UNIVERSITY OF CAPE TOWN
FACULTY OF HEATH SCIENCES

DIVISION OF BIOMEDICAL ENGINEERING

MASTERS THESIS

Title :

Geometric morphometrics for 3D dense surface correspondence: Population comparisons of shoulder bone morphology

By

Jean-Rassaire Fouefack (FFCJEA001)

Supervisor :

Dr. Tinashe Mutsvangwa

Co-Supervisors :

Prof. Tania Douglas and Dr. Adijat Omowumi Inyang

Submitted in fulfilment of the requirements for the degree of MSc in Biomedical Engineering by dissertation

October 31, 2018

Declaration of Authorship

• I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.

• I have used the Harvard convention for citation and referencing. Each contribution to, and quotation in, this report from the work(s) of other people has been attributed, and has been cited and referenced.

• This masters thesis report is my own work.

• I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Name: Jean-Rassaire Fouefack Date: .....October 31, 2018

Abstract

Background: Comparisons in morphological shape/form across population groups could provide population differences that might assist in making decisions on diagnosis and prognosis by the clinician. Geometric morphometrics (GM) is one of the fields that help to provide such population comparisons. In medical imaging and related disciplines, GM is commonly done using annotated landmarks or distances measured from 3D surfaces (consisting of triangular meshes). However, these landmarks may not be sufficient to describe the complete shape. This project aimed to develop GM for analysis that consider all vertices in the triangular mesh as landmarks. The developed methods were applied to South African and Swiss shoulder bones (scapula and humerus) to analyse morphological differences.

Methods: The developed pipeline required first establishing correspondence across the datasets through a registration process. Gaussian process fitting was chosen to perform the registration since it is considered state-of-the-art. Secondly, a novel method for automatic identification of vertices or areas encoding the most shape/form variation was developed. Thirdly, a principal component analysis (PCA) that addressed the high dimensionality and lower sample size (HDLSS) phenomenon was adopted and applied to the dense correspondence data. This approach allowed for the stabilisation of the distribution of the data in low-dimensional form/shape space. Lastly, appropriate statistical tests were developed for population comparisons of the shoulder bones when dealing with HDLSS data in both form and shape space.

Results: When the mesh-based GM analysis approach was applied to the training datasets (South African and Swiss shoulder bones), it was found that the anterior glenoid which is often the site of the shoulder dislocation is the most varied area of the glenoid. This has implications for diagnosis and provides knowledge for prosthesis design. The distribution of the data in the modified PCA space was shown to converge to a stable distribution when more vertices/landmarks are used for the analysis. South African and Swiss datasets were shown to be more distinguishable in a low-dimensional space when considering form rather than shape. It was found that left and right South African scapula bones are significantly different in terms of shape.

Discussion: In general, it was observed that the two populations means can be significantly different in shape but not in form. An improved understanding of these observed shape and form differences has utility for shoulder arthroplasty prosthesis design and may also be useful for orthopaedic surgeons during surgical preoperative planning.

Acknowledgments

I am heartily thankful to my supervisor Dr. Tinashe Mutsvangwa, for choosing me for this project. I appreciate his energy, advice, patience and words of motivation that helped me to understand the work and inspired me many times to give my best.

I would like to express my sincere gratitude to my co-supervisors Prof. Tania Douglas and Dr. Adijat Omowumi Inyang for their advice and suggestions, which were very helpful for the accomplishment of this project.

A special thanks to all my mates from the medical image inferencing & distributed diagnostics (Mi^2d^2) research group and Prof. Marcel Luethi of the Gravis research group, University of Basel. Their comments have enabled me to reflect upon this research from a practical perspective.

I would like to express my sincere appreciation to the honorable staff and friends from the Division of Biomedical Engineering at the University of Cape Town for a great and motivating work atmosphere.

Thanks to my co-supervisor for sponsoring this project through the South African Research Chair (NRF-SARChi) in Biomedical Engineering and Innovation.

Last but not least, I would like to thank my parents who have supported me in all possible ways throughout my life.

Contents

List of Figures

List of Tables

Abbreviations

PCA	Principal components analysis
PCs	Principal components
PCA$_{tr}$	Traditional PCA
PCA$_{nr}$	PCA based on NR method
FEM	Finite element modelling
CT	Computed tomography
GPA	General Procrustes analysis
GM	Geometric morphometrics
SA	South African
SW	Swiss
ORS	Optimal representation of shape
NR	Noise reduction
SVD	Single value decomposition
GPM	Gaussian process models
GP	Gaussian process
FFD	free-form deformation

Chapter One

Introduction

1.1 Background

The number of people with bone disorders incidents has increased over the world, especially for people with obesity and poor physical activity; this number is expected to double by 2020 (Amini et al., 2012). Several studies (Schlecht et al., 2015; Hennessy and Stringer, 2002; Fakhry et al., 2013; Perez et al., 2006; Kimmerle et al., 2008) have shown that sex, population and age are major factors of bone diseases and variation of dislocations. Moreover, studies have shown that clinical diagnosis can benefit from morphological analysis and classification of healthy and defective bones (Key and Funk, 2013; Boileau and Walch, 1997; Shahrani, 2012). Comparisons in morphological shape/form across population groups could provide population differences that might assist in making decisions on diagnosis by the clinician such as identifying whether a biological structure is defective or not through its classification in one of the existing healthy biological structures.

Morphological variation of human scapula and humerus are of interest to prosthetic designers and orthopedic surgeons (Chhabra et al., 2015). For example, the understanding of the population shape/form variation of these bones could assist in improving surgical intervention outcomes and prosthesis design of the total shoulder arthroplasty (TSA) and reverse total shoulder arthroplasty. Finite element modelling (FEM) of the shoulder joint could also benefit from providing useful information on shape/form differences required to estimate the load applied to the bone.

Geometric morphometrics (GM) (Bookstein and Rohlf, 1990) may provide a way to understand the morphological variation of entire biological structures such as bone shapes (Bookstein, 1997). Geometric morphometrics is the study of form in two-dimensional (2D) or three-dimensional (3D) spaces allowing in-depth investigation of morphological differences (Bookstein, 1997). This type of analysis aims, firstly, to outline underlying structure in the population. Secondly, it helps classify the differences in shape between and within populations. Lastly, it helps to classify population shape differences into subgroups. Researchers have developed GM methods that compare bone shape in 2D (Bookstein, 1997; Adams et al., 2013; Sheets et al., 2006) and 3D (Hennessy and Stringer, 2002; Fakhry et al., 2013; Key and Funk, 2013).

1.2 Problem statement

To the author's best knowledge, no work has been done on 3D comparisons of shoulder bone morphological variation in different population groups. Moreover, most work on bone morphometry uses sparse landmarks and semi-landmarks which are limited in terms of surface coverage leading to analysis that may not represent the complete anatomy of the bones in question. Furthermore, in GM, the instability of the distribution of the data in a low-dimensional space when there are more vertex coordinates than objects prevents meaningful clinical or biological visualization. Another limitation in the GM framework is that the significant mean shape/form difference tests fail on application of common statistical methods when dealing with more vertex coordinates than the number of objects.

1.3 Aims and objectives

This project aimed to develop GM methods for 3D dense surface correspondence analysis of shoulder bones between and within different populations, to enable population comparisons of shoulder morphology. In particular, humerus and scapula bones from South African and Swiss populations were compared. In order to achieve this aim, the following objectives were outlined:

- To develop a method to minimize operator bias in determining shape representation.
- To develop a protocol to stabilise the distribution of the data in a low-dimensional space when the number of landmarks representing samples increases.
- To apply statistical methods for population comparisons of the shoulder when dealing with more landmarks than objects.

1.4 General overview

The four main tasks of this project are dense correspondence establishment across the datasets; identification of areas encoding the most shape/form variation across the datasets; stabilisation of the distribution of the data in low-dimensional space in a high dimensionality low sample size (HDLSS) data context; and quantitative comparisons of the population mean shapes/ forms in a HDLSS context.

The first task regarding the establishment of dense correspondence was done through a modified Gaussian process fitting protocol. This novel protocol followed three steps: 1) obtaining an unbiased reference shape; 2) first round registration which allowed to obtain a free-form deformation (FFD) specific to each target shape and 3) second round registration, where each target specific FFD was used to register its target shape.

