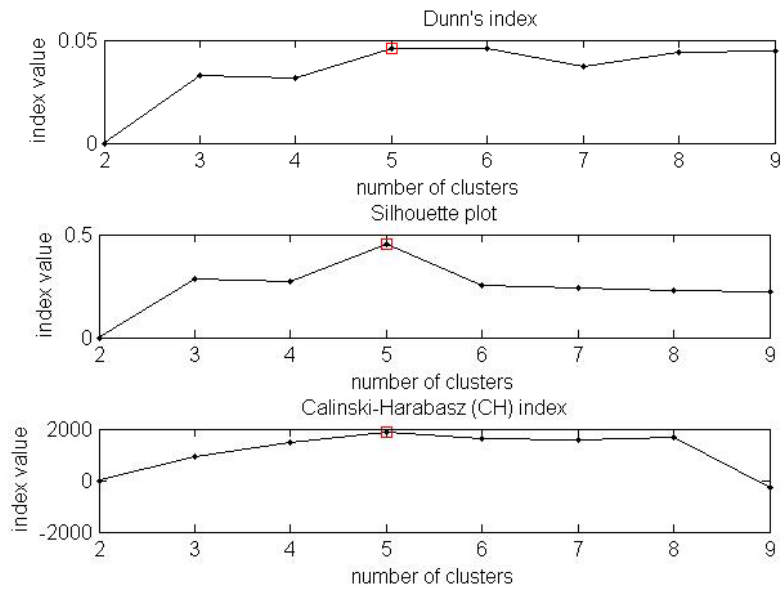
Lower lip- chin shape
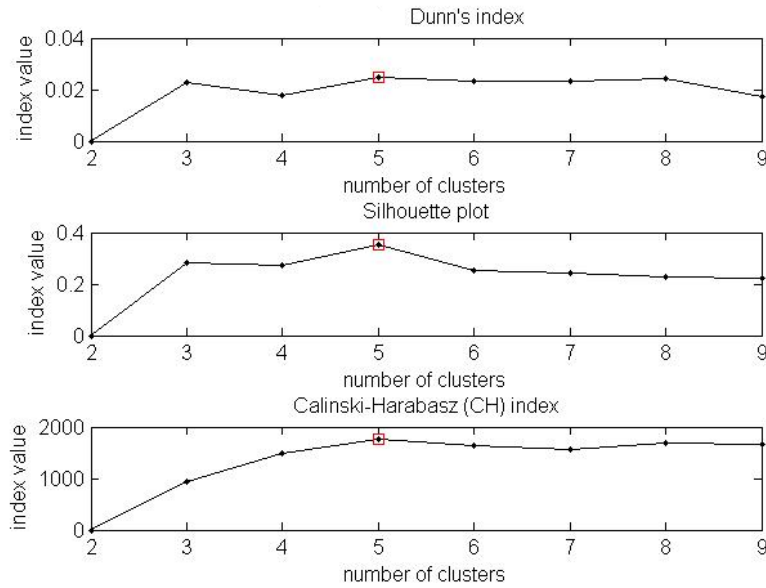
Lower lip- Tone shape

Figure 5.8: Validation index

Table 5.6: The percentage of the number of times the validation methods chose a certain number of clusters. Different numbers of clusters were found to be optimum using different validity indices: Dunn's index (DI), silhouette index (SI) and Calinski–Harabasz index (CH).

Traits	DI					SI					CH				
	3	4	5	6	7	3	4	5	6	7	3	4	5	6	7
Philtrum shape	5	10	35	20	30	0	10	45	10	35	0	0	30	15	55
Cupid's bow shape	100	0	0	0	0	100	0	0	0	0	80	20	0	0	0
Upper lip vermilion contour shape	70	30	0	0	0	45	55	0	0	0	30	70	0	0	0
Lower lip vermilion contour shape	40	60	0	0	0	55	45	0	0	0	75	25	0	0	0
Lower lip-chin shape	0	25	65	10	0	0	25	60	15	0	0	35	50	15	0
Lower lip tone shape	0	35	55	10	0	5	45	50	0	0	20	35	45	0	0

The best classification accuracy was in the range of 72% − 79% for the manual labels Table 5.5. The boosting classification approach was repeated but this time with the automatic labels. Table 5.7 shows classification accuracies for the combined descriptor from the geodesic curvature features and the geodesic distances, using both the manual and automatic lips morphological traits labels. In addition, the lower lip tone classification accuracy using the geodesic curvatures features and the au-

tomatic categorisation labels was 92.7%, while this trait classification accuracy was 76.8% when the same features and the manual categorisation labels are used. It is clear the automatic labels classification accuracies *outperform* that of the manual labels.

Table 5.7: Boosting classification results using the manual and automatic lips area traits labels

Trait name	Manual categories	Automatic categories
Philtrum shape	76.8%	89%
Upper lip contour	78.5%	87.6%
Cupid's bow	72%	84.8%
Lower lip contour	76.4%	90%
Lip-chin area	78%	86%

5.2.6   Experiment 6: Visualisation of the effect of traits categories on the lip area

PLS regression was used to characterise the effects of the traits categories on the regularised 3D faces. All statistical significance tests were based on 1000 permutations [16]. The partial effects (one variable is independent of the others) in the multivariate regression were coded by the partial regression coefficients. These coefficients define label weights at the mesh vertices, which were visualised as a heat map; cooler colours represent weaker effect on the vertex, while warmer colours represent stronger effect.

As an example, Fig. 5.9 illustrates the regression results for the manual philtrum shape labels, while Fig. 5.10 shows those for the automatic philtrum shape labels. In these figures the "partial coefficients" correspond to the magnitude of the vertex displacement in 3D; the proportion of variance that the predictor variable predicts at each individual vertex is shown in partial R^2. The effect of labels was displayed as colour maps of statistical significance using two colours, with yellow indicating highly significant results (p-value < 0.001) and green showing less significant or insignificant results (p-value ≥ 0.001).

In Figures 5.11 and 5.12 the detailed results for all trait dummy variable are visualised, in addition to the partial effect for each trait dummy variable, the multiple effect for all independent variables combined is also visualised. As can be seen from the multiple effect visualisation, all traits categories effect is concentrated and significant in the lips area.

partial coefficients R^2 $p < 0.001$ (yellow)
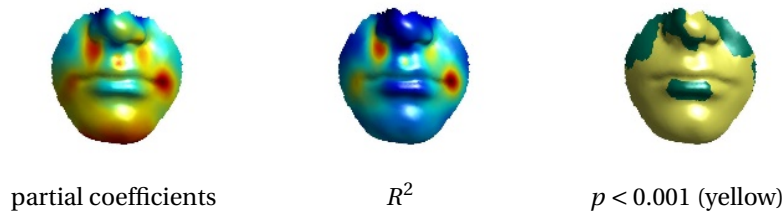
Figure 5.9: Regression results: effect of a manual label (philtrum shape) on the lower face

partial coefficients R^2 $p < 0.001$ (yellow)
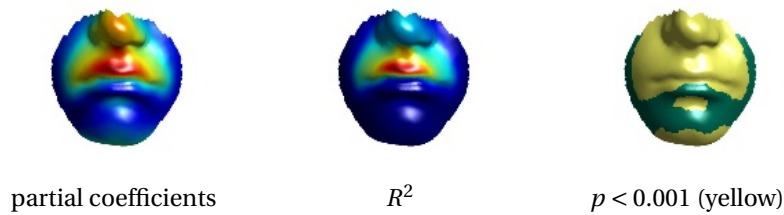
Figure 5.10: Regression results: effect of an automatic label (philtrum shape) on the lower face

5.3   Discussion

The first three experiments in the present study aimed to determine which facial features were the most effective in automatic classification of the lip traits using non-regularised and regularised 3D meshes. Although fairly effective for both the high and low resolution data, the Euclidean distances alone did not produce good enough classification accuracies. Experiment 2 showed that the geodesic distances between landmarks provided better classification accuracies. We attribute this to geodesic distances being more informative than Euclidean distance in describing facial surfaces; the classification accuracies are higher with only small differences between the non-regularised and regularised data. Experiment 3 used 3D geometric curvatures of the shortest geodesic path between two anthropometric landmarks as features in the classification experiments. The accuracies were found to improve markedly, which is due to the curvature features taking into account local facial geometry and, hence, characterising the lip shapes much better than the Euclidean and geodesic distances.

Most of the previous studies [110, 119] have focused on using combinations of facial features to achieve higher classification accuracies. In

the present study, we also adopted this approach and used a combination
of Euclidean and geodesic distances with *geodesic curvature features* (see
Table 5.5) to produce feature descriptors for the lips morphology classifi-
cation. The highest accuracy was obtained when geodesic distances were
combined with geodesic curvatures. The SVM (Support Vector Machine)
method failed to classify the lip traits efficiently, in contrast to the boost-
ing method, because the data are highly imbalanced in more than one
class; this finding seems to confirm the inference made in [45].

In spite of using different types and combination of features to clas-
sify the lip traits, the classification accuracies were not very high for the
lips traits manual categories provided in [295]. This encouraged us to
cateogrise the lip traits automatically using Kmeans++ algorithm. Three
internal validation techniques were used to select an optimum number
of clusters for each lip trait. The selection process was repeated 50 times
to find stable numbers of clusters. For the lower and upper lip contours,
these techniques were found to show quite contradicting results for three
and four clusters. For this reason, we used the categorisation results for
both alternatives in Experiment 6.

In Experiment 6, a new method was used to visualise the automatic
and manual categorisation (labeling) regression results for the lips area.
All previous studies [254], [185], and [61] dealt with continuous variables.
By contrast, the present study seems to be the first to deal with discrete
variables (automatic or manual categories) for face morphology classifi-
cation. Using dummy variables is a way to deal with the discrete variable
problem; for example, six dummy variables were used to represent the
philtrum shape trait which fall into seven classes. Figures 5.9 and 5.10
show the regression results for the philtrum shape trait with one dummy
variable.

Figure 5.11 details the regression results for the manual categories
showing the multivariable effect and contribution of each individual
dummy variable. Figure 5.12 details the regression results for the auto-
matic categories.

On careful examination of the results, one can see that our automatic
categories are fairly similar to the manual categories [295]. However, for
three out of seven traits, the lower lip-chin, lower lip contour and lower lip
tone shape, the automatic categorisation approach provide much better

Partial coeffs R^2 $p < 0.001$ Partial coeffs R^2 $p < 0.001$ Partial coeffs R^2 $p < 0.001$

Multiple effect

Upper lip border (3 clusters)

Lower lip-chin shape

Philtrum shape

Upper lip contour (4 clusters)

Lower lip tone shape

Lower lip border (3 clusters)

Cupid's bow shape

Lower lip contour (4 clusters)
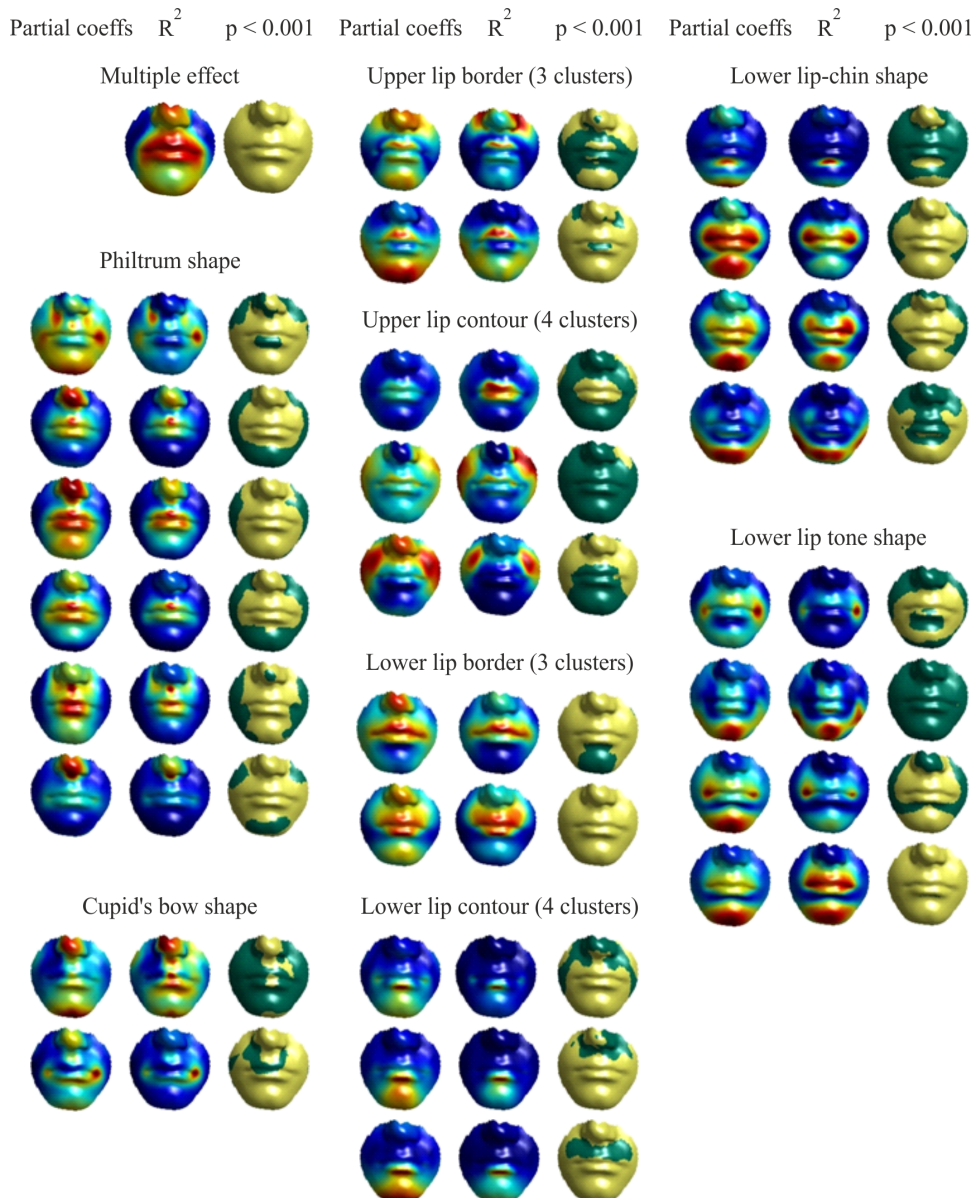
Figure 5.11: Visualisation of the effect of manual categories on the lower face based on the regression results for dummy variables. The 'Partial coeffs' columns display heat maps of the partial regression coefficients associated with mesh vertices (warmer colours correspond to stronger effects). The 'R^2' columns display heat maps of proportion of the variance. The '$p < 0.001$' columns show two-colour maps of the statistical significance of the effect: yellow for p-value < 0.001 and green for p-value ≥ 0.001.

categorisation results, as the automatic categories effect appears at the right areas (compare Figures 5.11 and 5.12). In particular, for the lower lip tone, the manual categories show a strong effect (red colours) near the oral commissures rather than in the mentolabial sulcus area, which is under the lower lip contour.

The cluster validity methods did not give us a clear indication of the optimum number of clusters for the upper and lower lip contours. However, it is apparent from Fig.5.12 that using four clusters would be preferable in both cases as the colours maps for respective dummy variables highlight the correct areas better.

In general, the classification accuracy using the automatic categories has outperformed the manual categories classification accuracy by at least 8%. All this testifies that the approach for automatic categorisation of 3D facial morphology proposed in this study has a considerable potential and gives an indication if these traits categories able to produce genes association when GWAS is performed as will be explained in Chapter 8.

Partial coeffs R^2 p < 0.001 Partial coeffs R^2 p < 0.001 Partial coeffs R^2 p < 0.001

Multiple effect · Upper lip border (3 clusters) · Lower lip-chin shape

Philtrum shape

Upper lip contour (4 clusters)

Lower lip tone shape

Lower lip contour (3 clusters)

Cupid's bow shape · Lower lip contour (4 clusters)

Figure 5.12: Visualisation of the effect of automatic categories on the lower face based on the regression results for dummy variables. The 'Partial coeffs' columns display heat maps of the partial regression coefficients associated with mesh vertices (warmer colours correspond to stronger effects). The 'R^2' columns display heat maps of proportion of the variance. The '$p < 0.001$' columns show two-colour maps of the statistical significance of the effect: yellow for p-value < 0.001 and green for p-value $\geq$ 0.001.

5.4   Summary

In this chapter an automatic approach for lips traits supervised and un-supervised classification using geodesic curvature features are proposed, these approaches can be used for any face morphology traits. The PLSR algorithm was utilised in a new application, for visualising the affect of manual or automatic face traits categories on the 3D face mesh. In the next chapter, an approach for gender analysis and discrimination is presented using the same 3D geometric features.

A 3D MORPHOMETRIC PERSPECTIVE FOR FACIAL GENDER ANALYSIS AND CLASSIFICATION

Gender identification plays a remarkable role in social communication. Humans find this task relatively easy, they are remarkably accurate at determining the gender of subjects from their facial appearance. Even with the hairstyle altered, men's facial hair removed, and no cosmetic cues, humans can still determine subjects' genders from their faces with more than 95% accuracy [2, 36, 41, 46, 50, 183, 206]. However, achieving similar accuracy in automatic gender classification using computers remains a challenge. It is crucial in many applications, for instance making human-computer interaction (HCI) more user friendly, conducting passive surveillance and access control, and collecting valuable statistics, such as the number of women who enter a store on a given day. Researchers have focused on developing various techniques for gender classification since the 1990s, when the first automated system (SexNet) capable of gender recognition using the human face was created [112].

Another essential topic of research, for more than two decades, has been studying *the relationship between facial traits and gender classification* or face recognition. Enlow et al. [93] contend that men have wider and longer noses compared to women and that the male forehead is more bowed and slanting than the female forehead, while Shepherd et al. [306] argue that the female nose is less pointed than the male nose. Establishing which portions of the face and facial morphology features are most

effective for gender classification remains an open research topic due to the strong dependency on the person's ethnicity and age.

Direct Euclidean and geodesic distance measures between 3D facial landmarks have been shown to be popular as local geometric gender classification features [110, 119]. In this Chapter, the new 3D geodesic path features which were described in Chapter 3 were assessed in a gender classification application using the ALSPAC dataset, which was a challenge as gender discrimination in young subjects is much more difficult than in adults [59]. The results were then compared to the gender classification results for the same dataset obtained by Toma [276], where he used the Euclidean distances and angle measures between face landmarks for gender classification purpose. The important contribution of this chapter was determining the most discriminative portions of the face for gender discrimination in teenage Caucasian populations.

The remainder of this chapter is organised as follow: In Section 6.1. the proposed gender analysis approach is described and the computational experiments for gender classification and face morphology gender discriminating ability are presented in Section 6.2. In Section 6.3 discussion of the acquired results and comparing them with other research results are highlighted. Finally, a general summary is described in Section 6.4.

6.1   The proposed gender analysis approach

An overview of the proposed gender analysis approach is provided in Fig.6.1, and the algorithm's different components are explained below.
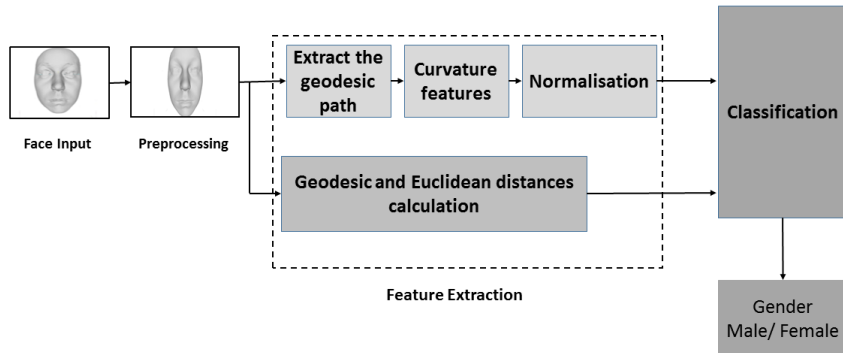
Figure 6.1: Block diagram of the proposed gender analysis system using novel and traditional 3D geometric features

6.1.1   Preprocessing

This work used $4,745$ samples (all British adolescents: 2,512 females and 2,233 males) from the ALSPAC dataset; all of the 3D faces had neutral expressions and were posed as frontal views. For each face, all 21 of the 3D landmarks were used to extract the features descriptors. The non-regularised preprocessing method was used only for preprocessing the data for gender analysis since the previous chapter showed the lips morphology classification accuracies in case the non-regularised preprocessing were higher than regularised preprocessing method accuracies.

6.1.2   Feature extraction

A: Geodesic path extraction

Gabriel Peyre's MATLAB fast-marching toolbox [216] was used to extract the geodesic paths and geodesic distances between landmarks. These paths were selected depending on Toma et al. [276] gender classification results, he concluded that only 24 distances provided gender recognition efficiency of over 70% from the 250 Euclidean distances which they extracted between ALSPAC data facial landmarks, thus in current work the landmarks pairs which their distance recorded highest classification ac-

curacies provided to us from [276] was used to extract the geodesic curvature features between them. Testing more landmarks pairs such as paths between the nose tip and eye corners will be planed for future work.

