



McNeil, Kathryn Helen (2012) Analysis of three-dimensional facial shape. MSc(R) thesis

<http://theses.gla.ac.uk/3519/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.



Analysis of Three-Dimensional Facial Shape

Kathryn Helen McNeil

*A Dissertation Submitted to the
University of Glasgow
for the degree of
Master of Science in Statistics*

School of Mathematics & Statistics

October 2011

© Kathryn Helen McNeil, October 2011

Abstract

Funding from the Wellcome Trust has been awarded to a consortium of institutions including the University of Glasgow to carry out research into the analysis of three-dimensional facial dysmorphology. The *Face3D Project* is the name given to this study and the aim is to develop methods of extracting information about the facial shape from three-dimensional images. There are numerous medical advantages in being able to accurately map the shape of a face in three-dimensions. This thesis is structured around the control facial shapes collected for the purpose of two medical applications; firstly the success of orthognathic surgery and secondly, the characterisation of the biological processes which underlie schizophrenia. This thesis covers the initial capture of a control data set of facial images and investigates methods of identifying and analysing information from the three-dimensional control facial shapes.

In Chapter 2 the theory of various statistical methods used in shape analysis is described in detail.

In Chapter 3 the stages of data collection are detailed including the recruitment process, exclusion criteria and image capture and data collection process itself.

Chapter 4 details a subsequent validation study where the accuracy of image capture and the accuracy of anatomical landmark allocation is investigated.

The design of the analysis takes the form of a four-layered hierarchical model where the source of variation in the model is calculated across each level. A measure of accuracy for individual landmark identification is found and an overall measure of accuracy in identifying a set of anatomical landmarks was found to be 1.81 ± 0.84 mm on average.

In Chapter 5 the landmark configurations of all control faces are analysed and evidence of sexual dimorphism based on the shape of landmark coordinates after adjustment for scale is found. Males were found to have larger faces than females on average. The measure of participant's body fat percentage on the position of anatomical landmarks is found to have a significant effect. In analysis of the landmarks on the midsagittal profile evidence of sexual dimorphism is found between the male and female landmark configurations after adjustment for scale. Subsequent Principal Component Analysis on the landmark configurations of males and females show little evidence of differences between males and females, despite formal analyses indicating evidence of sexual dimorphism

In Chapter 6 a method of curve identification is detailed for the philtrum ridges on the upper lip. Evidence of sexual dimorphism for the curves on the upper lip ridges was not identified. A general method of curve identification is then established. Curve identification of the midsagittal profile is described in detail. A formal test indicates that evidence sexual dimorphism is present in the shape of the midsagittal curves identified, however the result is borderline significant. After adjustment for body fat no difference is found between the mean shape of the midsagittal curves for males and females.

Acknowledgements

Firstly, I would like to thank Prof. Adrian Bowman for his invaluable guidance and support throughout the past year. I would also like to thank Dr. Stanislav Katina for the time he has given to explain concepts with unlimited patience.

I am very grateful to the Wellcome Trust and the School of Mathematics and Statistics for funding my research and for providing me with the opportunity. For making this year so enjoyable and for providing distractions when they were needed most, I thank my fellow post-grads.

For keeping me sane and for giving me the best goals to work towards, I thank David. I owe a big thank to my family for their encouragement and endless support throughout my education and this year of research. Finally, thank you to my nephew Lewis who made the decision to return to Scotland very easy.

Contents

1	Introduction	1
1.1	Motivation and Background	1
1.2	Data	2
1.3	Landmarks	3
2	Theoretical Considerations	6
2.1	Similarity Transformations & Shape Space	6
2.2	Full Ordinary Procrustes Analysis	9
2.3	Full Generalized Procrustes Analysis	11
2.4	Estimating Mean Shape	12
2.4.1	Coordinates in Tangent Space	12
2.4.2	Two-Sample Hotelling's T^2 Test	13
2.5	Methodology for Landmark Analysis	14
2.5.1	Permutation test	14
2.5.2	Principal Component Analysis	15
2.5.3	Landmark Distances	16
3	Data Collection	18
3.1	Exclusion Criteria	19
3.2	Questionnaire	20
3.3	Clinical Measurements	22
3.4	Image Capture Protocol	24
3.4.1	Loss of detail in 3D image	26

3.5	Participants	27
3.6	Problems with the Data	28
3.7	Problems with Landmarking	29
4	Validation Study	30
4.1	Motivation	30
4.2	Design of Validation Study	30
4.3	Analysis of Validation Study	34
4.3.1	Single Observer Consistency	34
4.3.2	A General Model For Reproducibility	38
4.4	Results of Validation Study	42
5	Landmark Analysis of Control Data	49
5.1	Sexual Dimorphism	50
5.2	Sexual Dimorphism in Control Data	51
5.3	Body Fat	60
5.4	Landmark Analysis of Midsagittal Profile	62
5.5	Landmark Analysis Conclusions	67
6	Curve Identification in Control Data	69
6.1	Methodology for Curve Identification	69
6.1.1	Principal Curves on Facial Surface	70
6.2	Curve Identification of Philtrum	71
6.2.1	Analysis of Philtrum Curves	75
6.3	A General Method of Facial Curve Identification	79
6.4	Curve Identification of Midsagittal Profile	80
6.4.1	Analysis of Midsagittal Curves	81
6.5	Curve Identification - Problems Encountered	88

6.6	Curve Identification Conclusions	89
7	Discussions & Conclusions	91
7.1	Summary	91
7.2	Limitations in the Data	92
7.3	Limitations in the Curve Identification	93
7.3.1	Future Analysis	94
A		96
A.1	Information Sheet	97
A.2	Data Capture Protocol	103
A.3	Permission Form	106

List of Tables

1.1	Anatomical Landmarks for Control Participants	5
4.1	Anatomical Landmarks Allocated for Validation Study	32
4.2	TEM Scores of Repeat Landmarks on 20 Images in Validation Study	37
4.3	Average Standard Deviation over all landmarks (mm).	42
5.1	Facial distances without adjustment for scale (mm).	56
5.2	Facial distances with adjustment for scale (mm).	57
5.3	Comparing landmark positions between Males and Females. . .	59

List of Figures

1.1	Visualisation tool for landmarking.	3
3.1	Electronic sensors of ©Bodystat machine.	22
3.2	©DI3D Camera Equipment	24
3.3	Stills from the bottom left and right cameras.	26
3.4	Distorted ears due to sideburns.	27
3.5	Age and % Body Fat for Males and Females.	28
4.1	Hierarchical Model of Validation Study Design	33
4.2	View of a singular configuration rotated to ‘face’ directly for- wards. Landmark number is shown as indicator of landmark position.	35
4.3	Frontal view of two repeat landmark configurations.	36
4.4	Frontal view of two repeat landmark configurations with the average landmark position superimposed.	39
4.5	Frontal view of average configurations <i>at Image level</i> , coloured by participant.	40
4.6	Standard Deviation values for all landmarks on the <i>x</i> -axis. . .	45
4.7	Standard Deviation values for all landmarks on the <i>y</i> -axis. . .	46
4.8	Standard Deviation values for all landmarks on the <i>z</i> -axis. . .	47
4.9	Face with spheres representing level of variation in identifying each landmark in the validation study.	48
5.1	Size of configurations for males and females.	51

5.2	Frontal view of all control data landmarks and the mean positions for males and females without adjustment for scale. . .	52
5.3	Frontal view of all control data landmarks and the mean positions for males and females with adjustment for scale.	53
5.4	Cumulative proportion of variability explained by Principal Components.	55
5.5	Boxplots of 1st and 2nd Principal Components scores split by Sex.	58
5.6	Plot of the first against the second Principal Component scores, split by Sex.	58
5.7	Plots of first and second Principal Components scores against % Body Fat, split by Sex.	60
5.8	Female % Body Fat plotted against Weight (kg).	61
5.9	Side view of shape-and-size of control data midsagittal landmarks and the mean midsagittal positions for males and females.	63
5.10	Side view of shape of control data midsagittal landmarks and the mean midsagittal positions for males and females with adjustment for scale.	63
5.11	Cumulative proportion of variability explained by Principal Components.	64
5.12	Plots of first and second Principal Components scores for midsagittal landmarks against % Body Fat, split by Sex.	65
5.13	Boxplots of first and second Principal Components scores for midsagittal landmarks split by Sex.	66
5.14	Plot of the first against the second Principal Component scores for midsagittal landmarks, split by Sex.	66
6.1	Horizontal Principal Curves on upper lip from <i>subnasale</i> to the lowest <i>christa philtri</i> in the <i>y</i> dimension.	72
6.2	Curvature versus arc length (mm) of a horizontal Principal Curve on the upper lip.	73

6.3	Maximum points of curvature on horizontal Principal Curves on upper lip.	74
6.4	Vertical Principal Curves on upper lip calculated from maximum points of curvature.	74
6.5	All registered configurations and the mean configurations for males and females with no adjustment for scale.	76
6.6	All registered configurations and the mean configurations for males and females after adjustment for scale.	76
6.7	Cumulative proportion of variability in philtrum ridge curves explained by Principal Components.	77
6.8	Boxplots of first and second Principal Components scores for philtrum ridge curves split by Sex.	78
6.9	First against second Principal Component, split by Sex.	79
6.10	Midsagittal strip of subset data and the Principal Curve fitted to this data.	81
6.11	Landmarks and pseudo landmarks on midsagittal Principal Curve.	82
6.12	Principal Curves of midsagittal landmarks split by Sex.	83
6.13	Principal Curves of midsagittal landmarks with adjustment for scale split by Sex.	83
6.14	Cumulative proportion of variability in midsagittal curves explained by Principal Components.	84
6.15	Midsagittal curves with adjustment for % Body Fat, split by Sex.	84
6.16	Boxplots of first and second Principal Components scores for midsagittal Principal Curves, split by Sex.	85
6.17	Plot of the first against the second Principal Component scores for midsagittal Principal Curves, split by Sex.	86
6.18	Mean midsagittal curves with lower and upper extremes of the first and second Principal Components.	87

6.19 Example of philtrum ridges converging on upper lip. 89

Chapter 1

Introduction

1.1 Motivation and Background

A person's face is their first impression presentation to the world and it can tell you a lot about them; their gender, their age and often their ethnic background. It is however, often difficult to pin-point exactly what it is about a person's facial structure that tells you this information. The ability to accurately map the surface of a three-dimensional facial shape has numerous advantageous applications, especially in the medical field. The *Face3D Project*, funded by the Wellcome Trust is a consortium of several institutions including the University of Glasgow (Statistics, Computing Science, Dental School), the Royal College of Surgeons in Ireland, Dublin City University, the University of Limerick and the Institute of Technology in Tralee, which aims to pursue research into developing methods of extracting information about facial shape from three-dimensional images. The statistical and computational methods developed will be applied by the Face3D project to two applications areas in the characterisation of the biological processes which underlie schizophrenia and in the quantitative assessment of the outcome of orthognathic surgery - (surgery of the jaw).

One of the first stages that facilitates the analysis of the two medical applica-

tions is the construction of a database to characterise control facial shape. A new set of control facial images was required in which participants satisfied the control criteria for both research areas of interest.

This thesis can be split into three sections. The first section describes the process of data collection including a validation study which explores the sources of variance in identifying specific features of a facial image, by placing anatomical landmarks. The second section focuses on analysing landmarks placed on all control faces, specifically investigating evidence of sexual dimorphism. The midsagittal profile and upper lip are focused on. In the final section of this thesis, a method of extracting curves from the facial images is explored and the curves on the midsagittal profile and upper lip are analysed.

1.2 Data

The process of data collection makes up a substantial part of this thesis. In addition to facial images being captured, clinical and demographic information was obtained so that a participant's eligibility for the control data set could be ascertained. It is not unreasonable to anticipate that facial shape may be affected by an individual's body composition therefore participants' body fat was also measured as part of the data collection process. Data has been collected from a total of 59 participants, 19 males and 40 females. Of these 59, 51 people fitted the criteria for the control data set; 16 males and 35 females.

If we are to think mathematically, the human face is after all just a three-dimensional shape. This three-dimensional shape can be translated from the three-dimensional image captured into a large data set of coordinates. It is these coordinates which form the starting point of all statistical shape analysis performed in this thesis. All 59 of the facial images captured have

been allocated 23 anatomical landmarks to produce 59 (3×23) configurations, with 51 of these making up the control data set of landmark configurations.

1.3 Landmarks

Points that are allocated to specific locations on a surface are called landmarks. A widely recognised set of anatomical landmarks for a human face is detailed in Farkas (1994). These anatomical landmarks each pin-point a significant location on the surface of the face, (e.g. the corner of the eye). The 23 landmarks that have been allocated to all faces in the control data set are listed in Table 1.1. Each facial image captured from the camera system is loaded into a software package called ©Landmark which has been developed by the IDAV (*Institute for Data Analysis and Visualization*) which allows the facial image to be visualised in three dimensions, (Wiley (2007)). This virtual three-dimensional image can be spun in every direction and zoomed

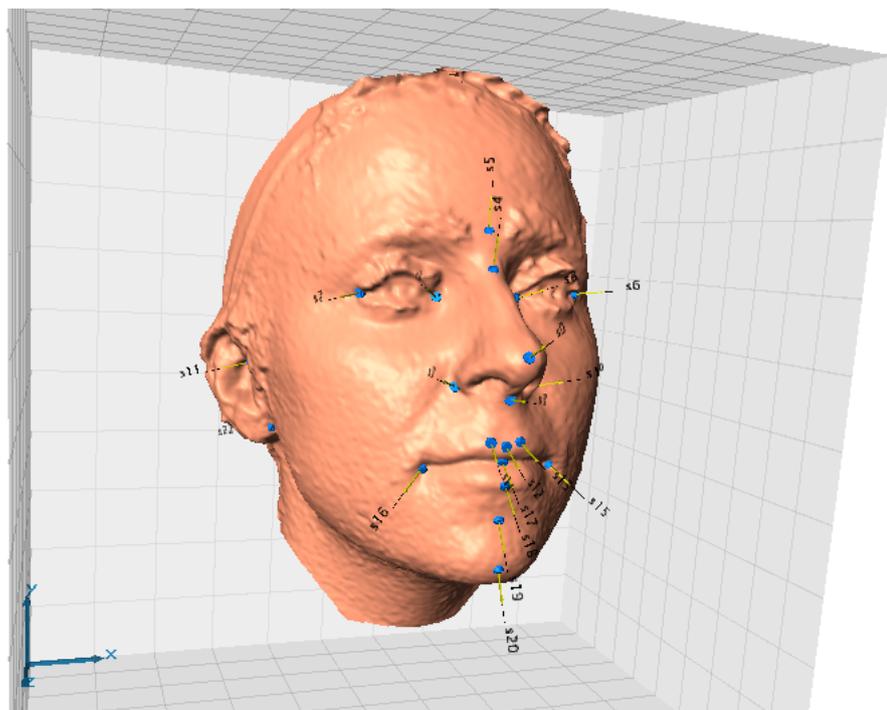


Figure 1.1: Visualisation tool for landmarking.

in on so that the facial shape can be viewed from every angle. When identifying anatomical landmarks, the ability to view the facial shape in this way is highly advantageous especially when identifying points of maximum curvature, such as the *pronasale* landmark on the nose.

There were problems encountered in identifying certain landmarks which will be discussed in detail in Section 3.7. Due to these issues, from the anatomical landmarks listed in Farkas (1994), a subset of only 23 landmarks are identified on the facial images for the control data. The three-dimensional coordinates of each landmark configuration on a face shape translates to a (3×23) matrix of data. It is this data that will be the backbone for the analysis in this thesis.

Table 1.1: Anatomical Landmarks for Control Participants

ID	Code	Region	Side	Abr.	Name
1	S0	nose	middle	pm	<i>pronasale</i>
2-3	S1-S2	nose	left/right	ac	<i>alare crest</i>
4	S3	nose	middle	sn	<i>subnasale</i>
5	S4	nose	middle	se	<i>sellion</i>
6	S5	nose	middle	n	<i>soft tissue nasion</i>
7-8	SS6-S7	eye	left/right	ex	<i>exocanthion</i>
9-10	S8-S9	eye	left/right	en	<i>endocanthion</i>
11-12	S10-S11	ear	left/right	t	<i>tragion</i>
13	S12	lips	middle	ls	<i>labiale superius</i>
14-15	S13-S14	lips	left/right	cph	<i>crista philtri</i>
16-17	S15-S16	lips	left/right	ch	<i>cheilion</i>
18	S17	lips	middle	sto	<i>stomion</i>
19	S18	lips	middle	li	<i>labiale inferius</i>
20	S19	chin	middle	sl	<i>sublabiale</i>
21	S20	chin	middle	pg	<i>soft tissue gnathion</i>
22-23	S21-S22	ear	left/right	a	<i>otobasion inferius</i>

Chapter 2

Theoretical Considerations

The shape analysis that is covered in this thesis will primarily be based on configurations of landmark coordinates, unique to each facial image. The landmark coordinates for each person do not share a common origin and are therefore independent of and non-comparable to each other. Procrustes methods can be used to standardise three-dimensional configurations such that they are centered at the origin, rotated and resized. These methods are detailed at length in Dryden and Mardia (1998) and the same notation in the explanation of the various Procrustes methods will be used in this chapter.

2.1 Similarity Transformations & Shape Space

A *configuration matrix* X , is the $(k \times m)$ matrix of Cartesian coordinates of k landmarks in m dimensions. The *configuration space* is the space in which all possible $(k \times m)$ landmark coordinates of a particular object lie. (For the purpose of facial shape analysis, $m = 3$.)

The definition of shape in Dryden and Mardia (1998) is given as “*all the geometrical information that remains when location, scale and rotational effects are filtered out from an object*”. The act of removing these location, rotation

and scale effects is referred to as removing the similarity transformations. It is often convenient that these similarity transformations are removed one at a time. The *original configurations* in this instance refers to the configurations of raw landmark coordinates before any similarity transformations have been removed.

The *centered configurations* describes the configurations once location has been removed and all configurations share the same shape origin. The Helmert sub-matrix H is the $(k - 1) \times k$ Helmert matrix without the first row, and can be used to remove the location of a configuration X . The full Helmert matrix H^F is a $(k \times k)$ orthogonal matrix with all elements in its first row equal to $\frac{1}{\sqrt{k}}$ and all remaining rows orthogonal to the first row such that

$$H^F = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{bmatrix}$$

and the Helbert sub-matrix is

$$H = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{bmatrix}$$

when $k = 3$.

Location is removed by centering all configurations to the origin, (i.e. a common pole), such that

$$X_C = CX \tag{2.1}$$

where X_C is a centered configuration and C is the transformation matrix such that

$$C = H^T H = I_k - \frac{1}{k} \mathbf{1}_k \mathbf{1}_k^T \quad \text{and so} \quad CX = H^T HX = H^T X_H \tag{2.2}$$

where k is the number of landmarks, H is the Helmert sub-matrix and X_H is the Helmertised coordinates such that $X_H = HX \in \mathbb{R}^{(k-1)m}$.

The centroid size $S(X)$ is the square root of the sum of squared Euclidean distances from each landmark in configuration X to the centroid such that

$$S(X) = \sqrt{\sum_{j=1}^k \|(X)_j - \bar{X}\|^2} \quad (2.3)$$

where $(X)_j$ is the j th row of X ($j = 1, \dots, k$) and $(\bar{X}) = (\bar{X}_1, \dots, \bar{X}_m)$ is the centroid. The *pre-shape* of a configuration X is the term given when all information about location and scale has been removed. Size can be standardised by rescaling to unit centroid size which is given by

$$S(X) = \|CX\| = \sqrt{\text{trace}(X^T CX)} = \sqrt{\text{trace}(X^T H^T H X)} = \|X_H\|, \quad (2.4)$$

since $H^T H = C$ is idempotent.

The *Size-and-shape* configuration is the landmark configurations once location and rotation have been removed. The size-and-shape configuration $[X]_S$ can be represented as

$$[X]_S = \{X_H \Gamma : \Gamma \in SO(m)\} \quad (2.5)$$

where X_H are the Helmertised coordinates and Γ is the $(m \times m)$ rotation matrix in $SO(m)$, the special orthogonal group of rotations. Finally, if size is removed from the size-and-shape of a configuration X by re-scaling to unit centroid size, then the resulting configuration describes the shape of X . The shape can be described by the set $[X]$ given by

$$[X] = \{Z\Gamma : \Gamma \in SO(m)\} \quad (2.6)$$

where Z is the pre-shape of X . In Dryden and Mardia (1998), the shape of configuration matrix X is described as “*all the geometrical information about*

X that is invariant under location, rotation and isotropic scaling (Euclidean distance transformations)”.

Procrustes analysis is an extremely useful tool for analysing landmark data. It is a method of standardising by matching the location, scale and rotation of multiple configurations so as they can be relatable for analysis.

2.2 Full Ordinary Procrustes Analysis

Full Ordinary Procrustes Analysis (*full OPA*) is used in a situation where we wish to compare the shapes of two independent configurations, say, of landmark coordinates. In order for these configurations to be comparable each of the landmarks in the first configuration is lined up, as closely as possible, to the corresponding landmark in the second configuration without losing information about the overall shape of either configuration. This is achieved by applying similarity transformations described in Section 2.1 such that both configurations are as close as possible according to Euclidean distance, using least squares techniques.

More formally, consider that it is of interest to compare two shapes which are defined by the configuration matrices X_1 and X_2 . Both are $(k \times m)$ coordinate matrices with k points and m dimensions. Full OPA applies least squares methods to “match up” configurations X_1 and X_2 using the similarity transformations such that the Euclidean distances between X_1 and X_2 are minimised. This is achieved by rotating, scaling and shifting configuration X_1 so that it is as similar as possible to configuration X_2 . The Euclidean distance used can be expressed as

$$D_{OPA}^2(X_1, X_2) = \|X_2 - \beta X_1 \Gamma - \mathbf{1}_k \gamma^T\|^2 \quad (2.7)$$

where $\|X\| = \sqrt{\text{trace}(X^T X)}$ is the Euclidean norm, β is a scale parameter ($\beta > 0$), Γ is an $(m \times m)$ rotation matrix ($\Gamma \in SO(m)$) and γ is an $(m \times 1)$

location matrix. The minimum of Equation 2.7 is the Ordinary (Procrustes) Sum of Squares, written as $OSS(X_1, X_2)$ such that

$$OSS(X_1, X_2) = \|X_1\|^2 \sin^2 \rho(X_1, X_2) \quad (2.8)$$

where $\rho(X_1, X_2)$ is the Procrustes distance between two points on the pre-shape surface.