The second task was to develop an automatic method to identify vertices/landmarks encoding the most shape/form variation. These vertices were identified by computing the volume of the ellipsoid defined by their standard deviations (in x, y and z-directions)

across the datasets. The method was used to reduce the size of the covariance matrix of the data by removing vertices/landmarks with a volume below a fixed threshold. The parametrization of the method allows for arbitrary selection of the threshold and for the operator to decide which vertices they want in their morphological analysis.

The third task was to apply a PCA in a HDLSS context to enable the stabilisation of the distribution of the data in a low-dimensional space. The noise reduction methodology (NR) was used to resolve the inconsistency of the eigenvalues in this context resulting in a PCA (called PCA_{nr}) that is more robust and performs better than the traditional PCA (called PCA_{tr}). The optimal representation of shape developed was combined with PCA_{nr} to evaluate how the variation of the number of vertices encoding shape/form differences can impact the distribution of the data in low-dimensional space.

The last task was to apply statistics for significant mean shape and form difference tests between two groups and classify specimens in classes. The multivariate analysis of the variance (MANOVA) test was applied and the classification of the specimens was performed using linear discriminant analysis (DA).

1.5 Ethical consideration

This project involved the use of images of human bones from fresh cadavers. Ethical approval was obtained from the Faculty of Health Sciences Human Research Ethics Committee (HREC) of the University of Cape Town (HREC REF NO: 546/2017). Computed tomography (CT) images obtained from fresh cadavers from the Division of Clinical Anatomy and Biological Anthropology, Faculty of Health Sciences, University of Cape Town and CT images from the SICAS Medical Repository in Switzerland (SICAS: http://www.si-cas.com/) were used.

1.6 Structure of thesis

The rest of the dissertation is organised as follows: Chapter 2 presents a review of the literature relevant to the project. Chapter 3 outilnes overall methodology of the project. Additionally, a schematic of the pipeline is provided. Chapter 4 presents the methodology for establishing dense correspondence across the 3D surface data. Chapter 5 presents a method to obtain the optimal representation of shape for morphological analysis. Chapter 6 presents the application of principal component analysis methods in a HDLSS context. Chapter 7 presents the statistical analysis for population comparisons of shape/form. Chapter 8 presents the summary of the findings, general discussion and recommendation for the future work.

Chapter Two

Review of relevant literature

2.1 General concepts

2.1.1 Human shoulder bones

The human shoulder consists of joints that combine with muscles and tendons to provide the arm with a wide range of motion, and three bones form the human shoulder (scapula, humerus and clavicle) (Good and MacGillivray, 2005). However, there is a price due to the range of motion allowed by the shoulder joint that leads to instability, resulting in dislocations and pain (Krøner et al., 1989; Hopkins, 2016; Koester et al., 2005). Three-dimensional (3D) morphometry has been used to understand the morphological variation of biological structures such as bones (Shahrani, 2012). For example, a better understanding of shoulder morphology has been provided by conducting geometric morphometric (GM) analysis of the shoulder based on diameter and height (Key and Funk, 2013).

2.1.2 Geometric morphometrics

Shape refers to all the geometrical information about an object that remain after removing scale/size, rotation, and translation while of **form** is shape including size (Kendall, 1977). Morphometry makes reference to the analysis forms, a concept that includes size and shape (Adams et al., 2013). Geometric morphometrics provides a way to understand the morphological variation of entire biological structures such as bone shapes (Bookstein and Rohlf, 1990). The study of shape changes in morphology has seen a significant re-emergence, more precisely, the development of GM methods based on coordinates (Cartesians) of points (Bookstein, 1997). As the morphological objects are incomparable in general when using the original points (coordinates points), a representation of the morphology by landmarks (shape variables) is the first step of the numerical GM analysis, which are 2D or 3D points unambiguously identifiable between specimens.

Geometric morphometrics integrates the geometry of organisms, thus allows researchers to study different aspects of shape (Zelditch et al., 2012). Various human biological structures have been analysed using GM methods such as palates (Bejdová et al., 2012), adult crania (Bigoni et al., 2010) , and human teeth (Shahrani, 2012), and faces (Hennessy and Stringer, 2002). Shape differences between healthy (control) and defective structures can be visualised in a low-dimensional space using GM analysis such as teeth of hypodontia patients (Shahrani, 2012), cleft lips (Bejdová et al., 2012). Moreover, age or sex variations between biological structures can also be investigated through GM analysis as reported

for adult crania in Bigoni et al. (2010).

2.2 Shape representation

Shapes are often represented as a set of landmarks (Ansari and Delp, 1990). A landmark is a point on a biological structure in which correspondences between objects to be studied are preserved (Mardia and Dryden, 1989). Landmark-based GM is a powerful approach for quantifying covariation of shape, biological shape and shape variation (Webster and Sheets, 2010). There are three basic types of landmarks, namely anatomical landmarks, semi-landmarks and pseudo-landmarks; these are described below.

Anatomical landmarks are biologically-meaningful points in an organism. These are points on one object which correspond unambiguously to points on another object. Formally, a shape is represented as a set of x, y or x, y, z cartesian coordinates, for 2D or 3D, respectively. However, many biological structures cannot be completely described using only anatomical landmarks (Gunz and Mitteroecker, 2013). **Pseudo-landmarks** are constructed points located between anatomical landmarks (Mardia and Dryden, 1989). **Semilandmarks** are defined by positions relative to other landmarks (pseudo-landmarks and anatomical landmarks) (Gunz and Mitteroecker, 2013). Pseudo-landmarks and semi-landmarks are useful during shape matching; when the matching process requires a large number of points. Bookstein (1997) considers semi-landmarks as corresponding members of a series of points that are located relative to one another by equal linear spacing from preceding points or equi-angular spacing associated to a radius vector originating from the centroid of a nearest curve, with the set collectively describing the geometry of a curve. Semi-landmarks allow for the quantification of 2D or 3D homologous surfaces and curves; associating them with traditional landmarks for statistical analysis (Gunz and Mitteroecker, 2013). Dense correspondence points on the other hand are all the vertices of a mesh (reference) that have a one-to-one correspondence with others vertices of the meshes (targets) across datasets. Since anatomical landmarks, pseudo-landmarks and semi-landmarks may be insufficient for well-defined surfaces, dense correspondence points can be used to cover the complete surface.

2.3 Correspondence

Correspondence has to be found before any statistical analysis of shape can be performed for a sample of objects. Landmarks that are used for analysis have to identifiable without ambiguity across the sample of objects and be homologous i.e. landmarks for different objects in the training data should be linked by a one-to-one correspondence. This can be achieved in four ways.

1) By visualisation and manual landmarking if only a few points are to be used for shape analysis (Gunz and Mitteroecker, 2013).

2) By manual landmark annotation and registration if dense points have to be used for shape analysis (Vaclav and Pelikan, 2015).

3) Through manual landmark annotation and semi-landmarks as corresponding members of a series of points that are located relative to one another (Green et al., 2015*a*).

4) In a fully automated way using, for example, registration to obtain dense correspondence (Mutsvangwa et al., 2015).

2.4 Registration

2.4.1 Rigid registration

Rigid-registration is the process that removes translation and rotation among different sets of data to get them into one coordinate system. Once landmarks are selected, object alignment should be performed, which transforms all the 3D datasets into the same coordinate system so as to align overlapping components of these sets. The iterative closest point (ICP) method (Besl and McKay, 1992) is one of the methods used for such transformations. However, rigid registration is not appropriate when objects in the sample population exhibit morphological variation; the most common approach in this context is to perform non-rigid registration.