Figure 6.2 illustrates the paths that were extracted for each region of the 3D face: the forehead/eyes region, the nose, the upper lip, and finally the lower lip/chin region. For each face, eight paths were extracted from the forehead/eyes region, nine paths from the nose region, ten paths from the upper lip region, and six paths from the lower lip/chin region.
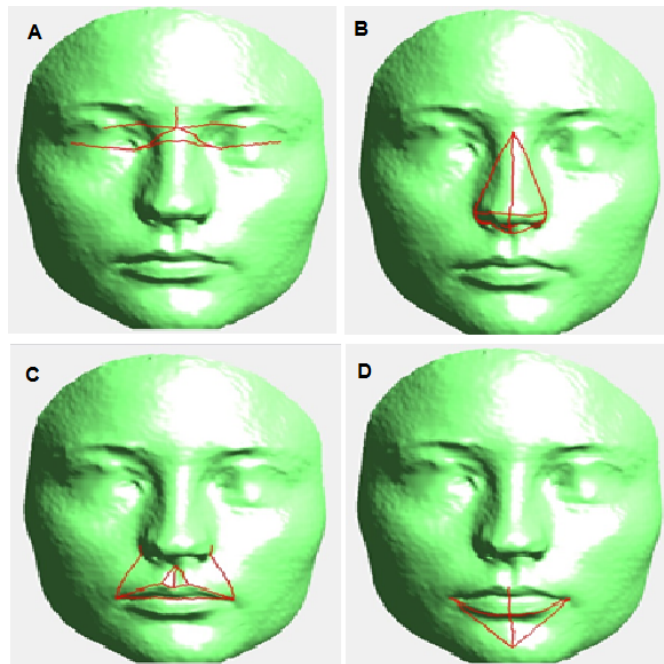
Figure 6.2: Geodesic paths used in the algorithm. The curvature features were extracted for these paths' surface points. Each face trait or region has a different number of geodesic paths. A: forehead/eyes paths; B: nose paths; C: upper lip paths; D: lower lip/chin paths

Curvature features

The principal curvatures were first computed for each path's surface point; the other features (mean curvature, Gaussian curvature, shape index, and curvedness) are then calculated on the basis of those curvature calculations, following the outlined procedure in Sections 4.1.1 and 4.2.1. The ring size was selected experimentally to acquired better classification results, the ring size 2 was provided better classification accuracy.

Normalisation of curvature features

The number of vertices included in a geodesic path varies for different faces. A normalisation procedure was required to ensure that the respective paths of all facial images have the same number of points at which curvature features are measured. To this end, a histogram distribution was calculated for each feature; the number of bins selected was 5, 10, 15, 20, or 25, depending on the maximum and minimum number of points in a path. This procedure provided us with four vectors of equal length: mean curvature, Gaussian curvature, shape index, and curvedness. These vectors were then concatenated to produce a single features descriptor, for each path in a face, then the paths features for each region of face k by concatenating its path descriptors:

$$\mathbf{D}^k_{\text{nose}} = [\mathbf{D}^k_1, \ldots, \mathbf{D}^k_8] \tag{6.1}$$

$$\mathbf{D}^k_{\text{eyes–forehead}} = [\mathbf{D}^k_1, \ldots, \mathbf{D}^k_8] \tag{6.2}$$

$$\mathbf{D}^k_{\text{upper lip}} = [\mathbf{D}^k_1, \ldots, \mathbf{D}^k_{10}] \tag{6.3}$$

$$\mathbf{D}^k_{\text{lower lip–chin}} = [\mathbf{D}^k_1, \ldots, \mathbf{D}^k_6] \tag{6.4}$$

6.1.3 Euclidean and geodesic distances calculation

Past researchers have frequently used geodesic and Euclidean distances as features for 3D facial recognition, 3D facial morphology analysis, and gender identification. In [95], the 3D Euclidean distance was used to measure the deviation of the morphological face traits from the normal face; these distances have been used also to delineate syndromes in [11]. Studies have shown, however, that geodesic distance is more appropriate for gender identification and for measuring levels of facial masculinity/femininity [109, 110].

For that, this work uses the Euclidean and geodesic distances as features for gender classification in order to compare their performance with the proposed geodesic curvature features (i.e. mean curvature, Gaussian curvature, shape index, and curvedness for the geodesic path between landmark). For each face, 33 geodesic distances and 33 Euclidean distances were calculated between the same landmarks that were utilised to extract the geodesic paths shown in Fig. 6.2. The fast marching algorithm and Eq. (4.6) were used to produce the geodesic and Euclidean distances,

respectively.

6.1.4   Classification

The LDA classifier was used to determine gender using a five-fold vali-
dation process, as suggested in [105] for large datasets and our prelim-
inary experiments the LDA classifier provide better classification result
comparing to SVM. Using this process, the $4,745$ samples were first parti-
tioned into five equally sized segments or folds; five iterations of training
and validation were subsequently performed in such a way that within
each iteration, a different fold of the data was held for validation while
the remaining four folds were used for training purposes.

LDA classifier

This work uses linear discriminant analysis (LDA) as a binary classifier to
predict the gender of $4,745$ 3D facial meshes since this classifier is easy
to implement and does not require the adjustment of any specific tuning
parameters. LDA has been successfully used for gender classification in
the past [31, 110].

In essence, LDA attempts to maximise the ratio of between class scat-
ters to within class scatter.

Assume we have m dimensional samples $\{x_1, x_2, ..., x_n\}$, N_1 of which
belong to W_1(first class), and N_2 belong to W_2 (second class). Then, to
compute the linear discriminant projection for these two classes, the fol-
lowing steps should be followed:

Calculate the classes' mean:

$$\mu_1 = \frac{1}{N_1} \sum_{X_i \in W_1} X_i, \quad \mu_2 = \frac{1}{N_2} \sum_{X_i \in W_2} X_i \tag{6.5}$$

Next, calculate the classes' covariance matrix:

$$S_1 = \sum_{X_i \in W_1} \left(X_i - \mu_1\right)\left(X_i - \mu_1\right)^T, \quad S_2 = \sum_{X_i \in W_2} \left(X_i - \mu_2\right)\left(X_i - \mu_2\right)^T \tag{6.6}$$

From that calculation, the within-class scatter matrix $S_W = S_1 + S_2$ and
the between-class scatter will be $S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$. The LDA projec-
tion is then obtained as the solution to the generalised eigenvalue prob-

lem.

$$S_W{}^{-1}S_B = \lambda W \qquad (6.7)$$

$$\left| S_W{}^{-1}S_B - \lambda I \right| = 0 \qquad (6.8)$$

Gender recognition is then established based on the calculation of the Euclidean distance between the query or the tested and extracted 3D facial feature descriptor after projection to the LDA space and the two classes' means, as well as the following projection to the LDA space $\mu_1^* = (W^*)^T \mu_1$ and $\mu_2^* = (W^*)^T \mu_2$ [109, 110, 298].

Verification and validation techniques

Three measures are used to access the performance of the LDA classifier: accuracy, sensitivity, and specificity [314]:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} * 100 \qquad (6.9)$$

$$Sensitivity = \frac{TP}{TP + FN} * 100 \qquad (6.10)$$

$$Specificity = \frac{TN}{FP + TN} * 100 \qquad (6.11)$$

where TP is the number of true positives (i.e., LDA identifies a man who was labeled as such), TN is the number of true negatives (i.e., the classifier recognises a woman who was labeled as such), FP is a false male classification, and FN is a false female identification. Accuracy indicates overall detection performance.

6.2 Experimental results

Five computational experiments were designed for this study in order to determine an optimal set of features for gender classification; an investigation was also conducted on the influence of different portions of the face on gender classification accuracy within the ALSPAC dataset. Experiments 1 through 3 determined which facial features were best suitable for gender classification; it was also investigated which facial portions (eyes, forehead, chins, lips, and nose) were most important for the task. Ex-

periments 1 and 2 used the traditional Euclidean and geodesic distances as classification features, while Experiment 3 utilised our novel feature descriptors. Experiment 4 then examined the effect of combining the Euclidean distance, geodesic distances and geodesic path curvature features on the classification scores. In these experiments, the classification performance was measured using classification accuracies, sensitivity and specificity values. These values are the average results for 5-folds cross validation runs.

Finally, Experiment 5 sought the most discriminatory features for gender recognition in each facial region. The evaluation criterion for all of the experiments was the average gender classification accuracy using the five fold cross validation. In addition, sensitivity and specificity measures were also used. A detailed explanation of these experiments is provided in the following subsections.

6.2.1   Experiment 1: Classification using 3D Euclidean distances

The gender was classified by using 33 3D Euclidean distances extracted from the 21 biologically significant landmarks. For each part of face, the Euclidean distances between its landmarks are concatenated to form the classification descriptor, for example, descriptor contain 8 Euclidean distances were used to classify the gender based on nose region only. The proposed algorithm classified 79.4% of faces correctly as either male or female. Table 6.1 shows the gender identification accuracies facial portions, in addition to the sensitivities and specificities. The Euclidean distances were calculated because of previous work [276] on the ALSPAC dataset used these as features for gender recognition purposes.

Table 6.1: Euclidean distance gender classification results. Accuracy measures the gender classification performance of the 3D Euclidean distance features, based on the different facial portions, while sensitivity and specificity measures the features' accuracy in identifying a male and specificity indicates the features' ability not to identify a false male

Face parts	Accuracy	Sensitivity	Specificity
Eye and forehead	67.3	0.69	0.66
Nose	**69.7**	0.64	0.71
Upper lip	65.6	0.6	0.68
Lower lip and chin	65	0.61	0.63
All parts	**79.4**	0.72	0.77

6.2.2 Experiment 2: Classification using 3D geodesic distances

The second experiment used geodesic distances to predict facial gender scores; many previous studies [43, 119, 125] suggested that geodesic distances may represent 3D models better than 3D Euclidean distances, thus in this work they have been used. Using the fast marching algorithm, 33 geodesic distances were calculated; these distances were extracted from 3D facial region landmarks to form the classification features descriptor, as shown in Fig. 6.2. The acquired gender classification results using these features are shown in Table 6.2.

Table 6.2: Geodesic distance gender classification results. Accuracy measures the gender classification performance of the 3D Geodesic distance features, based on the different facial parts, while sensitivity and specificity measures the features' accuracy in identifying a male and specificity indicates the features' ability not to identify a false male

Face parts	Accuracy	Sensitivity	Specificity
Eye and forehead	73.2	0.76	0.71
Nose	**75.4**	0.65	0.77
Upper lip	72.4	0.61	0.7
Lower lip and chin	69	0.7	0.67
All parts	**82.6**	0.81	0.78

6.2.3   Experiment 3: Classification using 3D geodesic path curvature

The previous two experiments utilised the Euclidean and geodesic distances as features for gender identification, both of which are traditional features for this task. In contrast, the third experiment uses the proposed features descriptor based on the *geodesic path curvatures.* As explained in chapter 4, the basic idea of this features descriptor relies on extracting the surface points of a geodesic path between landmarks Fig. 6.2, then determining the mean curvature, Gaussian curvature, shape index, and curvedness features for those points. The overall gender recognition accuracy was 87.3% this is *much better* than the results from Experiments 1 and 2. Table 6.3 shows the accuracy results using the geodesic path curvature feature descriptor for each 3D facial region, as well as the sensitivity and specificity values.

Table 6.3: Geodesic path curvature gender classification results. Accuracy measures the gender classification performance of the novel Geodesic path curvature, based on the different facial parts, while sensitivity and specificity measures the features' accuracy in identifying a male and specificity indicates the features' ability not to identify a false male

Face parts	Accuracy	Sensitivity	Specificity
Eye and forehead	81.5	0.85	0.76
Nose	**83.4**	0.79	0.88
Upper lip	79.4	0.73	0.81
Lower lip and chin	78.7	0.76	0.72
All parts	**87.3**	0.9	0.85

6.2.4   Experiment 4: Classification using a combination of features

After performing the previous experiments, we were able to rank the features according to their classification accuracy. The best-ranking result was achieved when the geodesic path curvature features were used, whereas the poorest result came about from utilising the Euclidean distance. In the fourth experiment, the robustness of the gender recogni-

tion performance was explored by using the combination from the Euclidean distances, geodesic distances (after scaling their values), and the geodesic path curvature features. The total 3D facial gender recognition rates are shown in Table 6.4. Overall, an improvement in classification accuracies were observed. The best performance, 88.6% accuracy, was achieved when the geodesic distance and the geodesic path curvature features were concatenated.

Table 6.4: Features combination results. The table shows that in general, the gender classification validation scores increased when a combination of 3D geometric features was used

Face parts	Accuracy	Sensitivity	Specificity
Geodesic distance and geodesic path curvature	**88.6**	0.87	0.9
Euclidean distance and geodesic path curvature	87.9	0.88	0.8
Geodesic and Euclidean distances and Geodesic path curvature	88.3	0.87	0.88

6.2.5   Experiment 5: Landmark discrimination ability

In this work, the proposed gender classification approach depends on anthropometric landmarks as the basic points to extract different classification features. Certain landmarks were thus selected for each 3D facial region in order to extract the Euclidean and geodesic distances and the geodesic path curvature features following the recommendations from [276]. The aim of this experiment is to determine for each face part which landmark pairs delineate the best geodesic curvatures based features for discriminating gender. The paths between pairs of landmarks were ranked depending on their individual classification accuracies using the best features combination (geodesic distance and geodesic path curvature) from Experiment 4. In other words, each path curvature features used as classification features descriptor in LDA classifier, the paths ranked according to their classification accuracies.

In spite of using feature selection methods is a significant part of many machine learning applications dealing with small-sample and high-dimensional data, but still raise challenges about the interpretability and stability of feature selection techniques. Additionally, different features selection method may provide different results [55, 85, 120]. Consequently, in this work the classification accuracy used to rank the landmarks paths and analysis the most discriminative landmark paths ac-

cording to their classification accuracies. The presented work follows [276] in using the classification accuracy to capture those landmarks paths relevant to gender variations.

Table 6.5 illustrates the landmarks' paths ranks for each 3D face region, while Fig. 6.3 shows the highest three ranks of landmark paths for each region.

Table 6.5: Landmark path rankings. This table illustrates the separated landmarks' gender classification abilities when their path descriptor features were applied using the LDA classifier; 1 represents the highest rank, while 6 represents the lowest rank

Eyes and forehead		Nose		Upper lip		Lower lip and chin	
Landmark	Rank	Landmark	Rank	Landmark	Rank	Landmark	Rank
g-n	2	n-prn	3	alL-chL	5	li-ls	2
exL-enL	5	Prn-sn	6	ahR-alR	5	**li-pg**	**1**
enL-enR	**1**	alL-n	4	chL-cphL	3	chL-li	4
enR-exR	6	alR-n	4	cphL-ls	2	chR-li	4
PsL-n	3	**alL-prn**	**1**	cphR-ls	2	chL-pg	3
psR-n	3	**alR-prn**	**1**	cphR-chR	3	chR-pg	3
enL-n	4	alL-sn	5	cphL-sn	6		
enR-n	4	alR-sn	5	cphR-sn	6		
		alL-alR	2	**Sn-ls**	**1**		
				chL-chR	4		
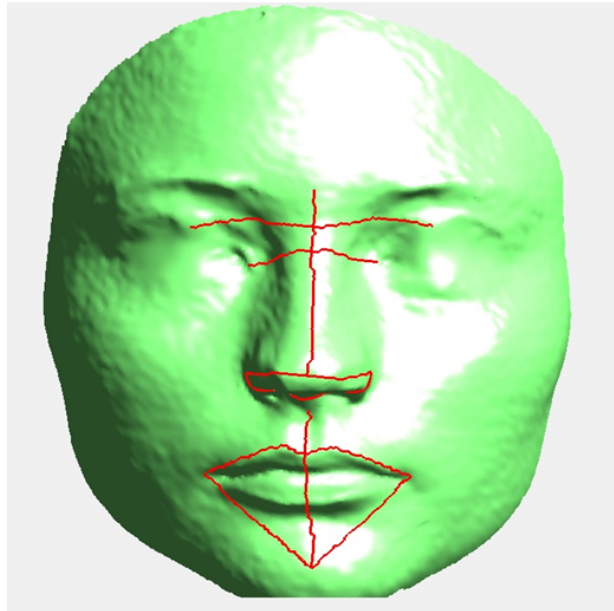
Figure 6.3: The highest-ranked landmarks' geodesic paths

6.3   Discussion

The first three computational experiments in this study aimed to determine which facial features are the most effective in gender classification. Using *only* the 3D Euclidean distance (Experiment 1), we found the gender classification accuracy to be 79.4%, which is well below human perceptual accuracy but close to the results 80% by Toma et al. [276]. Experiment 2 demonstrated that the geodesic distance between facial landmarks provides a better gender recognition score; this is intuitively clear since geodesic distance is a better measure of face shape than Euclidean distance. However, the classification accuracy 82.6% determined using this measure was still below the human accuracy threshold.

The feature descriptors (geodesic path curvatures) were subsequently evaluated in Experiment 3 and produced a classification accuracy of 87.3%. The proposed geometric descriptor is an amalgamation of the mean curvature, Gaussian curvature, curvedness, and shape index at the vertices of the path and thus represents a richer description of the surface than simple Euclidean or geodesic distance measures. Hence, the obvious improvement in the classification accuracy.

As shown in the previous studies [110, 119], a richer combination of

facial features can further improve classification accuracies. We explored the various combinations of Euclidean or geodesic distance with our new geodesic path features. We achieved a further improvement in the gender classification accuracy (88.6%) using a combination of the geodesic distance between landmarks and our geodesic path features.

This result compares favorably with other methods for gender classification. To the best of our knowledge, this is the best published result based on an anthropometric landmark approach, which was achieved for a *credible* large sample of 4,745 facial meshes. Several studies [50, 98, 110] used geodesic and Euclidean distances (or their combinations) as geometric gender classification features for both 2D or 3D facial images. The reported classification accuracies were generally higher than ours, but these were achieved for much smaller samples. Burton et al. [50] reported 96% accuracy for a sample of 179 faces, while Fellous et al. [98] obtained 90% accuracy for 109 facial images. Gilani et al. [110] achieved 98.5% accuracy for 64 3D facial scans.

Other studies have only utilised global facial features for gender classification. For example, Wu et al. [298] used raw shape from shading depth features to achieve a gender classification accuracy of 69.9% with the FRGCv1 data set comprising 200 subject faces. Lu et al. [180] obtained a gender classification rate of 85.4% using the vertices of a generic facial mesh fitted to the raw 3D data as a classification feature descriptor for the same FRGCv1 data set. Ballihi et al. [24] achieved a classification accuracy of 86% using a combination of radial and circular curves as classification features and specified the curves at the nose, forehead and cheeks regions, as a compact signature of a 3D face for face recognition and gender selection. However, it should be noted that the authors used a small sample of 466 subject faces. It should also be noted none of the above global methods is suitable for the investigation of specific relationships between individual facial regions and gender classification accuracy, which was the aim of this work. The present study operated on a *large* population cohort of 4,745 fifteen-year-old Caucasian adolescents; so the gender recognition efficiency identified in this study is likely to be more robust than that of the other studies based on small samples.

Physiological and psychological research [2, 50, 208] supports the idea that facial and gender recognition in the human brain is based more

on individual regions than on the whole face. For example, Edelman et al. [88] compared human performance against a computer model in the classification of gender in 160 adult individuals (80 males, 80 females) using frontal facial images. The study revealed that humans were better than the computer model at discriminating females based on the upper face, whereas for males the human accuracy was better for the lower face. It was also highlighted that males have thicker eyebrows and larger noses and mouths than females. Several forensic and anthropometric studies also showed that the female face, mouth, and nose are smaller than those of males [95].

Based on this information, the first three experiments conducted in the present study concentrated on using individual facial parts to determine the gender recognition capability. Fig.6.4 shows an annotated view comparing the classification performance among the 3D facial parts for each feature type (Euclidean and geodesic distances and geodesic path curvatures).
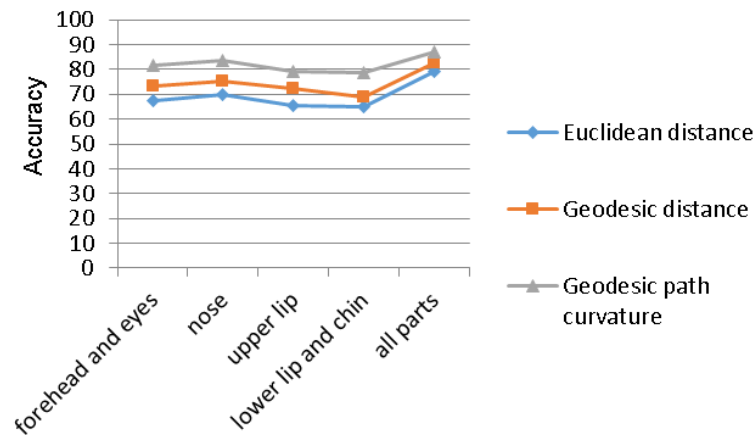
Figure 6.4: Classification performance using different types of 3D geometric features

As can be seen from Fig. 6.4, the nose is the most important facial part for gender discrimination in the ALSPAC dataset. In addition, the sensitivity and specificity results shown in Tables 6.2, 6.3 and 6.4 identify the nasal morphological areas that are most effective in discriminating the gender of young Caucasian people. This finding is in agreement with the medical studies [1,8], which addressed changes in nasal shapes and sizes

in groups of 11 to 17 year old subjects in relation to gender discrimination. These studies have found that nasal height and nasal bridge length become fully mature at 15 years of age in males and 12 years of age in females.

After establishing a set of strong gender differentiating 3D geometric features, the discrimination capabilities of pairs of landmarks and their curvature features along the geodesic shortest path between them that have the best influence on classification are evaluated. This performed by finding prime determinants of classification accuracy using the LDA classification method. Such landmarks can then form the basis of a more efficient focused selection of specific manual landmarks or even assist in developing a suitable directed automated landmark detection approach. The results indicate that the landmarks that describe 3D facial profile curves are important in gender classification, as shown in Fig. 6.3. These findings validate other studies that have relied solely on 3D profile curves. For example, Lei et al. [172] extracted the central vertical profile and the nasal tip transverse profile, and located the face feature points by analysing the curvature of profiles to obtain ten 3D geometric face features with an accuracy of 98.8% using the ZJU-3DFED dataset and an accuracy of 100% with the 3DFACE-XMU dataset. Also, Ter et al. [274] performed 3D face matching and evaluation using profile and contour of facial surface to achieve a mean average precision (MAP) of 0.70 and 92.5% recognition rate (RR) on the 3D face Shape Retrieval Contest data set (SHREC'08) and a MAP of 0.96 and 97.6% RR on the University of Notre Dame (UND) data set.