It is worth noting that if the roles of X_1 and X_2 were reversed such that X_2 was superimposed onto X_1 , the similarity transformations would not be the same. If the estimates for the reversed case are written as $(\hat{\gamma}^R, \hat{\beta}^R, \hat{\Gamma}^R)$ it can be seen that the location matrix $\hat{\gamma}^R$ is the negative equivalent of $\hat{\gamma}$ ($\hat{\gamma}^R = -\hat{\gamma}$), the rotation matrix $\hat{\Gamma}^R$ is the transposed version of $\hat{\Gamma}$ ($\hat{\Gamma}^R = (\hat{\Gamma})^T$) but that in general the scale parameter $\hat{\beta}^R$ is not equal to the inverse value of $\hat{\beta}$, ($\hat{\beta}^R \neq 1/\hat{\beta}$). In particular

$$OSS(X_2, X_1) \neq OSS(X_1, X_2)$$

with the exception when both configurations have the same size.

The *full Procrustes coordinates* of X_1 once it has been matched onto X_2 can be written as

$$X_1^P = \hat{\beta} X_1 \hat{\Gamma} + \mathbf{1}_k \hat{\gamma}^T, \quad (2.9)$$

where the superscript P denotes the Procrustes superimposition of X_1 onto X_2 . The *residual matrix* after Procrustes matching can be defined as

$$R = X_2 - X_1^P. \quad (2.10)$$

The residual matrix can give a good indication of the differences in shape between the two configurations. If there are large residuals limited to one region of the shape, say for example, the nasal area, then this can indicate that there is a larger shape difference between the configurations at this area.

2.3 Full Generalized Procrustes Analysis

Generalized Procrustes Analysis, (*GPA*), is similar to full OPA but for the case where we wish to analyse more than two configurations. We may wish to estimate the shape of a population mean μ , with the average shape of a random sample from a population with mean μ . Full GPA achieves the optimal position for ($n \geq 2$) configurations by translating, rotating and rescaling each and so superimposing them all onto one another and minimising the sum of squared Euclidean distances.

Consider the case where there are multiple, $n \geq 2$, configurations of landmarks available X_1, \dots, X_n . Assume that these configurations have been sampled randomly from a population with mean μ . Full GPA can be used to minimise the total sum of squares

$$G(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \|(\beta_i X_i \Gamma_i + \mathbf{1}_k \gamma_i^T) - (\beta_j X_j \Gamma_j + \mathbf{1}_k \gamma_j^T)\|^2 \quad (2.11)$$

by translating, rescaling and rotating the configurations X_1, \dots, X_n relative to each other. This is done while adopting the constraint

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n (\beta_i X_i \Gamma_i + \mathbf{1}_k \gamma_i^T) = 1 \quad (2.12)$$

so that the size of the configuration average is equal to 1.

The resulting *full Procrustes coordinates* for each X_i can be defined by the minimising parameters such that

$$X_i^P = \hat{\beta}_i X_i \hat{\Gamma}_i + \mathbf{1}_k \hat{\gamma}_i^T, \quad i = 1, \dots, n \quad (2.13)$$

where $\hat{\beta}_i > 0$ is the scale parameter, $\hat{\Gamma}_i \in SO(m)$ is the rotation matrix in the special orthogonal group of rotations and $\hat{\gamma}_i^T$ denotes the location parameters.

2.4 Estimating Mean Shape

In the situation where there is a random sample of configurations available w_1, \dots, w_n it is useful to be able to obtain an estimate of an average shape. There are two ways in which to calculate the mean shape of multiple configurations. The simplest method of calculating the *full Procrustes mean shape*, $[\hat{\mu}]$ is to take the arithmetic mean of each of the *full Procrustes coordinates* such that

$$[\hat{\mu}] = \frac{1}{n} \sum_{i=1}^n w_i^P, i = 1, \dots, n \quad (2.14)$$

where w_1^P, \dots, w_n^P denote the *full Procrustes coordinates* of w_1, \dots, w_n . For each landmark in a configuration, the mean coordinate across all configurations gives the mean location of the landmark.

The Procrustes residuals can be calculated as

$$r_i = w_i^P - \left(\frac{1}{n} \sum_{i=1}^n w_i^P \right), i = 1, \dots, n \quad (2.15)$$

The Procrustes residuals are useful for calculating the variability of a shape. They are also a useful way to gauge the accuracy of repeated landmark allocation which will be discussed in detail in Chapter 4.

If the mean Procrustes shape for all configurations is considered to be the pole within a particular shape space then all configurations depend on the location of this pole. In the vicinity of a pole, a linearized version of shape space is given by the concept of *tangent space*.

2.4.1 Coordinates in Tangent Space

In Dryden and Mardia (1998) a suitable tangent coordinate is defined as “a *tangent projection to the pre-shape which does not depend on the original rotation of the figure*”. When dealing with shape data, using the Procrustes

tangent projection to the shape space is a very useful tool as it simplifies the task of shape analysis to standard multivariate analysis. The shape space within which the three-dimensional coordinates lie, can be linearised into a tangent space equivalent. Within the tangent space, the coordinates are centered on the (Procrustes) mean shape and are represented by tangent projections to this pole.

2.4.2 Two-Sample Hotelling's T^2 Test

It is useful to be able to compare the mean shape of two independent populations, say for example the mean facial shape of a sample of males versus a sample of females. A Hotelling's T^2 Test tests the null hypothesis that no difference is present between the mean shapes of two independent populations of configurations.

Consider that it is of interest to compare the mean shape of two independent populations with configurations X_1, \dots, X_{n_1} and Y_1, \dots, Y_{n_2} with mean shapes $[\mu_1]$ and $[\mu_2]$ respectively. A Hotelling's T^2 two-sample analysis tests the following hypotheses;

$$H_0 : [\mu_1] = [\mu_2] \text{ versus } H_1 : [\mu_1], [\mu_2] \text{ unrestricted.}$$

This test is performed after the *Procrustes mean shape* $\hat{\mu}$ is obtained for all $n_1 + n_2$ configurations such that the two populations of configurations lie within the same tangent space centered at $\hat{\mu}$, the Procrustes mean shape. The assumptions of a Hotelling's T^2 test are normality of the data and homogeneity of the covariance matrices for the tangent coordinates v_1, \dots, v_{n_1} and $\omega_1, \dots, \omega_{n_2}$. The model proposed in the tangent space for each population's tangent coordinates is a multivariate normal distribution such that

$$v_i \sim N(\xi_1, \Sigma) \text{ and } \omega_j \sim N(\xi_2, \Sigma) \text{ for } i = 1, \dots, n_1; j = 1, \dots, n_2. \quad (2.16)$$

A common covariance matrix is assumed and v_i and ω_j are mutually independent. The Mahalanobis squared distance between sample means \bar{v} and $\bar{\omega}$

is

$$D^2 = (\bar{v} - \bar{\omega})^T S_u^- (\bar{v} - \bar{\omega}), \quad (2.17)$$

where $S_u = (n_1 S_v + n_2 S_\omega) / (n_1 + n_2 - 2)$ and S_v and S_ω are the sample covariance matrices for the two sample shapes being compared. S_u^- is the Moore-Penrose generalised inverse of S_u . The test statistic of the Hotelling's T^2 test is

$$F = \frac{n_1 n_2 (n_1 + n_2 - M - 1)}{(n_1 + n_2)(n_1 + n_2 - 2)M} D^2 \quad (2.18)$$

and this has an $F_{M, n_1 + n_2 - M - 1}$ distribution under the null hypothesis where M is the number of variables measured. For high values of F , the null hypothesis is rejected.

2.5 Methodology for Landmark Analysis

Methods of identifying trends in the landmark configurations identified from the control data (such as evidence of sexual dimorphism) and methods of describing the variation in the landmark data are detailed in this section.

2.5.1 Permutation test

In Section 2.4.2 the method of testing for significant differences between two mean shapes using the Hotelling's T^2 test was introduced. Where the assumptions of a Hotelling's T^2 test are questionable, an alternative method of testing mean shapes of populations is to use a *permutation test* (Good (2011)). For comparing the mean shape between male and female faces from the control data set, we have a situation where a two sample permutation test could be applied. The null hypothesis states that the mean shapes of the two populations are equal. In this instance, the data are permuted into two groups of equal size to the groups in the data. The test statistic is then evaluated for all possible permutations T_1, \dots, T_P . The observed test statistics are then ranked, within this list and from this the p -value of the permutation

test can be calculated such that

$$p - \text{value} = 1 - \frac{r - 1}{P}$$

where r is the rank and P is the number of permutations. At the 5% significance level, the null hypothesis is then rejected when the p -value is less than 0.05. In Ferrarini (2010), the robustness and sensitivity of this Hotelling's T^2 permutation test method is explored at length.

2.5.2 Principal Component Analysis

Principal Component Analysis (PCA) is a method of reducing dimensionality in multivariate data. The landmark coordinates for the control data is in the form of 51 (23×3) matrices. Consider an $(n \times p)$ matrix of tangent co-ordinates X . PCA seeks out the linear combinations of the columns of X which explains maximal variance in the data. By continuing with the notation in Venables and Ripley (2002), if we use S to denote the covariance matrix of X , it can be defined by

$$nS = (X - n^{-1}\mathbf{1}\mathbf{1}^T X)^T (X - n^{-1}\mathbf{1}\mathbf{1}^T X) = (X^T X - n\bar{\mathbf{x}}\bar{\mathbf{x}}^T) \quad (2.19)$$

where $\bar{\mathbf{x}} = \mathbf{1}^T X/n$ is the row vector of means of the variables. The sample variance of a linear combination $\mathbf{a}^T \mathbf{x}$ of a row vector \mathbf{x} is $\mathbf{a}^T \Sigma \mathbf{a}$ and can be used to obtain the linear combinations of the columns of X . This can then be achieved by maximising (or minimising) $\mathbf{a}^T \Sigma \mathbf{a}$ subject to $\|\mathbf{a}\|^2 = \mathbf{a}^T \mathbf{a} = 1$. The principal components are identified by the eigenvectors and eigenvalues of the sample covariance matrix S of X . The eigendecomposition of Σ can be defined as $\Sigma = C^T \Lambda C$ since it is a non-negative definite matrix, where C contains the eigenvectors and Λ is a diagonal matrix of (non-negative) eigenvalues in decreasing order.

The idea of principal components is to succinctly summarize the main patterns of variation in a data set and also when describing the variance in shape

analysis. In the context of shape analysis the aim is to find a set of orthogonal axes of variation which captures the majority of the variability within the shape. Commonly, the first few principal components explain most of the shape variability in the data set.

PCA can be used to identify differences between sub-populations within the data set, such as differences between males and females. By plotting the first and second principal component scores against each other, and highlighting which points come from each sex, visual evidence of sexual dimorphism can be seen when clusters of male and female observations are well separated. This indicates that some of the variation may be attributed to differences between males and females, suggesting a difference in landmark coordinate location between sexes. Similarly by plotting the principal component scores against continuous measurements, such as body fat, and distinguishing which observations are male and females, the relationship between body fat and shape can be explored while adjusting for sex.

2.5.3 Landmark Distances

Distances between landmarks can be calculated to provide simple height and width measurements of the facial shape for each control face. The distances between landmarks are the Euclidean distances in three dimensions such that

$$Distance = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (2.20)$$

for two landmarks with coordinates (x_1, y_1, z_1) and (x_2, y_2, z_2) . The calculated mean distances can then be compared between sexes.

The correlation between landmark distances and % Body Fat for all control participants can be calculated using Pearson's Product Moment Coefficient, commonly referred to as *Pearson's correlation*. Pearson's correlation ranges

between -1 to +1 and reflects the degree of linear relationship between a specific landmark distance and % Body Fat, (Pestman and Alberink (1991)).

Chapter 3

Data Collection

A new set of control facial images was required in which participants satisfied the control criteria for both research areas of interest; the success of orthognathic surgery and the characterisation of the biological processes which underlie schizophrenia. A data set of facial shapes has been built by compiling photographic images of participants who volunteered for this study. In addition to the facial images captured, information about the participants have also been collected in order to establish whether or not a participant meets the control criteria.

The planning stage of data collection was fundamental to the success of producing a high quality and representative control data set. In addition to the photographic image, it needed to be established what other information was required from each participant. This chapter will cover the numerous aspects that were considered prior to and during the process of data collection.

The data collection process can be split up into three sections; the questionnaire, clinical measurements of the participant's body and the capture process of the participant's three-dimensional image.

The University of Glasgow requires that all research involving human data

or material is subject to formal ethical review and therefore ethical approval was obtained from the School of Education Ethics Committee before any data was collected.

3.1 Exclusion Criteria

In order to address specific research questions using this control data set, biological and medical information needed to be gathered from each of the participants. Criteria for acceptable control data for this study includes:

- parents and grandparents of *white British, white Scottish, white Irish* or *white European* origin
- no history of cleft lip/palate themselves or in a first degree relative
- no history of facial surgery or a substantial injury that affected the facial shape
- no history of psychotic disorders themselves or in a first degree relative
- no history of suicide themselves or in a first degree relative, (successful or unsuccessful)

It was decided that each participant's ethnic background should originate from the British Isles or be *white European* (as defined by the office of National Statistics), as far back as grandparents.

It was essential all control faces had no structural defects to the facial shape so any participant who was born with a cleft lip/palate could not be included in the control population. A history of cleft lip/palate in a first degree relative is informative about the genetic make-up of the participant, with possible associated shape effects. This is the reasoning behind excluding any participant with a first degree relative with a cleft lip/palate, even where the

participant themselves does not have full blown cleft lip/palate. Participants who had sustained an injury or surgery which had altered the facial tissue or underlying bone structure were not suitable as controls.

The criteria for a control participant also included no history of mental illness, specifically no history of psychotic disorders themselves or in a first degree relative. Psychosis is a symptom of mental health conditions, such as schizophrenia or bipolar disorder (manic depression). For the purpose of characterising of the biological processes which underlie schizophrenia in a future study, any participants with a mental disorder of this type, or a history of such in a first degree relative could not be included as a control. As well as psychotic illness, it was also a criterion that no participant had a history of suicide themselves or in a first degree relative. This was deemed appropriate given that suicide is often associated with a psychotic illness.

This information was gathered by asking participants to answer a series of questions. This ensured that participants included in the control data set represented a baseline against which population groups of interest can be compared. The participants that did not fit the inclusion criteria were rejected from the control data set retrospectively.

3.2 Questionnaire

All participants were asked to provide information about themselves in the questionnaire detailed below. The purpose of the questionnaire was to collect basic demographic information and to establish whether or not a participant fitted the criteria for the control data set. The questions which were asked are listed below:

1. Date of birth
2. Sex

3. E-mail address
4. Have you ever had any serious facial injury?
5. Have you had, or are you seeking, any serious facial surgery, cosmetic or otherwise?
6. Is there a history of cleft lip and/or palate in your family?
7. Were both of your biological parents born in the British Isles; (*England, Scotland, Wales, Northern Ireland or Republic of Ireland*)? If not, where were they born?
8. Were all of your grandparents born in the British Isles; (*England, Scotland, Wales, Northern Ireland or Republic of Ireland*)? If not, where were they born?
9. What best describes your ethnic group? (*white British, white Scottish, white Irish, white European, other*)
10. Is there a history of psychotic illness or suicide in a first degree relative? (*A first degree relative refers to mother, father, sibling, son or daughter.*)

These questions were asked in a one-on-one environment between the analyst and the participant and responses were entered directly into a secure database. Given the sensitive nature of some of the questions, it was expected that some participants may wish not to answer some of the questions. All of the questions and a list of the exclusion criteria were included in the information sheet distributed during the recruitment process. (A copy of the Information Sheet is included in Appendix A.1.) This provided an opportunity for prospective participants to prepare for the more sensitive questions or to choose not to take part in the study. It was made clear that all questions were optional in the information sheet and during the data collection

session. The wording of the questions put to the participants was exactly as printed above.

3.3 Clinical Measurements

It was of interest to find out to what extent body fat affects the soft tissue structure of facial shape. It is known that individuals will have varying fat on their faces depending on their own amount of body fat. In order to standardise for this effect, body fat measurements were taken for each participant. A ©Bodystat machine was used to provide an estimate of participant's percentage body fat. To accurately obtain a body fat reading from the ©Bodystat machine, measurements such as an individual's height, weight and hip and waist circumference were obtained. These measurements were then entered into the ©Bodystat machine.

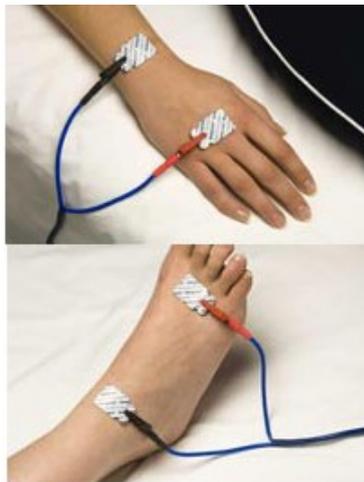


Figure 3.1: Electronic sensors of fat data will also be used in Section ©Bodystat machine. 5.3 where the relationship between facial shape and percentage body fat will be investigated. The ©Bodystat machine is an electronic device that uses a very small electric current to es-

The results that were recorded from the ©Bodystat output were Body Fat, (as a percentage) and Impedance, (measured in micro Ohms). Impedance is a measure of electrical current resistance through the body. All body measurements and results from the ©Bodystat output were entered into the secure database, along with the answers from the questionnaire. The body fat data will also be used in Section 5.3 where the relationship between

estimate a person's body composition. Sensor pads are placed at four sites on the individual's body; the back of the hand, the wrist of the same hand, the top of the foot and the the front of the ankle on the same foot as seen in Figure 3.1. Four crocodile clips attach the wires from the ©Bodystat machine to these sensors. Once all demographic information has been entered into the machine such as height, weight, hip and waist circumference, the machine is then ready to pass a small electric current through the subject's body from the sensors on the hand to the sensors on the foot. Participants were asked to lie down on a clinical chair for this part and ensure that the limbs of their body were separated as best as possible, (to prevent the electrical current from "short circuiting"). The ©Bodystat machine uses a measure of resistance, Impedance, measured in micro-Ohms, to calculate an estimate of an individual's body composition.

A free standing device was used to measure the participant's height. ©SECA medical scales were used to measure the participant's weight. A soft tape measure was used to measure participant's waist and hip circumferences.

Each clinical measurement was entirely optional and this was stated on the information sheet and reinforced during the data collection session. For the weight measurements, participants were asked to remove their shoes and any heavy outdoor clothing as well as any heavy items from their pockets, (such as phones, keys and spare change etc.). The measurements around the waist and hips needed only to be an estimate to the nearest centimetre (cm) so no further removal of clothing was required. For the ©Bodystat measurements, access was required to the hands and feet. On occasions where participants were wearing inappropriate clothing for this, (such as tights), they were invited to return at a later date.

3.4 Image Capture Protocol

For each participant, the analyst maintained the same structure and protocol for each image capture process. The camera used to capture images of participants can be seen in Figure 3.2. It consists of four independent cameras



attached to a static frame. The cameras capture images simultaneously to

Figure 3.2: ©DI3D Camera Equipment

give four stills in any one image capture and the ©Di3D software builds these stills into a single three-dimensional image. Throughout a capture session with multiple people, the camera equipment was positioned in the same location, or as close to it as possible. The camera system was calibrated at the beginning of each day and again throughout the day as required, at the discretion of the analyst. A static chair was positioned in front of the camera. The cameras were positioned so that the participant's eyes were lined up at the centre of each camera view before the image was taken. As an additional precaution, the distance from the centre of the camera system and the participant's cheek was roughly measured using a measuring tape, with the objective being that the distance be approximately 95 cm. This value was recommended by senior software engineer Ewan Borland from ©Di3D, who set the camera system up at its optimal position prior to the beginning of the data collection process.

Each subject was seated in the static chair in front of this camera. All participants were asked to pull their hair away from their face using hair clips or hair bands provided. Prior to their arrival participants were also asked to refrain from using heavy make-up and to shave any facial hair. These requests were also listed in the information sheet given to each prospective

participant. A shiny or hairy facial surface can produce an image where the facial structure or boundaries of facial features are inaccurate. Where possible, participants were asked to remove facial jewelery, if it was easy and painless for them to do so. Embedded jewelery such as studs were not required to be removed in view of the complications which may be involved.

Once in the static chair, before capturing an image, participants were asked to tilt their heads slightly backwards so that the area around the nose and chin could be captured accurately. Participants were asked to try to relax their jaw and mouth area so that there was no muscle tension or pursing of the lips. Often, participants were asked to say “Mississippi”, followed by the letter “n” and then asked to let their mouth rest, in order to relax the mouth and jaw into a natural position before capturing the image. If a participant is clenching their jaw or holding their mouth in an unnatural position, this will give an inaccurate measurement of their true facial structure at rest. Also if the mouth is open slightly, this can lead to inaccurate capture of the surface area around the lips (Farkas (1994)).

Some participants had ‘insufficient lips’ which means that the upper and lower lips do not naturally meet when the mouth is in a relaxed position. On occasion that a participant had ‘insufficient lips’, an additional photograph was taken where the participant closed their lips together in an unrelaxed state.

Approximately, 3 to 5 sets of ‘jpeg’ stills were captured for each participant and from these, the best set was selected. The selected image (or images, should the participant have ‘insufficient lips’), was built into 3D images and then checked to ensure that the image is useable for the study. The computer package ©DI3D uses standard digital stills cameras to capture simultaneously one or more stereo pairs of images of a subject. Figure 3.3 is an

example of two stills captured from the bottom left and right cameras. Four stills in total are used to build a three-dimensional image of the subject using the ©DI3D software. A copy of a Data Capture Protocol document can be found in Appendix A.2.

3.4.1 Loss of detail in 3D image

Loss of detail in image capture is a difficult problem to overcome. There are numerous factors that can lead to the detail of a participant's face being lost upon capture:

- Any of the four cameras being out of focus.
- Flashes being too bright, whitening out the skin.
- The skin being transparent. Younger



Figure 3.3: Stills from the bottom left and right cameras.

people, e.g. babies have very light, thin and transparent skin. The light easily penetrates the skin and bounces off the lower skin surface causing subsurface scattering. This will cause an ‘orange peel’ effect. In contrast, older people tend to have thicker skin which is less transparent and therefore less penetrative to light creating an overall smoother surface.

The area which proved most difficult to capture was the ears. This is thought to be due to two reasons. The first is because of hair around the ears such as sideburns. As the ©DI3D software is trying to build a three-dimensional surface, hairs captured by the camera give the illusion that a surface is present. An example of a case where the surface of the ears has been distorted because of hair around the ear area can be seen in Figure 3.4. The second

reason is the position of a participant's ears in relation to the cameras. It is possible that the four static cameras, at the angle they were positioned, failed to capture the detail of the ears compared to the rest of the face.

3.5 Participants



Figure 3.4: Distorted ears due to sideburns.