2.4.2 Non-rigid registration

The process of obtaining correspondence often involves the non-rigid registration (except when using semi-landmarks) of one object into another. Some non-rigid registration methods are described below.

a) **Thin plate spline (TPS):** The TPS method (Duchon, 1977) finds the interpolation function passing through a given set of points by minimizing the so-called "bending energy". Bending energy is defined as the integral over the squares of the second derivatives of the interpolation function. For the case of two surfaces (reference and target), the interpolation function provides a one-to-one correspondence between the two sets of homologous landmarks (Gunz and Mitteroecker, 2013; Green et al., 2016). Dense correspondence is established when there is a one-to-one correspondence between all the vertices on each surface of the dataset and all the vertices on the reference surface.

b) **Coherent point drift (CPD):** The CPD is a probabilistic method for non-rigid registration of point sets. The registration is treated as a maximum likelihood (ML) estimation problem with a motion coherence constraint over the velocity field such that one point set moves coherently to align with the second set. The method finds meaningful correspondence between two surfaces (target and reference) and recovers the underlying transformation that maps one point surface (reference) to the second (target) (Myronenko et al., 2006; Myronenko and Song, 2010).

c) **Registration using B-spline deformation model:** The registration problem for this method is to find a match matrix M between two point sets, A and B, such that a cost function, consisting of a shape distance is minimized. The elements of the matrix M are 1 and 0 respectively, for two points in correspondence and two points not

in correspondence. The shape distance is a linear combination of the elements of M and the Euclidean distance between two points (Kybic and Unser, 2003; Lee et al., 2003). Two surfaces (target and reference) are in correspondence when the match matrix is found between them. This method has been used to established correspondence on a statistical shape model construction of the scapula (Mayya et al., 2013).

d) **Gaussian process fitting:** The central problem of this method is to find point-to-point correspondence between a reference surface and a given surface, such that the new surface can be explained in terms of the reference (Gerig et al., 2014; Lüthi et al., 2013). This method uses a deformable model to fit the target surface. The goal is to find a deformation field from a space of deformations, which maps the corresponding points from the reference surface to a target surface (Gerig et al., 2014). The Gaussian process model generates the deformation field and the one that best maps the target is selected. This leads to an optimization problem, where the objective is to find the deformation field that minimizes the difference between the transformed and original points.

2.4.3 Establishment of dense correspondence

Once objects are locally registered using non-rigid registration, there are two ways to obtain the dense correspondence. These are:

1) **Nearest neighbour search**: Here, each vertex on the reference shape finds its closest projections on the sample surface, which define new mesh vertices on the surface. Dense correspondence is established after all the sample surfaces are re-meshed using the reference, see figure 2.1.

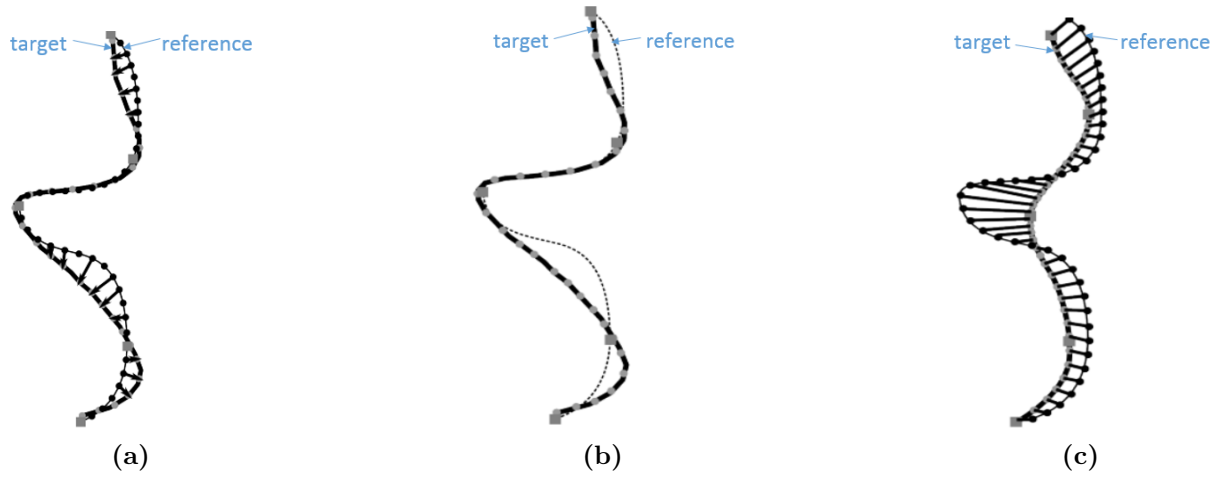


Figure 2.1 An illustration of nearest neighbour search. (a) Black dots on the transformed reference object search for their nearest points on the target object (black shape) (b) Nearest points located: Gray dots on the target are the closest points. (c) Dense correspondence is established. Image adapted from (Hutton et al., 2001).

2) **Replacement of the target with deformed reference**: Here the same reference is used to register to all the target shapes. Since the "fit" (the transformed reference fitted

to the target) is a deformed reference surface, each point on it is in correspondence with one point on the target surface. If the local registration between the surfaces is good, the deformed reference can be considered to represent the target anatomically.

Comparison of different non-rigid registration methods

Vaclav and Pelikan (2015) used three of the above registration techniques namely the TPS, CPD and B-spline methods for human face registration. The outcome showed that the use of TPS-based registration to obtain dense correspondence is fast, but does not establish consistent correspondence. The CPD algorithm exploits fuzzy correspondence in exchange for computationally expensive processing and while considered state-of-the-art in point set registration, it does not allow landmarks set by a suser to be incorporated to guide the registration. B-spline-based non-rigid registration seems to be the most adaptable since it can be used for surface meshes as well as for volume image registration, and also allows incorporating various criteria suh as landmarks. However, the corresponding pairs can change during the registration process due to the complexity of the point cloud shapes and the dissimilarity of the registered shapes (Vaclav and Pelikan, 2015). This means some homologous points may not be in correspondence.

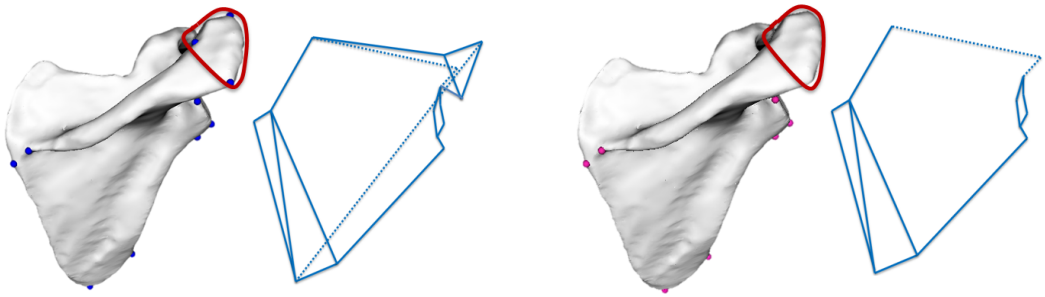
Lüthi et al. (2013) used Gaussian process registration and B-spline registration method to perform non-rigid registration of femur bone surfaces. The outcome showed that the B-spline registration is more accurate than Gaussian process fitting when a single scale kernel model (Gaussian process model based on example of shapes) is used. However, the authors point out that Gaussian process approach results can be more accurate than B-spline registration when there is more flexibility added to the model i.e. when a multi-scale kernel model is used.

2.4.4 Bias in landmark-based GM analysis

The landmarking process is critical in landmark-based GM, because these points encode all the biological shape variation information (Klingenberg, 2013). These points, however, may be defined or detected differently by different observers (Laga et al., 2014). An example of this disparity in landmarking for the scapula is when Püschel and Sellers (2016) considered that the acromion of the scapula may have features that were of importance biologically and therefore some landmarks should be selected on that part to perform GM analysis as shown in figure 2.2a. Young et al. (2015), however, supposed that the acromion does not carry important shape variability information and concluded that it is not necessary to select landmarks on that part as indicated in figure 2.2b. In a typical, purely anatomical, landmark biased GM, a wireframe allows for a visual representation of the shape. The wireframe representation (in blue for both figures) in both situations shows different topology. From this, it is evident that the analysis made using predefined landmarks may suffer from bias.

An additional concern is that manually selected landmarks for GM analysis may not be sufficient from a biological and geometrical point of view because they may not fully describe some biological forms and patterns (Perez et al., 2006). For example, in exploring the differences in biological shape, researchers may not know exactly where the differences

are a-priori. In order to gain more insight into shape differences, further information may be obtained by including as much of the actual object's surface as possible (Perez et al., 2006). Normally the TPS method enables the interpolation of the whole surface and establishes correspondence between the dataset, by providing a one-to-one mapping between every point of the space in which the starting form is embedded and space of the target shape (Klingenberg, 2013). This interpolation however, only enables the establishment of the transformation grid that transforms the grid line in the space of the starting point to the target one. This means significant information encoded by vertices or areas between landmarks remains excluded in the analysis. It would be beneficial to use dense correspondence and non-biased landmark selection because each vertex in a dense mesh of the shape can be considered as a landmark in the analysis. The shape-space is thus of sufficient high dimensionality to capture more biological shape signal (Hutton et al., 2003).