Moreover, the path between the internal eyes landmarks (enR-enL), the Ala shape path (alL-prn-alR), and the Cupid bow path (cphL-ls-cphR) yield the best characteristic paths for gender classification. These results have been corroborated by the results from previous study that were conducted on the same ALSPAC data set Toma [276], who worked on whole 3D faces (PCA analysis of a small set of Anthropometric landmarks and Euclidean distance measures). This study identified approximately the same face regions but with less accurate classification. Following the procedure described in Chapter 5 for determining the categorisation labels statistical effect using PLSR, Fig. 6.5 displays color maps of facial regression on gender labels, as can be seen, the effect of sex is mainly located in

the nose area, upper lips, chin, and eyes area and this is compatible with Fig. 6.4 and Fig. 6.3 results.
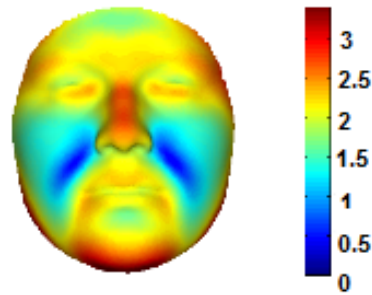
(a) partial coeffs
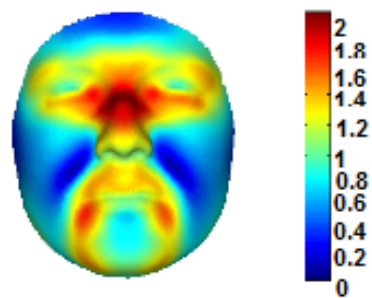
(b) (R^2)

(c) $p < 0.001$

Figure 6.5: Face regressed on sex. Shows the effect magnitude (effect), Partial R^2 indicates the amount of variability in the location of each vertex. $P < 0.001$ indicates the areas where partial R^2 was significant

Finally, the sensitivity and specificity showed little difference between the above results for facial parts' gender classification using the geodesic

and Euclidean distances and the geodesic path curvature features with the *exception* of the nose trait. In general, the first four experiments yielded good specificity and sensitivity results, particularly Experiment 4, in which the geodesic distance and geodesic path curvature features were integrated. In that experiment, the sensitivity value was 0.87, while the specificity value was 0.9.

6.4   Summary

Numerous studies have investigated gender classification approaches at 2D or 3D faces but a little of them highlighted the relationship between face morphology and gender discrimination at certain age group or ethnic. In this chapter, a new approach for gender analysis and discrimination is implemented using the proposed geodesic curvature features. The performance of these features in gender classification is compared with the Euclidean and geodesic distance performance. The gender classification accuracies using the geodesic curvature features are higher than these measurements accuracies. These features will be used to categorise the nose morphological traits in the next chapter.

UNSUPERVISED NOSE MORPHOLOGICAL TRAITS CATEGORISATION

Due to the distinct shape and symmetrical property of the nose, it is often used as a key feature point in the 3D facial identification, analysis, automatic landmarking, registration, etc. Many automatic 3D face landmarking techniques rely on the fact that the nose shape is usually invariant under variations in the pose and expression. Nose detection is the first stage in these methods then other landmarks such as eyes and mouth corners are specified according to the nose landmark position [258, 312]. The nose is used as a reference point to compute the geodesic distance from it to other points in the 3D face surface [6] for 3D face identification, or it was used as reference point to extract the radial and iso-level geodesic paths for gender classification and face recognition [25]. Recently, Emambakhsh et al. [91] addressed the problem of expression invariant face recognition by introducing an algorithm utilised the 3D shape of the nose as discriminative feature descriptors.

From the medical aspect, the nose morphological traits play an essential role in diagnosing many anomalies such as Noonan and 22q11 syndromes [122]. Their most striking face morphology characteristics are large nasal height, with relatively narrow nares and nasal base but with fullness above the tip of the nose [122]. In addition, the patients with asthmatic related problems have different nose dimension and traits specifications from non-asthmatics [9]. It is easy to conclude the importance of nose morphology analysis in different applications, especially in medical fields. Several attempts (Section 2.6. Literature Review) have been

made to classify noses. In such research, the nose traits categorisation was mainly performed manually in clinical practice and on small subject groups. Consequently, the aim of this chapter is to work on developing a method for automatic nose morphological traits categorisation *without ground truth manual categories*. In other words, the nose traits variation is automatically discovered. Five nose traits (nose shape, ridge shape, ala shape, tip shape and nasal tip shape) Fig. 7.1 are defined by a medical expert, we wish to categorise them automatically without previous knowledge about their categories.
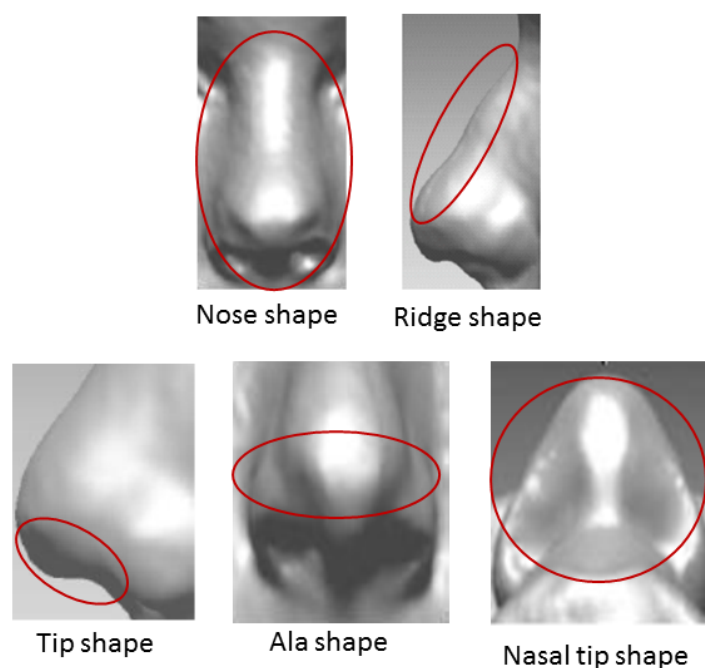
Figure 7.1: Nose traits

The rest of this chapter is divided into the following sections: Section 7.1. describes the proposed nose clustering approach, how the nose traits geodesic curvature features are extracted, normalised and categorised. The clustering operation, the number of cluster validation and categories affect visualisation are presented in Section 7.2. Section 7.3 contains a general discussion of the computational experiments that are conducted. Finally, this chapter summary is made in Section 7.4.

7.1   The proposed nose categorisation approach

The nose categorisation approach uses different components which are explained below, these are essentially the same components that have been used for lips morphological traits categorisation in Chapter 5. Figure 7.2 shows a block diagram of the proposed approach. Each block and its input parameters are described below:
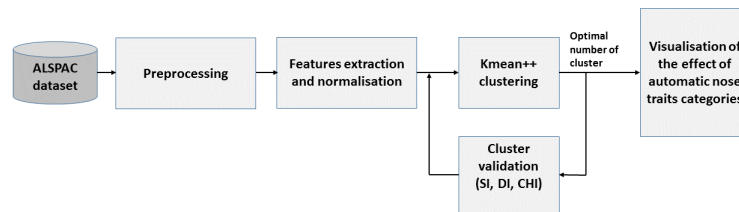
Figure 7.2: Block diagram of the proposed automatic categorisation approach for the nose traits

7.1.1   Preprocessing

This work used ALSPAC dataset which consists of 4,747 samples; all of the 3D faces had neutral expressions and were posed as frontal views. For each face, six anthropometric landmarks (g, n, prn, sn, alL and alR) were localised on the nose area are used to extract the feature descriptors Fig. 7.3. In this work, the non-regularised preprocessing method was used.
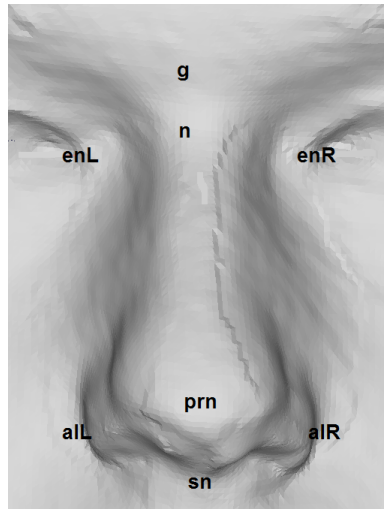
Figure 7.3: Nose landmarks

7.1.2   Feature extraction and normalisation

The categorisation approach was implemented using a combination of geodesic curvature features and geodesic distance.

Geodesic paths extraction

Gabriel Peyre's MATLAB fast-marching toolbox [216] was used to extract the geodesic paths and distances between the nose landmarks. Each nose's traits will be clustered according to curvature features and geodesic distances of certain geodesic paths. Figure 7.4 and Table 7.1 illustrates the paths used for each nose traits. The nose shape, the ala shape, the ridge shape, the tip shape and the nasal shape are the nose morphological traits that we are looking to categorise them automatically in this work.
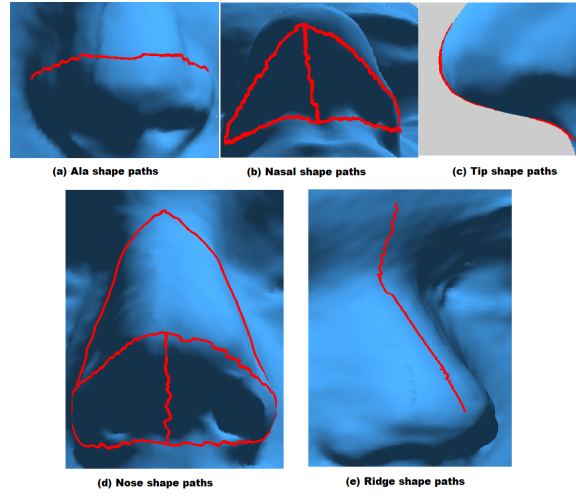
Figure 7.4: Nose traits geodesic paths. For each nose trait there are certain geodesic paths (colored with red), their curvature features are utilised for clustering purpose

Table 7.1: List of geodesic paths defining morphological nose traits

Trait name	List of related geodesic paths
Nose shape	n-alL, n-alR, sn-prn, alL-prn, alR-prn, alL-sn, alR-sn
Tip shape	prn-sn
Ridge shape	g-n, n-prn
Ala shape	alL-prn, alR-prn
Nasal tip shape	sn-prn, alL-prn, alR-prn, alL-sn, alR-sn

Curvature features

The principal curvatures were first computed for each nose path's surface point; the other features (mean curvature, Gaussian curvature, shape index, and curvedness) are then calculated on the basis of those curvature calculations. In this work, a ring size of 2 is used to calculate curvatures since this value recorded best classification accuracy in Chapter 5.

Normalisation of curvature features

Each nose's geodesic path has a different number of surface points (vertices). To deal with this, the same procedure that used for normalising the lips traits geodesic curvature features (Section 5.1.3) is followed, the normalised histogram distribution for each path features is calculated;

the selected number of bins was also 5, 10, 15, 20, and 25. Since a different number of geodesic paths are extracted for nose traits, the feature descriptor for each nose morphology trait can be calculated by concatenating its paths' descriptors.

7.1.3   Categorisation, cluster validation, and categories effect visualisation

The nose traits morphology is automatically categorised by using the Kmeans++ [19, 152] clustering technique as in Chapter 5. Determining the number of clusters relies on determining the best cluster validity indexes. In this work, three cluster validation methods are used to specify the number of classes for each nose traits: Silhouette index (SI), Dunn index and Calinski-Harabasz index (CH). These indices were explained in Chapter 5. The nose traits categorisation approach was performed without ground truth about traits categories. Therefore, analysing the traits clusters structure by using the PC plot [297] and Silhouette plot [243] was helpful in selecting the optimal number of clusters for nose traits.

Finally, Partial Least Squares regression [224], was used for investigating the significant influence of these traits categories on the nose area.

7.2   Experimentation and results

Three experiments were designed for this study in order to automatically categorise nose traits (nose shape, ridge shape, ala shape, tip shape and nasal tip shape) and to investigate the optimal number of clusters for each nose trait. A visualisation method was also conducted to visualise the influence of automatic nose morphological traits categories on 3D face nose area.

7.2.1   Experiment 1: Clustering and cluster validation

The Kmeans++ was used for the unsupervised classification scheme. The Kmeans++ clustering procedure was run 100 times. From 100 runs the clustering result that produced the minimal sum of squared distances score is chosen.

An analysis of the best number of clusters is accomplished by us-

ing the internal cluster validation techniques. The validity indexes were tested with Kmeans++ clustering results, a number of clusters are determined to correspond to the maximum value of (CH, Dunn, SI) indexes. The performance of validity indexes CH, Dunn and SI are compared for clustering parameter (C) ranging between 2 and 9. The tests carried out with 50 repetitions to choose a stable number of clusters for nose morphological traits.

Figure 7.5 is an example of using validation indexes to specify the number of clusters (C) for the nose shape trait. The selection the number of cluster to this trait was challenging because each internal validation method indicates different number of clustering (DI indicates 5, SI indicates 7 and CHI indicates 6). Therefore, Silhouette analysis is also used to specify the optimal number of clusters by studying the separation distance between the resulting clusters. The silhouette plot provides a way to assess a number of clusters visually, it displays how close each point in one cluster is to points in the neighboring clusters [243]. According to Silhouette analysis, the nose shape trait has seven clusters.

In the same manner, the correct number of clusters for other nose morphological traits are obtained: Ala shape 3 clusters, Nasal Tip shape 5 clusters, Ridge shape 3 clusters, and Tip shape 5 clusters.
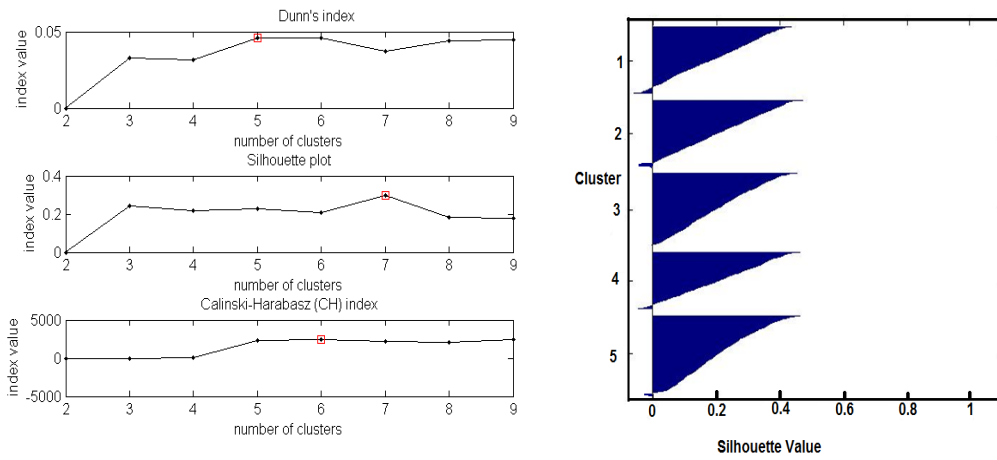
Dunn, SI and CH

Silhouette Plot 5 cluster
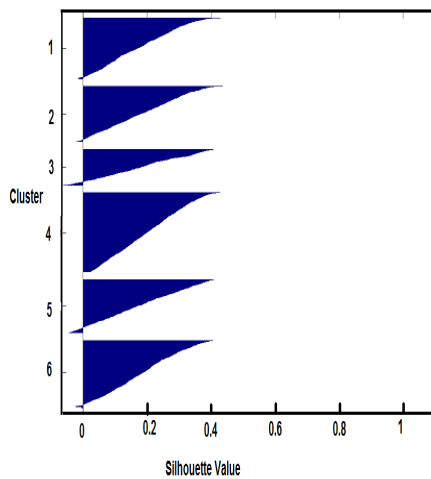
Silhouette Plot 6 cluster
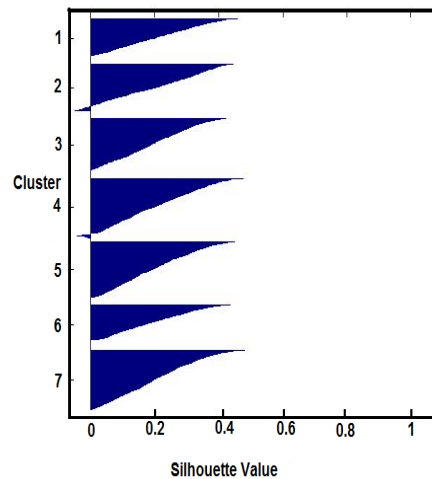
Silhouette Plot 7 cluster

Figure 7.5: Validation indexes for nose shape trait

7.2.2 Experiment 2: PCA and cluster analysis

The aim of this experiment is to study the effectiveness of principal components (PCs) in capturing cluster structure. Principal component analysis (PCA) deals with the relatively high number of variables by using an orthogonal linear transformation to convert the original set of correlated data into a set of variables that are linearly uncorrelated, represented by the principal components (PC) in a new coordinate system. The first coordinate axis runs along the direction of the greatest variance in the original set, the second greatest variance along the second coordinate axis

and so on [297]. It is often of interest to visually inspect how well the data points are separated in 2D or 3D space based on principal components. It is a method to assess similarities and differences between samples and determine whether samples can be grouped [303]. Fig. 7.6 plots the nose morphological traits geodesic curvatures feature clusters in the space of the first two PCs, which contain from 75% to 83% of the variation in the data. Each of the nose traits clusters is represented by the different color or different shape.

Nose Shape

Ala Shape

Nasal Tip Shape

Ridge Shape

Tip Shape
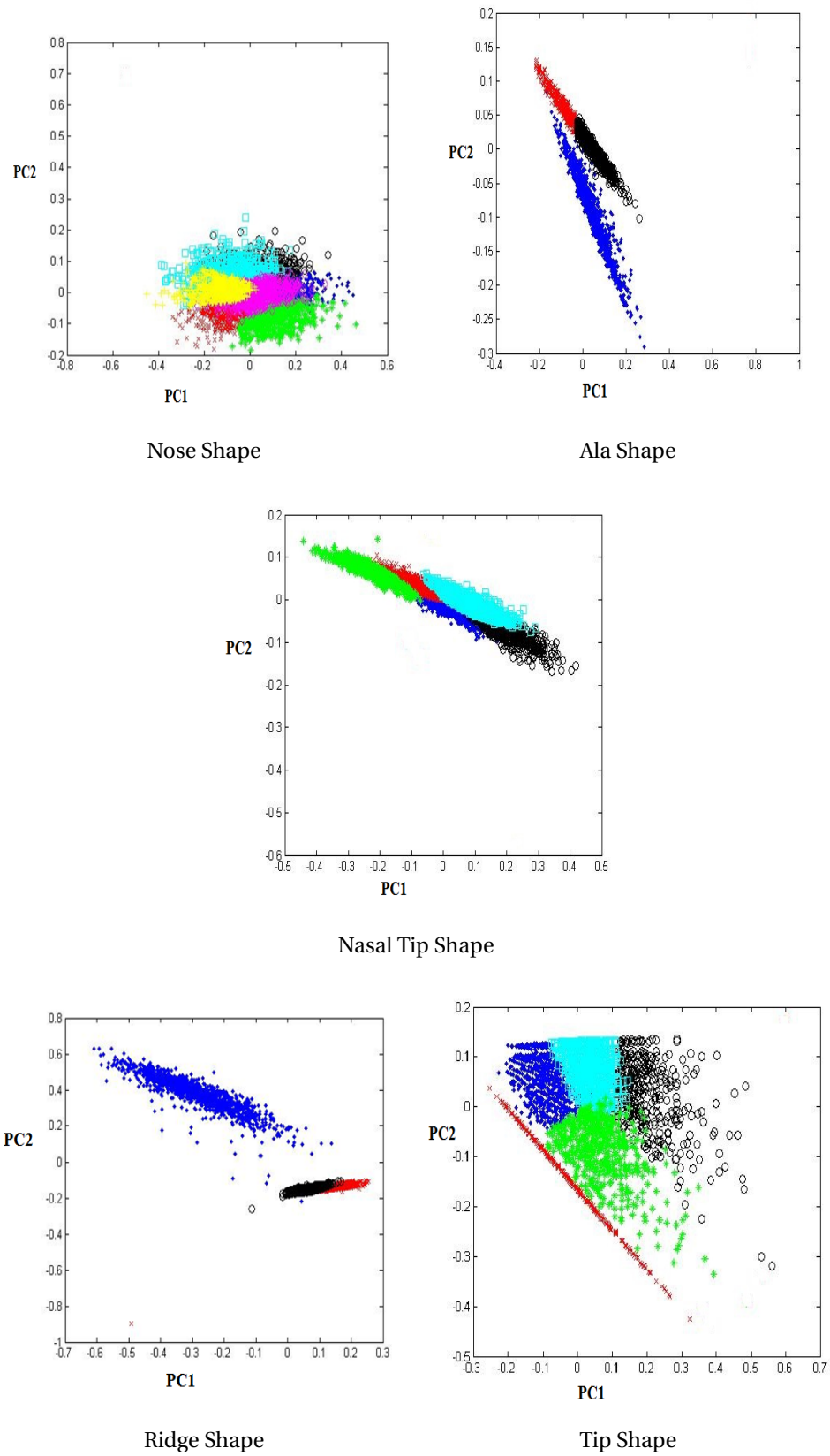
Figure 7.6: Visualisation of nose traits categories in the subspace of the first two PCs

7.2.3 Experiment 3: Visualisation the effect of traits categories on the nose area

Shape regression is an important technique in geometric morphometrics for investigating the effects of various independent variables on morphology dense data (dependent variables). In this experiment, PLSR was used to find the association between the automatic nose traits categories dummy variables and the nose area.