Participants were recruited mostly from students from the University of Glasgow. This decision was based primarily on convenience in terms of advertising the study to large audiences and location. All prospective participants were given an Information Sheet detailing background information on the Face3D study and the motivation for collecting a control data set.

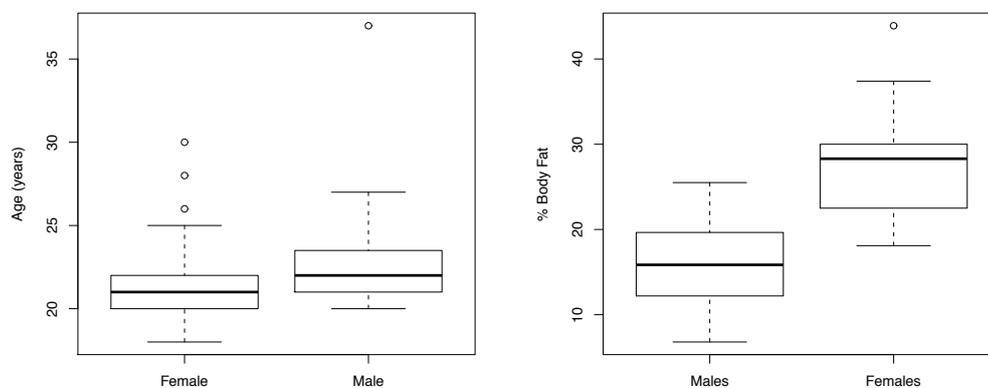
The Information Sheet also let prospective participants know exactly what data would be collected from them and how. It also instructed participants on how to prepare for their data collection session.

A web-site was set up with a diary of session times available for data capture. Participants could log in and select a session time which suited them. Due to the nature of some of the questions that were asked in the questionnaire, it was decided that only one participant should have their data collected at any one time.

Each participant signed a permission slip which agreed that the data collected from them could be used for the purposes of the research involved, and that the facial image and the non-medical personal information may be made available freely to others thereafter. A copy of the Permission Form can be found in Appendix A.3.

3.6 Problems with the Data

There are limitations to the control data which was collected. Many factors attribute to the lack of diversity in the participants' demographics. Firstly, as the majority of participants were students at the University of Glasgow the range of ages is narrow and the mean age of participants is low at 21.8 years, (23 for males, 21 for females). The spread of ages for males and females can be seen in Figure 3.5(a). The ratio of males to females is low; there are 16 male and 35 female participants. Most of the participants originate from a 1st year course in statistics for psychology students, of which there are over 300 members. Of these students, the majority are female which may contribute to the low number of male volunteers in this study.



(a) Age Ranges for each Sex

(b) % Body Fat for each Sex.

Figure 3.5: Age and % Body Fat for Males and Females.

It should also be noted that the range of body fat measurements of all participants is at the lower end of the scale. Figure 3.5(b) illustrates the ranges of body fat percentages observed from participants, split by sex. It was observed that few 'larger' people came forward to have their data collected. This could be attributed to the elective nature of the participant recruitment

process. Given that it was advertised that body composition measurement would be taken as part of data collection, it is plausible that people with a higher body fat would be less forthcoming.

3.7 Problems with Landmarking

As previously mention in Section 3.4.1, the lack of detail around the ear area made some of the landmarks difficult to identify, specifically the *tragion* and *otobasion inferius* landmarks. Due to the nature of facial shape, every participant's face is different. The accuracy of identifying landmarks differs from person to person. A measure of accuracy in identifying landmarks will be discussed in Chapter 4.

Chapter 4

Validation Study

4.1 Motivation

Landmark analysis is widely used to characterise the soft-tissue facial shape of participants such as in White (2004) where a measure of accuracy in landmark identification is discussed. The purpose of this validation study is primarily to validate the use of the camera equipment used to generate three-dimensional images and to provide a measure of within-analyst accuracy of landmark identification. Various sources of variation in landmark identification will be investigated. This validation study required to ensure that facial shapes generated from the camera are consistent. The other element is to ensure that landmark configurations generated are reliable and consistent so that well-founded and accurate conclusions can be drawn.

4.2 Design of Validation Study

Five people were chosen to take part in the validation study; 2 males and 3 females between the ages of 22 and 24. The study took place over two consecutive days. Four sets of images from the four cameras were taken of each subject; two sets on Day 1 and an additional two on Day 2. These sets of images were built into four independent three-dimensional images for each

of the five participants; two from Day 1 and two from Day 2, creating 20 three-dimensional facial images in total.

Day 1 The camera equipment is calibrated. Two sets of images are taken of each of the five subjects independently.

Day 2 The same procedure is carried out, exactly as on Day 1.

For any one day, between the two image captures, individuals were invited to readjust their facial positions by standing up and sitting down again or simply by talking. Hence, the two images captured on the same day were not taken in quick succession reflecting the same pose. The image capture procedure for the validation study followed that of the standard protocol for the image capture administered for the control data set. Environmental circumstances of the study were kept constant over the two days. All calibrations were performed and images captured by the same analyst to avoid between-observer error. No information other than age, sex and the three-dimensional facial image was collected from the five volunteers in the validation study.

More landmarks were identified for the validation study than for the control facial images. In total, thirty anatomical landmarks were identified on each of the 20 three-dimensional facial images. The landmark positions are detailed in Table 4.1. The landmarks were visually identified in the Landmark software seen in Figure 1.1, by a single analyst. Landmark identification of the same person in the validation study was non-consecutive to ensure that repeat landmarks of the same image were independent and not affected by the analyst's memory.

For each of the four independent three-dimensional images for each participant, anatomical landmarks were identified twice. Therefore, for a single participant eight landmark configurations were created. Figure 4.1 illustrates the structure of the hierarchical model for any one of the five participants

Table 4.1: Anatomical Landmarks Allocated for Validation Study

ID	Code	Region	Side	Abr.	Name
1	S0	nose	middle	pm	<i>pronasale</i>
2-3	S1-S2	nose	left/right	a	<i>alare</i>
4-5	S3-S4	nose	left/right	ac	<i>alare crest</i>
6	S5	nose	middle	sn	<i>subnasale</i>
7-8	S6-S7	nose	left/right	nt	<i>nostril top point, columella breakpoint</i>
9	S8	nose	middle	cc	<i>columella constructed point</i>
10-11	S9-S10	nose	left/right	nb	<i>nostril base points</i>
12	S11	nose	middle	se	<i>sellion</i>
13	S12	nose	middle	n	<i>soft tissue nasion</i>
14-15	S13-S14	eye	left/right	ex	<i>exocanthion</i>
16-17	S15-S16	eye	left/right	en	<i>endocanthion</i>
18-19	S17-S18	ear	left/right	t	<i>tragion</i>
20	S19	lips	middle	ls	<i>labiale superius</i>
21-22	S20-S21	lips	left/right	cph	<i>crista philtri</i>
23-24	S22-S23	lips	left/right	ch	<i>cheilion</i>
25	S24	lips	middle	sto	<i>stomion</i>
26	S25	lips	middle	li	<i>labiale inferius</i>
27	S26	chin	middle	sl	<i>sublabiale</i>
28	S27	chin	middle	pg	<i>soft tissue gnathion</i>
29-30	S28-S29	ear	left/right	a	<i>otobasion inferius</i>

in the validation study, (Pinheiro and Bates (2000)). Each level in the hierarchical model represents a source of variability across the eight landmark configurations for a single participant.

The first level of the model represents the variability present in repeat landmark configurations on a single image. Variability between configuration

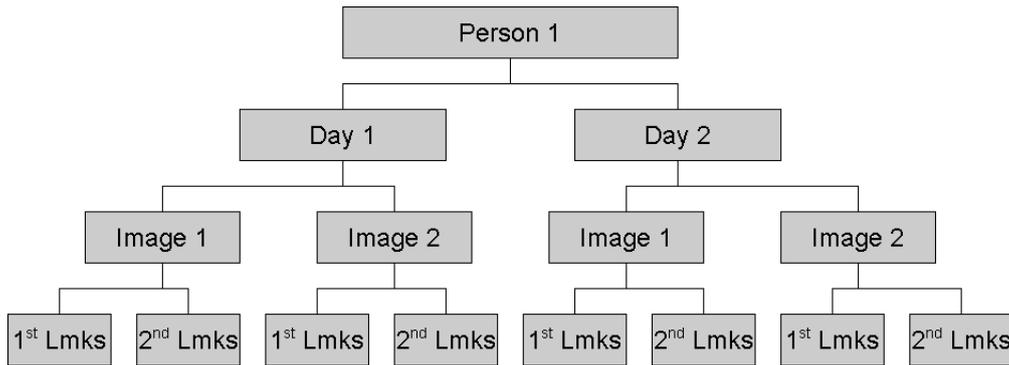


Figure 4.1: Hierarchical Model of Validation Study Design

coordinates at this level is as a result of intra-observer error.

The second level of the model represents the variability of repeat landmark configurations between two different images of the same person. Variability present between configuration coordinates at this level is a result of reproducibility error in representing facial shape of a participant. The error source could be from the cameras or possibly the software in building the three-dimensional image.

The third level of the hierarchical model represents the variability of landmark configurations between the two days. It is a reasonable assumption that the condition of a person's skin or facial expression or their general health varies between the two days of image capture. To give an example; a female subject may be wearing make-up on day 1 but not on day 2. This in turn will have an effect on the surface of the skin and therefore the Day on which an image is captured may provide a source of variability in landmark configurations.

Finally, the fourth level of the hierarchical model represents the variability in landmark configurations that exists between participants. Other than to provide a benchmark, it is not of interest to assess this variability as it is uninformative in the context of the validation study as participants' faces naturally differ in shape.

4.3 Analysis of Validation Study

For each level of the hierarchical model, sources of variation across the 40 landmark configurations will be investigated. For the purpose of consistent two-dimensional graphical displays, all of the configurations have been positioned as if 'facing' toward the observer. This is achieved by aligning the *soft tissue nasion* and the *sublabiale* landmarks, (landmarks 13 and 27 respectively) vertically on the y -axis. The x -axis is anchored perpendicular to the plane on the y -axis obtained by aligning all landmarks identified on the midline of the face, (landmarks 1, 6, 9, 12, 13, 20, 25, 26, 27 & 28). The z -axis is hence mutually perpendicular to the x and y -axis. The final positioning of the rotated landmarks can be seen in Figure 4.2. This method of displaying the landmark coordinates will be used throughout this thesis. It is thought that being able to visualise the data in this way can enable better understanding of the analyses applied, (Tukey (1974)).

4.3.1 Single Observer Consistency

We wish to identify the margin of human error incurred when allocating landmark points to a three-dimensional facial image. This can be identified by comparing the repeat landmarks of each of the four images per subject. For each subject there are 8 (30×3) configurations, two configurations per image.

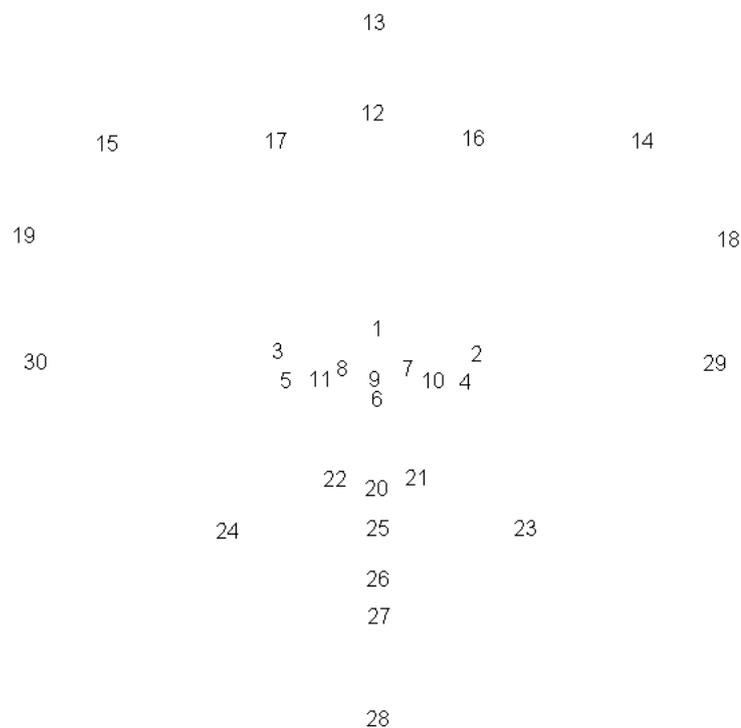


Figure 4.2: View of a singular configuration rotated to ‘face’ directly forwards. Landmark number is shown as indicator of landmark position.

At the first level of the hierarchical structure, repeats of the landmark configurations on the same image of the same person will be compared. A snapshot of two repeat landmark configurations allocated to the same image of a subject can be seen in Figure 4.3. At this stage, Procrustes analysis is not necessary as these landmarks were identified independently on the same three-dimensional facial image and are therefore already in the same shape space.

An indication of within-image error of landmark allocation can be quantified by calculating the Technical Error Measurement (TEM), (Lewis (1999)). TEM is calculated using the equation:

$$TEM = \sqrt{\frac{\sum_{i=1}^n d_i^2}{2N}} \quad (4.1)$$

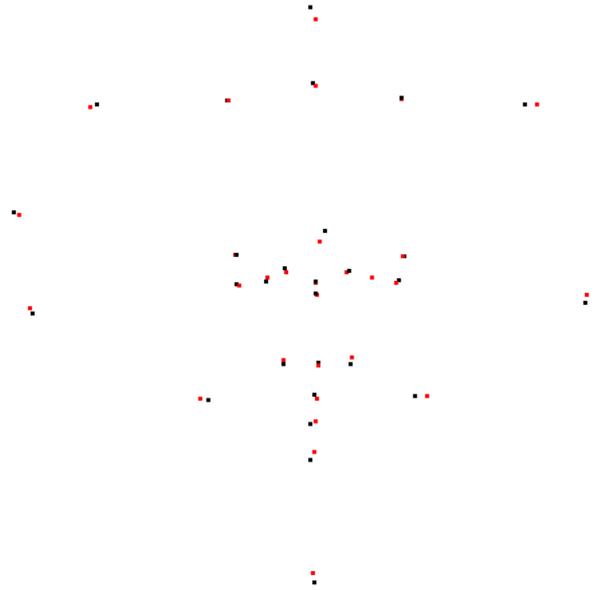


Figure 4.3: Frontal view of two repeat landmark configurations; the first (*black*) and second (*red*).

where d_i is the difference in two measurements for landmark i and N is the total number of duplicate measurements made, 40 in this case, (4 images on 5 people, landmarked twice). For the purpose of calculating the TEM, the images taken of the same person will be considered independent of each other as only the accuracy of repeat landmarks is being assessed. Hence the TEM scores are purely a reflection of the analyst's error in identifying repeat landmarks on the same image.

The TEM scores listed in Table 4.2 provide an estimate of the analyst's human error demonstrated while placing the 30 anatomical landmarks in the validation study. The measure of analyst's error is given for all three dimensions. For landmark 13, the *soft tissue nasion*, a high TEM score of 3.55mm is present on the y dimension which indicates that repeated identification of this landmark position had a high variation. Repeat landmark identification

Table 4.2: TEM Scores of Repeat Landmarks on 20 Images in Validation Study

ID	Landmark	x (mm)	y (mm)	z (mm)
1	<i>pronasale</i>	0.616	0.406	0.123
2	<i>left alare</i>	0.363	0.369	0.282
3	<i>right alare</i>	0.521	0.510	1.315
4	<i>left alare crest</i>	1.586	0.895	0.867
5	<i>right alare crest</i>	0.479	0.826	0.532
6	<i>subnasal</i>	0.271	0.040	0.238
7	<i>left columella breakpoint</i>	0.238	0.056	0.556
8	<i>right columella breakpoint</i>	0.219	0.162	0.053
9	<i>columella constructed point</i>	0.015	0.011	0.593
10	<i>left nostril base point</i>	0.455	0.263	0.369
11	<i>right nostril base point</i>	0.026	0.505	0.636
12	<i>sellion</i>	0.559	1.118	0.494
13	<i>soft tissue nasion</i>	1.230	3.548	0.027
14	<i>left exocanthion</i>	0.848	0.395	0.764
15	<i>right exocanthion</i>	0.098	0.338	0.937
16	<i>left endocanthion</i>	0.420	0.224	0.300
17	<i>right endocanthion</i>	0.184	0.105	0.314
18	<i>left tragion</i>	0.736	2.422	0.020
19	<i>right tragion</i>	0.145	0.789	0.762
20	<i>labiale superius</i>	0.138	0.274	0.199
21	<i>left crista philtri</i>	0.279	0.676	0.009
22	<i>right crista philtri</i>	1.012	0.128	0.429
23	<i>left cheilion</i>	1.041	0.438	0.235
24	<i>right cheilion</i>	0.202	0.659	0.166
25	<i>stomion</i>	0.023	0.241	0.004
26	<i>labiale inferius</i>	0.641	0.646	0.008
27	<i>sublabiale</i>	0.383	0.094	0.235
28	<i>soft tissue gnathion</i>	1.059	1.111	0.599
29	<i>left otobasion inferius</i>	0.170	1.306	0.240
30	<i>right otobasion inferius</i>	0.012	1.155	0.479
Average Score		0.466	0.657	0.393

of the left and right *tragion* also generates high variation which is reflected in the TEM scores in the y dimension. The high variability present in the identification of these landmarks may be explained by poor surface quality due to hair around the ear area as discussed previously. It is possible that for landmarks that are not located at obvious points of change on the facial surface, such as the corner of an eye, the location is more difficult to identify. It could be argued that for such landmarks, identification is subjective to the analyst. It can therefore be expected that high variation is present for such landmarks around the forehead and chin area where the surrounding surfaces are fairly homogenous. The average TEM scores for each dimension indicate that repeated identification over all landmarks varies most in the y dimension (i.e. on the vertical plane) with an overall average TEM score of 0.66mm.

4.3.2 A General Model For Reproducibility

Before further analysis, the 40 configurations of landmarks require to be on the same shape space. This could be achieved by using General Procrustes Analysis as described in Section 2.2, however, information between Person, Day and Image from which the configuration originates would be lost. The information between levels of the hierarchical model can be preserved by obtaining the rotation matrices and centroid matrices at the Image, Day and Person levels of the hierarchical model.

At the second level of the hierarchical model the landmarks allocated to any one image are not registered to the same origin between images. This can be overcome at this level, by calculating the mean configuration for each image using the two repeat landmark configurations and registering the mean configurations simply by using using full GPA.

The mean landmark configuration of the two repeats for each image per

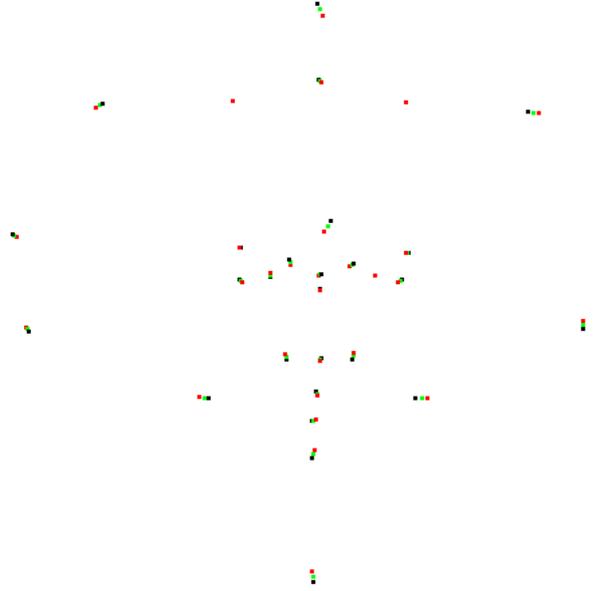


Figure 4.4: Frontal view of two repeat landmark configurations; the first (*black*) and second (*red*), with the average landmark position (*green*) superimposed.

person were obtained by calculating $\sum_{i=1}^5 \sum_{j=1}^4 \frac{(X_{ij1} + X_{ij2})}{2}$ where X_{ij1} is the landmark configuration for the 1st repeat and X_{ij2} represent the second repeat for the i^{th} person and the j^{th} image. Two landmark configurations placed on the same image of a subject and the corresponding average of the repeat landmarks can be see in Figure 4.4. The average configurations calculated are used to represent the landmark positions for each of the 20 images and will be referred to as the landmark configurations *at the Image level* for simplicity.

For each Person, Day and Image a unique centroid can be calculated from the average configuration *at Image level* by averaging all 30 landmark values for each dimension to give a single 3-dimensional coordinate. By this means, 20 unique centroids are calculated, one centroid for each pair of repeat configurations originating from any one image. In Figure 4.5 the landmark

configurations at Image level after being registered by full GPA, are shown and are colour-coordinated by Person.

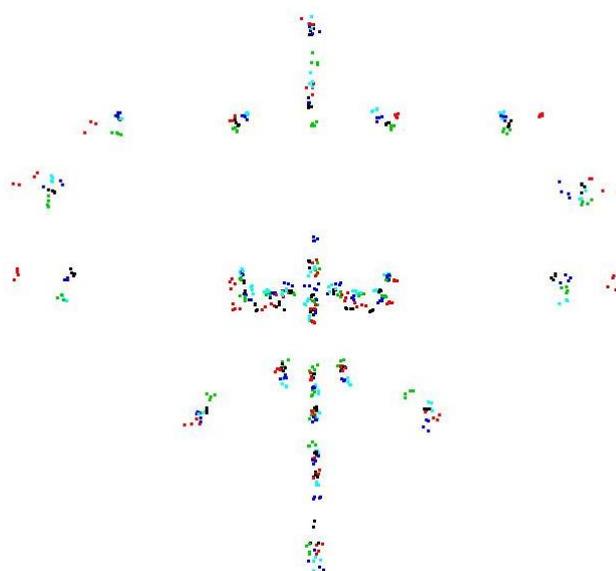


Figure 4.5: Frontal view of average configurations *at Image level*, coloured by participant.

For each Person and each Day, the two average configurations *at Image level* are registered using full GPA and the average of the pairs are calculated. These new average configurations represent the landmark positions *at Day level* from which 10 unique centroids specific to each Day within each Person can be calculated. By matching the average configuration *at Day level* to each of the registered configurations *at Image level* individually using Ordinary Procrustes Analysis, unique rotation matrices can be obtained for each Image within each Day and Person.

Similarly the same procedure is performed at the next level of the hierarchical model so that for any one person, the two average configurations *at Day*

level are registered by GPA and then the average configuration calculated to represent a landmark configuration *at Person level*. By the same method as previously described, unique rotation matrices for each Day within each Person along with 5 unique centroid coordinates specific to each Person are obtained.