(a) **(b)**

Figure 2.2 Some landmarks on the dorsal of the scapula. (a) Püschel and Sellers (2016) considered that there are three useful landmarks on the acromion; (b) Young et al. (2015) considered that there are no useful landmarks on the acromion. As can be seen by the different wireframed shapes, this subjectivity in the landmarking process would affect the analysis further downstream, leading to different results for the same shape species. Note that the objective of the two analyses may have been different and this is just an example of the disparity in representation of the same object.

2.4.5 Alignment of in-correspondence objects in shape analysis

The alignment of specimens is a crucial step during comparative shape analyses and aims to eliminate differences among specimens due to translations and rotations. After removing translation and rotation, specimens are mathematically defined in a space called **form** space (figure 2.3). Once specimens are in **form** space, size can be removed by normalizing them to unit size. After removing scale, specimens are in a new space called **shape** space (figure 2.3).

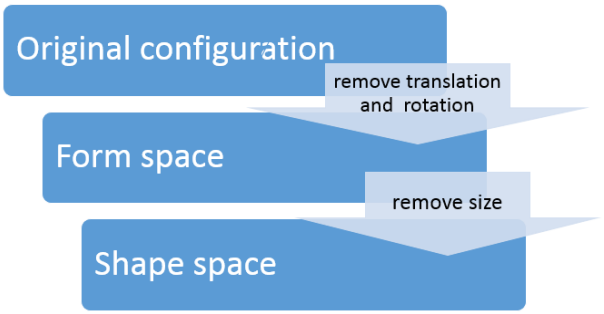


Figure 2.3 Hierarchy of various spaces in GM

Some differences among specimens remain after their alignment, which is indicated by a dispersion of landmarks around a reference configuration. This difference corresponds, by definition, to the landmark configuration differences. In order to use any linear statistical method for further analysis, specimens need to belong in a linear space; an example of this kind of space is tangent space. The tangent space is the linearized space obtained by projecting points of the shape space to a tangent plane before applying any linear method such as principal component analysis (PCA) (Stegmann and Gomez, 2002). The general Procrustes analyses (GPA) is a commonly used method to align objects which are in correspondence (Dryden and Mardia, 1998). The process starts by arbitrarily choosing a reference shape (one of the available instances). All instances are then superimposed onto the current reference shape. The mean shape of the current set of superimposed shapes is then computed and set as the new reference. This process is repeated until the square root of the sum of squared Procrustes distances between the mean shape and the current reference is below a predefined threshold. The algorithm is shown in figure 2.4 and an example of GPA algorithm is shown in figure 2.5.

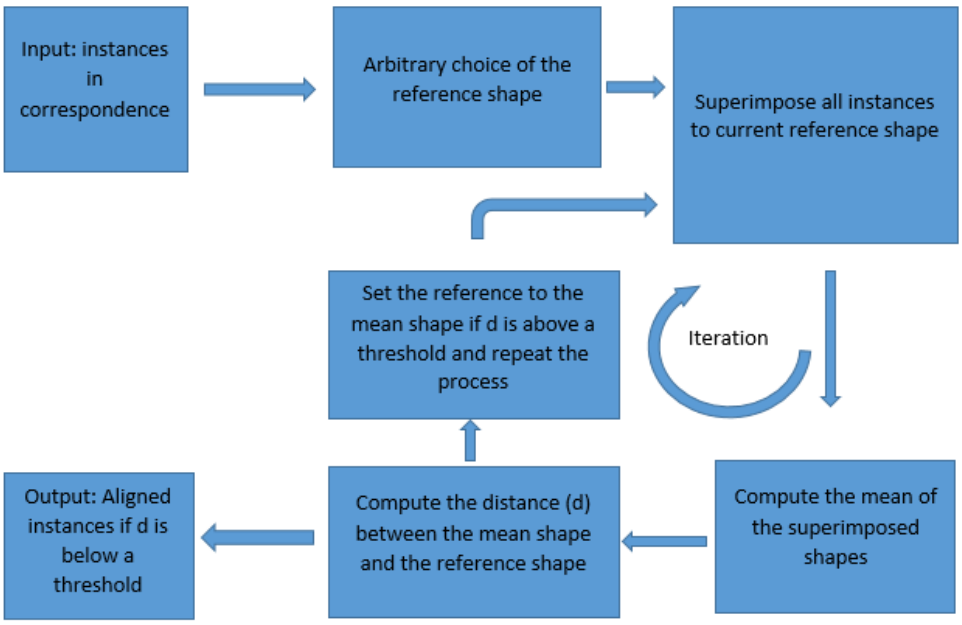
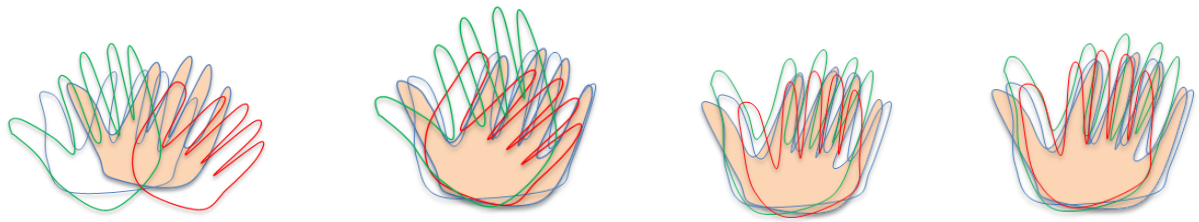


Figure 2.4 The GPA algorithm.



(a) (b) (c) (d)

Figure 2.5 Alignment of different hands in dense correspondence. From the left to right are different steps of the GPA method. (a) Non-aligned data in correspondence; (b) Translation removed; (c) Rotation removed; (d) Data aligned after performing the GPA.

2.5 Dimensionality and shape analysis

Once specimens have been brought into correspondence, aligned and projected into a linearized space, the dimensionality of data has to be analysed. This aims firstly to reduce the dimensionality of the data by finding an orthogonal space where the data can be represented. Secondly, it enables visualisation of the data in a common coordinate system. The PCA is the most popular method used in GM for this process (Fakhry et al., 2013; Green et al., 2015*b*; Franklin et al., 2006).

2.5.1 Principal component analysis

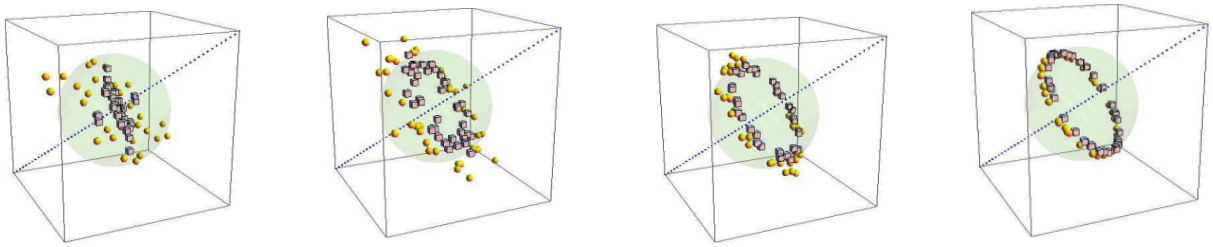
The objective of PCA is to extract the important information from the data; to represent it as a set of new orthogonal coordinates called principal components (PC), where greatest variance lies on the first PC, the second greatest variance on the second pc and so on (Abdi and Williams, 2010). This method requires data to have a normal distribution; a requirement which maybe a restriction on the application of PCA. Recently, Copula component analysis (CCA) (Egger et al., 2016) has been developed to transform empirical data to a normally distributed one. It maybe an advantage to include the CCA in the GM framework in order to make the PCA more general and not only for normally distributed data. The PCA enables us to have only a few principal components (PCs) account for the shape variability of most of the data. These variations can be visualized using a low-dimensional space. Many studies have made use of PCA for GM analysis of bone shapes (Fakhry et al., 2013; Green et al., 2015*b*; Franklin et al., 2006). However, the PCs do not behave as expected when the dimensionality of the data becomes high and the sample size is small. This is the case when using dense correspondence for shape analysis and it leads to failure when applying PCA as detailed below.