In Figure 7.7 the regression results for each trait dummy variable are visualised, in addition to the partial effect for each trait dummy variable, the multiple effect size for all independent variables combined is also visualised. The regression coefficients (partial coefficients) is the change in 3D vertices in relation to the change in the independent variable at that particular point, and the effect size (R^2) shows the proportion of the variance explained by the effect. The effect per vertex was visualised for significant p-value < 0.001, multiple and partial statistical significance tests were based on 1000 permutations.

The Nose Traits

partial
Coeffs

R^2

p < 0.001

Multi effect
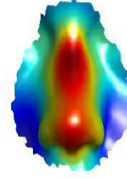
Nose shape
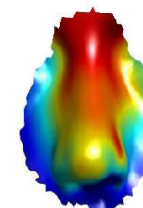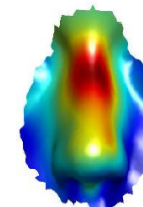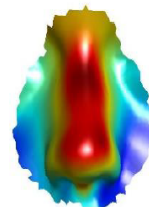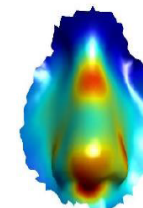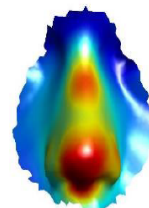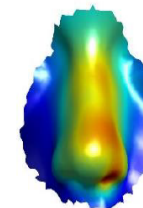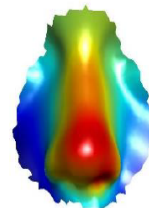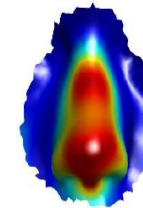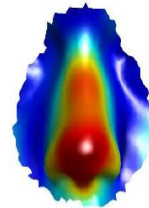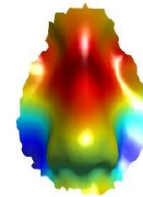
Ala shape
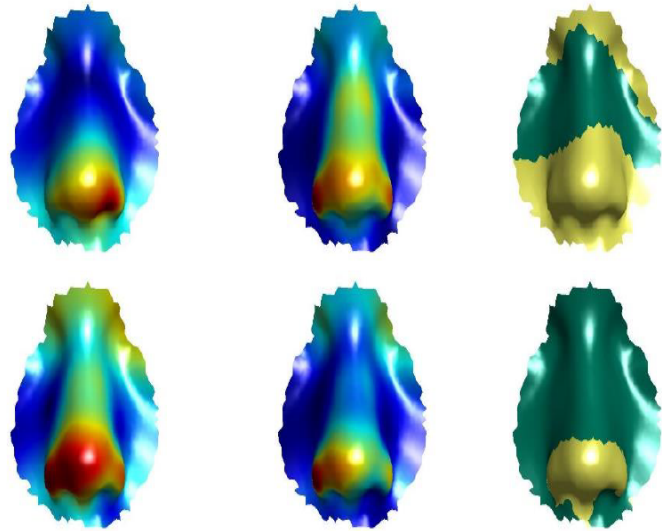
Nasal Tip shape
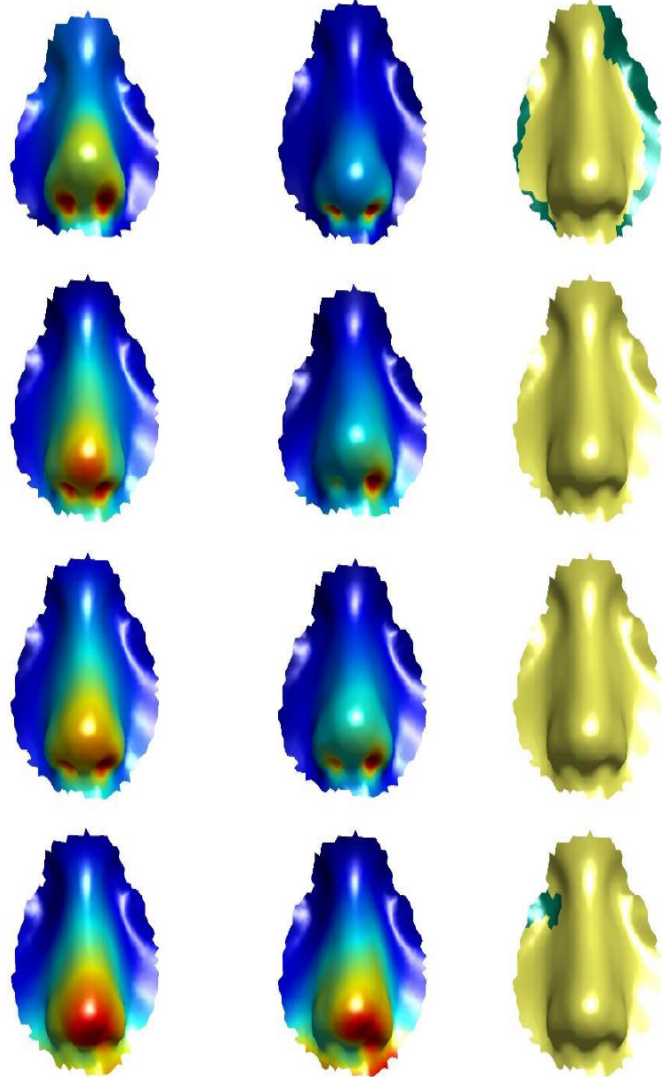
Ridge shape
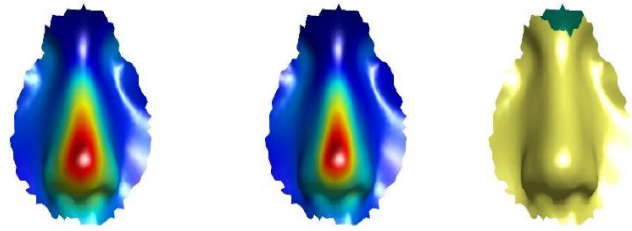
Tip Shape
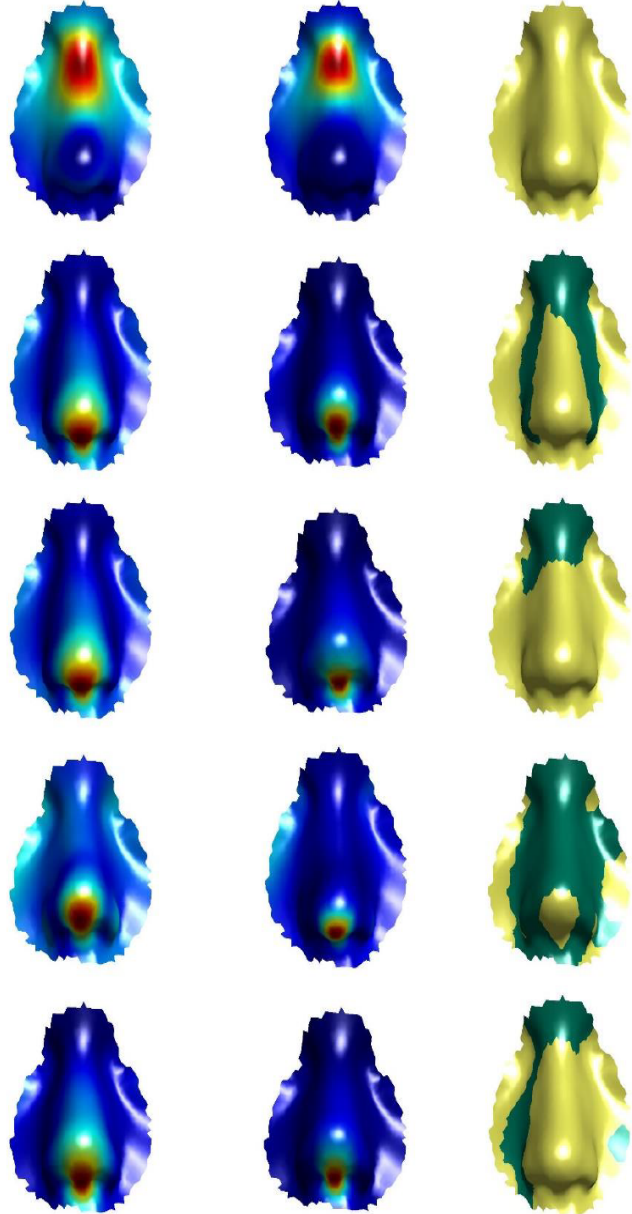
Figure 7.7: Automatic nose traits categories regression results

7.3   Discussion

Linear measurements taken on the nose give an indication of height, width, and thickness but do not provide sufficient information of the surface topography or shape of the various nasal structures [21]. Therefore, the automatic categorisation approach proposed in this chapter was based on external morphological characteristics of the nose and the surface shape of the nose. Different racial groups often have different characteristic nasal features [20]. Therefore, nose traits categories found in this work are restricted to a 15 year old Caucasian population. These categories may be different in other ethnic groups such as Middle Eastern and Asian populations.

Cluster validity checking is aimed to evaluate the clustering results and the selection of the scheme that best fits the underlying data. The majority of algorithms are based on specific criteria for defining the clusters in which a dataset can be partitioned. Figure 7.5 illustrates the validation indices for nose shape clustering. Dunn index selects five clusters and CH selects six clusters while Silhouettes selects seven clusters as the best partitioning. Figures 7.5b, 7.5c and 7.5d shows the Silhouettes graphical plots for the three number of clusters. Silhouette coefficients near +1 indicate that the sample is far away from the neighboring clusters. A value of 0 indicates that the sample is on or very close to the decision boundary between two neighboring clusters and negative values indicate that those samples might have been assigned to the wrong cluster. Consequently, the Silhouette plot is a helpful method for selecting the best number of clusters in addition to the internal validation method.

In this work, the automatic categorisation approach is applied to investigate nose traits variation without any ground truth about its categorisation. Therefore, using at least one more validation method to assess the effectiveness of clusters structure was essential. Principal components analysis (PCA) is a multivariate data analysis tool that can be used to recombine the variables of a large multivariate dataset in such a way that the first few variables of the reconstructed dataset account for the majority of the variance in the data. The hope for using PCA prior to cluster analysis is that PC's help in extracting the cluster structure in the dataset [226].

Figure 7.6 shows the PC plot of nose morphological traits clusters, some of these clusters are well separated and the clusters centroids are far from each other as in ala shape and tip shape clustering plots, while in nose shape clustering plot the clusters are overlapped. In unsupervised classification, it is rare to get an output that is one hundred percent accurate because real world data is rarely that simple and the density is not linearly separable [140].

Regardless of that, as can be seen in Fig. 7.7 automatic categorisation method seems work properly since the magnitude of the traits categories effect and the proportion of their variance at 3D face vertice R^2 appear significant (red) on the nose traits related to them. For example, the effect of the tip shape categories is mainly located in nose tip curve. The multiple effect regression results demonstrated that all traits categories effect is concentrated and significant in the nose area.

Finally, the nose morphological traits are assessed by two medical expert examiners. Figure 7.8 illustrates the examiner's rudimentary classification scale for the characterisation of nose traits. Their results indicate that the proposed method for automatic nose traits categorisation is acceptable from the medical expert view and these automatic categories can be used in much medical application such as medical problem diagnosing for example asthma condition and finding the gene's association.

Triangular   Funnel   Pug   Parallel bridge   Flat   Bulbous   Rhomboid

Nose shape images

Little Ala   Rhomboid   Clear Ala

Ala shape images

Round up   Flat   Round pointed up   Drop Tip   Slight Round up

Tip shape images

Bump   Convex   Straight

Ridge shape images

Normal rounded   Wide columella   Narrow columella   Flat   Pointed

Nasal tip shape images
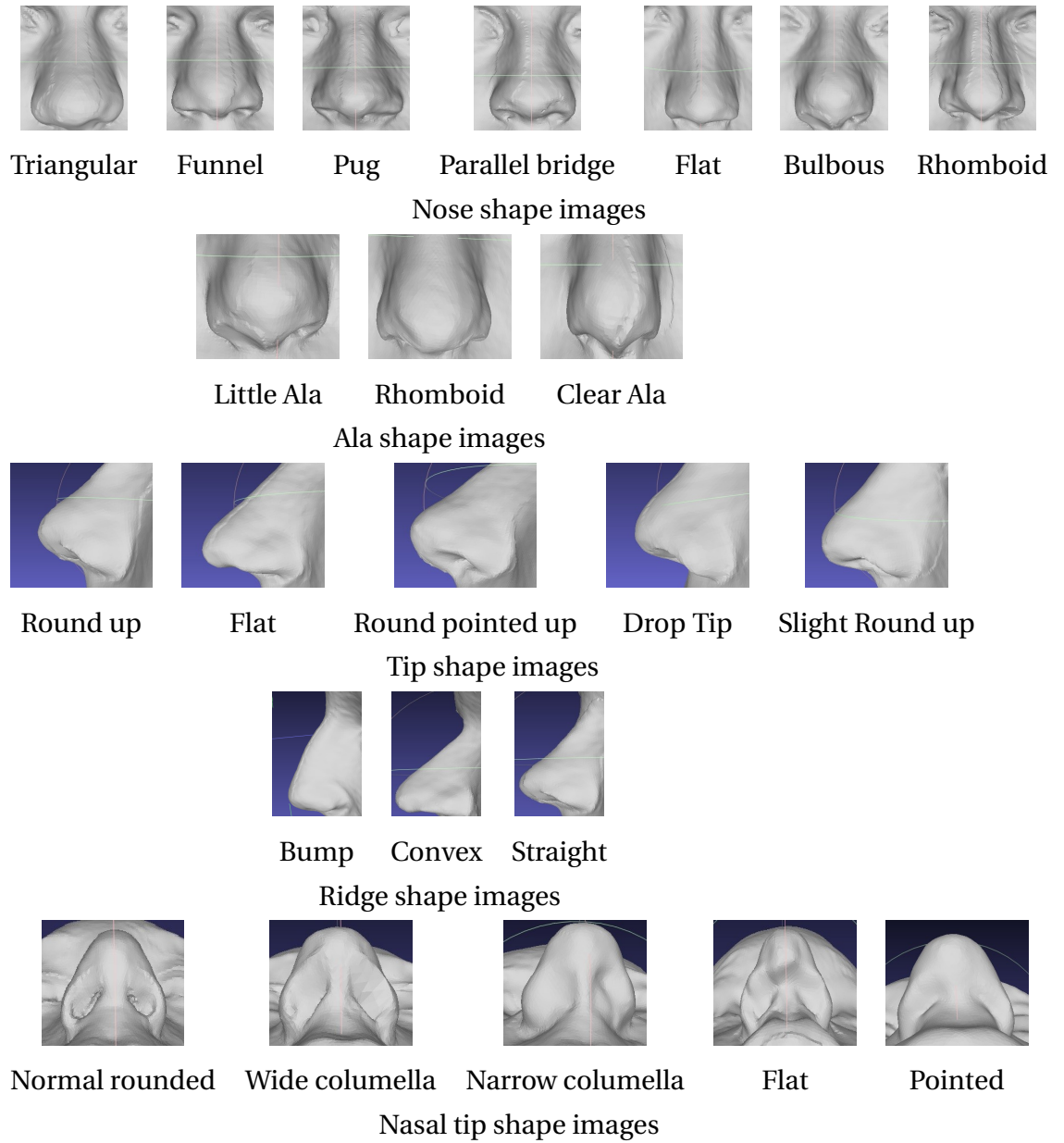
Figure 7.8: Nose traits variation assessment

7.4   Summary

The proposed approach in Chapter 5, which was used for lips traits categorisation, was used here for categorising the nose traits automatically. Additional validation methods were used because of the categorisation work in this case, posed more challenges, due to the absence of any ground truth traits categories. The results of this chapter indicate the ability of this approach in categorisation any face traits without previous

knowledge about this trait categories. The traits categories results of this chapter and Chapter 5 are used in the next chapter for finding genes association by applying GWAS.

Chapter 8

EXPLORING THE ASSOCIATION BETWEEN FACIAL MORPHOLOGICAL TRAITS AND GENES

Facial traits display gene-environment interactions, which indicates that facial characteristics are due to a mixture of environmental factors and genetic factors. Genetic effects are obviously dominant during embryonic facial morphogenesis, while the environment may also influence facial morphology during facial growths. Humans share most of their genomes. Consequently, only a comparatively small number of genetic differences have resulted in noticeable variation, that can be seen among individuals of our species, this variation such as differences in height, weight, face shape, and skin color. The complicated interaction between genes and the environment, as well as between multiple genes, makes it hard to understand and identify human traits variations. Therefore, instead of looking at complex human traits, several researchers have gone straight to the source and looked for nucleotide sequences, small variations in the individual nucleotides of the genomes (SNPs) that may cause the change in an individual's traits [37,176]. Despite intensive research on genetics of the facial morphology using human models, the genetic variation that upholds normal human facial morphological traits is still mostly elusive. The previous chapters demonstrated that it is possible to assess and classify facial morphology traits automatically, specifically lips and nose traits of normal human faces. This chapter explores the prospect of assigning genetic associations (nucleotide sequences) to the variations of

lips and nose morphological traits of normal faces. This chapter results will add interesting findings to normal face morphological traits variation and genes association research field, because it is the first GWAS conducted on face morphological traits which were categorised automatically. Another research contribution in this chapter is a simple method to visualise morphological traits variation influence on the 3D face.

This chapter is organised as follows. The genetic data, the facial traits and the GWAS are explained in Section 8.1. The results are demonstrated in Section 8.2. Section 8.3 presents method to analyse face morphological traits using average face. General discussion and comparison with other studies are highlighted in Section 8.4. Finally, the chapter summary is recorded in Section 8.5.

8.1   Methodology

The objective of this chapter was to perform a genome wide association study (GWAS) for lips and nose traits, which were produced automatically in Chapters 5 and 7. These traits are philtrum shape, Cupid's bow shape, upper and lower lip contour shape, lower lip chin shape, lower lip tone shape, nose shape, ridge shape, nasial shape, Ala shape and tip shape Fig. 5.2 and Fig. 7.1. A sample of individuals from the ALSPAC dataset were used for GWAS. Genotype was available for 8,365 individuals and 3D face scans for 4,747 individuals. A total of 3,687 had both genotype and 3D face scan available for testing.

8.1.1   Genome-wide Association Study (GWAS)

Genome wide association studies have demonstrated successes in assigning genetic association to facial morphology population variations (Table 2.2. Literature Review). Genome wide association studies involve two parts, the first one is the discovery part, by which the genetic associations are assessed for a population, and the second part is the replication, where results from the first part are tested. Ideally, the replication subjects should be from the same population group, as the discovering population in order to confirm the same effect. Also, the number of subjects should be larger than the discovering samples, so as to confirm any false positive results [51].

Many genes have been reported as master genes for head and face development [155]. Research attempts have been made to determine how common these genes' SNPs may affect normal facial morphological traits and specifically on the lips and nose traits [128, 176, 195]. Example genes and SNPs are:

- The (IRF6) gene on chromosome 1 and (ABCA4) on another part of chromosome 1 have been associated with cleft lip and cleft palate [28].

- The (PAX) genes have been shown to influence facial development, and PAX proteins are important in developing the specification of tissues. The (PAX) gene family consists of nine members [52]. PAX3 and PAX9 have been shown to influence facial development. PAX3 has been identified with ear, eye, facial development and facial deafness hand syndrome [107, 228]. PAX9 has been associated with a number of organs and other skeletal developments, particularly teeth. It has been also linked to congenital absence of all teeth syndrome and non-syndromic cleft lip with or without cleft palate [169, 288].

In this work, 64 SNPs are selected for performing the GWAS, to replicate the finding of many face traits and genes association studies (Section 2.8). These genes SNPs have been selected under the guidance of my Ph.D. co-supervisor Prof. Stephen Richmond (medical expert) from the Dental School, Cardiff University.

8.2   GWAS results

Genome associations were performed by a Ph.D. student Laurence Howe at Bristol University, under the supervision of Dr. Sarah Lewis. Bonferroni correction is a method for multiple testing analysis and finds statistical significance of an SNP association p-value. The p-value, which is the probability of seeing a test statistic equal to or greater than the observed test statistic if the null hypothesis is true, is generated for each statistical test. Statistical tests are generally called significant and the null hypothesis is rejected if the p-value falls below a predefined alpha value, which

is nearly always set to 0.05 [51]. In this GWAS there were around 64 independent SNPs, the statistical significance of an SNP association will be 0.05/64=0.0008.

Table 8.1 records the association results, the SNPs, and genes which their p-values less than or near the hit value 0.0008. While Table 8.2 reports the medical effect of these SNPs and genes on facial morphological traits. Some of these SNPs have proven relationship with facial deformities syndromes. Waardenberg syndrome (Fig. 8.1) has an effect on the face by, pale or brilliantly blue eyes, eyes of two different colors, or eyes with one iris having two different colors. Also the appearance of wide set eyes due to a prominent, and broad nasal root [106]. On the other hand, the Pierre Robin syndrome Fig. 8.2 causes three main abnormalities characteristics in humans faces: cleft palate, abnormal positioning of the jaw or mandible and airway obstruction caused by backward displacement of the tongue base [141].