The 5 configurations *at Person level* are registered using GPA and the average is calculated to give a single (30×3) configuration, the Procrustes Mean, representing the landmark positions over all people in the Validation Study.

A new set of 40 configurations can be now be calculated from the original configurations by applying a series of transformations using the unique centroid matrices and rotations matrices obtained. For each Person, the unique centroid coordinate is subtracted from each of the 8 configurations so that 5 sets of 8 configurations are now registered to the same origin within any one Person. For each Person, each Day and each Image the unique centroid coordinates are subtracted from the configurations and then multiplied by the unique rotation matrices obtained. By applying this transformation to the original 40 configurations information about the source of variability is preserved.

This transformed set of 40 configurations are comparable with origin zero, and can now be analysed formally in a hierarchical model. The structure of variation for any single landmark in one particular dimension can be expressed as

$$y_{ijk} = \mu + \alpha_i + \beta_{ij} + \gamma_{ijk} + \delta_{ijkl} \quad (4.2)$$

The meaning of the terms in this model is listed below

- μ the mean landmark position over the population
- α_i the adjustment in position for person i
- β_{ij} the adjustment in position for day j for person i
- γ_{ijk} the adjustment in position for image k for day j for person i
- δ_{ijkl} the adjustment in position for repeat l for image k for day j for person i

All of these terms with the exception of μ are treated as random variables, each with its own associated standard deviation. Specifically $\alpha_i \sim N(0, \sigma_{person}^2)$, $\beta_{ij} \sim N(0, \sigma_{day}^2)$, $\gamma_{ijk} \sim N(0, \sigma_{image}^2)$ and $\delta_{ijkl} \sim N(0, \sigma_{repeat}^2)$. The statistical model applied is known as a *hierarchical* or *random effects* model. From the hierarchical model of a configuration we obtained standard deviation values at Repeat, Image, Day and Person level. These measurements are used to quantify the amount of error attributed to each level in the hierarchical model.

4.4 Results of Validation Study

Table 4.3: Average Standard Deviation over all landmarks (mm).

	x	y	z
Person	1.659	4.551	4.324
Day	1.133	0.494	1.301
Image	0.577	0.668	0.816
Repeat	0.601	0.659	0.673

Table 4.3 shows the average standard deviations across all landmarks. The large standard deviations at *Person level* are uninformative as we already know that high variability in people’s faces exists naturally. Across all sources of variation, the highest variability occurs in the z dimension. This result is likely to be attributed to the accuracy in being able to landmark

to areas around the ears. Consider the shape of a face looking directly at you. The rate at which the z dimension changes as a function of the x and y dimensions on the surface of the shape at some places is high which may account for large variability observed in the z dimension.

The variability of the analyst's repeat identification of the landmarks on average is approximately between 0.60mm and 0.67mm across all dimensions. At *Image level*, the variability is higher than that at the *Repeat level* in the y and z dimensions. This result suggests that a higher variability exists between images of the same person than between repeat landmarks identified on a single image by the analyst. The average standard deviation across all landmarks at *Day level* is higher still. This indicates that there does appear to be a difference in the facial shape captured between the two days of image capture. The error measurements at *Repeat level* differ from those seen in Table 4.2. This difference is plausible as the TEM scores simply measure the error of repeat landmark allocations per single image and assume that all images are independent of each other, whereas the hierarchical model does not make this assumption.

For a single landmark identification on one image, the relevant measure of reproducibility combines the variability at the Day, Image and Repeat levels. For any one landmark i , the cumulative standard deviation S , for the x dimension is such that

$$S_{x_i} = St.Dev(x_i) = \sqrt{\sigma_{Repeat,i}^2 + \sigma_{Image,i}^2 + \sigma_{Day,i}^2} \quad \forall i = 1, \dots, 30 \quad (4.3)$$

where $\sigma_{Repeat,i}$, $\sigma_{Image,i}$ and $\sigma_{Day,i}$ are the standard deviations at the Repeat, Image and Day level respectively from the hierarchical model of landmark i . The standard deviation of the combined coordinates which represents the variance in terms of Euclidean distance can be expressed as

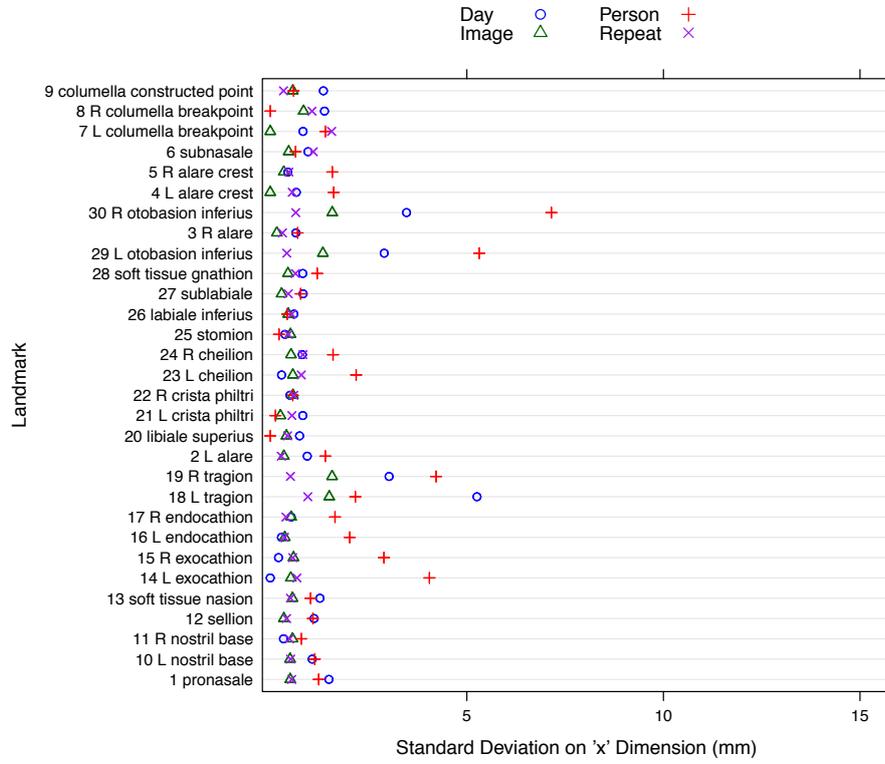
$$St.Dev(\theta_i) = \sqrt{S_{x_i}^2 + S_{y_i}^2 + S_{z_i}^2} . \quad (4.4)$$

By this method the overall standard deviation at Repeat, Image and Day level over all landmarks is calculated as $1.81 \pm 0.84\text{mm}$.

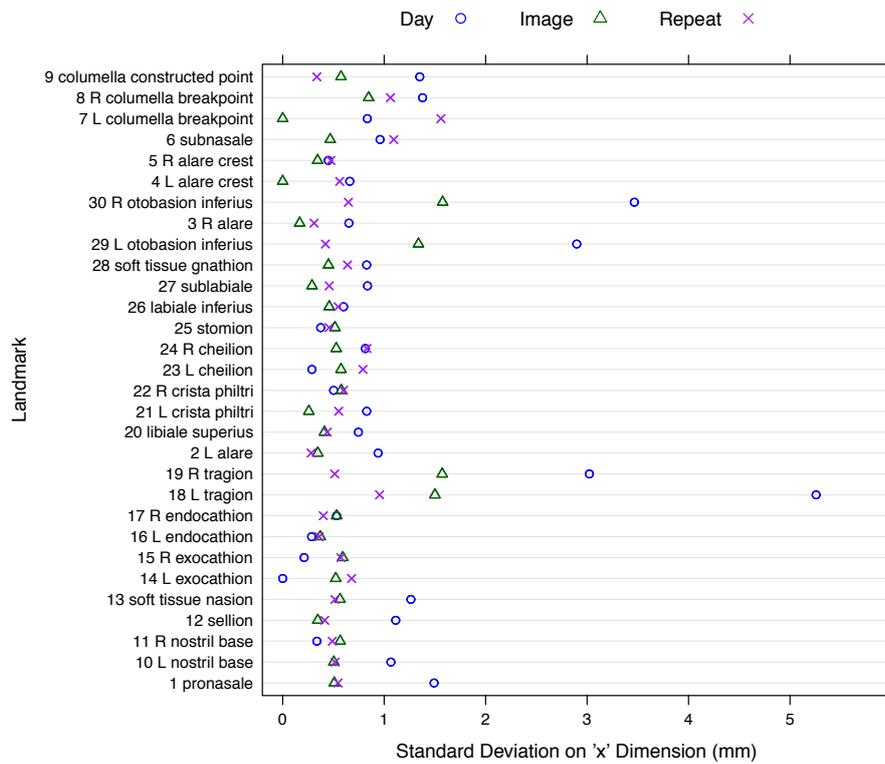
Figures 4.6, 4.7 and 4.8 show dot-plots of the standard deviations across all landmark configurations for x , y and z dimensions respectively. For each pair of graphs, the source of variance has been split by each level of the hierarchical model; firstly standard deviations across Repeats, Images, Day and Participants and secondly standard deviations across Repeats, Images and Day. As expected, the majority of variation in the data originates between participants simply because people's facial shape naturally differs. It is for this reason that the variability is shown across only the Repeats, Images and Day in addition.

For the standard deviation values between Repeats, Images and Days only, there is an indication of high variation in identifying landmarks on the ears. In all three dimensions, landmarks identified on the *Tragions* and on the left and right *Otobasion Inferius* have high variation between Days and between Images of the same person. With a few exceptions, the standard deviation of landmark identification between Repeat landmarks and between Images is between 0mm and 2mm. These error measurements are higher than the landmark identification errors observed by Ayoub et al. (2003) in a cohort study of infants with cleft palates, aged between 3 and 6 months. The errors obtained from the validation study are still minor and therefore it is not unreasonable to conclude that this method of obtaining data from images is valid.

In Figure 4.9 the positions of landmarks for a participant have been identified with spheres. The radius of the spheres directly corresponds to the level of variation in repeat landmark allocation at the Repeat, Image and Day levels combined and for all x , y and z dimensions combined for all participants in

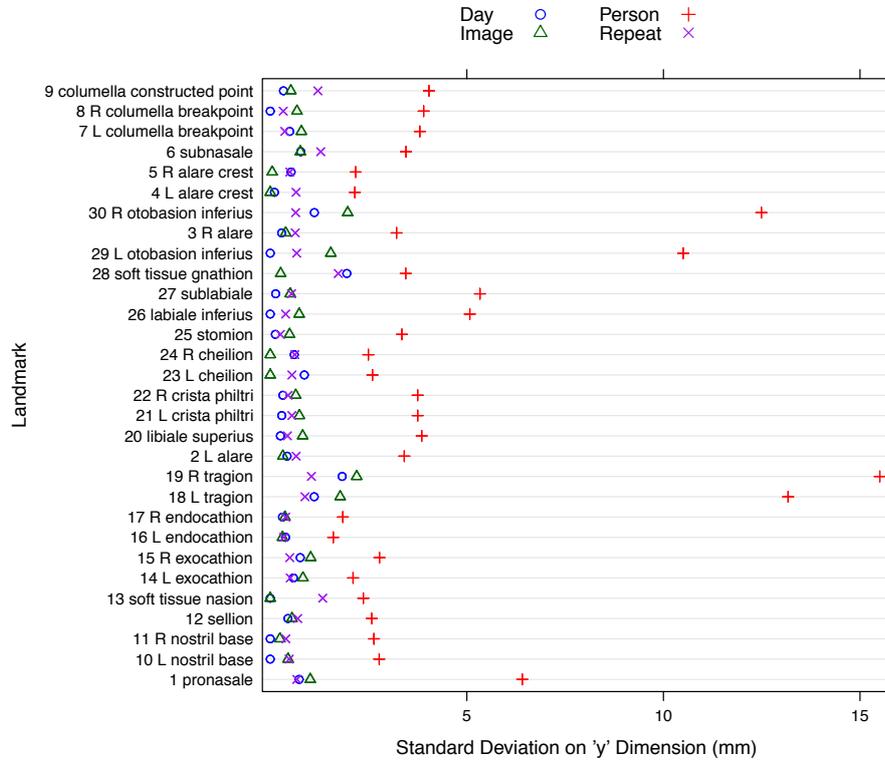


(a) Standard Deviation values for all landmarks *Repeat, Image, Day* and *Person* level

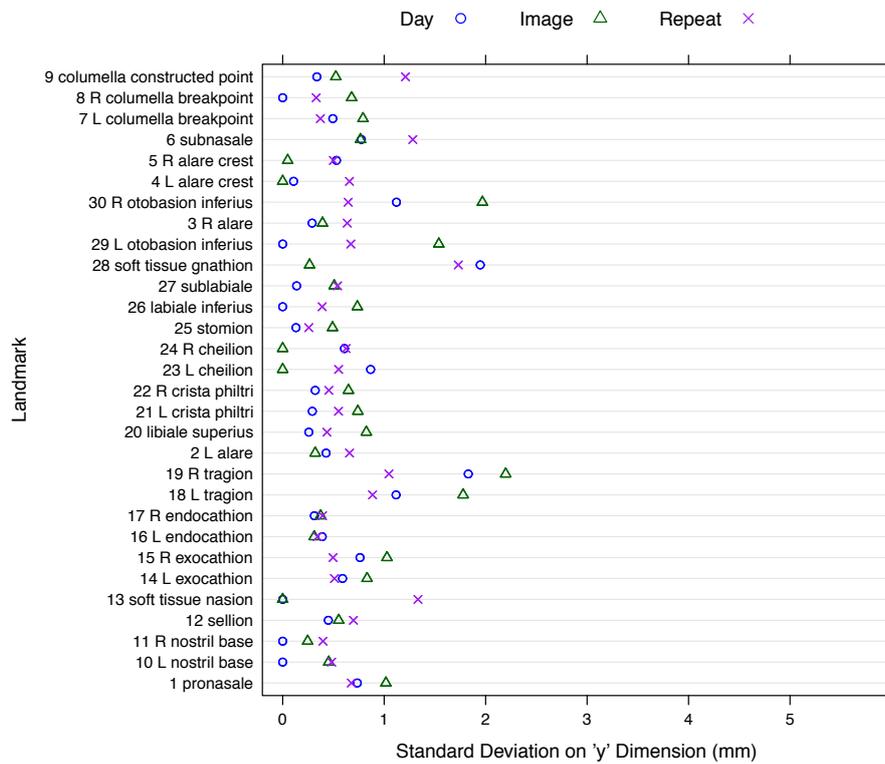


(b) Standard Deviation values for all landmarks *Repeat, Image* and *Day* level

Figure 4.6: Standard Deviation values for all landmarks on the *x*-axis.

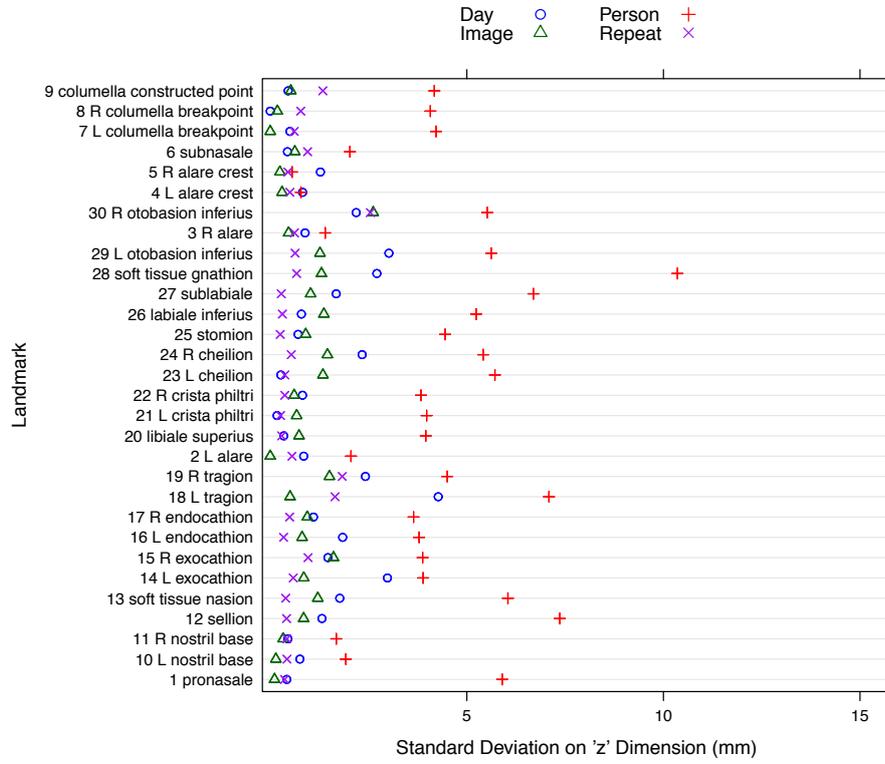


(a) Standard Deviation values for all landmarks *Repeat, Image, Day* and *Person* level

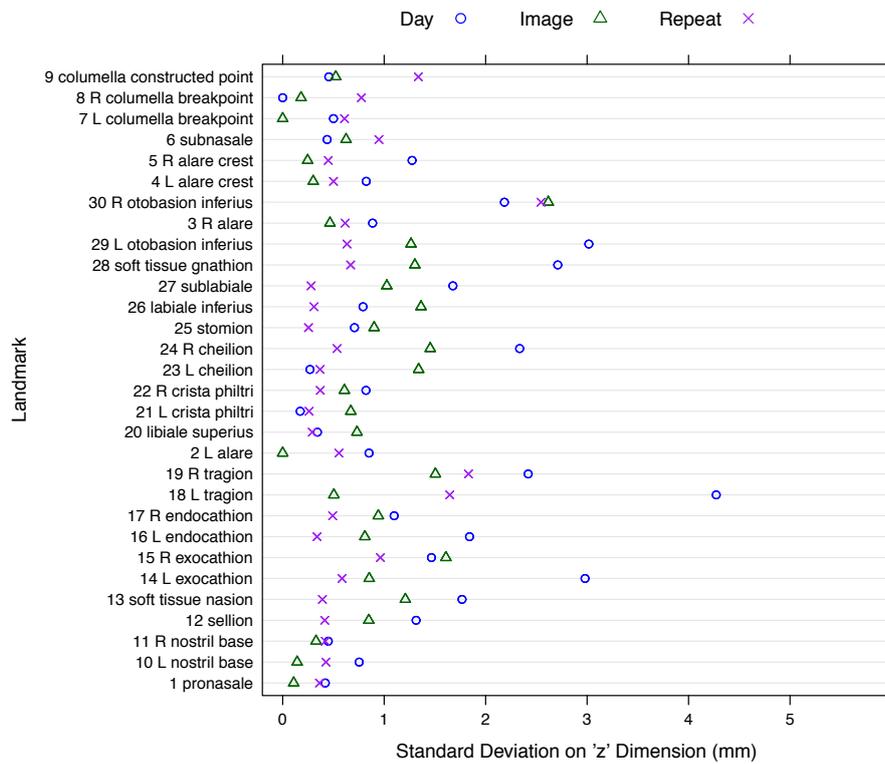


(b) Standard Deviation values for all landmarks *Repeat, Image* and *Day* level

Figure 4.7: Standard Deviation values for all landmarks on the *y*-axis.



(a) Standard Deviation values for all landmarks *Repeat*, *Image*, *Day* and *Person* level



(b) Standard Deviation values for all landmarks *Repeat*, *Image* and *Day* level

Figure 4.8: Standard Deviation values for all landmarks on the *z*-axis.

the validation study, (not just for this pictured individual). The spheres are also coloured from a palette ranging from light yellow (*low variation*) to red (*high variation*).

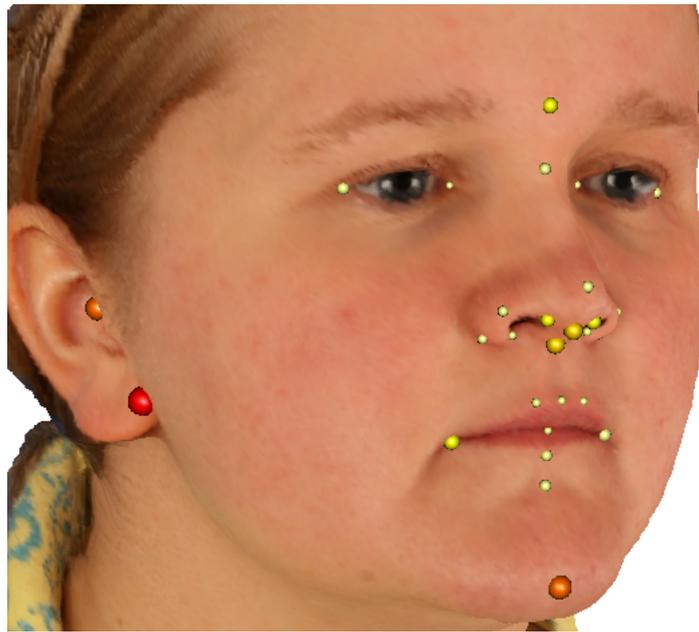


Figure 4.9: Face with spheres representing level of variation in identifying each landmark in the validation study.

Chapter 5

Landmark Analysis of Control Data

Analysis of the landmarks identified on Control participants will now be explored. In this chapter demographic information obtained from the participants - such as sex and % body fat - will be included in the analysis of facial shape. It is of interest to establish firstly whether there is a statistically significant difference in facial shape between males and females in the Control data. The relationship between facial shape and body fat will also be explored for the Control data. Principal Components Analysis as described in Section 2.5.2, will be used to describe variability in the landmark coordinates identified.

Before any analysis, all 51 configurations in the Control dataset are registered using full Generalized Procrustes Analysis. Initially adjustment for scale between configurations will not be implemented so that the size-and-shape configurations are analysed first. The configurations will then be re-scaled to unit centroid size so that the shape of the control configurations can be analysed. Similarly, as in the configurations in the Validation Study, the configurations are rotated in such a way that they are ‘facing’ directly forwards for the purpose of visually representing the landmarks in two dimensions.

5.1 Sexual Dimorphism

Evidence of sexual dimorphism is widely documented for a range of ages and for populations of different ethnicity. A study of three-dimensional gender differences is documented in Bugaighis (2011) of Caucasian children living in the North-East of England and aged 8 to 12 years old. The data in the study included 39 males and 41 females. The facial images were captured using a 3D stereophotogrammetric system and were characterised by 39 anthropometric landmarks. The results focus on inter-landmark height and width distances on the captured facial images and the same Euclidean method of calculating distances as explained in Section 2.5.3 is adopted. Significant differences observed between sexes were mainly from linear width measurements, with wider faces occurring in males. Facial ratios were also calculated between sexes for the intra-landmark distances with no significant differences between sexes observed. It was concluded in Bugaighis (2011) that differences observed between males and females were attributed to size rather than shape. The results from the study described in Bugaighis (2011) are indicative of the type of results that should be expected in this study, but cannot be directly comparable as the participants are children aged between 8 and 12 years.