2.5.2 Complication of using PCA with HDLSS data

One of the features of modern data is that the data dimensionality is extremely high while the sample size is often relatively low (Yata and Aoshima, 2010). High dimensionality, low sample size (HDLSS) data, which has more variables than independent sampling units,

has become a ubiquitous phenomenon in medical imaging and biomedical research (Chi and Muller, 2013*a*). The PCA in GM is based on the $d \times n$ configuration matrix $\mathbf{X}$ which consists of landmark coordinates, where d is the dimensionality of the data and the n the sample size. The sample covariance matrix of the data $\mathbf{S} = n^{-1}\mathbf{X}\mathbf{X}^T$ is proportional to the product of the configuration matrix $\mathbf{X}$ and its transpose. Since the dimension of the sample covariance matrix $\mathbf{S}$ is very large, it is a challenge to handle with $\mathbf{S}$ directly. A useful approach is to work with the "dual" of $\mathbf{S}$ as reported in (Seshamani et al., 2011; Jung et al., 2009; Ahn, 2006). The dual approach reduces the number of columns and rows of the data matrix, by replacing $\mathbf{X}$ by $\mathbf{X}^T$; and is defined as $\mathbf{S}_D = n^{-1}\mathbf{X}^T\mathbf{X}$. However, when the dimensionality increases and the sample size is fixed, the $n \times n$ matrix $\mathbf{S}_D$ converges to the unit matrix i.e. when the dimensionality increases while the sample size is fixed, it is almost certain that the covariance matrix will tend to the unit matrix (Ahn, 2006).

To illustrate the manner in which the PCA collapses in the context of HDLSS, consider the example presented in (Ishii et al., 2014), where two sets of samples were generated from a model, one from the first principal component (PC) score and the other from the second PC score. This example shows how the dimensionality can negatively affect the PCA by making the PC direction inconsistent (figure: 2.6). The samples from the first PC score are represented by the sphere and those from the second PC score are represented by the square, and the projections of the data onto the subspace generated by the PC directions (PC1 and PC2) are shown. Scatter plots of 20 random samples with different values of $d(= 4, 40, 400$ and $4000)$ are shown. The dotted plot line denotes the unit vector in 3D $I = (1, 1, 1)^T$. Observe that all the objects appear around the unit circle of the orthogonal complement of $I = (1, 1, 1)^T$ when d becomes large i.e. data behaves like it has a spherical distribution, therefore any underlying structure can no longer be determined from the data.



(a) d=4 **(b)** d=40 **(c)** d=400 **(d)** d=4000

Figure 2.6 Scatterplots of PC scores of a dataset $\mathbf{X}$ with $n = 2$ and different values of d (dimension): (a) when $d = 4$, the distribution of data in the space is original and the underlying structure in the data can be provided through visualisation; (b) when $d = 40$, the data started behaving differently and all the data represented by the square have moved away from the line of the unit vector i.e. the objects are no longer in their original distribution; (c) when $d = 400$, almost all the data appear around the unit sphere i.e. they behave as spherically distributed data. (d) when $d = 4000$, all the data appear around the unit sphere and any underlying structure that is provided through the visualisation of these data no longer describes the original data. Image obtained from (Ishii et al., 2014).

Jung et al. (2009) showed that when the covariance matrix tends to the unit matrix, the PC directions are inconsistent. Therefore the approximation of the data with the first few principal components can no longer provide important underlying structure in the data through visualisation. This issue has been addressed in applied statistics in two ways; by providing consistency properties of the eigenvalue estimation; as well as its limiting distribution. The first approach is to use the cross data matrix methodology proposed by Yata and Aoshima (2010) to estimate the eigenvalues of the sample covariance matrix, and the second is to use the noise reduction (NR) methodology proposed by Yata and Aoshima (2012). Both approaches have been compared in Yata and Aoshima (2012) and the NR approach seemed to perform better than the cross data matrix. According to the author's best knowledge, this issue has not been addressed in GM analysis. In the GM framework, the PCA only enables dimensionality and visualisation of the data. Once the PCA has been performed, multivariate analysis needs to be carried out in order to test shape differences by measuring differences between population mean shapes.

2.5.3 Measuring differences between group means

Multivariate analysis is essentially the statistical process of simultaneously analysing multiple independent variables with multiple dependent variables using matrix algebra. Several methods can be considered depending on the desired type of analysis (Morrison, 1998). Let us assume that we have two populations $\pi_i : d \times n_i$ matrix $\mathbf{X}_i = [X_{1(i)} \dots X_{n_i(i)}], i = 1, 2$; where $X_{k(i)} = (x_{k1(i)} \dots x_{kd(i)})^T$, $k = 1 \dots n_i$ are independent and identically distributed as a d-dimensional multivariate distribution with mean μ_i, covariance matrix Σ_i and $\mathbf{X}_i, i = 1, 2$ are the configuration vectors of shape landmarks, d is the dimensionality of the data, n the number of samples and [,] is the horizontal concatenation of vectors. Let us consider the following two-sample mean test,

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \qquad vs \qquad H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$$

The multivariate analysis of variance (MANOVA) is the most powerful method to perform the above test by testing in a probabilistic way whether the null hypothesis (H_0) has to be rejected or not (French et al., 2008). The conventional test for the above hypothesis is Hotelling's $\mathbf{T}^2$ test (Hotelling, 1931). Aoshima and Yata (2011) however, showed that this test fails when dealing with HDLSS data. Indeed Hotelling's $\mathbf{T}^2$ test is appropriate for the hypothesis above when d is fixed and is less than $n = n_1 + n_2 - 2$ and when $\Sigma_1 = \Sigma_2$ (Chen et al., 2010). Bai and Saranadasa (1996) evaluated the performance of this test in the context of HDLSS data and reported that the power statistic decreases as $d \longrightarrow \infty$. This negative effect is due to the inversion of the sample covariance matrix that is involved in Hotelling's $\mathbf{T}^2$ test. Hotelling's $\mathbf{T}^2$ test is not defined when $d > n$, since the sample covariance matrix does not have an inverse. This makes power statistics in GM analysis liable to collapse when dealing with HDLSS data.

The permutation test (Brombin and Salmaso, 2009; Brombin et al., 2009)has been used to overcome the high dimensionality effect in GM. However, permutation test still encounters difficulties as the dimensionality increases. Another difficulty is selecting the optimal block size for block-wise permutation tests. This approach also assumes the test of null hypothesis, which specifies a group permutation under which the distribution of

data, wouldn?t be affected (Zhou et al., 2014; Wei et al., 2016). Finally, the permutation approach is only used for statistical tests and does not address the failure of the PCA in HDLSS context.

Ishii et al. (2017) proposed two tests from the field of applied statistics to perform power statistics using NR method and cross data matrix methodologies in HDLSS context. Ishii et al. (2017) addressed this issue by defining an appropriate statistic that enables performing power statistics using the F-test in the HDLSS context. However, it is yet to be used to address the HDLSS issue in GM analysis when using dense correspondence points.

The F-test proposed in Ishii et al. (2017) helps to avoid the issue above arising when inverting the covariance matrix in the MANOVA. However, the inverse is still needed to perform discriminant analysis (DA) (Fisher, 1936) in order to classify the population into sub-groups. Some authors (Srivastava, 2006; Srivastava and Kubokawa, 2007) proposed different approaches to estimate the inverse of the sample covariance matrix which is the estimator of Σ_i^{-1}. They showed that the one proposed by Srivastava and Kubokawa (2007) is better than the one in Srivastava (2006). Yata and Aoshima (2012) have proposed a new estimator of the inverse of the sample covariance matrix which improves the one defined in Srivastava and Kubokawa (2007) using the noise reduction method.

2.6 Conclusion

Most of the previous studies have focused in landmark-based GM, which may lead to the failure of the analysis since the operator defined landmarks may not be points that encode salient signals related to morphological differences. This may not provide an adequate analysis on morphological variability, because there may be other points, or some areas encoding major information, that are not encompassed in the analysis. Moreover, there is limited literature on shoulder analysis using landmark-based (or any low-dimensional primitives) GM or using 3D dense correspondence points. Little attention has been paid on PCA when dealing with HDLSS data for GM. Such analysis may be useful for the comparison of shape/form variation in populations due to its capacity to 1) address the lack of reliability in providing relevant unrevealed structure in data by visualisation in low-dimensional space and 2) address the lack of precision in classification.