Many medical studies have investigated in depth the relationship between these syndromes and face morphology variation and growth, especially in infant and children by using the anthropometric measurements (e.g., see [76, 165, 250]). For that, using the geodesic curvature features presented in Chapter 4 for diagnosing these syndromes through studying face traits variation is planning for future work.

The SNPs rs9995821 and rs1852985 located at genes DCHS2 and SUPT3H/RUNX2 respectively have shown an association with nasion position and nose shape facial morphological trait (columella inclination, nose bridge breadth, and nose wing breadth) in approximately 6,000 Latin Americans [4]. In contrast, in this work, they report an association with the lower lip in approximately 4747 Caucasian people. Thus, this result is interesting and should be examined further by carefully replicating this finding in another population group, this planned as a part of our future work beyond this thesis.
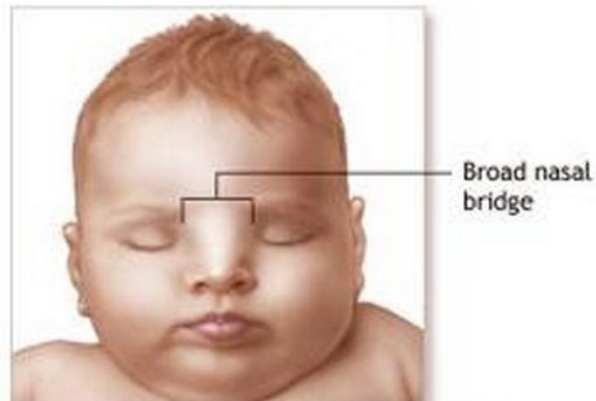
Figure 8.1: The Waardenberg syndrome [267]

Figure 8.2: The Pierre Robin syndrome [266]

Table 8.1: Results of the discovery phase genome wide association study(GWAS)

Feature	p-value	Reference SNP	Gene
Lower lip contour shape	0.0004187	rs9995821	DCHS2
	0.0006972	rs1852985	SUPT3H/RUNX2
Lower lip tone shape	3.447e-06	rs7011739	FAM84B
Nose shape	0.0001763	rs974448	DCHS2
	1.318e-06	rs10176525	PAX3
Nose tip shape	0.000999	rs11655006	BC039327

Table 8.2: Results of the discovery phase genome wide association study(GWAS) medical effects

SNP	Medical effect	Reference
rs9995821	Alae and subnasale Columella inclination	[4]
rs1852985	Nose bridge breadth	[4]
rs7011739	Mild alteration to facial morpholgy, sporadic Cleft Lip Plate and Lower lip shape	[283]
rs974448	Nasal root, Waardenberg syndrome	[64]
rs10176525	Nose width	[247]
rs11655006	Pierre Robin syndrome	[130]

8.3   Face morphological traits analysis using average face

Average face is an important method for studying facial anomalies morphology, evaluating facial asymmetries, average facial growth and comparing facial morphologies for different ages, gender and ethnicity etc [123]. An average face can be superimposed on to a face and the surface difference between the faces can be quantified in millimeters as mean differences between all points from one surface to another. Color maps can be used to visualise these differences [137].

In this chapter, the average face used to visualise the morphological traits influence on the 3D face[1]. The GWAS results indicate four traits have genes association, these traits are the nose shape (7 categories), Lower lip contour shape (4 categories), the nose tip shape(5 categories), and lower lip tone shape (5 categories). Average faces were constructed for each trait category subjects, (e.g. nose shape has seven average faces). Then the superimposition results between these average faces are visualised as a color map which indicates the surface distances between two average faces, the red color is highlighted the maximum difference while the blue color the minimum difference. Figures 8.3, 8.4, 8.5 and 8.6 show only the superimposition results of average face contains the minimum number of subjects against the others for nose shape trait, lower lip shape trait, lower lip tone shape trait and nose tip shape trait respectively.

---

[1]The average face is a simple method can be used to assess genes influence on the 3D face. Unfortunately, ALSPAC institution refused to provide us with the genetic information of the subjects
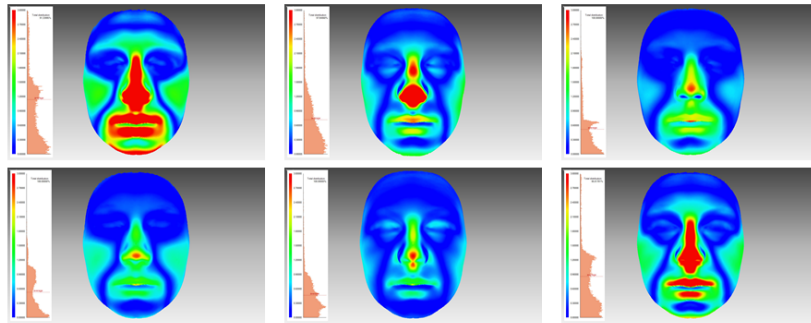
Figure 8.3: Superimposition of average faces for nose shape trait
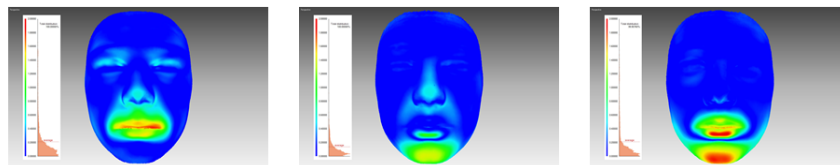
Figure 8.4: Superimposition of average faces for lower lip shape trait
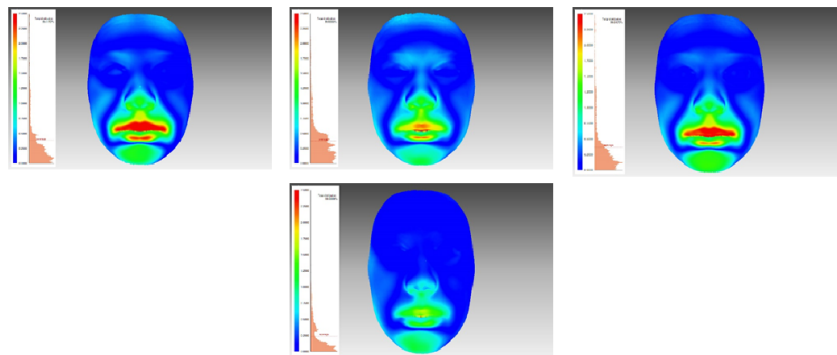
Figure 8.5: Superimposition of average faces for lower lip tone shape trait
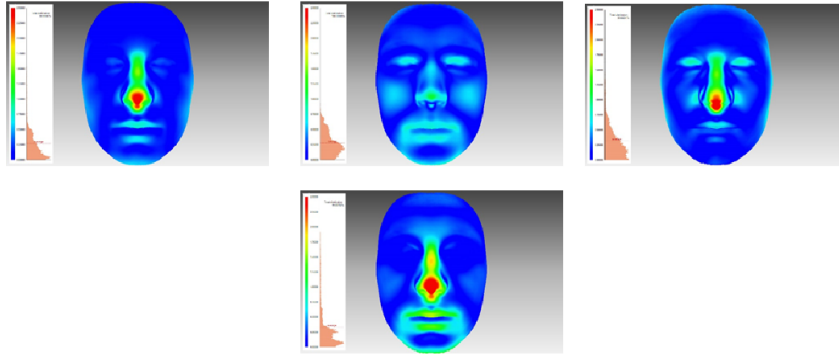
Figure 8.6: Superimposition of average faces for nose tip shape trait

8.4 Discussion

The majority of human studies in this field have focused on the genetics of various facial abnormalities such as cleft lip/palate, Downs syndrome, Floating-Harbor syndrome, and Noonan syndrome [75,124,132,202,253]. In spite of that, there is insufficient evidence in the literature on the association between facial morphological traits and genes in a normal population as this work did.

The Euclidean and angular measurements for 2D photographs or manual anthropometric have been widely used in genetic studies [4, 37, 128,276]. The current study, which uses novel 3D geometric features, provides the opportunity to better capture facial traits shape automatically and determine which genetic variants may influence these traits (the lips and nose morphological traits). Few studies have been carried out using the 3D geometric features for identifying the genes associations. For examples, Claes et al. [63] attempted modeling 3D facial shape from DNA. The spatially dense quasi-landmarks were used to measure face shape in population samples with mixed ancestry from three locations (United States, Brazil, and Cape Verde). They looked for the relationship between facial variation and the effects of sex, ancestry, and a subset of facial candidate genes, by using bootstrapped response-based imputation modeling (BRIM). The authors reported results on a set of 20 genes showing significant effects on facial features. Recently, Tsagkrasoulis et al. [282] placed thousands of landmarks throughout the facial surface of 952 twins recruited from the Twins UK and automatically established point-wise correspondence across faces. They used these landmarks to characterise

facial geometry through curvature features and measured the heritability maps of these faces (no genes association work). In general, these studies were conducted on small datasets and used the correspondences dense landmarks for calculating the geometric features for the whole face. In contrast, in this work, a much larger population group was tested and novel 3D geometric features were proposed, these features characterised the facial morphological traits efficiently for genes association purpose. Another interesting study was conducted on 694 Uyghurs, the authors investigated whether 3D faces could be predicted to certain degree by using the top associated SNPs. The approach was tested in hypothetical forensic scenarios, to evaluate the prediction approach practical potential [220]. This study has motivated us to work on prediction the lips and nose categories from the associated SNPs in the future.

Two other interesting studies have been conducted on ALSPAC dataset. Paternoster et al. [213] identified genetic association between the common 'intronic' SNP rs7559271 in PAX3 gene and the 3D facial distance 'nasion to mid-endocanthion'. Variation in $n - men$ distance reflects variation in the nasal bridge prominence trait. Their study analysed the association between the 21 landmarks principle components and 250 facial measurements (90 distances, 118 angles, and 42 ratios) with 500.000 SNPs. Wilson [296] performed a genome wide association study for 27 lips traits (manually categorised) with 500.0000 SNPs. Two trait/genotype associations were found in the discovery sample that reached GWAS significance; DOCK1 and chin dimple and CDH4 and mentolabial fold.

In this work, 11 lip and nose morphological traits (automatically categorised) were tested for the association with 64 SNPs only. The genes associations were identified between 6 SNPs and four lips and nose morphological traits in a population cohort of 15-year-old adolescents. The most significant p-values were found between the nose shape and rs10176525 and rs7011739 with lower lip tone.

Figures 8.3, 8.4, 8.5 and 8.6 illustrate the superimposition results of average faces. The superimposition method was used to validate the discovered face morphological traits categories. The average faces superimposition results are approximately compatible with regression results in Chapters 5 and 7, Fig. 5.12 and Fig. 7.7.

8.5   Summary

In this chapter eleven lips and nose traits are used for applying GWAS, four of them show significant association with genes' SNPs. A simple method to visualise face morphological traits variation is proposed, this method can be used also to visualise the genes variation influence on the 3D face. Generally, these results prove the *effectiveness* of the proposed geometric features and face morphological traits classification and caregorisation approach in the medical field application specifically in genetic association studies. Therefore, more research ideas are planned for the future.

Chapter 9

CONCLUSIONS AND FUTURE WORK

The automatic classification and categorisation approach and the 3D geometric features presented in this thesis is principally different from the previous approaches and geometric features in at least three respects:

- The proposed facial features are based on face surface variation, rather than simple anthropometric measurements.

- The morphological traits are classified to reflect the standard variation in healthy population rather than a population with a diseases/syndromes.

- The facial morphological traits variation analysis is performed automatically rather than manually.

9.1 Detailed conclusions

This thesis has made several contributions to the state of the art in 3D face morphological traits classification and categorisation methods:

- In Chapter 3 two methods for data preprocessing implemented (non-regularised and regularised), to find which method is more efficient in facial morphological traits classification. The regularisation method that used here was used in [63] for efficient face shape measurements. The two methods are used for lips morphology classification and the accuracies indicate the non-regularised method is more suitable for better classification accuracies using the 3D geometric features proposed in Chapter 4.

- In Chapter 4 a new set of 3D geometric features are derived from mean, Gaussian curvatures, shape indices, and curvedness measures obtained along the *geodesic path* between 3D facial anthropometric landmarks. These features describe the face surface shape variation relying on the biological landmarks rather than the simple anthropometric measurements which provide poor surface information. The impact of these features on face morphology classification is tested comparing to anthropometric measurements. The basic curvature features showed robust accuracies in classification face parts (philtrum area) but for strict traits classification, the geodesic curvatures were found more convenient. However, these features extraction suffers from two limitations: selecting of the suitable features descriptor normalisation method and the ring size for curvature features calculation.

- In Chapter 5 the first automatic approach for lips morphological traits classification and categorisation is proposed. The geodesic curvature features which were proposed in Chapter 4 were used for lips supervised classification. The classification accuracies were calculated using the Boosting and SVM methods for both types of preprocessing methods described in Chapter 3. The classification accuracies using the boosting method were outperformed those of SVM method because it handles the data balancing problem. The classification accuracies ($71\% - 77\%$) for non-regularised preprocessing method were higher than the regularised ones. Consequently, this chapter's experiments proved that the non-regularised preprocessing method is better in characterisation the face shape and facial morphological traits. The classification accuracies for regularised face is comparable to those achieved using non-regularised preprocessing ($67\% - 72\%$). This indicates the robustness of the proposed geodesic curvature features in analysis and classification face morphology traits in high and low resolution meshes. The state of art measurements (Euclidean and geodesic) distances are also calculated for lips traits classification purpose. Their classification accuracies are less than the geodesic curvatures features classification accuracies.

An automatic categorisation approach was implemented since the best obtained classification accuracies using manual categories performed by [295] ranged from 72% to 79%. The number cluster selection for Kmeans++ clustering method was a challenging problem, therefore different internal validation methods were tested for verification. The PLSR was used to determine the effect of the manual and the automatic traits categories on the 3D faces' lips area . This is the first work to produce a method to evaluate the performance of manual or automatic facial morphological traits categories visually. Generally, the classification accuracies based on the automatic lip categories were higher than those obtained using the manual categories by at least 8% and the automatic categories were found to be more significant in the lip area than the manual categories.

- In Chapter 6 the proposed novel 3D geometric features were applied to the problem of gender analysis and discrimination. Five experiments have been performed in this chapter that have explored in some detail aspects of facial traits based on key anthropometric landmarks. Their results have shown that the geodesic path curvature features extracted between the 3D facial landmarks have the ability to classify the gender of Caucasian teenagers with an accuracy of 87.3%. The combination of the new 3D geometric descriptor with more classical distance measures obtained the best classification accuracy of 88.6%. The hybrid geodesic path curvature features and geodesic distance demonstrated an improved ability not only in terms of accuracy but also in sensitivity and specificity. The sensitivity and specificity results show a noticeable variation between Caucasian teenagers in terms of both female and male nose morphology. Finally, the geodesic paths between certain facial landmarks were more discriminative for gender classification and were more significant in 3D facial-profile contours. The nose Ala path, Cupid bow path, and the path between the internal eyes landmarks were shown to be significant.

- In Chapter 7 the nose morphological traits were categorised automatically using the geodesic curvature features proposed in Chap-

ter 4 and the geodesic distance. The same approach that was used for categorising the lips morphological traits in Chapter 5 is utilised here. The results of this chapter indicate the potential of the proposed approach and the 3D geometric features to categorise any face traits without previous knowledge about these trait categories. The nose traits categories were assessed and described by the medical experts from dental school, Cardiff University (Prof.Stephen Richmond and Dr. Caryl Wilson).

- In Chapter 8 the Genome Wide Association Study (GWAS) was performed for the lips and nose morphological traits categories determined in Chapter 5 and Chapter 7. The associations between 11 facial morphological traits and 64 were analysed. The statistical significance of p-value was 0.0008 according to Bonferroni correction multiple testing analysis. The P-value was less than or near the significant value for six SNPs, two of them related to facial Waardenberg syndrome and Pierre Robin syndrome. In contrast, Wilson [296] performed a genome wide association study for 27 lips traits (manually categorised) with 500.0000 SNPs, two SNPs only have p-value less than the hit p-value $1e-7$. Moreover, in this chapter, a new method was proposed for visualising facial traits variation influence on 3D face. In general, this chapter results prove that the method proposed in the present thesis for automatic categorisation of 3D facial morphology is effective for genes association medical application and open the door for remarkable research in the future.

9.2   Future work

There are many research questions that have arisen as a result of this thesis, and the potential for future research includes:

- Investigate other methods for data clustering and curvature calculation such as Spectral clustering method and Voronoi curvature calculation method.

- Investigate several approaches that may be used to compute geodesic distances and paths on triangular meshes (e.g., see [300, 305]), to

determine if they affect face morphology classification accuracy and the computational time positively.

- Use other methods to normalise the geodesic curvature features. For example, using the mean, the variation, the max value, the min value, the mode value statistical information for describing the geodesic curvature feature descriptors.

- Use the geodesic curvature features in the judgment of the degree of masculinity or femininity of a face since facial gender is considered to be a continuum of masculinity or femininity.

- Investigative other facial traits shape variation. For example, categorise the chin shape automatically and replicate the GWAS for this face traits according to [296] finding.

- Replicate this thesis work (automatic lips and nose traits categorisation approach) on other datasets such as TwinsUK [80] and FaceBase [203] datasets.

- Use the geodesic curvature features to study the face morphology variation in subjects have craniofacial syndromes, such as Pierre Robin and Waardenberg syndromes.

- Analyse the relation between the facial morphological traits variation and age groups.

- Develop a quantitative model to predict 3D faces morphological traits based on the GWAS SNPs. In other words, built system to predicate the face traits shapes from the related SNPs.

- Analyse the association of lips and nose morphological traits with climate adaptation. In other words, study the effect of the environmental conditions on face traits shape and growth.

BIBLIOGRAPHY

[1] Maged Abdelkader, Samuel Leong, and Paul S White. Aesthetic proportions of the nasal aperture in 3 different racial groups of men. *Archives of facial plastic surgery*, 7(2):111–113, 2005.

[2] Hervé Abdi, Dominique Valentin, Betty Edelman, and Alice J O'Toole. More about the difference between men and women: evidence from linear neural networks and the principal-component approach. *Perception*, 24(5):539–562, 1995.

[3] Akinlolu Abdulazeez Adelaja et al. Nasal biometrics and nasofacial proportion in Hausas and Yorubas using Akinlolu-Raji image-processing algorithm. *CHRISMED Journal of Health and Research*, 3(2):112–130, 2016.

[4] Kaustubh Adhikari, Macarena Fuentes-Guajardo, Mirsha Quinto-Sánchez, Javier Mendoza-Revilla, Juan Camilo Chacón-Duque, Victor Acuña-Alonzo, Claudia Jaramillo, William Arias, Rodrigo Barquera Lozano, Gastón Macín Pérez, et al. A genome-wide association scan implicates *DCHS2, RUNX2, GLI3, PAX1* and *EDAR* in human facial variation. *Nature communications*, 7:1–25, 2016.

[5] Rachid Ahdid, El Mahdi Barrah, Said Safi, and Bouzid Manaut. Facial surface analysis using iso-geodesic curves in three dimensional face recognition system. *arXiv preprint arXiv:1608.08878*, pages 1–20, 2016.

[6] Rachid AHDID, SAFI Said, and Bouzid MANAUT. Three dimensional face surfaces analysis using geodesic distance. *Journal of Computer Sciences and Applications*, 3(3):67–72, 2015.

[7] Lale Akarun, B Gokberk, and Albert Ali Salah. 3D face recognition for biometric applications. In *Signal Processing Conference, 2005 13th European*, pages 1–5. IEEE, 2005.

[8] Muharrem Akgüner, Ali Barutçu, and Can Karaca. Adolescent growth patterns of the bony and cartilaginous framework of the nose: a cephalometric study. *Annals of plastic surgery*, 41(1):66–69, 1998.

[9] Ala Al Ali, Stephen Richmond, Hashmat Popat, Arshed M Toma, Rebecca Playle, Alexei I Zhurov, David Marshall, Paul L Rosin, and John Henderson. The influence of asthma on face shape: a three-dimensional study. *The European Journal of Orthodontics*, pages 67–85, 2012.

[10] Maisa O Al-Sebaei. The validity of three neo-classical facial canons in young adults originating from the Arabian Peninsula. *Head & face medicine*, 11(1):4–10, 2015.

[11] Kristina Aldridge, Ian D George, Kimberly K Cole, Jordan R Austin, T Nicole Takahashi, Ye Duan, and Judith H Miles. Facial phenotypes in subgroups of pre-pubertal boys with autism spectrum disorders are correlated with clinical phenotypes. *Molecular Autism*, 2(1):15–25, 2011.

[12] Pierre Alliez, David Cohen-Steiner, Olivier Devillers, Bruno Lévy, and Mathieu Desbrun. Anisotropic polygonal remeshing. *ACM Transactions on Graphics (TOG)*, 22(3):485–493, 2003.

[13] Mohammad Alqattan, J Djordjevic, AI Zhurov, and Stephen Richmond. Comparison between landmark and surface-based three-dimensional analyses of facial asymmetry in adults. *The European Journal of Orthodontics*, 37(1):1–12, 2015.

[14] ALSPAC. Avon longitudinal study of parents and children. www.bris.ac.uk/alspac/researchers/data-access/data-dictionary. Accessed: Jun 2017.

[15] Chouchane Ammar, Belahcene Mebarka, Ouamane Abdelmalik, and Bourennane Salah. Evaluation of histograms local features and dimensionality reduction for 3D face verification. *Journal of information processing systems*, 12(3):468–488, 2016.