Significant differences in overall soft tissue facial shape were found in Ferrario (1995) from a sample of 32 male and 30 female Caucasian adults aged 19-32 years. Boundaries of the facial outline were created on the soft tissue facial structure by joining up anatomical landmarks. Harmonic Fourier analysis was applied to refit the soft tissue facial outlines for both sexes and it was observed that the outline area of male faces is 14% larger than that of females on average. In Ferrario (1993) on the same sample of faces, similar methods were adopted for specific characteristics of the face such as the nose and eye boundaries. It was discovered that sexual dimorphism is more evident in

the lower third of the face compared to the upper part which showed less gender difference. In this study such methods will not be explored but it is of interest to be aware of such previous findings.

In White (2004) it was found that in a sample of 83, 3 month old infants significant differences were found between males and females in specific large facial dimensions such as the face height and ear-to-chin. It was found that the (ch-ch) width of females lips were narrower compared to males by 1.1mm on average.

5.2 Sexual Dimorphism in Control Data

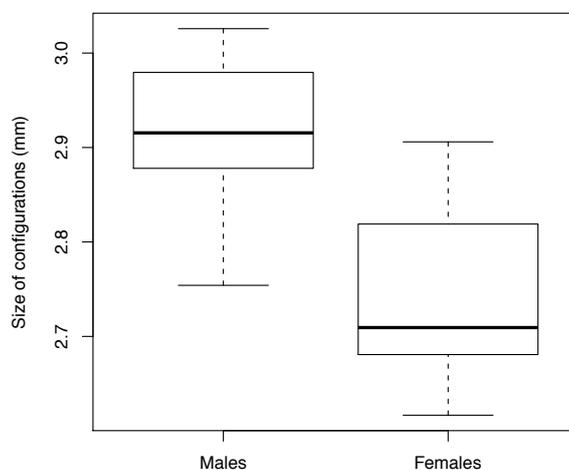


Figure 5.1: Size of configurations for males and females.

It is of interest to investigate the statistical difference in facial shape between males and females in the control data set. Firstly, in order to measure any difference in size between males and females, the centroid size is obtained for each configuration. The centroid size of a configuration is calculated as the square root of the sum of squared distances of a set of landmark coordi-

nates from their centroid. The results of the male and female configuration sizes are shown in Figure 5.1, scaled down by a factor of 100 for simplicity. There is a clear difference in the mean configuration sizes between sexes, with females having smaller configurations sizes than males. This could be attributed to females having smaller faces in general than males. The range of configuration sizes appears to be similar for both sexes. A two sample t-test found the difference in size between sexes to be statistically significant at the 1% level. This coincides with the findings in Ferrario (1995) where the mean area, delimited by a soft-tissue facial outline, was found to be significantly larger in men than women.

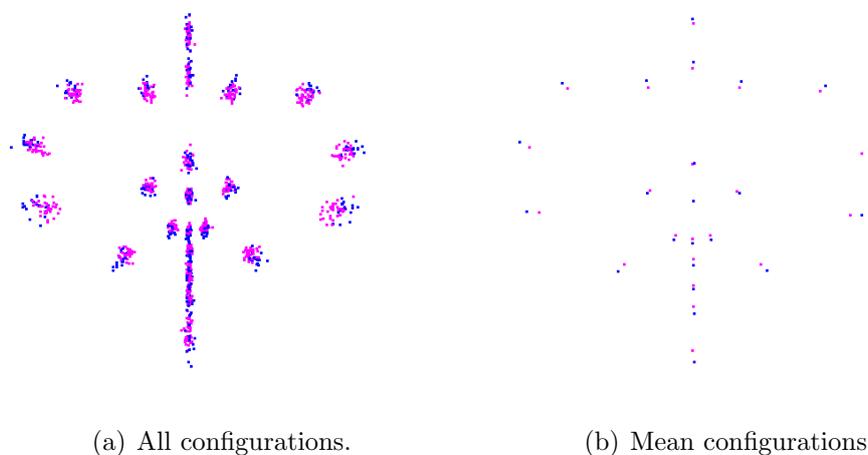


Figure 5.2: Frontal view of all control data landmarks and the mean positions for males (*blue*) and females (*pink*) without adjustment for scale.

In Figure 5.2 positioning of landmarks for both males and females after Generalized Procrustes registering without scale adjustment can be seen. These size-and-shape configurations illustrate that there is a clear difference in facial size between males and females. Females appear to have smaller faces than males within the control participants. It is of interest to establish if a difference in shape is present between males and females therefore GPA is ap-

plied again to the control landmark data, this time with a scale adjustment. The positioning of landmark configurations with scale adjustment applied can be seen in Figure 5.3.

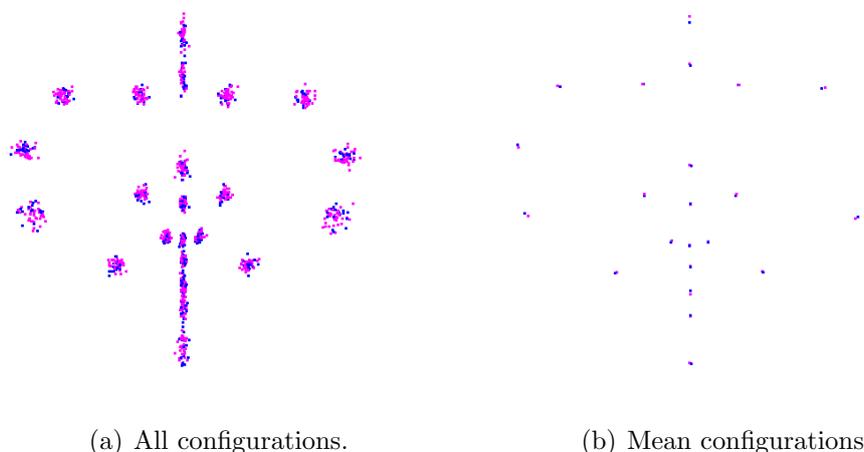


Figure 5.3: Frontal view of all control data landmarks and the mean positions for males (*blue*) and females (*pink*) with adjustment for scale.

On initial inspection of 5.3(b) a difference between the mean landmark locations of males and females is not obvious. The landmark at the top of the nose, (*sellion*), appears to be higher up the face for females than males on average. For the control data, the *soft tissue nasion* appears higher in females than males.

From a Hotelling's t-test based on re-sampling (based on 1000 permutations), which tests the null hypothesis that the mean facial shape of males is equal to that of females, a significant difference at the 5% level is found. It is worth noting that this result is only just significant as it is only just below the significant cut-off of 0.05 (p-value = 0.045) and should therefore not be considered a strong result.

Distances for width and height of the facial attributes were calculated for the size-and-shape configurations and for the size configurations in Tables 5.1 and 5.2 respectively. The mean distances between landmarks overall and the mean distances split by sex, are quantified along with a Pearson's correlation value for intra-landmark distances and % body fat. Calculating the Pearson's correlation provides a method of identifying if and which intra-landmark distances are likely to change with a change in % body fat in the wider population.

For the landmark configurations representing size-and-shape of the control faces, four of the eight intra-landmark distances are significantly different between sexes at the 5% level. The (chR-chL) mouth width is significantly narrower in females compared to males at the 1% level, with a difference of approximately 4.3mm on average. The outer (exR-exL) eye width is significantly narrower in females compared to males at the 1% level, with a width difference of approximately 4.9mm on average.

For the landmark configurations that have been adjusted for scale, i.e. those representing only shape of the control faces, only two intra-landmark distances are significantly different between sexes at the 5% level. As was suggested in Figure 5.3(b), the *soft tissue nasion* to *sellion* height is significantly longer in females, compared to males by approximately 1.6mm. It would not be unreasonable to suggest that the difference in location of the *soft tissue nasion* between sexes is a result of the way in which the configurations are superimposed onto one another in the GPA, rather than a true reflection of the differences in location of the *soft tissue nasion* itself between sexes. The (ls-li) lip height is borderline significant at the 5% level. The distance between the *labiale superius* and the *labiale inferius* is greater in females by approximately 1.6mm, compared to males. The (*chelion-chelion*) width of the lips appears to be very slightly narrower for females than males by 1.0mm

on average which is in agreement with White (2004).

There are no high correlations between any of the intra-landmark distances and % body fat in either Table 5.1 or 5.2. Higher correlations are prevalent in the shape-and-size configurations than in the intra-landmark distances for the configurations with adjustment for scale. The largest correlation in Table 5.1 is a negative correlation between (chR-chL) mouth width and % body fat of 0.38, which is not a strong relationship. As might be expected, the largest correlation exists in Table 5.2 between (n-se) upper nasal length and % body fat with a positive correlation of 0.28. Again, this is not a strong relationship.

The configurations of males and females can also be compared using Principal Component Analysis as described in Section 2.5.2. We perform this analysis on the set of control landmark configurations which have been Procrustes registered and adjusted for scale. Figure 5.4 shows the cumulative amount of variation explained by each additional principal component.

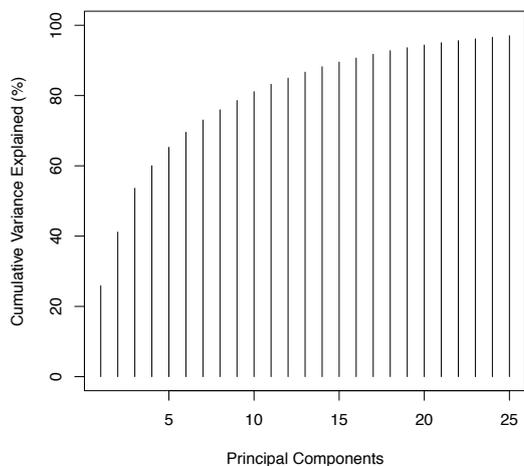


Figure 5.4: Cumulative proportion of variability explained by Principal Components.

The first two principal components explain 41% of the variation in the data. Over 90% of all the variation in the data is explained after the 16th principal component. Figures 6.16(a) and 6.16(b) show boxplots of the 1st and 2nd principal component scores, split by sex. In both figures there appears to be a difference between the sexes, however a set of two sample t-tests comparing the scores between males and females gives a non-significant results for both the

Table 5.1: Facial distances without adjustment for scale (mm).

<i>Measurement</i>	<i>Female</i>		<i>Male</i>		<i>Overall</i>		<i>Correlation</i>		<i>Significant</i> <i>Male/Female</i> <i>Difference</i>
	<i>Mean ± SD</i>	<i>Mean ± SD</i>	<i>Mean ± SD</i>	<i>Mean ± SD</i>	<i>With %Body Fat</i>	<i>Fat</i>			
Mouth									
Mouth width	chR-chL	48.5±3.1	52.8±3.1	49.9±3.7	-0.38				**
Philtrum width	cphR-cphL	11.8±1.5	12.7±1.1	12.1±1.5	-0.24				*
Lip height	ls-li	16.8±2.6	16.3±2.6	16.6±2.6	0.15				
Eyes									
Outer eye width	exR-exL	92.6±3.8	97.5±4.0	94.3±4.4	-0.33				**
Inner eye width	enR-enL	33.1±2.7	34.6±2.9	33.5±2.8	-0.18				
Nose									
Upper nasal length	n-se	15.9±2.4	15.4±2.3	15.8±2.4	0.17				
Lower nasal length	se-pm	41.2±3.0	43.4±4.6	41.9±3.7	-0.37				
Nasal base width	acL-acR	30.5±2.1	32.2±2.9	31.0±2.5	-0.14				*

^a* significant at 5% level.^b** significant at 1% level.

Table 5.2: Facial distances with adjustment for scale (mm).

<i>Measurement</i>	<i>Female</i>		<i>Male</i>		<i>Overall</i>		<i>Correlation</i>		<i>Significant</i> <i>Male/Female</i> <i>Difference</i>
	<i>Mean ± SD</i>	<i>Mean ± SD</i>	<i>Mean ± SD</i>	<i>Mean ± SD</i>	<i>With %Body Fat</i>	<i>Fat</i>			
Mouth									
Mouth width	chR-chL	49.6±3.2	50.6±2.5	49.9±3.0	-0.15				
Philtrum width	cphR-cphL	12.1±1.5	12.2±1.0	12.1±1.4	-0.09				
Lip height	ls-li	17.2±2.4	15.6±2.4	16.6±2.5	0.27				†
Eyes									
Outer eye width	exR-exL	94.9±3.7	93.4±2.7	94.4±3.4	0.08				
Inner eye width	enR-enL	33.8±2.6	33.1±2.7	33.6±2.6	0.04				
Nose									
Upper nasal length	n-se	16.3±2.4	14.7±2.2	15.8±2.4	0.28				*
Lower nasal length	se-pm	42.1±3.0	41.6±4.3	42.0±3.4	-0.17				
Nasal base width	acL-acR	31.2±2.2	30.9±2.6	31.1±2.3	0.10				

^a* significant at 5% level.^b† borderline significant at 5% level.

1st and 2nd principal components.

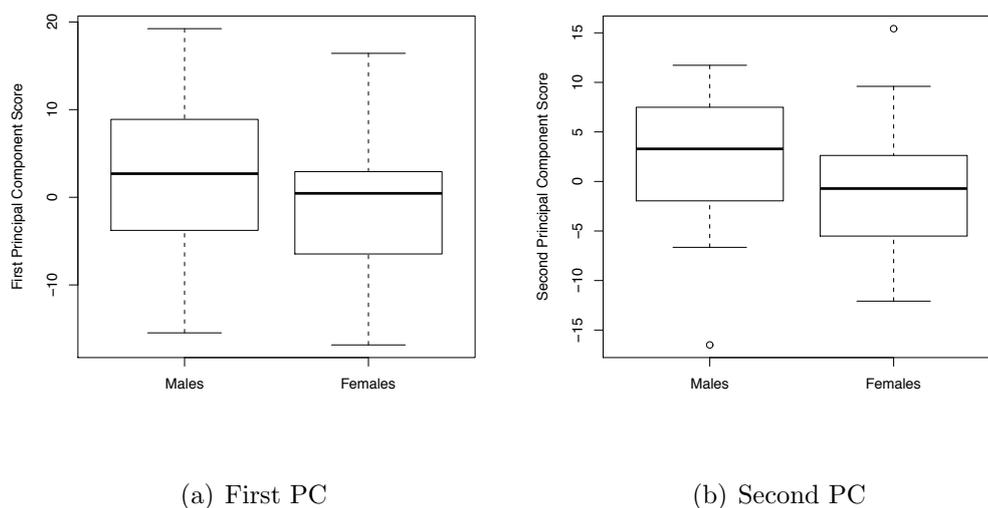


Figure 5.5: Boxplots of 1st and 2nd Principal Components scores split by Sex.

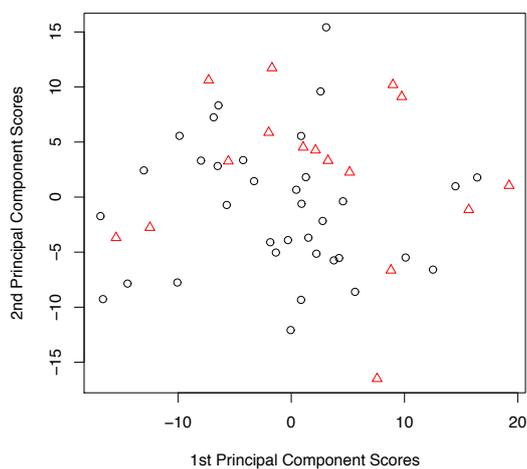


Figure 5.6: Plot of the first against the second Principal Component scores, split by Sex; males (*red triangles*), females (*black circles*).

The first and second principal component scores are plotted in Figure 5.6 for each sex. No clear separation between males and females is visible from this plot indicating that little or no difference is present in facial shape is present

between sexes.

As detailed in Section 2.4.1, the linearised version of the shape space in the vicinity of the Procrustes mean shape pole can be described as the tangent space. The Procrustes tangent coordinates of each configuration are obtained for the 16 males and 35 females with respect to the pre-shape Procrustes mean shape. These tangents can be used to represent the facial shapes of each participant and are independent of the location, orientation and size in space. A Hotelling's T^2 -test found differences between sets of tangent coordinates for males and females at the 5% significance level. The results of comparisons made between tangent coordinates for each sex by a two-sample t-test are summarised for each landmark in Table 5.3. Statistically significant differences between males and females were identified at 7 of the anatomical landmarks at the 5% level.

Table 5.3: Comparing landmark positions between Males and Females.

ID	Landmark	x	y	z	ID	Landmark	x	y	z
1	<i>pronasale</i>				13	<i>labiale superius</i>			
2	<i>left alare crest</i>				14	<i>left crista philtra</i>			
3	<i>right alare crest</i>				15	<i>right crista philtra</i>			
4	<i>subnasale</i>				16	<i>left chelion</i>			
5	<i>sellion</i>				17	<i>right chelion</i>			
6	<i>soft tissue nasion</i>		*		18	<i>stomion</i>			
7	<i>left exocanthion</i>			*	19	<i>labiale inferius</i>			
8	<i>right exocanthion</i>			*	20	<i>sublabiale</i>			
9	<i>left endocanthion</i>			*	21	<i>soft tissue gnathion</i>		*	
0	<i>right endocanthion</i>			*	22	<i>left otobasion inferius</i>			
11	<i>left tragion</i>				23	<i>right otobasion inferius</i>			
12	<i>right tragion</i>		*						

¹* significant at 5% level.

5.3 Body Fat

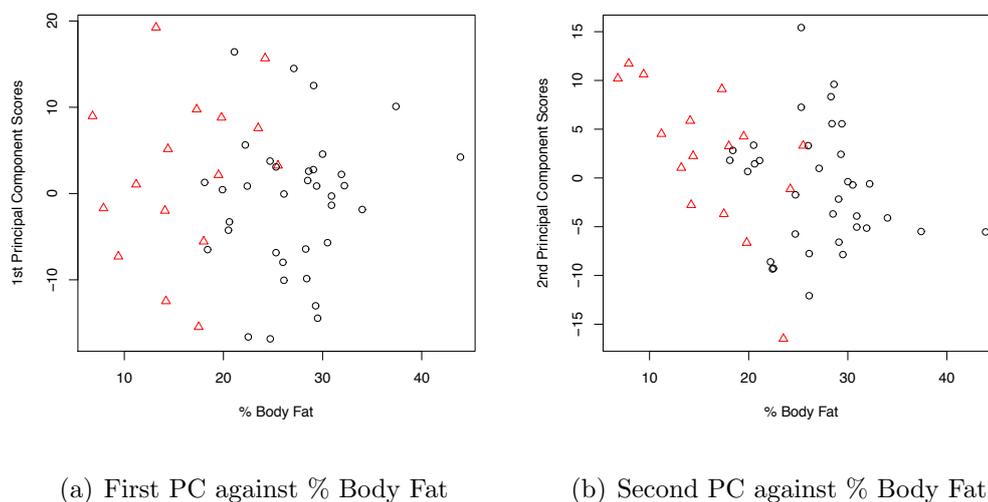


Figure 5.7: Plots of first and second Principal Components scores against % Body Fat, split by Sex; males (*red triangles*), females (*black circles*).

It is well documented and a commonly known fact that females carry more body weight than males in general. In Figure 3.5(b), a noticeable difference in percentage body fat can be seen between male and female control participants. It is therefore reasonable to assume that any difference in facial shape between males and females, may be attributed to a difference in % body fat. In the control data, a strong significant difference was found in % body measurements between the sexes, with females having a body fat between 7.7-14.7% higher than males on average. The first and second principal components plotted against % body fat are shown in Figure 5.7. In both figures, a clear separation is present between males and females.

For two females in the study, body fat measurements were missing. In order to adjust the data for body fat in this case, an estimate of of body fat was calculated from the participants weight measurements. Figure 5.8 illustrates the strong positive linear relationship present between female's Weight and %

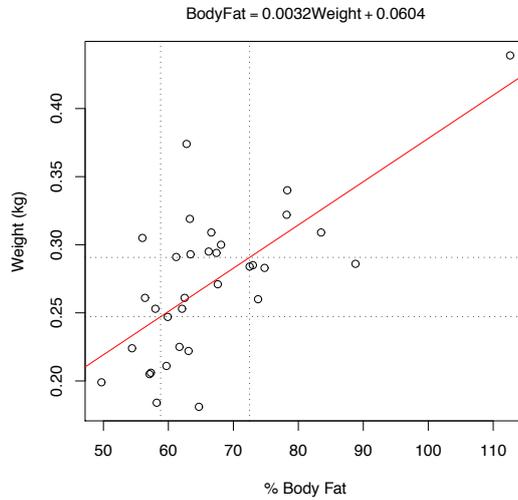


Figure 5.8: Female % Body Fat plotted against Weight (kg).

Body Fat measurements, (correlation of 0.69). Perpendicular sets of ‘dotted’ lines on the plot indicate the estimate of where the missing values of % Body Fat measurements would lie in relation to the rest of the data. Using the following linear model

$$\%BodyFat = 0.0032 \times Weight + 0.0604$$

estimates of missing %Body Fat values can be calculated.

It is of interest to establish whether the differences in facial shape found between males and females is a result of differences in % Body Fat between the sexes. The landmarks for all males and females separately are adjusted for % Body Fat. This is achieved by fitting a linear model, where the original configuration is the response and the corresponding body fat measurement for that participant is the variable for each dimension and landmark. A new adjusted landmark value for each dimension is created by adding the residual values from the linear model described to the corresponding original values in the configuration. From a Hotelling’s T^2 -test (based on 1000 permutations), comparing the mean shape between male and females after adjustment for body fat, no significant difference is found at the 5% level. This indicates

that the difference in shape between males and females observed previously in the landmark configurations, could be attributed to differences in % Body Fat between sexes.

5.4 Landmark Analysis of Midsagittal Profile

Previously we have investigated the landmark orientation from a front facing perspective. Variation in facial shape can also be visualised between participants in the control group by looking at the side-on view of their profiles. It is of interest to analyse the profile shape of control participants represented by the landmarks down the center of the face, (the *midsagittal line*).

A subset of nine landmarks, lying on the midsagittal line of participants faces were identified. From the *soft tissue nasion* to the *soft tissue gnathion* and all landmarks in-between which lie on the midline of the face. These subsets of landmarks will represent the facial profile of each participant. All 51 of the (9×3) configurations were registered by GPA, initially with no adjustment for scale made. To visualize the size-and-shape of the profile configurations, a side view of all profiles and a mean representation of participant's profiles can be seen in Figure 5.9.