Chapter Three

Overview of Methodology

This project aimed to develop a GM method for 3D dense surface correspondence analysis of shoulder bones, to enable population-wide shape comparisons. The motivation was to facilitate more insight into shoulder bone shape differences between and within different populations. The results were expected to be of interest for clinical diagnosis, prosthesis design, and classification of normal vs pathological shoulder bone shapes. In particular, humerus and scapula bones from South African and Swiss populations were compared.

In order to achieve the aim of this project, the following objectives were outlined:

- Develop a method to minimize operator bias in determining shape representation.
- Develop a robust PCA protocol for GM in a HDLSS context, by making the distribution of the data in low-dimensional space independent of dimensionality. This would allow the original distribution of data to be preserved.
- Apply a robust multivariate analysis for GM in the HDLSS context for performing population shape comparisons.

The schematic below summarises the methodology of the pipeline developed to achieve the objectives.

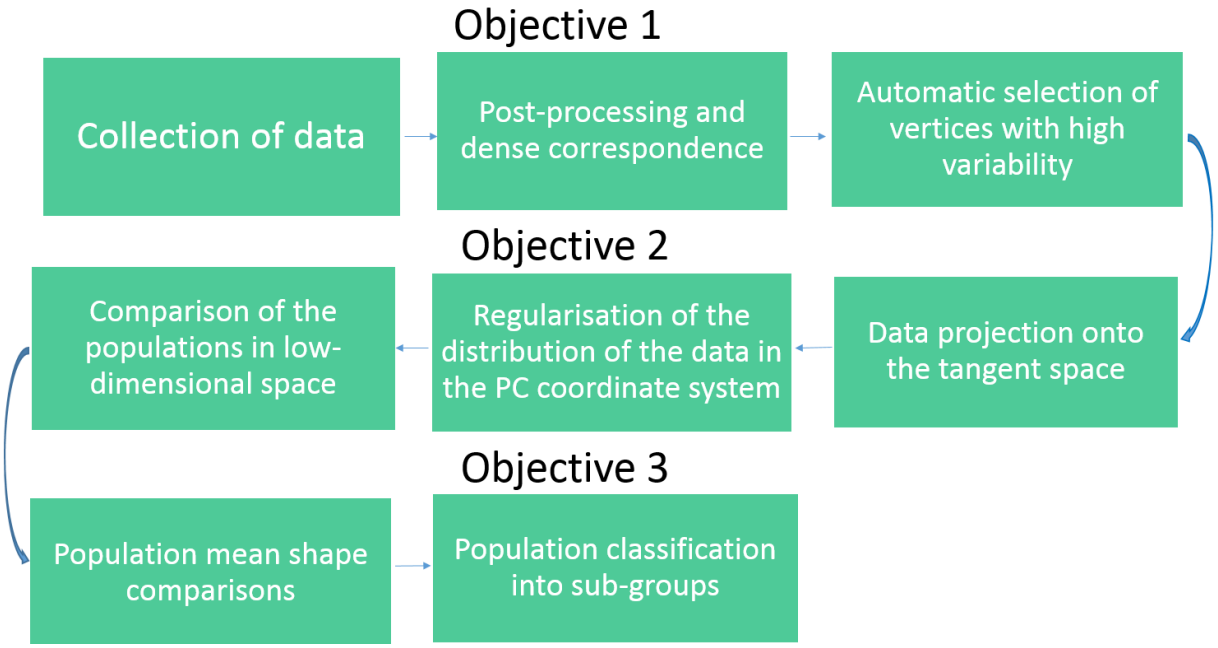


Figure 3.1 Summary of the project.

.

Chapter Four

Establishing dense correspondence

4.1 Introduction

One of the largest sources of complexity in morphological analysis is the establishment of one-to-one correspondence between different objects (Pontil and Verri, 1998). In GM analysis, the correspondence across different subjects has traditionally be established by: 1) manual annotation of landmarks by an expert (anatomical landmarking); 2) points located on an object according to some mathematical or geometrical property, i.e. high curvature or an extreme point (pseudo-landmarks); or 3) constructed points on an object either on the outline or between landmarks (pseudo-landmarks). The establishment of dense correspondence on 3D mesh surfaces is particularly complex as all vertices of the mesh are then considered as potential landmarks. As reported in the literature this requires an automated process as the vertices are often too numerous and most lie in visually unremarkable locations on the object (Guo et al., 2013). There are several existing methods presented in the literature for such process and the Gaussian process fitting method has been reported to be state of the art (Lüthi et al., 2013). This section presents the semi-automated Gaussian process fitting protocol used to establish correspondences between all vertices that define the 3D meshes of the South African and Swiss shoulder bones.

4.2 Description of the data

South African data: The data consisted of 3D surfaces segmented from CT images of scapulae and humeri. Eighteen scapulae and 18 humeri from fresh cadavers were collected from the Division of Clinical Anatomy and Biological Anthropology, Faculty of Health Sciences, University of Cape Town. Of these, one was from female and 8 were from male cadaveric specimens. The age of the imaged descendants ranged from 21 to 90 years.

Swiss data: The data consisted of 18 scapulae and 18 humeri from CT volumes obtained from the SICAS Medical Repository in Switzerland (SICAS: http://www.si-cas.com/).The CT volumes were from 1 female and 8 male descendants and all volume contained both shoulders. The recorded age at the time of death ranged from 21 to 90 years.

The CT volume voxel sizes ranged from 0.97 mm to 1.27 mm in x and y-directions and from 0.5 mm to 0.8 mm in z-direction. The segmentation and 3D reconstruction protocol to obtain 3D surfaces meshes of the shoulder bones is described below.

4.3 Segmentation and 3D reconstruction

The CT image volumes consisting of Digital Imaging and Communications in Medicine (DICOM) files were imported into the medical modelling tool; Amira v6.2.0 (http://www.fei.com/). The 3D reconstruction was performed using the intensity-based segmentation (Pham et al., 2000) to obtain the bone shapes (an example of the process is shown in figure 4.1 for the scapula). The objective of intensity-based segmentation is to separate one or more regions of interest in an image from the backgorund (regions that do not contain relevant information). It is assummed that pixels belonging to the region of interest occupy a different value range than those of the background pixels. Since the image dataset was from CT and the objects of interest were bone, this was a valid assumption. The labelled images were converted to 3D surfaces via the surface generation module in Amira.

The surfaces generated had 24882 and 24180 vertices (to carry much shape/form information), for the humerus and scapula, respectively. The vertices on the meshes were isotropic, meaning that there was no mesh density variation across the dataset. Tests performed in (Mutsvangwa et al., 2015) showed that isotropic vertex placement improved the process of establishing dense correspondence, via non-rigid registration.

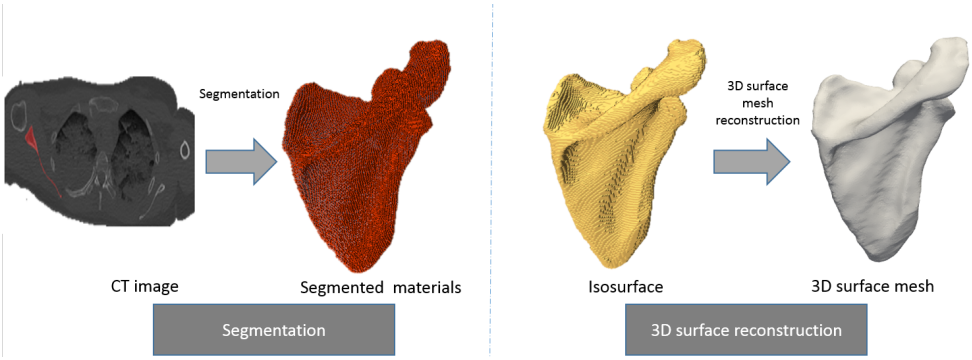


Figure 4.1 Segmentation and 3D reconstruction process from CT images. Left: single CT slice of shoulder showing the scapula in red and scapula materials separated from the background. Right: isosurface from the segmented materials in yellow and 3D surface reconstructed in grey.