[16] Marti J Anderson and Pierre Legendre. An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. *Journal of statistical computation and simulation*, 62(3):271–303, 1999.

[17] Per Rønsholt Andresen, Fred L Bookstein, K Couradsen, Bjarne Kjaer Ersboll, Jeffrey L Marsh, and Sven Kreiborg. Surface-bounded growth modeling applied to human mandibles. *IEEE transactions on medical imaging*, 19(11):1053–1063, 2000.

[18] Gidudu Anthony, Hulley Gregg, and Marwala Tshilidzi. Image classification using SVMs: one-against-one vs one-against-all. *arXiv preprint arXiv:0711.2914*, pages 1–20, 2007.

[19] David Arthur and Sergei Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035. Society for Industrial and Applied Mathematics, 2007.

[20] Saw Chit Aung, Foo Chee Liam, and Lee Seng Teik. Three dimensional laser scan assessment of the Oriental nose with a new classification of Oriental nasal types. *British journal of plastic surgery*, 53(2):109–116, 2000.

[21] SC Aung, RCK Ngim, and ST Lee. Evaluation of the laser scanner as a surface measuring tool and its accuracy compared with direct facial anthropometric measurements. *British journal of plastic surgery*, 48(8):551–558, 1995.

[22] Noyan Aynechi, Brent E Larson, Vladimir Leon-Salazar, and Soraya Beiraghi. Accuracy and precision of a 3D anthropometric facial analysis with and without landmark labeling before image acquisition. *The Angle orthodontist*, 81(2):245–252, 2011.

[23] Mohamed Baghdadi, Nacéra Benamrane, and Lakhdar Sais. Fuzzy generalized fast marching method for 3d segmentation of brain structures. *International Journal of Imaging Systems and Technology*, 27(3):281–306, 2017.

[24] Lahoucine Ballihi, Boulbaba Ben Amor, Mohamed Daoudi, Anuj Srivastava, and Driss Aboutajdine. Boosting 3-D-geometric features for efficient face recognition and gender classification. *IEEE Transactions on Information Forensics and Security*, 7(6):1766–1779, 2012.

[25] Lahoucine Ballihi, Boulbaba Ben Amor, Mohamed Daoudi, Anuj Srivastava, and Driss Aboutajdine. Geometric based 3D facial gender classification. In *Communications Control and Signal Processing (ISCCSP), 2012 5th International Symposium on*, pages 1–5. IEEE, 2012.

[26] Mark Barash, Philipp E Bayer, and Angela van Daal. Candidate gene scan for Single Nucleotide Polymorphisms involved in the determination of normal variability in human craniofacial morphology. *bioRxiv*, pages 1–22, 2016.

[27] Bülent Baydaş, Abdulvahit Erdem, İbrahim Yavuz, and İsmail Ceylan. Heritability of facial proportions and soft-tissue profile characteristics in Turkish Anatolian siblings. *American journal of orthodontics and dentofacial orthopedics*, 131(4):504–509, 2007.

[28] Terri H Beaty, Jeffrey C Murray, Mary L Marazita, Ronald G Munger, Ingo Ruczinski, Jacqueline B Hetmanski, Kung Yee Liang, Tao Wu, Tanda Murray,

M Daniele Fallin, et al. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near mafb and abca4. *Nature genetics*, 42(6):525–529, 2010.

[29] TH Beaty, MA Taub, AF Scott, JC Murray, ML Marazita, H Schwender, MM Parker, JB Hetmanski, P Balakrishnan, MA Mansilla, et al. Confirming genes influencing risk to cleft lip with/without cleft palate in a case–parent trio study. *Human genetics*, 132(7):771–781, 2013.

[30] George N Bebis, Michael Georgiopoulos, George M Papadourakis, and Gregory L Heileman. Increasing classification accuracy using multiple-neural-network schemes. In *Aerospace Sensing*, pages 221–231. International Society for Optics and Photonics, 1992.

[31] Juan Bekios-Calfa, Jose M Buenaposada, and Luis Baumela. Revisiting linear discriminant techniques in gender recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(4):858–864, 2011.

[32] Asa Ben-Hur and Jason Weston. A user's guide to support vector machines. In *Data mining techniques for the life sciences*, pages 223–239. Springer, 2010.

[33] Mohammed Bennamoun, Yulan Guo, and Ferdous Sohel. Feature selection for 2d and 3d face recognition. *Wiley Encyclopedia of Electrical and Electronics Engineering*, 2015.

[34] Paul J Besl and Neil D McKay. Method for registration of 3-D shapes. In *Robotics-DL tentative*, pages 586–606. International Society for Optics and Photonics, 1992.

[35] Raj Bhopal. The beautiful skull and Blumenbach's errors: the birth of the scientific concept of race. *BMJ*, 335(7633):1308–1309, 2007.

[36] Irving Biederman. Recognition-by-components: a theory of human image understanding. *Psychological review*, 94(2):115–147, 1987.

[37] Stefan Boehringer, Fedde Van Der Lijn, Fan Liu, Manuel Günther, Stella Sinigerova, Stefanie Nowak, Kerstin U Ludwig, Ruth Herberz, Stefan Klein, Albert Hofman, et al. Genetic determination of human facial morphology: links between cleft-lips and normal variation. *European Journal of Human Genetics*, 19(11):1192–1197, 2011.

[38] David Bommes and Leif Kobbelt. Accurate computation of geodesic distance fields for polygonal curves on triangle meshes. In *VMV*, volume 7, pages 151–160, 2007.

[39] Fred L Bookstein. *Morphometric tools for landmark data: geometry and biology.* Cambridge University Press, 1997.

[40] Dante Bresolin, Peter A Shapiro, Gail G Shapiro, Michael K Chapko, and Steven Dassel. Mouth breathing in allergic children: its relationship to dentofacial development. *American journal of orthodontics*, 83(4):334–340, 1983.

[41] John C Brigham and Paul Barkowitz. Do they all look alike? the effect of race. sex, experience, and attitudes on the ability to recognize faces. *Journal of Applied Social Psychology*, 8(4):306–318, 1978.

[42] Alexander M Bronstein, Michael M Bronstein, Leonidas J Guibas, and Maks Ovsjanikov. Shape google: Geometric words and expressions for invariant shape retrieval. *ACM Transactions on Graphics (TOG)*, 30(1):1, 2011.

[43] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. Three-dimensional face recognition. *International Journal of Computer Vision*, 64(1):5–30, 2005.

[44] Elizabeth Brown and David I Perrett. What gives a face its gender? *Perception*, 22(7):829–840, 1993.

[45] Iain Brown and Christophe Mues. An experimental comparison of classification algorithms for imbalanced credit scoring data sets. *Expert Systems with Applications*, 39(3):3446–3453, 2012.

[46] Vicki Bruce, A Mike Burton, Elias Hanna, Pat Healey, Oli Mason, Anne Coombes, Rick Fright, and Alf Linney. Sex discrimination: how do we tell the difference between male and female faces? *Perception*, 22(2):131–152, 1993.

[47] Vicki Bruce and Andy Young. Understanding face recognition. *British journal of psychology*, 77(3):305–327, 1986.

[48] Nicola Burns, Yaxin Bi, Hui Wang, and Terry Anderson. Sentiment analysis of customer reviews: Balanced versus unbalanced datasets. *Knowledge-Based and Intelligent Information and Engineering Systems*, pages 161–170, 2011.

[49] Charles J Burstone. Lip posture and its significance in treatment planning. *American journal of orthodontics*, 53(4):262–284, 1967.

[50] A Mike Burton, Vicki Bruce, and Neal Dench. What's the difference between men and women? evidence from facial measurement. *Perception*, 22(2):153–176, 1993.

[51] William S Bush and Jason H Moore. Genome-wide association studies. *PLoS computational biology*, 8(12):e1002822, 2012.

[52] Patrick Callaerts, Georg Halder, and Walter J Gehring. *PAX-6* in development and evolution. *Annual review of neuroscience*, 20(1):483–532, 1997.

[53] Ruth Campbell, Jane Walker, and Simon Baron-Cohen. The development of differential use of inner and outer face features in familiar face identification. *Journal of Experimental Child Psychology*, 59(2):196–210, 1995.

[54] John C Carey, M Michael Cohen, Cynthia JR Curry, Koenraad Devriendt, Lewis B Holmes, and Alain Verloes. Elements of morphology: standard terminology for the lips, mouth, and oral region. *American Journal of Medical Genetics Part A*, 149(1):77–92, 2009.

[55] Girish Chandrashekar and Ferat Sahin. A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1):16–28, 2014.

[56] Kyong 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*, 28(10):1695–1700, 2006.

[57] Jindong Chen and Yijie Han. Shortest paths on a polyhedron. In *Proceedings of the sixth annual symposium on Computational geometry*, pages 360–369. ACM, 1990.

[58] Li Chen and Soma Biswas. Digital curvatures applied to 3D object analysis and recognition: A case study. In *International Workshop on Combinatorial Image Analysis*, pages 45–58. Springer, 2012.

[59] Yi D Cheng, Alice J O'Toole, and Hervé Abdi. Classifying adults' and children's faces by sex: Computational investigations of subcategorical feature encoding. *Cognitive science*, 25(5):819–838, 2001.

[60] Peter Claes. *A robust statistical surface registration framework using implicit function representations-application in craniofacial reconstruction.* PhD thesis, K.U. Leuven, 2007.

[61] Peter Claes, Katleen Daniels, Dirk Vandermeulen, Paul Suetens, and Mark D Shriver. A PLS regression framework for spatially-dense geometric morphometrics to analyze effects on shape and shape characteristics: applied to the study of genomic ancestry and sex on facial morphology. In *BIOLOGICAL SHAPE ANALYSIS: Proceedings of the 3rd International Symposium*, pages 205–215, 2015.

[62] Peter Claes, Harold Hill, and Mark D Shriver. Toward DNA-based facial composites: preliminary results and validation. *Forensic Science International: Genetics*, 13:208–216, 2014.

[63] Peter Claes, Denise K Liberton, Katleen Daniels, Kerri Matthes Rosana, Ellen E Quillen, Laurel N Pearson, Brian McEvoy, Marc Bauchet, Arslan A Zaidi, Wei Yao, et al. Modeling 3D facial shape from DNA. *PLoS Genet*, 10(3):e1004224, 2014.

[64] Peter Claes and Mark D Shriver. New entries in the lottery of facial GWAS discovery. *PLoS genetics*, 12(8):e1006250, 2016.

[65] Peter Claes, M Walters, and J Clement. Improved facial outcome assessment using a 3D anthropometric mask. *International journal of oral and maxillofacial surgery*, 41(3):324–330, 2012.

[66] Peter Claes, Mark Walters, Mark D Shriver, David Puts, Greg Gibson, John Clement, Gareth Baynam, Geert Verbeke, Dirk Vandermeulen, and Paul Suetens. Sexual dimorphism in multiple aspects of 3D facial symmetry and asymmetry defined by spatially dense geometric morphometrics. *Journal of anatomy*, 221(2):97–114, 2012.

[67] Peter Claes, Mark Walters, Dirk Vandermeulen, and John Gerald Clement. Spatially-dense 3D facial asymmetry assessment in both typical and disordered growth. *Journal of anatomy*, 219(4):444–455, 2011.

[68] Laurent D Cohen and Ron Kimmel. Global minimum for active contour models: A minimal path approach. *International journal of computer vision*, 24(1):57–78, 1997.

[69] David Cohen-Steiner and Jean-Marie Morvan. Restricted delaunay triangulations and normal cycle. In *Proceedings of the nineteenth annual symposium on Computational geometry*, pages 312–321. ACM, 2003.

[70] ATLAS Collaboration et al. The evolution of boosting algorithms-from machine learning to statistical modelling. *Methods Inf Med*, 53(6):419–427, 2014.

[71] Timothy F Cootes, Andrew Hill, Christopher J Taylor, and Jane Haslam. The use of active shape models for locating structures in medical images. In *Biennial International Conference on Information Processing in Medical Imaging*, pages 33–47. Springer, 1993.

[72] Robert S Corruccini, Louisa B Flander, and Samvit S Kaul. Mouth breathing, occlusion, and modernization in a north Indian population: An epidemiologic study. *The Angle Orthodontist*, 55(3):190–196, 1985.

[73] Josanne Cox-Brinkman, Anouk Vedder, Carla Hollak, Linda Richfield, Atul Mehta, Kate Orteu, Frits Wijburg, and Peter Hammond. Three-dimensional face shape in fabry disease. *European Journal of Human Genetics*, 15(5):535, 2007.

[74] Keenan Crane, Clarisse Weischedel, and Max Wardetzky. Geodesics in heat: A new approach to computing distance based on heat flow. *ACM Transactions on Graphics (TOG)*, 32(5):152–162, 2013.

[75] Ellen A Croonen, Ineke van der Burgt, Livia Kapusta, and Jos M Th Draaisma. Electrocardiography in Noonan syndrome *PTPN11* gene mutation—phenotype characterization. *American Journal of Medical Genetics Part A*, 146(3):350–353, 2008.

[76] Elias O Da-Silva, Jose EM Batista, Marco AB Medeiros, and Sônia Fonteles. Craniofacial anthropometric studies in Waardenburg syndrome type I. *Clinical genetics*, 44(1):20–25, 1993.

[77] Mohamed Daoudi, Anuj Srivastava, and Remco Veltkamp. *3D face modeling, analysis and recognition.* John Wiley & Sons, 2013.

[78] Markus A de Jong, Andreas Wollstein, Clifford Ruff, David Dunaway, Pirro Hysi, Tim Spector, Fan Liu, Wiro Niessen, Maarten J Koudstaal, Manfred Kayser, et al. An automatic 3d facial landmarking algorithm using 2d gabor wavelets. *IEEE Transactions on Image Processing*, 25(2):580–588, 2016.

[79] M De Menezes, R Rosati, I Baga, A Mapelli, and C Sforza. Three-dimensional analysis of labial morphology: effect of sex and age. *International journal of oral and maxillofacial surgery*, 40(8):856–861, 2011.

[80] UK Department of Twin Research Genetic Epidemiology, King's College London. Twinsuk. http://www.twinsuk.ac.uk. 2017.

[81] Jelena Djordjevic, Alexei I Zhurov, Stephen Richmond, Visigen Consortium, et al. Genetic and environmental contributions to facial morphological variation: a 3D population-based twin study. *PloS one*, 11(9):e0162250, 2016.

[82] Gianluca Donato and Serge Belongie. Approximate thin plate spline mappings. In *European conference on computer vision*, pages 21–31. Springer, 2002.

[83] Chitra Dorai and Anil K Jain. COSMOS—a representation scheme for 3D free-form objects. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(10):1115–1130, 1997.

[84] Hassen Drira, Boulbaba Ben Amor, Mohamed Daoudi, and Anuj Srivastava. Pose and expression-invariant 3D face recognition using elastic radial curves. In *British Machine Vision Conference*, pages 1–11, 2010.

[85] Peter Drotár, Juraj Gazda, and Zdenek Smékal. An experimental comparison of feature selection methods on two-class biomedical datasets. *Computers in biology and medicine*, 66:1–10, 2015.

[86] Richard O Duda, Peter E Hart, and David G Stork. *Pattern classification*. John Wiley & Sons, 2012.

[87] Sahibsingh A Dudani. The distance-weighted k-nearest-neighbor rule. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6(4):325–327, 1976.

[88] Betty Edelman, DominiquE Valentin, and Hervé Abdi. Sex classification of face areas: how well can a linear neural network predict human performance? *Journal of Biological Systems*, 6(03):241–263, 1998.

[89] Boris Efraty, Emil Bilgazyev, Shishir Shah, and Ioannis A Kakadiaris. Profile-based 3d-aided face recognition. *Pattern recognition*, 45(1):43–53, 2012.

[90] Robert Kyle Eichelberger and Victor S Sheng. Does one-against-all or one-against-one improve the performance of multiclass classifications? In *AAAI*, 2013.

[91] Mehryar Emambakhsh and Adrian Evans. Nasal patches and curves for expression-robust 3D face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 39(5):995–1007, 2017.

[92] Reyes Enciso, Alex M Shaw, Ulrich Neumann, and James Mah. Three-dimensional head anthropometric analysis. In *Medical Imaging 2003*, pages 590–597. International Society for Optics and Photonics, 2003.

[93] DH Enlow. *Handbook of facial growth*. WB Saunders Company, Philadelphia, 1982.

[94] K Enver, Pakize Erdo, and Kemal Polat. Brain mri segmentation based on different clustering algorithms. *International Journal of Computer Applications*, 155(3):37–40, 2016.

[95] Leslie G Farkas. *Anthropometry of the Head and Face*. Raven Pr, 1994.

[96] Leslie G Farkas and John C Kolar. Anthropometrics and art in the aesthetics of women's faces. *Clinics in plastic surgery*, 14(4):599–616, 1987b.

[97] LG Farkas and JC Kolar. Anthropometric guidelines in cranio-orbital surgery. *Clinics in plastic surgery*, 14(1):1–16, 1987a.

[98] Jean-Marc Fellous. Gender discrimination and prediction on the basis of facial metric information. *Vision research*, 37(14):1961–1973, 1997.

[99] Lister Hill National Center for Biomedical Communications. Genetics home reference. www.ghr.nlm.nih.gov/primer/genomicresearch. July 11, 2017.

[100] Joseph HG Fu et al. Convergence of curvatures in secant approximations. *Journal of Differential Geometry*, 37(1):177–190, 1993.

[101] Suranjan Ganguly, Debotosh Bhattacharjee, and Mita Nasipuri. 3D face recognition from range images based on curvature analysis. *ICTACT Journal on image and video processing*, 4(03):748–753, 2014.

[102] Tim Gatzke and Cindy Grimm. Improved curvature estimation on triangular meshes. Technical Report WUCSE-2004-9, Washington University in St. Louis, Sep 1, 2004.

[103] Timothy D Gatzke and Cindy M Grimm. Estimating curvature on triangular meshes. *International journal of shape modeling*, 12(01):1–28, 2006.

[104] Silvère Gauthier, William Puech, Roseline Bénière, and Gérard Subsol. Analysis of digitized 3D mesh curvature histograms for reverse engineering. *Computers in Industry*, 92:67–83, 2017.

[105] Tobias Gehrig, Matthias Steiner, and Hazım Kemal Ekenel. Draft: evaluation guidelines for gender classification and age estimation. *Technical report, Karlsruhe Institute of Technology*, pages 1–12, 2011.

[106] Genetics Home Reference. Waardenburg syndrome. www.ghr.nlm.nih.gov/condition/waardenburg-syndrome. November 7, 2017.

[107] Marion Gerard, Marc Abitbol, Anne-Lise Delezoide, Jean-Louis Dufier, Jacques Mallet, and Michel Vekemans. *PAX*-genes expression during human embryonic development, a preliminary report. *Comptes rendus de l'Academie des sciences. Serie III, Sciences de la vie*, 318(1):57–66, 1995.

[108] Anthony Gidudu, Gregg Hulley, and Tshilidzi Marwala. Image classification using svms: One-against-one vs one-against-all. In *Proccedings of the 28th Asian Conference on Remote Sensing*, page 96, 2007.

[109] Syed Zulqarnain Gilani and Ajmal Mian. Perceptual differences between men and women: A 3D facial morphometric perspective. In *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pages 2413–2418. IEEE, 2014.

[110] Syed Zulqarnain Gilani, Kathleen Rooney, Faisal Shafait, Mark Walters, and Ajmal Mian. Geometric facial gender scoring: objectivity of perception. *PloS one*, 9(6):e99483, 2014.

[111] Jack Goldfeather and Victoria Interrante. A novel cubic-order algorithm for approximating principal direction vectors. *ACM Transactions on Graphics (TOG)*, 23(1):45–63, 2004.

[112] Beatrice A Golomb, David T Lawrence, and Terrence J Sejnowski. SEXNET: A neural network identifies sex from human faces. In *NIPS*, volume 1, pages 2–8, 1990.

[113] Troy Gor, Chung How Kau, Jeryl D English, Robert P Lee, and Peter Borbely. Three-dimensional comparison of facial morphology in white populations in Budapest, Hungary, and Houston, Texas. *American Journal of Orthodontics and Dentofacial Orthopedics*, 137(3):424–432, 2010.

[114] John C Gower. Generalized procrustes analysis. *Psychometrika*, 40(1):33–51, 1975.

[115] Nick Guenther, Matthias Schonlau, et al. Support vector machines. *Stata Journal*, 16(4):917–937, 2016.

[116] Yulan Guo, Mohammed Bennamoun, Ferdous Sohel, Min Lu, Jianwei Wan, and Ngai Ming Kwok. A comprehensive performance evaluation of 3D local feature descriptors. *International Journal of Computer Vision*, 116(1):66–89, 2016.

[117] Yulan Guo, Jun Zhang, Min Lu, Jianwei Wan, and Yanxin Ma. Benchmark datasets for 3d computer vision. In *Industrial Electronics and Applications (ICIEA), 2014 IEEE 9th Conference on*, pages 1846–1851. IEEE, 2014.

[118] Prashant Gupta. Cross-validation in machine learning. www.towardsdatascience.com/cross-validation-in-machine-learning-72924a69872f. 2017.

[119] Shalini Gupta, Mia K Markey, and Alan C Bovik. Anthropometric 3D face recognition. *International journal of computer vision*, 90(3):331–349, 2010.

[120] Isabelle Guyon and André Elisseeff. An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182, 2003.

[121] Bernd Hamann. Curvature approximation for triangulated surfaces. In *Geometric modelling*, pages 139–153. Springer, 1993.