The mean locations of the *sellion*, the *pronasale* and the *subnasale* indicate that on average, the nose is shorter for females than males for this control population. Large sex differences in mean locations of landmarks on the forehead (*soft tissue nasion* and *sellion*) and landmarks on the chin (*sublabiale* and *soft tissue gnathion*) seen in Figure 5.9(b) may be attributed to smaller faces in females on average compared to males. The landmark positions with scale adjustment for the profile configurations of landmarks can be seen in Figure 5.10.

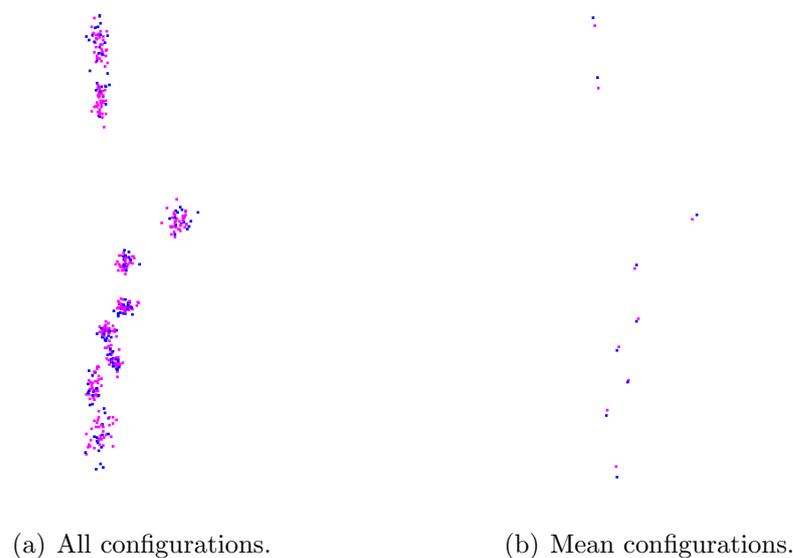


Figure 5.9: Side view of shape-and-size of control data midsagittal landmarks and the mean midsagittal positions for males (*blue*) and females (*pink*).

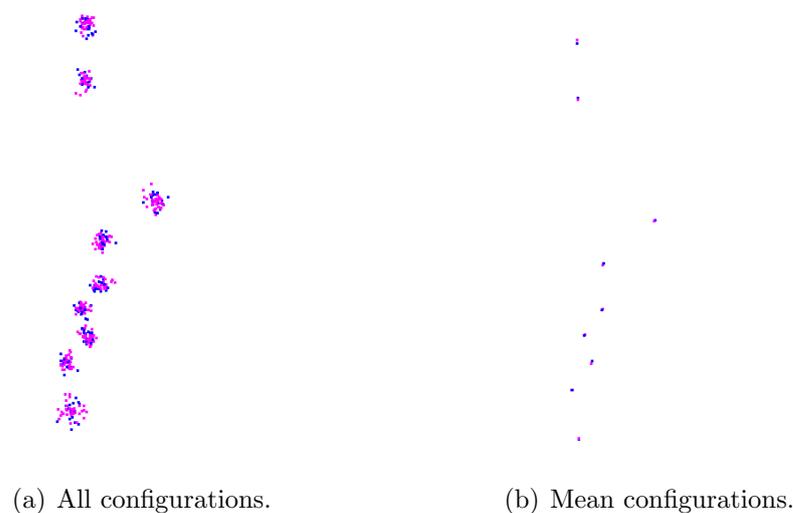


Figure 5.10: Side view of shape of control data midsagittal landmarks and the mean midsagittal positions for males (*blue*) and females (*pink*) with adjustment for scale.

The mean difference in profile landmark locations after GPA with scale adjustment is not immediately obvious between sexes on initial inspection of Figure 5.10(b). Again, the *soft tissue nasion* appears to be higher on the face for females, compared to males. The lower lip, *labiale inferius* appears

be lower on the facial profile in females, compared to males. The visual first impressions coincide with findings in Table 5.2, where significant differences in lip height and upper nasal length were found between sexes. After adjustment for scale is applied, an overall significant difference between the sexes is found from a Hotelling's T^2 test (based on 1000 permutations), on the subset of profile landmarks. This indicates that a difference in the profile shape is present between males and females in this control group, based on landmarks.

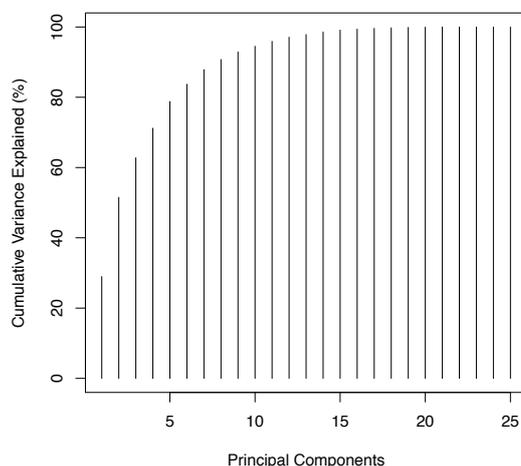


Figure 5.11: Cumulative proportion of variability explained by Principal Components.

The first and second principal component scores are plotted against % body fat in Figure 5.12, split by sex. For both plots, there is evidence of a difference between males and females which support our previous findings. As described in section 4.3, the configurations of profile landmarks are adjusted for by body fat. No obvious change from the positions of the landmarks to those seen in Figure 5.10 is observed. Another Hotelling's T^2 test (based on 1000 permutations) indicates that a significant difference is again present between the mean landmark profile shape of males and females. With a p-value of approximately 0.042, the difference is only just significant compared to the p-value of 0.036 before adjustment for body fat. This indicates that

PCA is used to describe the variability in the landmarks on the mid-sagittal profile. The first two principal components describe 51% of the variability in the data. Over 90% of the entire variability in this subset of landmarks coordinates for all control data is described by the first 8 principal components.

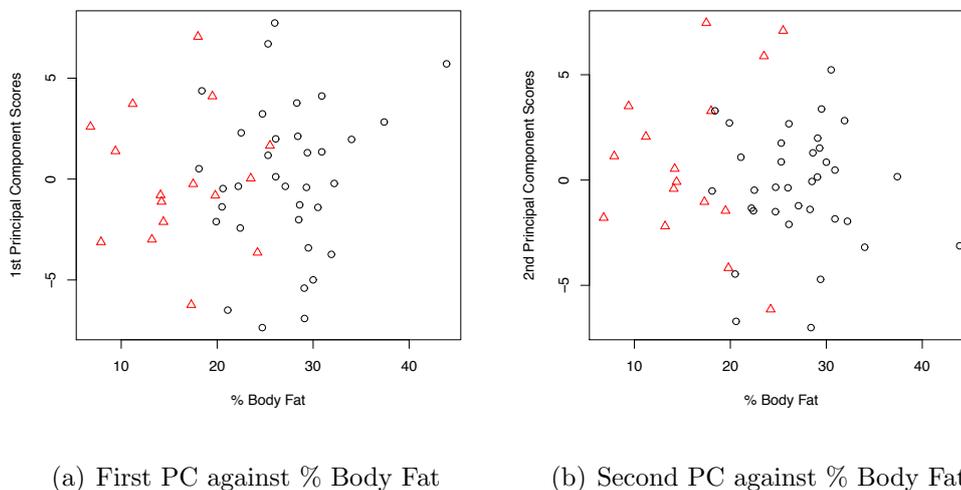


Figure 5.12: Plots of first and second Principal Components scores for midsagittal landmarks against % Body Fat, split by Sex; males (*red triangles*), females (*black circles*).

initial dimorphism witnessed for the midsagittal landmarks between males and females can be attributed partially to differences in % Body Fat. Given that the difference in the p-values is relatively small and that no obvious difference is present between plots of mean profile landmarks before and after body fat adjustment, analysis will proceed with no adjustment for body fat, the argument being that by correcting for this factor, information may be being lost about the variation of the landmark configurations.

Box plots of the first and second principal components for the midsagittal landmark configurations can be seen in Figure 5.14, split by sex. On first inspection there does not appear to be a large difference between the sexes for the first or second principal component scores. A pair of Student's t-tests confirm that no statistically significant difference is present in scores between sexes, for the first or second principal components.

The first and second principal components, obtained for the midsagittal land-

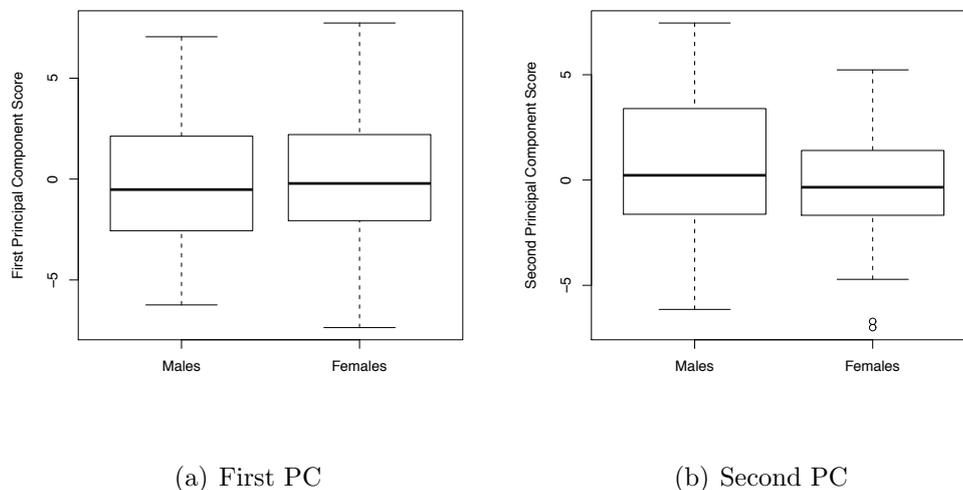


Figure 5.13: Boxplots of first and second Principal Components scores for mid-sagittal landmarks split by Sex.

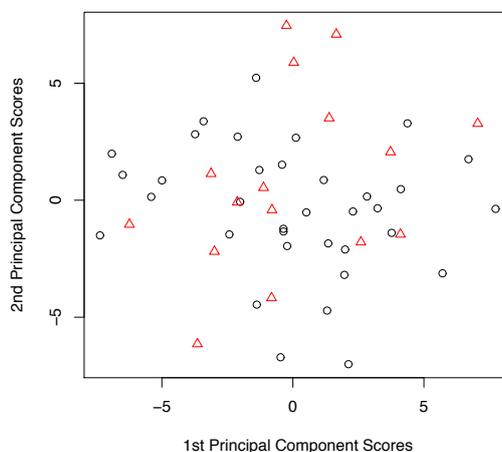


Figure 5.14: Plot of the first against the second Principal Component scores for midsagittal landmarks, split by Sex; males (*red triangles*), females (*black circles*).

marks are plotted against each other in Figure 5.13 and split by sex. From the first and second principal components, no obvious split between males and females is present. Clustering of sexes separately in Figure 5.13 would suggest that the principal components describe difference sources of the variance for males and females. In this instance however, all the variation in

the data appears to be described by a combination of both male and female profile shapes, for the first two principal components. Although the Hotelling's T^2 -test (with 1000 permutations) found a significant difference between sexes, the evidence from the principal component scores suggests that this difference is small.

5.5 Landmark Analysis Conclusions

Evidence of sexual dimorphism was found in the landmark configurations representing the size-and-shape and the shape of the population of control data.

It was found that after landmark adjustment for % Body Fat, no statistical evidence of sexual dimorphism was found in the shape of the control landmark configurations. It can be concluded that the sexual dimorphism witnessed prior to this adjustment may be a result of sex differences in % Body Fat rather than facial shape.

Four intra-landmark distances were observed to be significantly different in length between sexes for the size-and-shape configurations of landmarks. The largest differences were observed for the width of the eyes and mouth between sexes.

Two intra-landmark distances were observed to be significantly different in length between sexes for the shape of the landmark configurations. The largest difference in landmark position between sexes was that of the *soft tissue nasion*. The variation of identifying this particular landmark was seen to be high in the validation study, (with a TEM score of 3.55mm). It is therefore not unreasonable to speculate that this observation may be a result of

GPA matching and high variability in identification rather than evidence of true difference in location of the *soft tissue nasion* between sexes.

For the analysis of the nine midsagittal landmarks, evidence of sexual dimorphism was found in the shape of the configurations. After and adjustment for % Body Fat for the location of the midsagittal landmarks, sexual dimorphism was still prevalent between the shapes of the configurations. It is possible that the areas of the face which change most with higher or lower body fat are not represented by the nine landmarks on the midsagittal line. Hence, the effect of adjusting for % Body Fat was not witnessed in this subset of configurations. PCA suggested that very little difference was present between the midsagittal landmarks of males and females, despite formal analyses indicating evidence of sexual dimorphism.

It is worth remembering at this stage that the control data contains a sample of only 16 males and 35 females. In a larger study, evidence of sexual dimorphism may be more prevalent.

Chapter 6

Curve Identification in Control Data

It is of interest to map the facial shape by obtaining the curves lying upon the facial shape surface. This section will detail the method of identifying curves from the face. A specific method of identifying curves from the upper lip will be explored and the curves mapping the ridges of the philtrum will be analysed. A general method of curve identification will be described. The specific method of identifying the curve down the midsagittal profile of the facial shape will also be explored, and the midsagittal curves analysed. Evidence of sexual dimorphism present between the male and females curves will be examined. Finally, problems encountered whilst mapping the curves on the faces will be discussed.

6.1 Methodology for Curve Identification

The method of curve identification used is detailed extensively in Bowman et al. (2011). The specific method of identifying curves on the upper lip philtrum will be detailed in Section 6.2. The method of curve identification will be described in general terms in Section 6.3. Specific methods of curve identification for the midsagittal profile curve will then be detailed in Section

6.4. Principal curves are used extensively in identifying the curves of the facial surface in this thesis.

6.1.1 Principal Curves on Facial Surface

In Hastie and Stuetzle (1989) principal curves are described as “*smooth one-dimensional curves that pass through the middle of a p -dimensional data set, providing a nonlinear summary of the data*”. A one-dimensional curve in \mathbf{R}^p can be denoted as a vector $\mathbf{f}(\lambda)$ of p coordinate functions for which λ provides an ordering along the curve. (In this instance, $p = 3$ as the surface of a face is a three-dimensional, upon which we wish to place principal curves.) The basic idea of a principal curve is that for any parameter value λ , the mean is calculated for all observations that have $\mathbf{f}(\lambda)$ as their closest point on the curve. \mathbf{f} is called a principal curve if $\mathbf{f}(\lambda)$ is the average of all observations at that point along the curve, for all values of λ . A principal curve can therefore be described as a non-parametric generalisation of a linear principal component. For the purpose of this analysis principal curves are calculated using a specific function in \mathbf{R} called ‘princurve’ (Weingessel (2009)).

Suppose \mathbf{X} is an $(n \times p)$ matrix of n observations in \mathbf{R}^p with probability density $h(\mathbf{x})$. Let \mathbf{f} denote a class of differentiable 1-dimensional curves (C^∞) in \mathbf{R}^p , parameterized by λ . The projection index $\lambda_{\mathbf{f}}$ can be defined as

$$\lambda_{\mathbf{f}}(\mathbf{x}) = \max_{\lambda} \{ \lambda : \|\mathbf{x} - \mathbf{f}(\lambda)\| = \inf_{\mu} \|\mathbf{x} - \mathbf{f}(\mu)\| \}. \quad (6.1)$$

where $f(\mu)$ denotes the mean values of all \mathbf{x} observations which have $\mathbf{f}(\lambda)$ as their closest point on the curve. More simply, we can say that the projection index $\lambda_{\mathbf{f}}(\mathbf{x})$ of \mathbf{x} is the value of λ for which $\mathbf{f}(\lambda)$ is closest to \mathbf{x} .

For each λ value of a curve, there is a corresponding value of curvature at that point. For a curve in three-dimensional space, curvature can be defined

as follows by Cartesian coordinates x , y and z :

$$Curvature = \frac{\sqrt{(x'y'' - x''y')^2 + (x'z'' - x''z')^2 + (y'z'' - y''z')^2}}{(x'^2 + y'^2 + z'^2)^{3/2}} \quad (6.2)$$

where the prime represents differentiation with respect to λ . For any principal curve, locations of interest can be obtained by identifying points of minimum and maximum curvature using simple methods of differentiation.

6.2 Curve Identification of Philtrum

The method of curve identification will now be discussed for the area of the upper lip. It is of interest to extract the vertical ridges which map out the philtrum on the top lip. All faces have been rotated to face directly towards the observer, by the same method previously described.

The area located on the upper lip from which we wish to identify curves can be extracted as a subset of data from the entire set of face coordinates. The *subnasale* and the lowest *christa philtri* landmark in the y dimension are used as horizontal cutoff points for the subset of upper lip coordinates. Similarly, the x dimension location of the left and right *chellion* landmarks are used as the vertical cutoffs points for the subset of upper lip coordinates. Hence a new subset surface area of the entire facial shape is created for the area of the upper lip.

A new three-dimensional origin is created by standardising all the points in the subset of data which characterise the upper lip and philtrum. It is from this origin that we can set dimension y equal to zero and therefore create a plane on the x and z dimensions. From the *subnasale* to the lowest *christa philtri* of the upper lip surface area, thirty principal curves are fitted orthogonally to the surface of the upper lip at evenly spaced intervals on the y dimension. The principal curves are fitted to the three-dimensional strips of

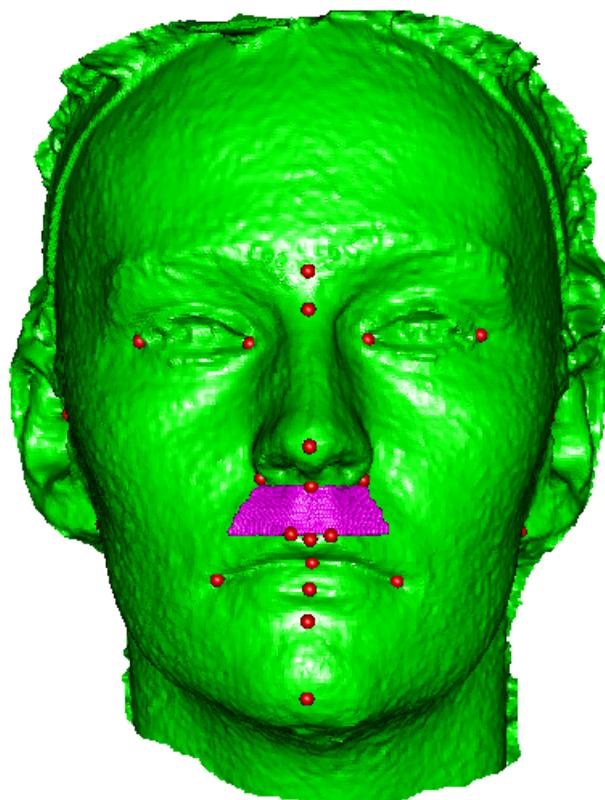


Figure 6.1: Horizontal Principal Curves on upper lip from *subnasale* to the lowest *chista philtri* in the y dimension.

data across the upper lip area at each interval. The thirty horizontal principal curves fitted to the subset of upper lip coordinates can be seen in Figure 6.1. It can be seen in Figure 6.1 that the orthogonal principal curves capture the information in the subset of upper lip data coordinates by describing the curvature across upper lip at thirty evenly spaced increments.

Points of maximum and minimum curvature are identified for each of the thirty principal curves running horizontally across the upper lip. Figure 6.2 shows the plot of curvature versus arc length of a single principal curve on the upper lip. Smoothing techniques are applied to each of the horizontal principal curves on the upper lip so as to limit the number of extremes iden-

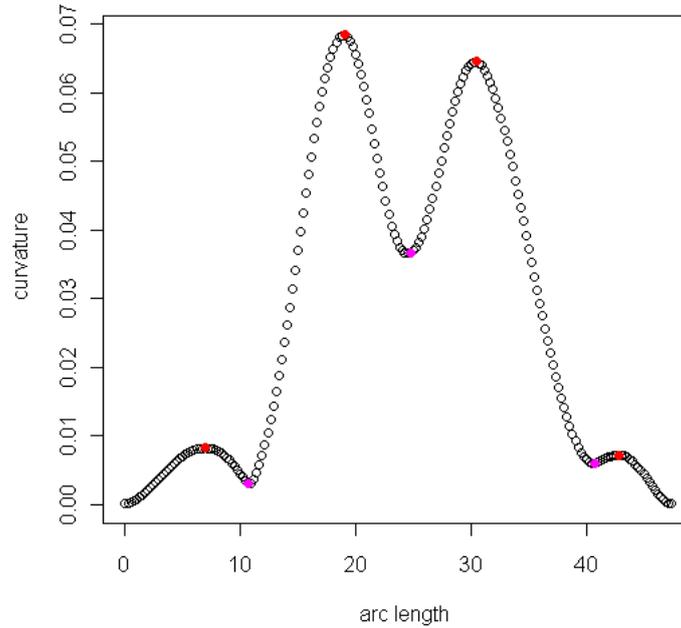


Figure 6.2: Curvature versus arc length (mm) of a horizontal Principal Curve on the upper lip.

tified, (Bowman and Azzalini (1997)). The points of maximum (*pink*) and minimum (*red*) curvature on the principal curve are highlighted. It is only of interest to obtain the maximum points of curvature which represent the philtrum ridges on the upper lip. The extreme coordinates are allocated to a $(k \times 3)$ matrix for either the left or right philtrum ridge depending on their location on the x -axis, where k is the number of extreme points allocated to that ridge and the number of dimensions is 3. A threshold is set up on the x dimension between the left and right *crista philtri*, minus and plus a constant c respectively. The location of the *labiale superius* landmark on the x -axis is the middle boundary used to determine which of the two ridges an extreme point is allocated to. A maximum extreme is allocated to the left ridge if, on the x -axis, it lies between $(chL - c)$ and ls and allocated to the right ridge if, on the x -axis, it lies between ls and $(chR + c)$.

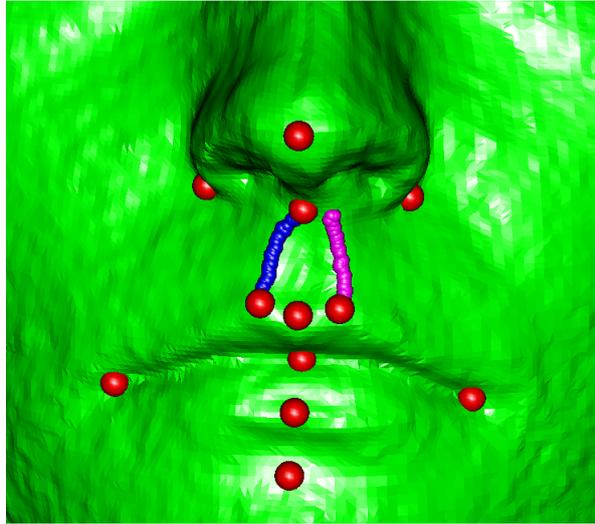


Figure 6.3: Maximum points of curvature on horizontal Principal Curves on upper lip.