After the 3D reconstruction, the quality of the meshes was evaluated in terms of smoothness and coverage of CT volume. The smoothness of the mesh is important as the bone objects that the meshes represent are naturally smooth. This smoothness criterion was evaluated by visualising the mesh. Assessment of the coverage of the original CT volumes by their representative 3D meshes is necessary to make sure that the meshes represent the true boundary of the bones. The assessment was done by checking that the 3D meshes (when aligned to their representative CT volumes) conformed to the boundary of the perceived CT bone edges in the 3D volumes.

4.4 Determining landmarks for guiding registration

The registration process used in this project (see section 4.6) to obtain dense correspondence across the 3D surfaces requires landmarks for guidance (see section 4.6). These points have to be accurately defined and precise; covering the full surface. As such, anatomical landmarks were placed on each scapula and humerus for all data sets (South African and Swiss). The annotation of landmarks was performed in Scalismo (an open source shape modelling platform developed at the University of Basel, Switzerland; (www.github.com/unibas-gravis/scalismo). The selected landmarks for the humerus and scapula are shown in figures 4.2 and 4.3, respectively. Descriptions of the landmarks are provided in table 4.1 for the humerus and table 4.2 for the scapula.

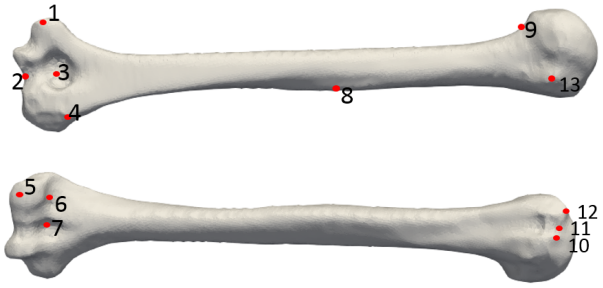


Figure 4.2 Landmark (red points) that were selected on 3D humeri surface meshes. The numbers refer to the points described in table 4.1.

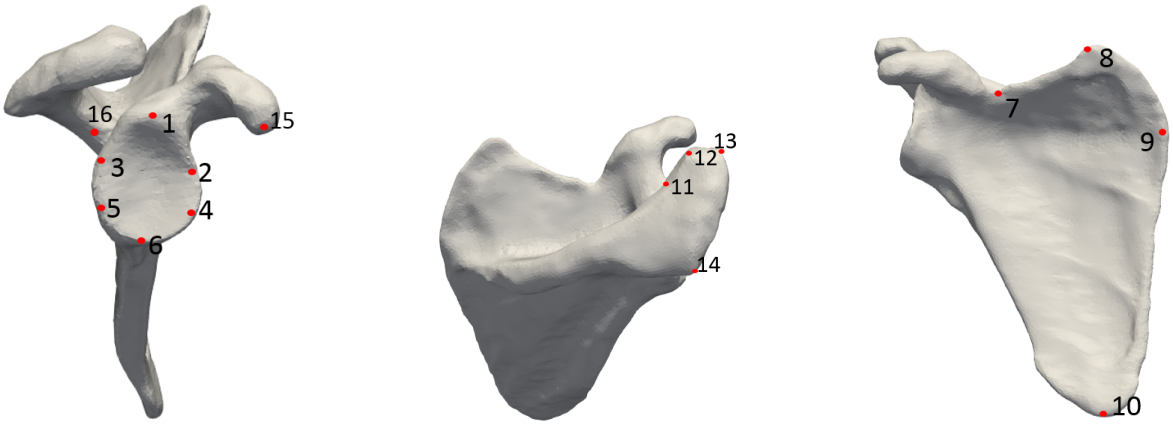


Figure 4.3 Landmarks (red points) that were selected on 3D scapulae surface meshes. The numbers refer to the points defined in table 4.2.

Table 4.1 Description of landmarks for the humerus bone

No	Description of the landmark
1	The most proximal tip on the medial epicondyle.
2	The point in the middle of the capitulum and the trochlea in the posterior view.
3	The most distal point in the middle of the olecranon fossa.
4	The most proximal tip of the lateral epicondyle.
5	The most proximal point of the capitulum in anterior view.
6	The most distal point in the radial fossa.
7	The most distal point in the coronoid fossa.
8	The most proximal point of the deltoid tuberosity.
9	The point on the anatomical neck in the prosterior view.
10	The tip of the lesser tubercle.
11	The point in middle of the lesser and greater tubercles.
12	The tip of the greater tubercle.
13	The tip above the surgical neck.

Table 4.2 Description of landmarks for the scapular bone

No	Description of the landmark
1	Superior middle point of the supraglenoid tubercle.
2	The furthest point on an imaginary horizontal line from landmark 3 to the posterior rim of the glenoid.
3	Anterior point of the neck of the glenoid rim.
4	The furthest point on an imaginary horizontal line from landmark 5 to the posterior rim of the glenoid.
5	The most anterior point of the glenoid rim.
6	The most caudal point of the glenoid rim.
7	The intersection of the coracoid process and the superior border of the blade.
8	The point (visually the highest point) that makes the superior angle of the scapular blade.
9	The point is located in the region where the spine of the scapula intersects the medial border of the blade.
10	The lowest point on the scapula blade.
11	The most posterior point on the medial border of the acromion.
12	The most anterior point on the medial border of the acromion.
13	The most anterior point on the lateral border of the acromion.
14	The most posterior point on the lateral border of the acromion.
15	The tip of the coracoid process with the highest level of curvature.
16	The scapula notch .

In order to evaluate the homology of these landmarks across the sample population, landmarking quality was assessed by performing landmark precision tests as described below.

4.5 Evaluating landmarking quality

The registration process relied on landmarks for guidance and the selection of landmarks depends on human observation. It was important to assess landmarking precision in the approximation of the original landmark since the landmarking affects the quality of the registration. The protocol for precision assessment requires the annotation of landmarks by an observer, performing the annotation at least twice. A time interval between the observer's annotation process is required to remove recall bias. The aim is to evaluate the percentage of landmarks selected with high precision. The precision levels shown in Table 4.3 refers to the ones defined in (Mutsvangwa et al., 2011).

Table 4.3 Precision levels with corresponding error intervals

Precision levels	**Error (mm) intervals**
Highly precise	$[0, 1]$
Precise	$]1, 1.5]$
Moderately precise	$]1.5, 2]$
Imprecise	$]2, \infty]$

The precision error is the difference made by the observer between each landmarking session. Equation 4.1 shows how the error e_l^j is computed for each landmark. Once the repeatability error is computed for each landmark across the population, the average e_l (equation 4.1) of these errors is computed for that specific landmark across the population. The global error is the average of those averages and denotes the error made by the observer for the dataset. The following derivation of the error is from (Mutsvangwa et al., 2011).

Let $l_i^j = (x_i, y_i, z_i), i = 1, 2, j = 1, 2$ be a landmark coordinate selected by an observer at time i on the j^{th} shape and $l = (x, y, z)$ be the average of the landmark select at the both times. The precision error is defined as

$$e_l^j = \frac{1}{2}\sum_{i=1}^{2}\sqrt{((x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2)} \text{ and } e_l = \frac{1}{n}\sum_{j=1}^{n}(e_l^j) \tag{4.1}$$

where e_l is the average error for the landmark l across the dataset, with e_l^j the error for that landmark on the j^{th} shape and n the number of the shape landmark.

4.5.1 Experiment 1: Determining landmarking precision

An observer selected landmarks twice for the humerus and scapula datasets, separately. The landmarking precision assessment was performed as described above. The results are shown in Table 4.4 for the scapula, and for the humerus. It can be observed for the scapula that landmarks were selected with a global error of 1.01 mm and 68.75% of them were selected with high precision. For humerus, the average error was 1.13 mm and 44.62% of the total landmarks were highly precise. It can be observed in general that more than

75% of the total landmarks were precise. However, in order to cover the whole surface of the shape, all the landmarks were then used to guide the registration in order to establish dense correspondence across the data.