[122] Peter Hammond, Tim J Hutton, Judith E Allanson, Bernard Buxton, Linda E Campbell, Jill Clayton-Smith, Dian Donnai, Annette Karmiloff-Smith, Kay Metcalfe, Kieran C Murphy, et al. Discriminating power of localized three-dimensional facial morphology. *The American Journal of Human Genetics*, 77(6):999–1010, 2005.

[123] Peter Hammond, Tim J Hutton, Judith E Allanson, Linda E Campbell, Raoul Hennekam, Sean Holden, Michael A Patton, Adam Shaw, I Karen Temple, Matthew Trotter, et al. 3D analysis of facial morphology. *American journal of medical genetics Part A*, 126(4):339–348, 2004.

[124] Peter Hammond, Michael Suttie, Raoul C Hennekam, Judith Allanson, Eileen M Shore, and Frederick S Kaplan. The face signature of fibrodysplasia ossificans progressiva. *American Journal of Medical Genetics Part A*, 158(6):1368–1380, 2012.

[125] A Ben Hamza and Hamid Krim. Geodesic matching of triangulated surfaces. *Image Processing, IEEE Transactions on*, 15(8):2249–2258, 2006.

[126] Xia Han, Hassan Ugail, and Ian Palmer. Gender classification based on 3D face geometry features using SVM. In *CyberWorlds, 2009. CW'09. International Conference on*, pages 114–118. IEEE, 2009.

[127] Simon Haykin and Neural Network. A comprehensive foundation. *Neural networks*, 2(2004):41, 2004.

[128] Huiyu He, Xue Mi, Jiayu Zhang, Qin Zhang, Yuan Yao, Xu Zhang, Feng Xiao, Chunping Zhao, and Shutao Zheng. Correlation between facial morphology and gene polymorphisms in the uygur youth population. *Oncotarget*, 8(17):1–28, 2017.

[129] Zahra Heidari, Hamidreza Mahmoudzadeh-Sagheb, Azam Asemi Rad, and Narjes Dahmardeh. Anthropometric measurements of the lips in 18-25-year-old men of Sistani and Baluch descent. *Bull. Env. Pharmacol. Life Sci*, 3:139–142, 2014.

[130] Katherine L Hill-Harfe, Lee Kaplan, Heather J Stalker, Roberto T Zori, Ramona Pop, Gerd Scherer, and Margaret R Wallace. Fine mapping of chromosome 17 translocation breakpoints ≥900 kb upstream of *SOX9* in acampomelic campomelic dysplasia and a mild, familial skeletal dysplasia. *The American Journal of Human Genetics*, 76(4):663–671, 2005.

[131] Reed A Holdaway. A soft-tissue cephalometric analysis and its use in orthodontic treatment planning. Part I. *American journal of orthodontics*, 84(1):1–28, 1983.

[132] Rebecca L Hood, Matthew A Lines, Sarah M Nikkel, Jeremy Schwartzentruber, Chandree Beaulieu, Małgorzata JM Nowaczyk, Judith Allanson, Chong Ae Kim, Dagmar Wieczorek, Jukka S Moilanen, et al. Mutations in *SRCAP*, encoding SNF2-related CREBBP activator protein, cause Floating–Harbor syndrome. *The American Journal of Human Genetics*, 90(2):308–313, 2012.

[133] Jeremy Horwood, Giovanni Salvi, Kate Thomas, Larisa Duffy, David Gunnell, Chris Hollis, Glyn Lewis, Paulo Menezes, Andrew Thompson, Dieter Wolke, et al. IQ and non-clinical psychotic symptoms in 12-year-olds: results from the ALSPAC birth cohort. *The British Journal of Psychiatry*, 193(3):185–191, 2008.

[134] Dennis Howitt and Duncan Cramer. *Introduction to statistics in psychology*. Pearson education, 2007.

[135] Kun Hwang, Dae Joong Kim, and Se Ho Hwang. Musculature of the pars marginalis of the upper orbicularis oris muscle. *Journal of Craniofacial Surgery*, 18(1):151–154, 2007.

[136] TIBCO Software Inc. Partial least squares (pls). www.statsoft.com/Textbook/Partial-Least-Squares. 2017.

[137] Syed MS Islam, Mithran S Goonewardene, and Mauro Farella. A review on three dimensional facial averaging for the assessment of orthodontic disorders. In *Innovations and Advances in Computing, Informatics, Systems Sciences, Networking and Engineering*, pages 391–397. Springer, 2015.

[138] Norhayati Jaffar. Application of 3D mapping of the face for forensic application using reconstructive anatomical techniques. Master's thesis, University of Melbourne, 2009.

[139] Sina Jahanbin, Hyohoon Choi, Yang Liu, and Alan C Bovik. Three dimensional face recognition using iso-geodesic and iso-depth curves. In *Biometrics:*

Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on, pages 1–6. IEEE, 2008.

[140] Anil K Jain, M Narasimha Murty, and Patrick J Flynn. Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3):264–323, 1999.

[141] Linda P Jakobsen, Mary A Knudsen, James Lespinasse, Carmen García Ayuso, Carmen Ramos, Jean-Pierre Fryns, Merete Bugge, and Niels Tommerup. The genetic basis of the Pierre Robin sequence. *The Cleft palate-craniofacial journal*, 43(2):155–159, 2006.

[142] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. Support vector machines. In *An Introduction to Statistical Learning*, pages 337–372. Springer, 2013.

[143] Thorsten Joachims, Thomas Finley, and Chun-Nam John Yu. Cutting-plane training of structural SVMs. *Machine Learning*, 77(1):27–59, 2009.

[144] Berglind Johannsdottir, Freyr Thorarinsson, Arni Thordarson, and Thordur Eydal Magnusson. Heritability of craniofacial characteristics between parents and offspring estimated from lateral cephalograms. *American Journal of Orthodontics and Dentofacial Orthopedics*, 127(2):200–207, 2005.

[145] Orhan Hakki Karatas and Ebubekir Toy. Three-dimensional imaging techniques: A literature review. *European journal of dentistry*, 8(1):132, 2014.

[146] Deepu Singh Kataria, Rakesh Kumar Ranjan, and SA Perwaiz. Study of variation in total facial index of north Indian population. *International Journal of Health Sciences and Research (IJHSR)*, 5(4):122–127, 2015.

[147] Chung How Kau, Stephen Richmond, Angela Incrapera, Jeryl English, and James Jiong Xia. Three-dimensional surface acquisition systems for the study of facial morphology and their application to maxillofacial surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 3(2):97–110, 2007.

[148] Chung How Kau, Stephen Richmond, Alexei Zhurov, Maja Ovsenik, Wael Tawfik, Peter Borbely, and Jeryl D English. Use of 3-dimensional surface acquisition to study facial morphology in 5 populations. *American Journal of Orthodontics and Dentofacial Orthopedics*, 137(4):S56–e1, 2010.

[149] L Keith and TV Moore. The developing human: clinically oriented embryology. *7th Elsever Science (USA)*, pages 158–185, 2008.

[150] András Kelemen, Gábor Székely, and Guido Gerig. Elastic model-based segmentation of 3-d neuroradiological data sets. *IEEE Transactions on medical imaging*, 18(10):828–839, 1999.

[151] Matthew J Kesterke, Zachary D Raffensperger, Carrie L Heike, Michael L Cunningham, Jacqueline T Hecht, Chung How Kau, Nichole L Nidey, Lina M Moreno, George L Wehby, Mary L Marazita, et al. Using the 3D Facial Norms Database to investigate craniofacial sexual dimorphism in healthy children, adolescents, and adults. *Biology of sex differences*, 7(1):23–30, 2016.

[152] Fouad Khan. An initial seed selection algorithm for k-means clustering of georeferenced data to improve replicability of cluster assignments for mapping application. *Applied Soft Computing*, 12(11):3698–3700, 2012.

[153] Hyun-Jin Kim, Sun-Wha Im, Ganchimeg Jargal, Siwoo Lee, Jae-Hyuk Yi, Jeong-Yeon Park, Joohon Sung, Sung-Il Cho, Jong-Yeol Kim, Jong-Il Kim, et al. Heritabilities of facial measurements and their latent factors in Korean families. *Genomics & informatics*, 11(2):83–92, 2013.

[154] Danil Kirsanov. Exact geodesic for triangular meshes. www.uk.mathworks.com/matlabcentral/fileexchange/18168 exact geodesic for triangular meshes. Updated 3 Mar 2008.

[155] Christian Peter Klingenberg, Larry J Leamy, Eric J Routman, and James M Cheverud. Genetic architecture of mandible shape in mice: effects of quantitative trait loci analyzed by geometric morphometrics. *Genetics*, 157(2):785–802, 2001.

[156] Paul GM Knoops, Caroline AA Beaumont, Alessandro Borghi, Naiara Rodriguez-Florez, Richard WF Breakey, William Rodgers, Freida Angullia, NU Owase Jeelani, Silvia Schievano, and David J Dunaway. Comparison of 3D scanner systems for craniomaxillofacial imaging. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 2017.

[157] LAP Kohn. The role of genetics in craniofacial morphology and growth. *Annual Review of Anthropology*, 20(1):261–278, 1991.

[158] John C Kolar, Elizabeth M Salter, and Seth M Weinberg. Preoperative craniofacial dysmorphology in isolated sagittal synostosis: a comprehensive anthropometric evaluation. *Journal of Craniofacial Surgery*, 21(5):1404–1410, 2010.

[159] Oscar Koller, Hermann Ney, and Richard Bowden. Deep learning of mouth shapes for sign language. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 85–91, 2015.

[160] W.P. Koppen. *Learning 3D face shape features from local coherence.* PhD thesis, Centre for Vision, Speech and Signal Processing Faculty of Electronic Engineering University of Surrey, 2014.

[161] WP Koppen, WJ Christmas, DJM Crouch, WF Bodmer, and JV Kittler. Extending non-negative matrix factorisation to 3D registered data. In *Biometrics (ICB), 2016 International Conference on*, pages 1–8. IEEE, 2016.

[162] Sotiris Kotsiantis, Dimitris Kanellopoulos, Panayiotis Pintelas, et al. Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering*, 30(1):25–36, 2006.

[163] Sotiris B Kotsiantis, I Zaharakis, and P Pintelas. Supervised machine learning: A review of classification techniques, 2007.

[164] Kevin Lai, Liefeng Bo, Xiaofeng Ren, and Dieter Fox. A scalable tree-based approach for joint object and pose recognition. In *Aaai*, volume 1, page 2, 2011.

[165] Sirpa H Laitinen, Arja Heliövaara, and Reijo E Ranta. Craniofacial morphology in young adults with the pierre robin sequence and isolated cleft palate. *Acta odontologica Scandinavica*, 55(4):223–228, 1997.

[166] Joseph S Lappin, J Farley Norman, and Flip Phillips. Fechner, information, and shape perception. *Attention, Perception, & Psychophysics*, 73(8):2353–2378, 2011.

[167] M Latha and R Surya. Brain tumour detection using neural network classifier and kmeans clustering algorithm for classification and segmentation. *Brain*, 1(01):27–32, 2016.

[168] Marc T Law, Nicolas Thome, and Matthieu Cord. Bag-of-words image representation: Key ideas and further insight. In *Fusion in Computer Vision*, pages 29–52. Springer, 2014.

[169] Won-Chul Lee, Tetsutaro Yamaguchi, Chiaki Watanabe, Akira Kawaguchi, Mayako Takeda, Yong-Il Kim, Shugo Haga, Yoko Tomoyasu, Hajime Ishida, Koutaro Maki, et al. Association of common *PAX9* variants with permanent tooth size variation in non-syndromic East Asian populations. *Journal of human genetics*, 57(10):1–12, 2012.

[170] Yonguk Lee, Hwanjong Song, Ukil Yang, Hyungchul Shin, and Kwanghoon Sohn. Local feature based 3d face recognition. In *Audio-and Video-Based Biometric Person Authentication*, pages 909–918. Springer, 2005.

[171] Yinjie Lei, Mohammed Bennamoun, Munawar Hayat, and Yulan Guo. An efficient 3d face recognition approach using local geometrical signatures. *Pattern Recognition*, 47(2):509–524, 2014.

[172] Yunqi Lei, Haibin Lai, and Qingmin Li. Geometric features of 3D face and recognition of it by PCA. *Journal of multimedia*, 6(2):207–216, 2011.

[173] Cuixia Li, Yingjun Tan, Dingbiao Wang, and Peijie Ma. Research on 3D face recognition method in cloud environment based on semi supervised clustering algorithm. *Multimedia Tools and Applications*, 76(16):17055–17073, 2017.

[174] Huibin Li, Di Huang, Jean-Marie Morvan, Yunhong Wang, and Liming Chen. Towards 3D face recognition in the real: A registration-free approach using fine-grained matching of 3D keypoint descriptors. *International Journal of Computer Vision*, 113(2):128–142, 2015.

[175] Yan Liang, Yun Zhang, and Xian-Xian Zeng. Pose-invariant 3D face recognition using half face. *Signal Processing: Image Communication*, 57:84–90, 2017.

[176] Fan Liu, Fedde Van Der Lijn, Claudia Schurmann, Gu Zhu, M Mallar Chakravarty, Pirro G Hysi, Andreas Wollstein, Oscar Lao, Marleen De Bruijne, M Arfan Ikram, et al. A genome-wide association study identifies five loci influencing facial morphology in Europeans. *PLoS Genet*, 8(9):e1002932, 2012.

[177] Hsien-Chang Liu, Chan-Hung Su, Yueh-Hsuan Chiang, and Yi-Ping Hung. Personalized face verification system using owner-specific cluster-dependent LDA-subspace. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 4, pages 344–347. IEEE, 2004.

[178] Yanchi Liu, Zhongmou Li, Hui Xiong, Xuedong Gao, and Junjie Wu. Understanding of internal clustering validation measures. In *2010 IEEE International Conference on Data Mining*, pages 911–916. IEEE, 2010.

[179] Rushi Longadge and Snehalata Dongre. Class imbalance problem in data mining review. *arXiv preprint arXiv:1305.1707*, pages 1–6, 2013.

[180] Xiaoguang Lu, Hong Chen, and Anil K Jain. Multimodal facial gender and ethnicity identification. In *International Conference on Biometrics*, pages 554–561. Springer, 2006.

[181] Anders Lundström and John McWilliam. Comparison of some cephalometric distances and corresponding facial proportions with regard to heritability. *The European Journal of Orthodontics*, 10(1):27–29, 1988.

[182] Anders Lundström and John S McWilliam. A comparison of vertical and horizontal cephalometric variables with regard to heritability. *The European Journal of Orthodontics*, 9(1):104–108, 1987.

[183] Roy S Malpass and Jerome Kravitz. Recognition for faces of own and other race. *Journal of personality and social psychology*, 13(4):330–334, 1969.

[184] Davide Maltoni, Dario Maio, Anil Jain, and Salil Prabhakar. *Handbook of fingerprint recognition*. Springer Science & Business Media, 2009.

[185] Harold Matthews, Tony Penington, Ine Saey, Jane Halliday, Evelyn Muggli, and Peter Claes. Spatially dense morphometrics of craniofacial sexual dimorphism in 1-year-olds. *Journal of Anatomy*, 229(4):549–559, 2016.

[186] Richard M McKearney, John V Williams, and Nigel S Mercer. Quantitative computer-based assessment of lip symmetry following cleft lip repair. *The Cleft Palate-Craniofacial Journal*, 50(2):138–143, 2013.

[187] Dereck S Meek and Desmond J Walton. On surface normal and Gaussian curvature approximations given data sampled from a smooth surface. *Computer Aided Geometric Design*, 17(6):521–543, 2000.

[188] Dominik Meier and Elizabeth Fisher. Parameter space warping: shape-based correspondence between morphologically different objects. *IEEE Transactions on Medical Imaging*, 21(1):31–47, 2002.

[189] Roberto Melotti, Jon Heron, Matthew Hickman, John Macleod, Ricardo Araya, and Glyn Lewis. Adolescent alcohol and tobacco use and early socioeconomic position: the ALSPAC birth cohort. *Pediatrics*, pages peds–2009, 2011.

[190] Ezgi Mercan, Linda G Shapiro, Seth M Weinberg, and Su-In Lee. The use of pseudo-landmarks for craniofacial analysis: A comparative study with L1-regularized logistic regression. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 6083–6086. IEEE, 2013.

[191] Merriam-Webster. Cluster analysis. www.merriam-webster.com/dictionary/cluster analysis, 2008–2017.

[192] L Levern Merrifield. The profile line as an aid in critically evaluating facial esthetics. *American journal of orthodontics*, 52(11):804–822, 1966.

[193] Mark Meyer, Mathieu Desbrun, Peter Schröder, Alan H Barr, et al. Discrete differential-geometry operators for triangulated 2-manifolds. *Visualization and mathematics*, 3(2):52–58, 2002.

[194] Ajmal S Mian, Mohammed Bennamoun, and Robyn Owens. Three-dimensional model-based object recognition and segmentation in cluttered scenes. *IEEE transactions on pattern analysis and machine intelligence*, 28(10):1584–1601, 2006.

[195] Steven F Miller, Seth M Weinberg, Nichole L Nidey, David K Defay, Mary L Marazita, George L Wehby, and Lina M Moreno Uribe. Exploratory genotype–phenotype correlations of facial form and asymmetry in unaffected relatives of children with non-syndromic cleft lip and/or palate. *Journal of anatomy*, 224(6):688–709, 2014.

[196] Luc Mioulet, Toby P Breckon, Andre Mouton, Haichao Liang, and Takashi Morie. Gabor features for real-time road environment classification. In *Industrial Technology (ICIT), 2013 IEEE International Conference on*, pages 1117–1121. IEEE, 2013.

[197] LC Morecroft. *A Statistical Approach to Facial Identification.* PhD thesis, University of Sheffield, 2009.

[198] Ana Belen Moreno, Angel Sánchez, José Fco Vélez, and Fco Javier Díaz. Face recognition using 3D surface-extracted descriptors. In *Irish Machine Vision and Image Processing Conference*, volume 2, pages 1–6, 2003.

[199] Ayako Mori, Tatsuo Nakajima, Tsuyoshi Kaneko, Hisashi Sakuma, and Yoshimitsu Aoki. Analysis of 109 Japanese children's lip and nose shapes using 3-dimensional digitizer. *British journal of plastic surgery*, 58(3):318–329, 2005.

[200] Iordanis Mpiperis, Sotiris Malassiotis, and Michael G Strintzis. 3-D face recognition with the geodesic polar representation. *IEEE Transactions on Information Forensics and Security*, 2(3-2):537–547, 2007.

[201] John B Mulliken, Ram Burvin, and Leslie G Farkas. Repair of bilateral complete cleft lip: intraoperative nasolabial anthropometry. *Plastic and reconstructive surgery*, 107(2):307–314, 2001.

[202] Akiko Nakamura, Masahira Hattori, and Yoshiyuki Sakaki. A novel gene isolated from human placenta located in Down syndrome critical region on chromosome 21. *DNA Research*, 4(5):321–324, 1997.

[203] Weinberg College of Arts and Sciences. Comprehensive craniofacial data and resources. www.facebase.org. 2017.

[204] Joseph O'Rourke. Computational geometry column 35. *ACM SIGACT News*, 30(2):31–32, 1999.

[205] Graziella Orru, William Pettersson-Yeo, Andre F Marquand, Giuseppe Sartori, and Andrea Mechelli. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neuroscience & Biobehavioral Reviews*, 36(4):1140–1152, 2012.

[206] Alice J O'Toole, Jennifer Peterson, and Kenneth A Deffenbacher. Another race effect for categorizing faces by sex. *Perception*, 25(6):669–676, 1996.

[207] Alice J O'Toole, Thomas Vetter, Nikolaus F Troje, and Heinrich H Bülthoff. Sex classification is better with three-dimensional head structure than with image intensity information. *Perception*, 26(1):75–84, 1997.

[208] Stephen E Palmer. Hierarchical structure in perceptual representation. *Cognitive psychology*, 9(4):441–474, 1977.

[209] Ajeet Kumar Panda, Mukesh Kumar, Mukesh Kumar Chaudhary, and A Amiya Kumar Gupta. Brain tumour extraction from mri images using kmeans clustering. *Brain*, 4(4):356—359, 2016.

[210] Sebastien Paris. Multiclass GentleAdaboosting. https://uk.mathworks.com/matlabcentral/fileexchange/22997-multiclass-gentleadaboosting. Updated 24 Nov 2011.

[211] Sang-Jun Park and Dong-Won Shin. 3D face recognition based on feature detection using active shape models. In *Control, Automation and Systems, 2008. ICCAS 2008. International Conference on*, pages 1881–1886. IEEE, 2008.

[212] Georgios Passalis, Ioannis A Kakadiaris, Theoharis Theoharis, George Toderici, and N Murtuza. Evaluation of 3d face recognition in the presence of facial expressions: an annotated deformable model approach. In *Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*, pages 171–171. IEEE, 2005.

[213] Lavinia Paternoster, Alexei I Zhurov, Arshed M Toma, John P Kemp, Beate St Pourcain, Nicholas J Timpson, George McMahon, Wendy McArdle, Susan M Ring, George Davey Smith, et al. Genome-wide association study of three-dimensional facial morphology identifies a variant in *PAX3* associated with nasion position. *The American Journal of Human Genetics*, 90(3):478–485, 2012.

[214] AJ Patil, CS Patil, RR Karhe, and MA Aher. Comparative study of different clustering algorithms. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng*, 3(7):10490–10497, 2014.

[215] Shouneng Peng, Jingze Tan, Sile Hu, Hang Zhou, Jing Guo, Li Jin, and Kun Tang. Detecting genetic association of common human facial morphological variation using high density 3D image registration. *PLoS Comput Biol*, 9(12):e1003375, 2013.