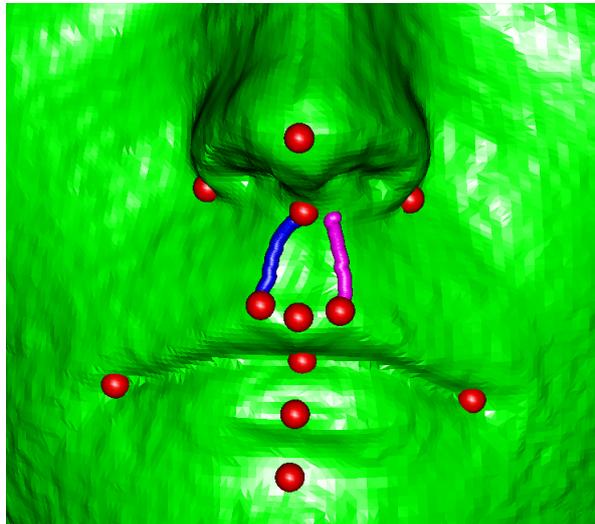


Figure 6.4: Vertical Principal Curves on upper lip calculated from maximum points of curvature.

The points of maximum curvature allocated to the left philtrum ridge (*blue*) and the right philtrum ridge (*magenta*), can be seen in Figure 6.3. Due natural variation between people's faces, the number of k coordinates of maximum extremes in each ridge differs between ridges and between faces in the control data.

From the extreme points identified on the philtrum ridges, vertical principal curves can be calculated. Figure 6.4 shows the principal curves calculated from sampling from the maximum extremes seen in Figure 6.3. The principal curves on the philtrum ridges consist of two (30×3) matrices. These principal curves can be used to represent the curves of the upper lip philtrum for each control face.

Difficulties encountered while identifying curves on the upper lip ridges are discussed in Section 6.5. Of the 51 control faces, 46 pairs of philtrum ridges were successfully identified; 15 males and 31 females.

6.2.1 Analysis of Philtrum Curves

The pairs of philtrum ridges will be analysed together rather than individually across all participants in the control data. The curves of the philtrum ridges are represented by the set of pseudo-landmarks. The vertical valley ridge between the philtrum ridges will not be analysed in this section as it is being covered in the analysis of the midsagittal profile curve, (Section 6.4.1).

All the configurations for the pairs of philtrum ridges were registered using Generalized Procrustes Analysis. In Figure 6.5(a) the size-and-shape configurations and the mean configurations for males and females can be seen. High variability is present across all participants and the curves identified appear wiggly. Immediately it is visible that one male participant has an upper lip that converges almost in the middle of his upper lip. No separation between sexes is immediately obvious from the figure of all 46 configurations. The mean upper lip ridges for males appear to be longer and more widely separated than those of females. A Hotelling's T^2 test (with 500 permutations) was used to test if the mean shape of the male and female philtrum ridges are equal. It gave a non-significant result indicating that statisti-

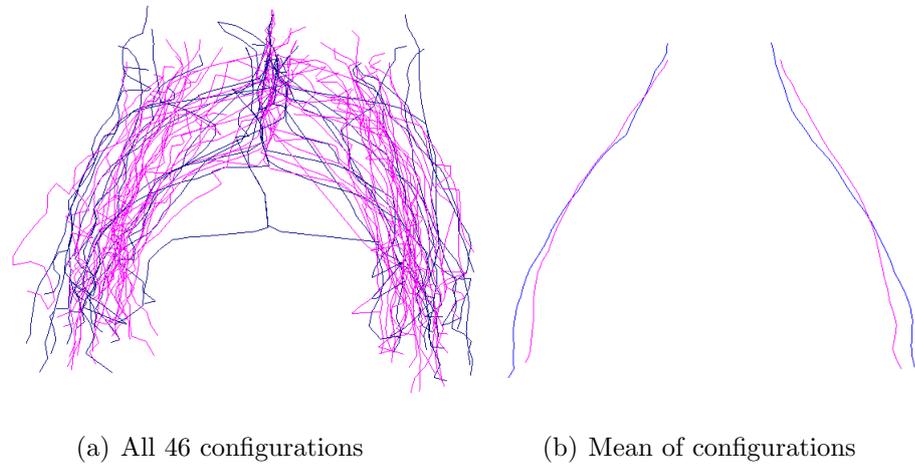


Figure 6.5: All registered configurations and the mean configurations for males (*blue*) and females (*pink*) with no adjustment for scale.

cally no difference in philtrum ridge shape is present between sexes before adjustment for scale. In order to adjust for the difference in size between males and females, GPA is used to register all the pairs of upper lip ridge configurations representing the curves and an adjustment for scale is applied.

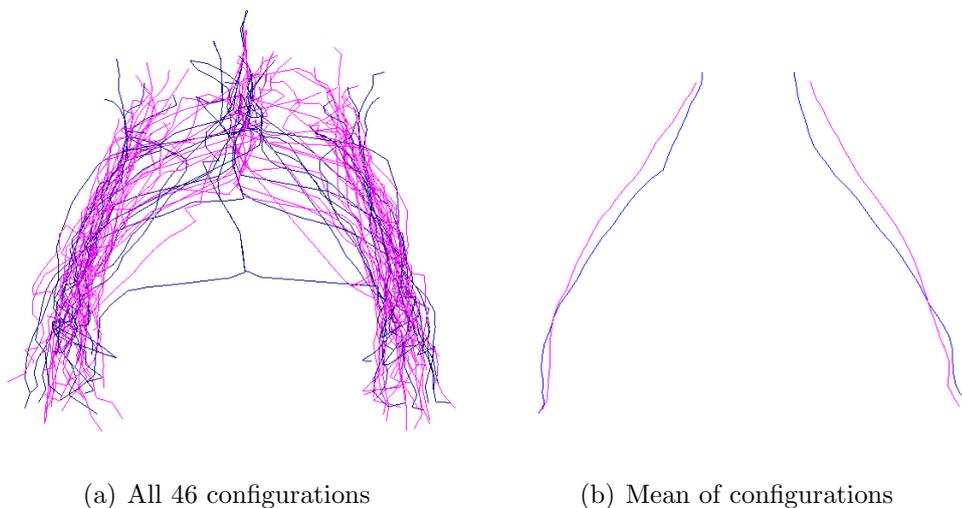


Figure 6.6: All registered configurations and the mean configurations for males (*blue*) and females (*pink*) after adjustment for scale.

In Figure 6.6, all configurations of upper lip ridges and the mean shapes for males and females can be seen. There appears to be less variation between configurations of upper lip ridges across all participants than that seen before scale adjustment. The mean shape of the male upper lip ridges are narrower at the top of the philtrum ridges than that for females. This result may be due to an outlier effect caused by the male participant with an upper lip that converges half way up his philtrum. In a larger sample of males it would become apparent whether or not this observation is rare and in which case could be removed. However, given that the sample upper lip ridges consists of only 15 males and 31 females, this set of philtrum ridges will not be removed from the data set. A Hotelling's T^2 test (with 500 permutations) is used to test if the mean shape of the male and female philtrum ridges is the same. This provides a non-significant result indicating that statistically no difference in philtrum ridge shape is present between sexes.

Principal Component Analysis can be performed on the upper lip curves to explain the variability in the data. The first two principal components explain, in total, over 78% of the variability in the philtrum ridge curves for the 46 participants and over 90% of the variability is explained by the first 5 principal components. In Figure 6.8, the first and second principal component scores are displayed in boxplots and split by sex. No obvious sex difference is visible between the scores of the first and second principal components for the philtrum ridge curves. Two two-sample t-tests provide non-significant results for differences between sexes for

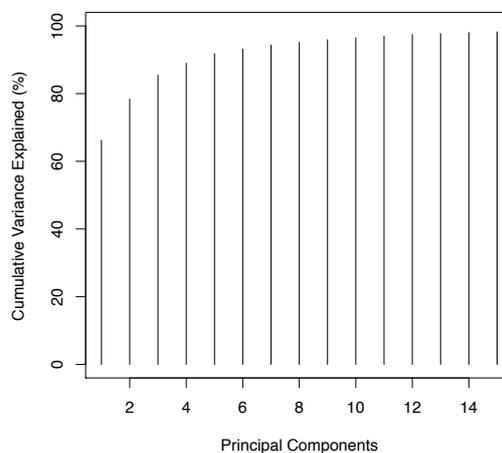


Figure 6.7: Cumulative proportion of variability in philtrum ridge curves explained by Principal Components.

the first and second principal component scores.

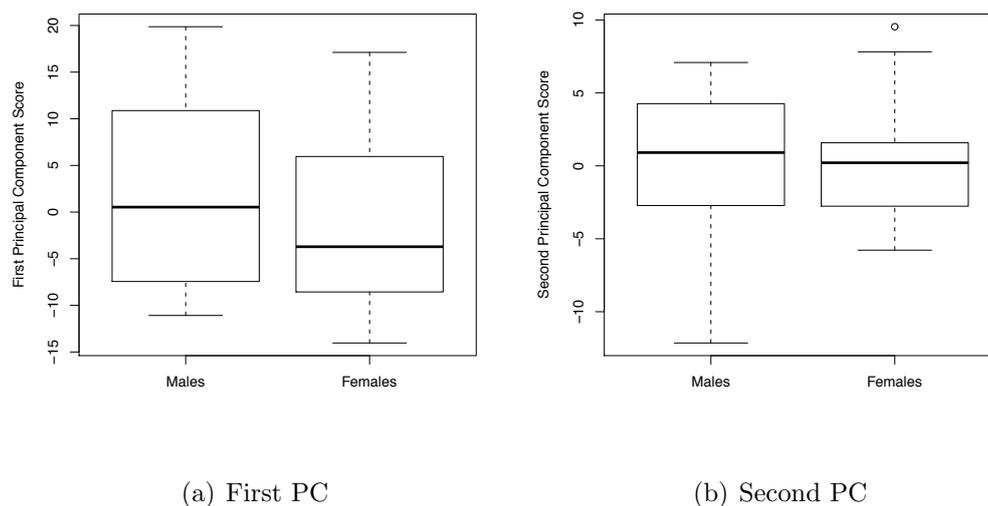


Figure 6.8: Boxplots of first and second Principal Components scores for philtrum ridge curves split by Sex.

To further investigate any presence of sexual dimorphism in the curves of the upper lip ridges, the first and second principal component scores are plotted against each other and split by sex. No separation or clustering is present between males (*red triangles*) and females (*black circles*) in Figure 6.9 indicating that little or no difference exists in the shape of the philtrum between males and females. The scores of the first and second principle component scores overlap each other between sexes indicating that the corresponding principal components describe the same variability in the data of the entire control population.

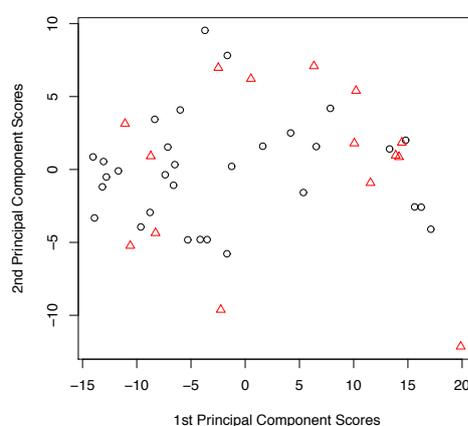


Figure 6.9: First Principal Component against second Principal Component, split by Sex; males (*red triangles*), females (*black circles*).

6.3 A General Method of Facial Curve Identification

A more general method of curve identification will now be described. This general method is implemented when identifying the curves on the philtrum in Section 6.2 and the midsagittal profile curve in Section 6.4. Each face in the control data set is made up of approximately 135000 coordinate values from which subsets of coordinates can be extracted for curve analysis. Each facial shape is different and for this reason, any method of curve identification is unlikely to be successful for every face. A general method of curve identification is described in the following steps:

Step 1 Identify the subset data coordinates of the facial area upon which you wish to identify curves, using anatomical landmarks to mark the boundaries.

Step 2 Establish the plane upon which the subset of coordinates lies and standardise all coordinates which lie on the subset of data so that the shape space of the coordinates has origin zero.

Step 3 On this subset of three-dimensional coordinates, fit a set of orthogo-

nal principal curves at evenly spaced increments so that all coordinates of the subset of data are captured by principal curves.

Step 4 Apply smoothing methods to the horizontal principal curves to identify the facial regions upon which the curves lie and to control the number of local and global points of extreme curvature that are identified.

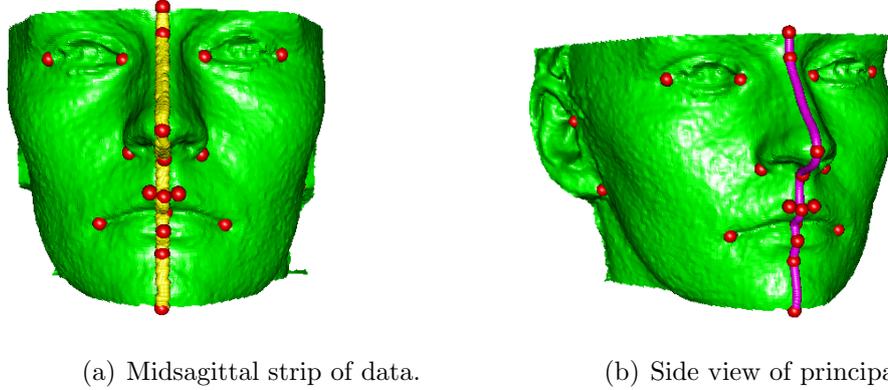
Step 5 Identify points of maximum (or minimum) curvature, (i.e. extremes) on each of the fitted horizontal principal curves, therefore identifying the specific facial location upon which the curve lies.

Step 6 Using the points of maximum (or minimum) curvature selected to sample from, fit a single principal curve. This new fitted principal curve can be used to represent the facial curve of interest.

6.4 Curve Identification of Midsagittal Profile

By applying similar methods described in Section 6.3, the curve on the Midsagittal profile of the face can also be obtained.

Using the landmarks on the midsagittal profile as a guide, a subset of the data coordinates on the midsagittal profile of facial surface obtained. The range of coordinates values in the x dimension of the landmarks on the midsagittal profile were calculated. The width of this three-dimensional midsagittal strip of coordinates is calculated by subtracting and adding a small amount to the ‘left’ and ‘right’ boundary of this range, respectively. The top and bottom boundaries of the midsagittal strip are bounded in the y dimension by the *soft tissue nasion* and the *soft tissue gnathion*. This midsagittal strip of data can be seen in Figure 6.10(a), highlighted in yellow. A principal curve is fitted to the data coordinates on the midsagittal strip. The estimated principal



(a) Midsagittal strip of data.

(b) Side view of principal curve.

Figure 6.10: Midsagittal strip of subset data and the Principal Curve fitted to this data.

curve fitted to the midsagittal subset of data can be seen in Figure 6.10(b). A midsagittal principal curve is obtained for the 51 control participants' facial shapes.

Each estimated midsagittal principal curve is evaluated at set points to produce a smaller (50×3) matrix of coordinates representing the principal curve. Using this re-evaluated principal curve, the original landmarks on the midsagittal profile are linearly interpolated by a designated number of pseudo-landmarks. Therefore, the same number of pseudo-landmarks are present between each landmark on the midsagittal profile. This is considered a more appropriate representation of the midsagittal curve as each of landmarks are directly comparable between participants in the control study. For a single participant, the landmarks, (*in red*) and the evenly spaced pseudo-landmarks, (*in blue*) can be seen in Figure 6.11.

6.4.1 Analysis of Midsagittal Curves

For all 51 control participants, the points representing the midsagittal principal curves are analysed. All configurations of points are registered us-

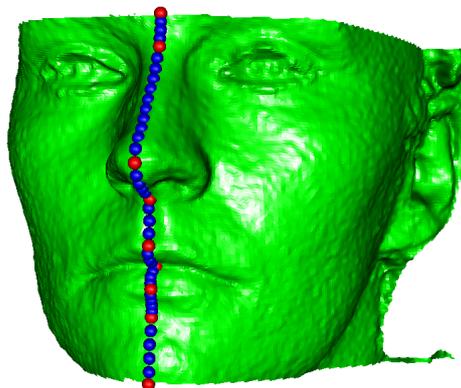


Figure 6.11: Landmarks and pseudo landmarks on midsagittal Principal Curve.

ing Procrustes methods. In Figure 6.12, the size-and-shape of all midsagittal curves and the mean midsagittal curves can be seen, split by sex. No adjustment for scale has been made at this point. Sexual dimorphism is apparent in Figure 6.12(b) in the size of the nose. Females appear to have smaller noses and shorter faces in general compared to males which reinforces our findings in Section 5.1. In order to establish any sexual dimorphism in profile shape of the midsagittal curves, an adjustment for scale is made.

The shape of all midsagittal curves and the mean curves with adjustment for scale can be seen in Figure 6.13, for both sexes. The presence of sexual dimorphism can be assessed by comparing the mean shape of the midsagittal principal curves for males and females. There is very little visual evidence of a difference between the mean midsagittal principal curves between males and females. On close inspection, there is a suggestion that females have slightly fuller lips and smaller noses than males, although this is not visually obvious.

The null hypothesis that no difference in shape exists between midsagittal curves for males and females can be tested using a Hotelling's T^2 test, (1000 permutations). The result of this permutation test is significant indicating that a statistically significant difference is present in the midsagittal shape

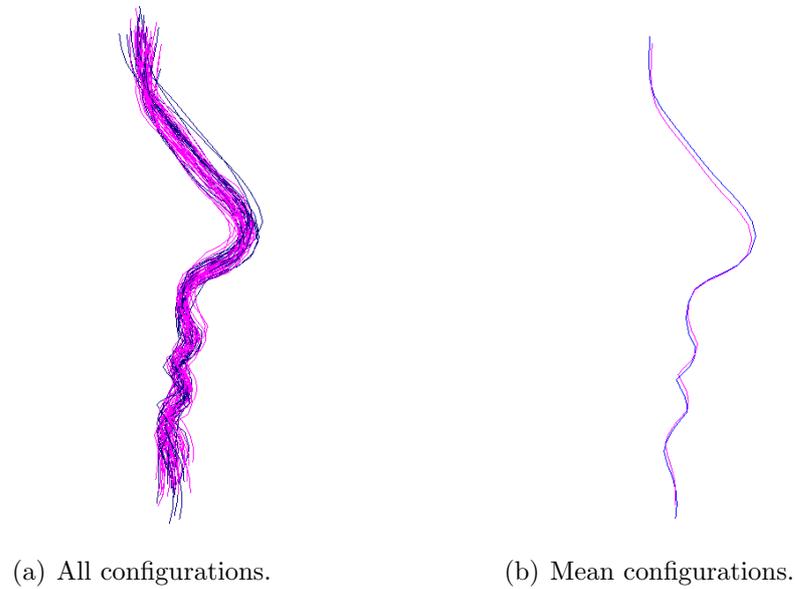


Figure 6.12: Principal Curves of midsagittal landmarks split by Sex; males (*blue*), females (*pink*).

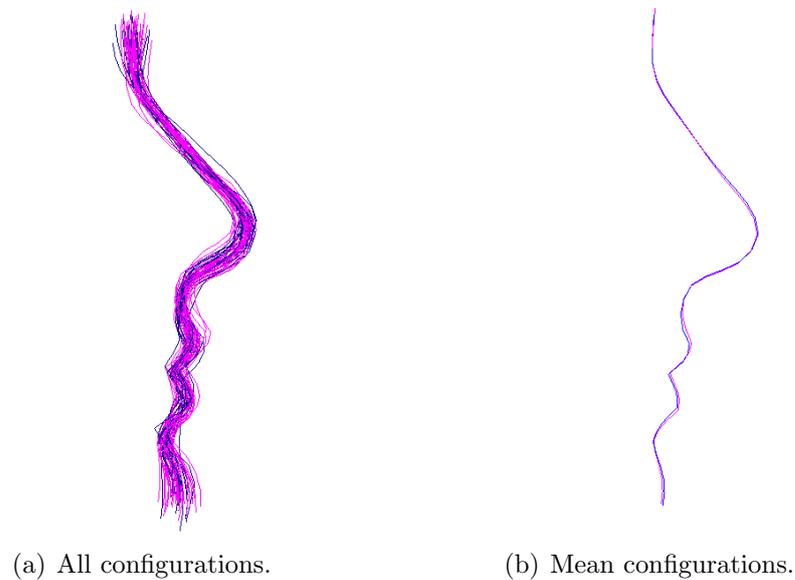


Figure 6.13: Principal Curves of midsagittal landmarks with adjustment for scale split by Sex; males (*blue*), females (*pink*).

between sexes at the 5% level. This however should not be considered to be a strong result as the p-value is 0.049, which lies very close to the threshold of significance.

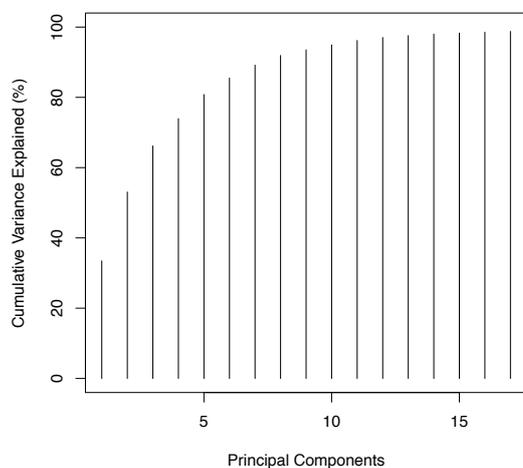


Figure 6.14: Cumulative proportion of variability in midsagittal curves explained by Principal Components.

To describe the variability in the midsagittal curves, principal component analysis can be performed. In Figure 6.14, the cumulative percentage of variability explained for each principal component can be seen. The first eight principal components explain over 90% of the variance in the data. The first and second principal components explain, in total, 53% of the variability in the data, 2% more variability than was explained using the landmarks on the midsagittal profile alone.



Figure 6.15: Midsagittal curves with adjustment for % Body Fat, split by Sex; males (*blue*), females (*pink*).

Plots of the first and second principal component scores, plotted against % Body Fat for the midsagittal curve data and split by sex are almost identical

to those seen in Figure 5.12. As previously observed in landmark analysis of the midsagittal curve, separation and clustering between males and females was witnessed in the plots of the first and second principal component scores against % Body Fat. Using previously described methods, the midsagittal curves for males and females are adjusted for by body fat. Figure 6.15 shows the mean midsagittal curves for males and females after adjusting the coordinates in the midsagittal principal curves for % Body Fat for each control.

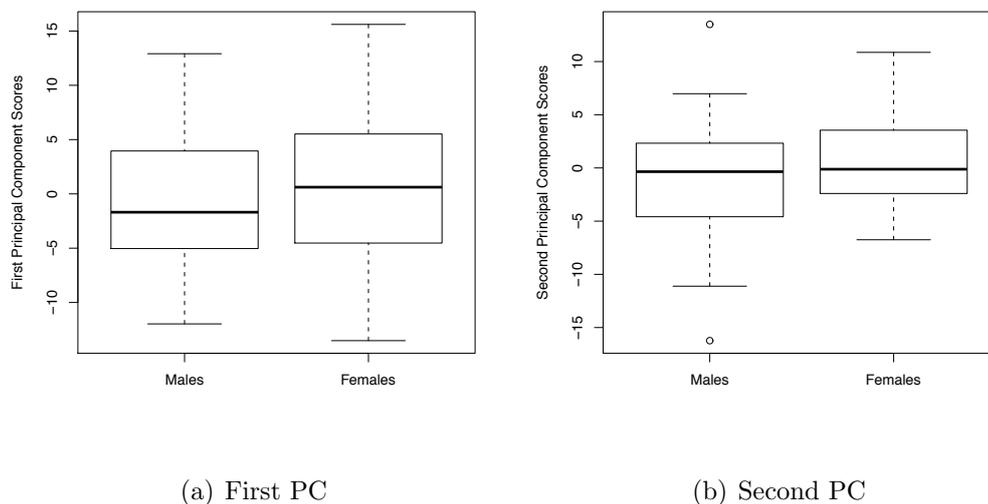


Figure 6.16: Boxplots of first and second Principal Components scores for midsagittal Principal Curves, split by Sex.

After adjustment for body fat, there are no obvious differences between the mean midsagittal curves for males and females. There is a suggestion of females having smaller noses, fuller lips and possibly a smaller chin. These differences between sexes are however very subtle. The mean midsagittal curves with adjustment for body fat are almost identical to those seen in Figure 6.13(b). A Hotelling's T^2 test (with 1000 permutations) gives a non significant result when testing the null hypothesis that the mean shapes of the midsagittal curves of males and females are the same after body fat adjustment. This result simply indicates that there is no difference in the mean

midsagittal curves is found between sexes even once the configurations of pseudo-landmarks representing the curves have been adjusted for by body fat. Analysis on the midsagittal curves will continue without adjustment for % Body Fat.

Box plots of the first and second Principal Components can be seen in Figure 6.16. For the first and second principal components, little difference between sexes is evident. A set of two-sample t-tests confirm that no statistically significant difference in scores is present between sexes for both the first and second principal components.

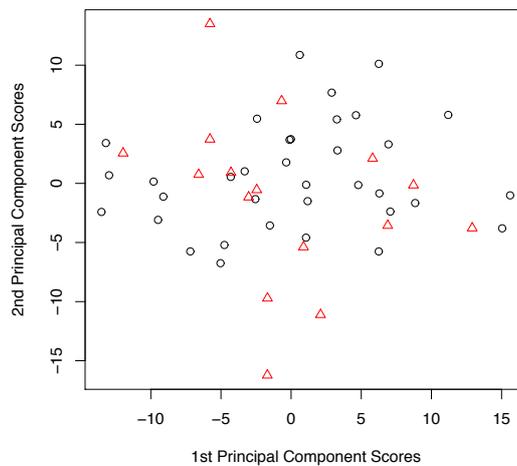
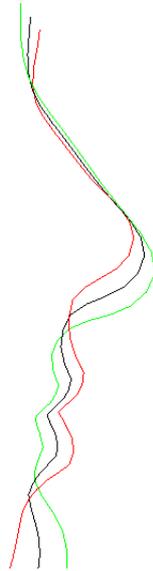


Figure 6.17: Plot of the first against the second Principal Component scores for midsagittal Principal Curves, split by Sex; males (*red triangles*), females (*black circles*).

The first and second principal component scores are plotted against each other in Figure 6.17 and colour coded by sex. No obvious pattern or clustering is visible for males or females, indicating that sexual dimorphism is not present in the shape of the midsagittal curves of control participants in this study.



(a) First PC extremes



(b) Second PC extremes

Figure 6.18: Mean midsagittal curves with lower extremes (*green*) and upper extremes (*red*) of the first and second Principal Components.

The variation that each principal component describes can be investigated by adding to the mean curve, the minimum and maximum principal scores multiplied by the eigenvectors of principal components, (Bowman and Bock

(2006)). For simplicity, the minimum score multiplied by the eigenvectors of a principal component and added to the mean midsagittal curve will be named the *lower extreme*. Similarly, the maximum score multiplied by the eigenvectors of a principal component and added to the mean midsagittal curve will be named the *upper extreme* of that principal component.

In Figure 6.18(a), the upper and lower extremes of the variation that the first principal component explains can be seen. The first principal component describes 33% of all variability of midsagittal curves in the control data set. The lower extreme for the first principal component describes the large noses, small lips and protruding chins and the upper extreme describes the small noses, fuller lips and small chins of the people in the control data set. Figure 6.18(b) shows the upper and lower extremes for the second principal component which describes 20% of all variability of midsagittal curves in the control data set. For this reason, the extremes lie closer to the mean shape of the midsagittal curves. Little evidence of sexual dimorphism was witnessed between midsagittal curves of males and females, therefore the principal curve extremes will not be investigated for males and females separately.

6.5 Curve Identification - Problems Encountered

Every facial shape is naturally different, which translates to a varied range of three-dimensional surfaces. For this reason, it is difficult to formulate a method of curve identification that is successful for each face. Take for example, the curves on the upper lip. The initial method of allocating the coordinates of the horizontal principal curves which are at maximum points of curvature to either the left or right ridge assumes that the ridges are in-

dependent of each other on the facial surface. This however is in many cases not true as philtrum ridges can often converge into a single ridge below the *subnasale*. In this instance, a single coordinate on the first few horizontal principal curves just below nose and often very close to the *subnasale* on the x -axis is identified as an extreme point of curvature. When this situation occurs, the single coordinate values are allocated to both the left and right ridge matrices. The two vertical principal components calculated from the $(k \times 3)$ matrices as before. An example of the philtrum ridges converging and the resulting principals curves on the upper lip can be seen in Figure 6.19.

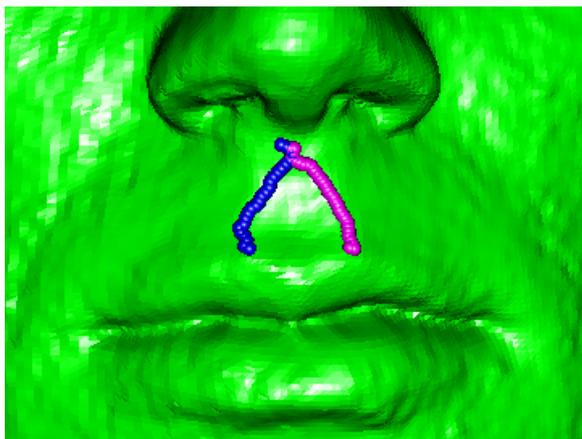


Figure 6.19: Example of philtrum ridges converging on upper lip.

6.6 Curve Identification Conclusions

High variability was prevalent in the curves identified on the upper lip ridges of the 46 control participants. No significant differences were observed between the philtrum ridge curves for males and females, either before or after adjustment for scale.

It was observed from the size-and-shape midsagittal curves that females had shorter noses and shorter faces in general compared to males. From the

shape of the midsagittal curves, very little evidence of sexual dimorphism was visually observed although a borderline significant difference was observed between the mean shape of the midsagittal curves between sexes.

After adjustment for % Body Fat, little difference in the shapes of the midsagittal curves was observed and no significant differences were found, between sexes. Once again, it is not unreasonable to suggest that this lack of effect of the adjustment for body fat could be due to the areas of the face which change with a person's body fat not being represented by the curve on the midsagittal profile.

Chapter 7

Discussions & Conclusions

7.1 Summary

The initial objective of this thesis was to obtain a database of control facial shapes for the purpose of two medical applications; firstly the success of orthognathic surgery and secondly, the characterisation of the biological processes which underlie schizophrenia. This thesis covered the initial capture of a control data set of facial images and in addition the collection of demographic and clinical information from control participants. Methods of identifying and analysing information from the three-dimensional control facial shapes were investigated. The landmark configurations identified on the control facial images were analysed at length and methods of curve identification were discussed, specifically for the curves on the midsagittal profile and on the upper lip. A subsequent validation study was included in this thesis which established the reliability of landmark identification and sources of variation across the configurations of landmarks obtained. The facial curves identified were analysed and problems encountered during identification were discussed. Much of the statistical shape analysis was performed using the ‘shapes’ package in R, (Dryden (2009)). Visualisation of the three dimensional data was achieved using the ‘rgl’ package in R, (Adler and Murdoch (2010)).

The validation study found that the overall measure of accuracy in identifying a set of anatomical landmarks was 1.81 ± 0.84 mm on average. This was considered an acceptable level of error and the method of obtaining data from images was deemed valid for further analysis. Evidence of sexual dimorphism was found in the shape of the landmark coordinates between male and female control participants, more specifically, males were found to have larger faces than females. The measure of control participant's body fat percentage on the position of anatomical landmarks was found to have a significant effect. This suggests that an individual's body fat percentage should be considered when using facial landmark positions alone to estimate facial shape. There was no obvious visual differences between the midsagittal curves between males and females but a borderline significant difference indicates that given a larger sample of controls, a larger difference may be prevalent. No significant difference was found between sexes when the midsagittal curves were adjusted for by body fat which suggests that the differences present prior to this adjustment could be attributed mainly to the participants body fat.

7.2 Limitations in the Data

In Chapter 2, the process of data collection was described at length. Of the 59 participants who volunteered to take part in the study, only 51 satisfied the control criteria. Most of the participants were recruited from the University of Glasgow, after the Face3D study was advertised during lectures to undergraduate students. This meant that the ages of the participants was young and the range of ages was narrow. Due to the lack of diversity in the ages of the participants, the sample collected represented facial shapes of the population of people meeting the control criteria, in their early twenties. As a result, it was not possible to investigate the effect of age on facial shape with this control data. The ratio of males to females in the control data is low. The reasoning behind this may be attributed to the lectures in

which the study was advertised. A large number of participants originated from 1st year psychology lectures, which was taken in 2010 by a majority of females. In this study, a sample of 15 males and 36 females represents the population of facial shapes, which has obvious drawbacks. In low numbers of participants, outliers, (e.g. someone with a particularly large nose) will have a larger effect on the mean facial shape. In general the facial shapes of the control participants do not describe that of the general population, especially for males. It could however be argued, that candidates for orthognatic surgery are likely to be young adults, in which case the control group would be a representative sample.

Limitations also exist in the way in which the images were captured. As discussed in Section 3.4.1, there are several sources which may contribute to the loss in detail of an image captured. Difficulty in landmarking on the ears was observed in the validation study and for the control images. It was observed that the highest variation in identifying a landmark was at the *soft tissue nasion*. An ‘orange peel’ effect was observed on numerous three-dimensional facial images which distorts the surface area and therefore possibly the curves identified on the surface of the face. By adjusting the settings when capturing the image using the ©DI3D software an option to increase the smoothing applied could help to eliminate this ‘orange peel’ effect on the facial surface. This however was not implemented as there is a danger that use of smoothing might lose information.

7.3 Limitations in the Curve Identification

In general, every person has a different facial shape which means that it is difficult to formulate a method of curve identification which is successful for each face in the control data. The curve identification was performed with the statistical software R (R Development Core Team (2005)). However,

principles of curve identification were adapted for individual faces, such that each control face was examined one at a time. As described in Section 6.2, thresholds were calculated for the region of the upper lips in which to accept an extreme point of curvature as part of the left or right ridge. A maximum extreme is allocated to the left ridge if, on the x -axis, it lies between $(chL - c)$ and ls and allocated to the right ridge if, on the x -axis, it lies between ls and $(chR + c)$, where c is a constant. It was observed that in order to accurately allocate points of maximum curvature to either ridge matrix, the value of c was altered slightly for each face. This was a lengthy procedure but unavoidable in order to obtain the curve of interest.

It was observed that the curves identified on the upper lip ridges were high in variability and wiggly. As a result inconclusive results were obtained as to whether or not sexual dimorphism existed. A method of reducing the variability of the curves by adding a smoothing parameter may be advantageous in future analysis.

7.3.1 Future Analysis

It would be advantageous to have a second analyst allocate landmarks to all control images. From this a study of intra-observer accuracy in landmark identification could be explored. It would be of interest to see if similar variability was achieved for each of the anatomical landmarks between analysts.

It would be of interest to collect more data under a similar protocol. This would allow broader opportunities for analysis of the facial shape and facial curves identified. Data collected from a larger range of age groups could be used to investigate the relationship of facial shape with age. Similarly, data collected from a sample of people with a larger range in % Body Fat could be used to investigate further the relationship of facial shape with %Body Fat.

It would be interesting to visually compare the quality of images collected by a second analyst under the same image capture protocol.

Identifying the facial curves a second time would enable an investigation into the accuracy of curve identification by comparing the two sets of curves from the same facial shape.

It would be interesting to experiment with applying a smoothing parameter to the upper lip curves identified.

It would be highly advantageous to develop a more sophisticated automated R-code which can identify facial curves for all faces without the lengthy step-by-step approach.

Appendix A

A.1 Information Sheet



Face3D project: a study on facial shape

Request for participation in the collection of control data from students and staff

A consortium of research partners involving the University of Glasgow (Statistics, Computing Science, Dental School), the Royal College of Surgeons in Ireland, Dublin City University, the University of Limerick and the Institute of Technology in Tralee has been awarded funding by the Wellcome Trust to pursue research in three-dimensional facial shape. Two medical applications are involved, one on the developmental processes associated with schizophrenia and the other on the success of orthognathic surgery.

Control data are required as part of this study and you are invited to participate. This would involve a brief 15 minute session where a stereo-camera system will be used to capture your facial shape in three dimensions. Some brief questions will also be asked. In order to address the specific research questions of the study, there is biological and medical information we must gather and the questions you will be asked are listed overleaf.

In particular, criteria for acceptable control data for this study include ethnic origin, namely *white British*, *white Scottish*, *white Irish*, *white European* as defined by the Office of National Statistics. Height, weight and a measurement of body fat will also be taken as these may be linked to facial shape. A simple electronic device will be used for body fat; this requires skin sensors to be placed on ankle and wrist. However, these measurements are optional.

Your data will be used in the construction of a database to characterise control facial shape. This will be invaluable information in identifying the special characteristics of the patient groups of interest. As a further incentive to participate, you will be given a virtual three-dimensional image of your face.

Your data will be held in a secure location and will be used only by those in Glasgow and Ireland who are involved in the research project. However, we are also seeking permission to make the data more widely available at the end of the project. Your image will not be identifiable by name and any personal medical information will be removed. However, due to the nature of a facial image anonymity cannot be completely preserved.

Guidance for participants

Data from those who have beards, moustaches or other types of

facial hair cannot be used because of the difficulty in constructing suitable images of facial shape.

Prior to arrival, participants are requested to refrain from using heavy make-up.

During image capture, participants will be asked to pull their hair away from their face. Hair clips or hair bands will be provided.

Questions to be asked

1. Date of birth
2. Sex
3. E-mail address
4. Have you ever had any serious facial injury?
5. Have you had, or are you seeking, any serious facial surgery, cosmetic or otherwise?
6. Is there a history of cleft lip and/or palate in your family?
7. Were both of your biological parents born in the British Isles?
(England , Scotland, Wales, Northern Ireland or Republic of Ireland)
8. Were all of your grandparents born in the British Isles?
(England , Scotland, Wales, Northern Ireland or Republic of Ireland)
9. What best describes your ethnic group?
(white British, white Scottish, white Irish, white European, other)
10. Is there a history of psychotic illness or suicide in a first degree relative?

(Psychotic illness refers to conditions such as schizophrenia and not to conditions such as depressive disorders. A first degree relative refers to mother, father, sibling, son or daughter.) [This question is optional.]

Permission

Participants will be asked to sign the following statement, giving permission for the data to be used.

I have read the information sheet for this study and I agree that the data collected from me in this study can be used for the purposes of the research involved, and that the facial image, together with information on age, sex and ethnic background (questions 1, 2, 7–9), may be made available freely to others in an anonymised manner, thereafter.

Participants have the right to withdraw retrospectively any permission given and to require that their own data be destroyed. However, this cannot be effected once the data becomes available beyond the research project.

Further information on the research project

Questions about participation, or about the research project more generally, may be directed to adrian.bowman@glasgow.ac.uk.

The research project has a web site (address below) where further information is available and where the results of the research will be posted in due course.

www.Face3D.ac.uk

A.2 Data Capture Protocol

Data Capture Protocol

This document is intended to outline a standard format for obtaining information from participants in a 3D image capturing project.

Questionnaire

All participants will be asked a series of questions, as documented in the participant information sheet. The analyst will ask these questions and input the answers directly into a secure database.

Equipment

Throughout a capture session with multiple people, the camera equipment will be positioned in the same location, or as close to it as possible.

The camera will be calibrated at the beginning of each day and again throughout the day as required, at the discretion of the analyst.

A static chair will be positioned in front of the camera. The cameras will be positioned so that the participant's eyes are

lined up at the centre of each camera view before the image is taken.

It is expected that three images will be captured and from these, the best set will be selected. The selected image will then be built into a 3D image and then checked to ensure that the image is useable for the study.

Participants

Each of the participants will be allocated a set time in which to arrive at the location of the data collection. Due to the nature of some of the questions that will be asked during the questionnaire, only one participant will be in the room along with the analyst at one time.

Each subject will be seated in the static chair. All participants will be asked to pull their hair away from their face using hair clips or hair bands provided. Prior to their arrival participants will also be asked to refrain from using heavy make-up and to shave any facial hair. A shiny or hairy facial surface could produce an image where the facial structure or boundaries of facial features that are inaccurate. Facial jewellery should be removed, if that is easy to do, but embedded jewellery such as studs should not be removed in view of the complications which may be involved. Assessment of image quality can be made after

the capture.

Participants will be asked to tilt their heads slightly backwards so that the area around the nose and chin can be captured accurately. The participant's eyes should be open. Participants will be asked to say "Mississippi", followed by the letter "n" and then asked to let their mouth rest, in order to relax the mouth and jaw into a natural position before capturing the image. If a participant is clenching their jaw or holding their mouth in an unnatural position, this will give an inaccurate measurement of their true facial structure.

Where rest position involves an open mouth, participants will additionally be asked to close their lips lightly together and a further image will be captured.

Timing

It is estimated that collection of information will take approximately 15 minutes per participant.

A.3 Permission Form



Face3D project: a study on facial shape

Permission

I have read the information sheet for this study and I agree that the data collected from me in this study can be used for the purposes of the research involved, and that the facial image, together with information on age, sex and ethnic background (questions 1, 2, 7–9), may be made available freely to others in an anonymised manner, thereafter.

I understand that I am free to withdraw from this study at any stage and to request that my data be deleted (if this is before the data from the study is made more widely available).

Study number.....

Name.....

Signature

Date.....

Bibliography

- Adler, D. and D. Murdoch (2010). *rgl: 3D visualization device system (OpenGL)*. R package version 0.91.
- Ayoub, A. F., A. Garrahy, C. A. Hood, J. E. White, M. Bock, J. P. Siebert, R. Spencer, and A. Ray (2003). Validation of a vision-based, three-dimensional facial imaging system. *Cleft Palate-Craniofacial Journal* (40), 523–529.
- Bowman, A. W. and A. Azzalini (1997). *Applied Smoothing Techniques for Data Analysis - The Kernel Approach with S-Plus Illustrations*. Oxford Science Publications.
- Bowman, A. W. and M. T. Bock (2006). Exploring variation in three-dimensional shape data. *Journal of Computational and Graphical Statistics* 15(3), 524–541.
- Bowman, A. W., D. Brown, S. Katina, and J. Smith (2011). Anatomical curve identification. (In Preperation).
- Bugaighis, I. (2011). Three-dimensional gender differences in

- facial form of children in the north east of england. *The European Journal of Orthodontics*.
- Dryden, I. (2009). *shapes: Statistical shape analysis*. R package version 1.1-3.
- Dryden, I. L. and K. V. Mardia (1998). *Statistical Shape Analysis*. John Wiley and Sons, Chichester.
- Farkas, L. G. (1994). *Anthropometry of the Head and Face* (2nd ed.). New York: Raven Press.
- Ferrarini, L. (2010). Robustness and sensitivity of hotelling's t^2 -based permutation tests for selection of 3d morphological markers. *International Symposium on Biomedical Imaging (ISBI)*, 57–60.
- Ferrario, V. F. (1993). Sexual dimorphism in the human face assessed by euclidean distance matrix analysis. *Anatomical Journal* 183, 593 – 600.
- Ferrario, V. F. (1995). Fourier analysis of human soft tissue facial shape: Sex differences in normal adults. *Anatomical Journal* 187, 593 – 602.
- Good, P. I. (2011). *Permutation Tests in Analyzing the Large Number of Variables in Biomedical and Satellite Imagery*. Hoboken, NJ, USA.: John Wiley and Sons, Inc.

- Hastie, T. and W. Stuetzle (1989). Principal curves and surfaces. *Journal of the American Statistical Association* 84(406), 502 – 516.
- Lewis, S. J. (1999). Quantifying measurement error. *Current and recent research in osteoarchaeology: Proceedings of the 4th, 5th and 6th meetings of the Osteoarchaeological Research Group*, 54–55.
- Pestman, W. R. and I. B. Alberink (1991). *Mathematical statistics: problems and detailed solutions*. Berlin, Germany: WB-Druck GmbH and Co.
- Pinheiro, J. C. and D. M. Bates (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer.
- R Development Core Team (2005). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
- Tukey, J. W. (1974). Mathematics and the picturing of data. *Canadian Mathematical Congress*, 523–531.
- Venables, W. N. and B. D. Ripley (2002). *Modern Applied Statistics with S* (fourth ed.). Springer.
- Weingessel, A. (2009). *princurve: Fits a Principal Curve in Arbitrary Dimension*. R package version 1.1-10.

White, J. E. (2004). Three-dimensional facial characteristics of caucasian infants without cleft and correlation with body measurements. *Cleft Palate - Craniofacial Journal* 41(6), 593 – 602.

Wiley, D. F. e. a. (2007, November). *Landmark: Institute for Data Analysis and Visualization* (3.6 ed.). University of California, Davis: Visualisation and Graphics Research Group.