[216] Gabriel Peyre. Fast marching MATLAB toolbox. uk.mathworks.com/matlabcentral/fileexchange/6110 toolbox fast marching. Updated 27 Jun 2009.

[217] Gabriel Peyre. Toolbox graph — a toolbox to process graph and triangulated meshes. www.uk.mathworks.com/matlabcentral/fileexchange/5355-toolbox-graph/content/toolbox_graph. Updated 19 Jul 2009.

[218] Gabriel Peyré and Laurent Cohen. Geodesic computations for fast and accurate surface remeshing and parameterization. In *Elliptic and Parabolic Problems*, pages 157–171. Springer, 2005.

[219] Anika Pflug, Adrian Winterstein, and Christoph Busch. Ear detection in 3D profile images based on surface curvature. In *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2012 Eighth International Conference on*, pages 1–6. IEEE, 2012.

[220] Lu Qiao, Yajun Yang, Pengcheng Fu, Sile Hu, Hang Zhou, Jingze Tan, Yan Lu, Haiyi Lou, Dongsheng Lu, Sijie Wu, et al. Detecting genome-wide variants of Eurasian facial shape differentiation: DNA based face prediction tested in forensic scenario. *bioRxiv*, pages 50–68, 2016.

[221] TC Rae. The Neanderthal face is not cold adapted. *Journal of Human Evolution*, 60(2):234–239, 2011.

[222] Javier Ramírez, JM Górriz, Diego Salas-Gonzalez, A Romero, Míriam López, Ignacio Álvarez, and Manuel Gómez-Río. Computer-aided diagnosis of alzheimer's type dementia combining support vector machines and discriminant set of features. *Information Sciences*, 237:59–72, 2013.

[223] D Ramyachitra and P Manikandan. Imbalanced dataset classification and solutions: A review. *International Journal of Computing and Business Research*, 5(4):1–15, 2014.

[224] Stefan Rännar, Fredrik Lindgren, Paul Geladi, and Svante Wold. A PLS kernel algorithm for data sets with many variables and fewer objects. Part 1: Theory and algorithm. *Journal of Chemometrics*, 8(2):111–125, 1994.

[225] Sebastian Raschka. About feature scaling and normalization. https://sebastianraschka.com/Articles/2014_about_feature_scaling.html#about-standardization, Jul 11, 2014.

[226] Soumya Raychaudhuri, Joshua M Stuart, and Russ B Altman. Principal components analysis to summarize microarray experiments: application to sporulation time series. In *Biocomputing 2000*, pages 455–466. World Scientific, 1999.

[227] Anshuman Razdan and MyungSoo Bae. Curvature estimation scheme for triangle meshes using biquadratic Bézier patches. *Computer-Aided Design*, 37(14):1481–1491, 2005.

[228] Andrew P Read and Valerie E Newton. Waardenburg syndrome. *Journal of medical genetics*, 34(8):656–665, 1997.

[229] Eréndira Rendón, Itzel Abundez, Alejandra Arizmendi, and ElviaM Quiroz. Internal versus external cluster validation indexes. *International Journal of computers and communications*, 5(1):27–34, 2011.

[230] Eréndira Rendón, Itzel M Abundez, Citlalih Gutierrez, Sergio Díaz Zagal, Alejandra Arizmendi, Elvia M Quiroz, and H Elsa Arzate. A comparison of internal and external cluster validation indexes. In *Proceedings of the 2011 American Conference, San Francisco, CA, USA*, volume 29, pages 1–10, 2011.

[231] Robert M Ricketts. Esthetics, environment, and the law of lip relation. *American journal of orthodontics*, 54(4):272–289, 1968.

[232] Neil Risch, Kathleen Merikangas, et al. The future of genetic studies of complex human diseases. *Science*, 273(5281):1516–1517, 1996.

[233] Meredith Stacy Robinson and Mubarak Ariyo Bidmos. The skull and humerus in the determination of sex: reliability of discriminant function equations. *Forensic science international*, 186(1):86–e1, 2009.

[234] Antonio Rosas and Markus Bastir. Thin-plate spline analysis of allometry and sexual dimorphism in the human craniofacial complex. *American Journal of Physical Anthropology*, 117(3):236–245, 2002.

[235] MC Ruiz and J Illingworth. Automatic landmarking of faces in 3D-ALF^{3D}. In *Visual Information Engineering, 2008. VIE 2008. 5th International Conference on*, pages 41–46. IET, 2008.

[236] Szymon Rusinkiewicz. Estimating curvatures and their derivatives on triangle meshes. In *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on*, pages 486–493. IEEE, 2004.

[237] Takeshi Saitoh, Mitsugu Hisagi, and Ryosuke Konishi. Japanese phone recognition using lip image information. In *MVA*, pages 134–137. Citeseer, 2007.

[238] Marko Sarstedt and Erik Mooi. Cluster analysis. In *A concise guide to market research*, pages 273–324. Springer, 2014.

[239] KPNV Satyasree and J Murthy. An exhaustive literature review on class imbalance problem. *Int. J. Emerg. Trends Technol. Comput. Sci*, 2:109–118, 2013.

[240] I Savoye, Ruth Loos, Carine Carels, Cathérine Derom, and Robert Vlietinck. A genetic study of anteroposterior and vertical facial proportions using model-fitting. *The Angle Orthodontist*, 68(5):467–470, 1998.

[241] Katrin Schaefer, Bernhard Fink, Karl Grammer, Philipp Mitteroecker, Philipp Gunz, and Fred L Bookstein. Female appearance: facial and bodily attractiveness as shape. *Psychology Science*, 48(2):187–204, 2006.

[242] Katja Schwenzer-Zimmerer, J Haberstok, Laszlo Kovacs, BI Boerner, N Schwenzer, P Juergens, H-F Zeilhofer, and C Holberg. 3D surface measurement for medical application—technical comparison of two established industrial surface scanning systems. *Journal of medical systems*, 32(1):59–64, 2008.

[243] scikit. Selecting the number of clusters with silhouette analysis on kmeans clustering. scikit-learn.org. 2007–2017.

[244] Hyekyung Seo, Yeunkun Song, Choongbuem Kim, and Hyunwook Kim. Characteristics of Korean children's facial anthropometry evaluated by three-dimensional imaging. *Journal of the International Society for Respiratory Protection Vol*, 33(1), 2016.

[245] Chiarella Sforza, Fadil Elamin, Riccardo Rosati, Marco Alberto Lucchini, Davide G Tommasi, and Virgilio F Ferrario. Three-dimensional assessment of nose and lip morphology in North Sudanese subjects with Down syndrome. *The Angle Orthodontist*, 81(1):107–114, 2011.

[246] Chiarella Sforza, Gaia Grandi, Miriam Binelli, Claudia Dolci, Marcio De Menezes, and Virgilio F Ferrario. Age-and sex-related changes in three-dimensional lip morphology. *Forensic science international*, 200(1):182–194, 2010.

[247] John R Shaffer, Ekaterina Orlova, Myoung Keun Lee, Elizabeth J Leslie, Zachary D Raffensperger, Carrie L Heike, Michael L Cunningham, Jacqueline T Hecht, Chung How Kau, Nichole L Nidey, et al. Genome-wide association study reveals multiple loci influencing normal human facial morphology. *PLoS genetics*, 12(8):e1006149, 2016.

[248] MRI Shah, S Anwar, DK Mondal, S Yesmin, and S Ahmed. Anthropometry of the nose: a comparative study between adult male Santhals and Bengalis in Bangladesh. *Mediscope*, 2(2):28–32, 2015.

[249] Rutuja G Shelke and SA Annadate. Face recognition and gender classification using feature of lips. *International Journal of Innovation and Scientific Research (IJISR), Innovative Space of Scientific Research Journals (ISSR)*, 10(2):1–12, 2014.

[250] Yoshi F Shen, Karin Vargervik, Snehlata Oberoi, and Radhika Chigurupati. Facial skeletal morphology in growing children with Pierre Robin sequence. *The Cleft Palate-Craniofacial Journal*, 49(5):553–560, 2012.

[251] Jia-Yu Shi, Han Zhou, Run-Yi Mao, Yi Chen, Jing-Tao Li, and Hai-Ying Huo. A preliminary study on the key factors contributing to the attractive lips of Chinese children. *Asian Pacific journal of tropical medicine*, 5(4):318–322, 2012.

[252] Jiazheng Shi, Ashok Samal, and David Marx. How effective are landmarks and their geometry for face recognition? *Computer vision and image understanding*, 102(2):117–133, 2006.

[253] Mahdi A Shkoukani, Michael Chen, and Angela Vong. Cleft lip—a comprehensive review. *Frontiers in pediatrics*, 1:1–15, 2013.

[254] Sarah Shrimpton, Katleen Daniels, Sven De Greef, Francoise Tilotta, Guy Willems, Dirk Vandermeulen, Paul Suetens, and Peter Claes. A spatially-dense regression study of facial form and tissue depth: towards an interactive tool for craniofacial reconstruction. *Forensic science international*, 234:103–110, 2014.

[255] Alan Slater and Paul C Quinn. Face recognition in the newborn infant. *Infant and Child Development*, 10(1-2):21–24, 2001.

[256] Jonatan Snyders, Peter Claes, Dirk Vandermeulen, and Paul Suetens. Non-rigid surface registration algorithms: Technical details and comparison. Technical Report KUL/ESAT/PSI/1404, KU Leuven, 2014.

[257] Beni Solow and Antje Tallgren. Natural head position in standing subjects. *Acta Odontologica Scandinavica*, 29(5):591–607, 1971.

[258] Mingli Song, Dacheng Tao, Shengpeng Sun, Chun Chen, and Stephen J Maybank. Robust 3D face landmark localization based on local coordinate coding. *IEEE Transactions on Image Processing*, 23(12):5108–5122, 2014.

[259] Olga Sorkine, Daniel Cohen-Or, Yaron Lipman, Marc Alexa, Christian Rössl, and H-P Seidel. Laplacian surface editing. In *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 175–184. ACM, 2004.

[260] Darryl Stewart, Adrian Pass, and Jianguo Zhang. Gender classification via lips: static and dynamic features. *IET biometrics*, 2(1):28–34, 2013.

[261] David W. Stockburger. Multiple regression with categorical variables. http://www.psychstat.missouristate.edu/multibook/mlt08m.html, 2016.

[262] Xiang Sun and Jean-Marie Morvan. Curvature measures, normal cycles and asymptotic cones. *Actes des rencontres du CIRM*, 3(1):3–10, 2013.

[263] Yanmin Sun, Mohamed S Kamel, and Yang Wang. Boosting for learning multiple classes with imbalanced class distribution. In *Sixth International Conference on Data Mining (ICDM'06)*, pages 592–602. IEEE, 2006.

[264] Vitaly Surazhsky, Tatiana Surazhsky, Danil Kirsanov, Steven J Gortler, and Hugues Hoppe. Fast exact and approximate geodesics on meshes. *ACM transactions on graphics (TOG)*, 24(3):553–560, 2005.

[265] Michael Suttie, Tatiana Foroud, Leah Wetherill, Joseph L Jacobson, Christopher D Molteno, Ernesta M Meintjes, H Eugene Hoyme, Nathaniel Khaole, Luther K Robinson, Edward P Riley, et al. Facial dysmorphism across the fetal alcohol spectrum. *Pediatrics*, 131(3):779–788, 2013.

[266] SyndromesPedia. Pierre Robin syndrome — symptoms, causes, treatment and complications. syndromespedia.com/pierre-robin-syndrome-treatment-pictures-symptoms-prognosis.html. 2011–2017.

[267] SyndromesPedia. Waardenburg syndrome — pictures, symptoms, diagnosis, treatment. syndromespedia.com/waardenburg-syndrome-pictures-symptoms-diagnosis-treatment.html. 2011–2017.

[268] 3D SYSTEM. Rapidform 3d scanning software recognized as a world class product by korea's ministry of knowledge economy. www.uk.3dsystems.com/press-releases/rapidform-3d-scanning-software-recognized-as-a-world-class-product-by-koreas-ministry-of-knowledge-economy. 2017.

[269] Leslie Talbert, Chung How Kau, Terpsithea Christou, Christos Vlachos, and Nada Souccar. A 3D analysis of Caucasian and African American facial morphologies in a US population. *Journal of orthodontics*, 41(1):19–29, 2014.

[270] Yinhang Tang, Huibin Li, Xiang Sun, Jean-Marie Morvan, and Liming Chen. Principal curvature measures estimation and application to 3D face recognition. *Journal of Mathematical Imaging and Vision*, pages 1–23, 2017.

[271] C Tanikawa, E Zere, and K Takada. Sexual dimorphism in the facial morphology of adult humans: a three-dimensional analysis. *HOMO-Journal of Comparative Human Biology*, 67(1):23–49, 2016.

[272] Gabriel Taubin. Estimating the tensor of curvature of a surface from a polyhedral approximation. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pages 902–907. IEEE, 1995.

[273] Rebecca Valerie Taylor. *Analysis of three-dimensional craniofacial images: applications in forensic science, anthropology and clinical medicine.* PhD thesis, University of Melbourne, 2008.

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

[275] Markus Tiirma. Comparison between three different clustering algorithms. *Photogrammetric Journal of Finland*, 13(2):0–2, 1993.

[276] Arshed M Toma. *Characterization of normal facial features and their association with genes.* Phd thesis, Cardiff University, 2014.

[277] Arshed M Toma, A Zhurov, R Playle, Egle Ong, and Stephen Richmond. Reproducibility of facial soft tissue landmarks on 3D laser-scanned facial images. *Orthodontics & craniofacial research*, 12(1):33–42, 2009.

[278] Arshed M Toma, Alexei I Zhurov, Rebecca Playle, David Marshall, Paul L Rosin, and Stephen Richmond. The assessment of facial variation in 4747 British school children. *The European Journal of Orthodontics*, 34(6):655–664, 2012.

[279] Federico Tombari, Samuele Salti, and Luigi Di Stefano. Performance evaluation of 3d keypoint detectors. *International Journal of Computer Vision*, 102(1-3):198–220, 2013.

[280] Yoko Tomoyasu, Tetsutaro Yamaguchi, Atsushi Tajima, Toshiaki Nakajima, Ituro Inoue, and Koutaro Maki. Further evidence for an association between

mandibular height and the growth hormone receptor gene in a japanese population. *American Journal of Orthodontics and Dentofacial Orthopedics*, 136(4):536–541, 2009.

[281] Dimosthenis Tsagkrasoulis, Pirro Hysi, Tim Spector, and Giovanni Montana. Heritability maps of human face morphology through large-scale automated three-dimensional phenotyping. *arXiv preprint arXiv:1608.08199*, 2016.

[282] Dimosthenis Tsagkrasoulis, Pirro Hysi, Tim Spector, and Giovanni Montana. Heritability maps of human face morphology through large-scale automated three-dimensional phenotyping. *Scientific Reports*, 7:10–30, 2017.

[283] Veli Vural Uslu, Massimo Petretich, Sandra Ruf, Katja Langenfeld, Nuno A Fonseca, John C Marioni, and François Spitz. Long-range enhancers regulating Myc expression are required for normal facial morphogenesis. *Nature genetics*, 46(7):753–758, 2014.

[284] Paul van den Akker. Petrus Camper on natural design and the beauty of Apollo's profile. *Petrus Camper in context: Science, the arts, and society in the eighteenth-century Dutch Republic*, pages 243–255, 2015.

[285] Florine Vegter and J Joris Hage. Clinical anthropometry and canons of the face in historical perspective. *Plastic and reconstructive surgery*, 106(5):1090–1096, 2000.

[286] Enrico Vezzetti and Federica Marcolin. Geometry-based 3D face morphology analysis: soft-tissue landmark formalization. *Multimedia tools and applications*, 68(3):895–929, 2014.

[287] Enrico Vezzetti, Federica Marcolin, and Giulia Fracastoro. 3D face recognition: An automatic strategy based on geometrical descriptors and landmarks. *Robotics and Autonomous Systems*, 62(12):1768–1776, 2014.

[288] Jing Wang, Yuanzhi Xu, Jing Chen, Feiyu Wang, Renhuan Huang, Songtao Wu, Linjing Shu, Jingyi Qiu, Zhi Yang, Junjie Xue, et al. *PAX9* polymorphism and susceptibility to sporadic non-syndromic severe anodontia: a case-control study in southwest China. *Journal of Applied Oral Science*, 21(3):256–264, 2013.

[289] Kathie H Wang, Carrie L Heike, Melissa D Clarkson, Jose LV Mejino, James F Brinkley, Raymond W Tse, Craig B Birgfeld, David A Fitzsimons, and Timothy C Cox. Evaluation and integration of disparate classification systems for clefts of the lip. *Frontiers in physiology*, 5:163–175, 2014.

[290] Yiding Wang, Meng Meng, and Qingkai Zhen. Learning encoded facial curvature information for 3D facial emotion recognition. In *Image and Graphics (ICIG), 2013 Seventh International Conference on*, pages 529–532. IEEE, 2013.

[291] Yongmei Wang, Bradley S Peterson, and Lawrence H Staib. Shape-based 3D surface correspondence using geodesics and local geometry. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 2, pages 644–651. IEEE, 2000.

[292] Seth M Weinberg, Trish E Parsons, Mary L Marazita, and Brion S Maher. Heritability of face shape in twins: a preliminary study using 3D stereophotogrammetry and geometric morphometrics. *Dentistry 3000*, 1(1), 2013.

[293] Jason Weston and Chris Watkins. Multi-class support vector machines. Technical report, Technical Report CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, May, 1998.

[294] Katarzyna Wilamowska, Jia Wu, Carrie Heike, and Linda Shapiro. Shape-based classification of 3D facial data to support 22q11.2DS craniofacial research. *Journal of digital imaging*, 25(3):400–408, 2012.

[295] Caryl Wilson, Rebecca Playle, Arshed Toma, Alexei Zhurov, Andy Ness, and Stephen Richmond. The prevalence of lip vermilion morphological traits in a 15-year-old population. *American Journal of Medical Genetics Part A*, 161(1):4–12, 2013.

[296] Caryl Wilson-Nagrani. *Matching genotype to phenotype in a detailed assessment of lip morphology*. PhD thesis, Cardiff University, 2016.

[297] Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52, 1987.

[298] Jing Wu, William AP Smith, and Edwin R Hancock. Facial gender classification using shape-from-shading. *Image and Vision Computing*, 28(6):1039–1048, 2010.

[299] Shi-Qing Xin and Guo-Jin Wang. Improving Chen and Han's algorithm on the discrete geodesic problem. *ACM Transactions on Graphics (TOG)*, 28(4):104–112, 2009.

[300] Chunxu Xu, Tuanfeng Y Wang, Yong-Jin Liu, Ligang Liu, and Ying He. Fast wavefront propagation (fwp) for computing exact geodesic distances on meshes. *IEEE transactions on visualization and computer graphics*, 21(7):822–834, 2015.

[301] Shicai Yang, George Bebis, Muhammad Hussain, Ghulam Muhammad, and Anwar M Mirza. Unsupervised discovery of visual face categories. *International Journal on Artificial Intelligence Tools*, 22(01):1–30, 2013.

[302] Ching-Chiang Yeh, Fengyi Lin, and Chih-Yu Hsu. A hybrid kmv model, random forests and rough set theory approach for credit rating. *Knowledge-Based Systems*, 33:166–172, 2012.

[303] Ka Yee Yeung and Walter L. Ruzzo. Principal component analysis for clustering gene expression data. *Bioinformatics*, 17(9):763–774, 2001.

[304] Liu Yilin. *3D facial model analysis for clinical medicine*. PhD thesis, National University of Singapore, 2013.

[305] Xiang Ying, Shi-Qing Xin, and Ying He. Parallel chen-han (pch) algorithm for discrete geodesics. *ACM Transactions on Graphics (TOG)*, 33(1):9–15, 2014.

[306] Andrew W Young and Hadyn D Ellis. *Handbook of research on face processing: The face and social attribution*. Elsevier, 2013.

[307] Xiaohui Yuan and Mohamed Abouelenien. A multi-class boosting method for learning from imbalanced data. *International Journal of Granular Computing, Rough Sets and Intelligent Systems*, 4(1):13–29, 2015.

[308] Georgios Vasileiou Zacharopoulos, Andreas Manios, Eelco De Bree, Chung How Kau, Markos Petousis, Ioanna Zacharopoulou, and Nikolaos Kouremenos. Neoclassical facial canons in young adults. *Journal of Craniofacial Surgery*, 23(6):1693–1698, 2012.

[309] Arslan A Zaidi, Brooke C Mattern, Peter Claes, Brian McEcoy, Cris Hughes, and Mark D Shriver. Investigating the case of human nose shape and climate adaptation. *PLoS genetics*, 13(3):e1006616, 2017.

[310] Miriam Leah Zelditch, Donald L Swiderski, and H David Sheets. *Geometric morphometrics for biologists: a primer*. Academic Press, 2012.

[311] Junli Zhao, Cuiting Liu, Zhongke Wu, Fuqing Duan, Minqi Zhang, Kang Wang, and Taorui Jia. 3D facial similarity measure based on geodesic network and curvatures. *Mathematical Problems in Engineering*, 2014:1–18, 2014.

[312] Wenyi Zhao, Rama Chellappa, P Jonathon Phillips, and Azriel Rosenfeld. Face recognition: A literature survey. *ACM computing surveys (CSUR)*, 35(4):399–458, 2003.

[313] Ji Zhu, Hui Zou, Saharon Rosset, Trevor Hastie, et al. Multi-class adaboost. *Statistics and its Interface*, 2(3):349–360, 2009.

[314] Wen Zhu, Nancy Zeng, Ning Wang, et al. Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. *NESUG proceedings: health care and life sciences, Baltimore, Maryland*, pages 1–9, 2010.