Table 4.4 Statistical analysis of the repeatability errors of landmarks

Global mean and standard deviation of mean errors of all landmarks.				
	humerus		Scapula	
	mean	std	mean	std
Global error (mm) for the dataset	1.13	0.55	1.01	0.48
% of landmark selected associated to precisions				
	Humerus		Scapula	
% Highly precise	44.62		68.75	
% Precise	28.67		6.25	
% Moderately precise	18.9		18.75	
% Imprecise	7.81		6.25	

4.6 Registration

4.6.1 Free-form deformation

A free-form deformation (FFD) is a geometric technique based on the idea of enclosing a rigid object within a structure object, and transforming the object within the structure as the structure is deformed (Sederberg and Parry, 1986). Given a reference shape in 3D, a FFD model can be defined as a deformation $u : \Omega \longrightarrow \mathbb{R}^3$ that allows the definition of other shapes belonging to the same family using the reference. One recently introduced FFD technique is based on Gaussian process (Lüthi et al., 2017). The deformation in this technique is modelled as:

$$u \sim GP(\mu, k) \tag{4.2}$$

where $GP(\mu, k)$ is a Gaussian process with the mean deformation $\mu : \Omega \longrightarrow \mathbb{R}^3$ and the covariance function kernel $k : \Omega \times \Omega \longrightarrow \mathbb{R}^3 \times \mathbb{R}^3$. The flexibility of this model depends on the kernel function. The Gaussian kernel K_{GP} is one such kernel and is defined as:

$$K_{GP}^{(s,l)}(x, y) = s \exp(-||x - y||^2/l^2) \tag{4.3}$$

where l is the parameter defining the range over which deformations are correlated and s is the parameter representing the scale of the deformation vector.

4.6.2 Description of the Gaussian process fitting method

As presented in the literature, Gaussian process fitting is a state-of-the-art registration method (Lüthi et al., 2013). The method works by fitting a reference shape (selected from the sample datasets with the prerequisite that it has the representative topology of that species of shape) to each of the other shapes. The protocol for using this registration algorithm is a three-step process involving: 1) rigid alignment of the FFD model (as

described in section 4.6.1) to the target shape; 2) obtaining a posterior model and 3) fitting the posterior model to the target shape.

For the first step, landmarks are used to guide the rigid-registration using a modified iterative closest point (ICP) algorithm, which allows alignment of the 3D surfaces for each bone. The modified ICP works by transforming the parameterized landmarks of the FFD model to the corresponding parameterized landmarks of each target object such that there is a clear correspondence between both landmark sets. Secondly, the actual correspondence between the sets of landmarks is used as a prior information to obtain a posterior model. The posterior model is itself a FFD model after taking into account the relevant evidence related to the particular points (landmarks) on the reference shape corresponding to other points on the target shape. Lastly, the posterior model that is rigidly aligned with target moves to fit as best as possible the target object. This is a non-linear transformation which locally deforms the posterior FFD model to best fit the target by finding a deformation field from a space of deformations, which maps the corresponding points from the reference surface to a target surface.

4.6.3 Developed registration protocol

Developing FFD models - The training datasets were registered using the Gaussian process fitting protocol. A reference mesh was chosen to build a FFD model for each dataset (humerus and scapula, separately). The parameters s and l were chosen so that the model was flexible enough to fit the targets. Three Gaussian kernels were combined to provide very flexible models as follows;

$$K_{GP}(x, y) = \sum_{i=1}^{3} K_{GP}^{(s_i, l_i)}(x, y) \tag{4.4}$$

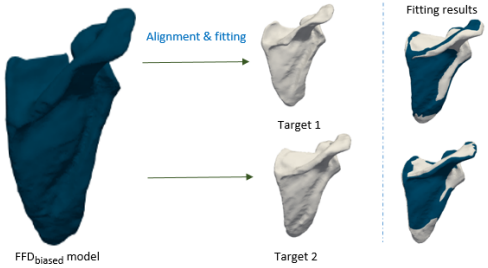
where s_i varying from 1 cm to 30 cm and l_i from 2 cm to 90 cm. The variation of s has the effect that the deformation vector is amplified at different levels with larger values leading to large deformations while small values yield smaller deformations. The variation of l provides different ranges over which deformations are correlated, where large values lead to global deformations while smaller values yield more local deformations.

Obtaining unbiased reference shape - The developed FFD models for each bone shape were fitted onto all the target meshes for each shape (as illustrated in figure 4.4*a* for the scapula) using the landmarks selected in section 4.4 as mentioned above for the fitting process. However, it was noted that the topology of the "fits" would be the same as the reference mesh (called FFD_{biased}) topology and this could bias the shape analysis. To mitigate against this bias, all the registered meshes from the first registration was used to create a new reference (mean) mesh (figure 4.4*b*); the new reference was used to develop a new FFD model (now called $\text{FFD}_{unbiased}$). The $\text{FFD}_{unbiased}$ models (one for each bone family) were desired as their topologies were not biased towards any of the shapes in their respective original datasets. A novel iterative fitting procedure was developed which involved two steps and is describe below.

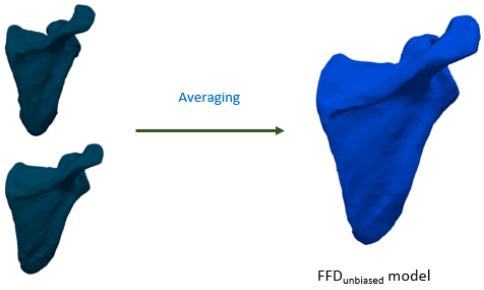
First fitting step- For each bone, the $\text{FFD}_{unbiased}$ model was registered to each mesh

across the dataset. It can be observed in figure 4.5*a* that the model did not fit the target well. For instance, the acromium in targets 1 and 2 did not fit well when the $\text{FFD}_{unbiased}$ model was used. The fitting was improved in the second registration step (figure 4.5*a*). In order to perform the second registration step, additional landmarking was done.

Second fitting step- The landmarks selected on the "fits" from the first fitting were used to guide the registration for each target by repeating step 1, replacing the $\text{FFD}_{unbiased}$ model with a new FFD model (called FFD_{target}) built from the first fit. This second step resulted in the improvement of the registration as can be observed in figure 4.5*b*. Observe that each FFD_{target} model (blue) fitted its target better compared to the one in figure 4.5*a*. The acromium fits very well compared to the one in the first fitting step. These two fitting steps were done for both bone (humerus and scapula) datasets.

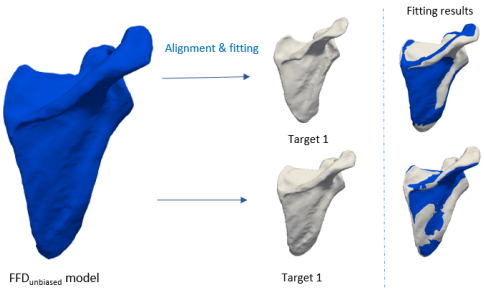


(a) Fitting with biased FFD model illustrated for two scapulae.

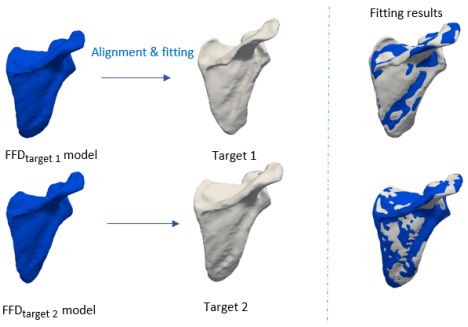


(b) Computation of unbiased FFD model illustrated for two scapulae.

Figure 4.4 Protocol to compute unbiased FFD model illustrated here using two scapulae example. a) A FFD_{biased} is used to fit targets. B) The "fits" (green) are averaged to create $\text{FFD}_{unbiased}$. This was done using all the shapes for each respective bone.



(a) First registration step illustrated for two scapulae.



(b) Second registration step illustrated for two scapulae.

Figure 4.5 Protocol to achieve dense correspondence across the population. a) $\text{FFD}_{unbiased}$ is used to fit the targets (target 1 and target 2). b) Each "fit" in a) is used to build target specific FFD model ($\text{FFD}_{target1}$ and $\text{FFD}_{target2}$) and then used to fit its target again.

4.6.4 Experiment 1: Testing the drift in landmark homology due to iterative registration

Due to its iterative nature, the fitting process described above may suffer from drift in landmark homology. A drift test was performed which aimed to measure how far the selected landmarks for the second registration step drifted from the expected ones. The drift of a landmark is the magnitude between the corresponding point in the "fit" of a landmark in the reference mesh and the landmark selected at the expected area of the "fit" by an observer. Figure 4.6 illustrates what is meant by drift for a selected landmark. The selected landmarks (white point) is projected onto the surface mesh to get its corresponding point in the mesh. The target point or original landmark (red point) is retrieved using its index. The drift error is the difference in location between the red and the white points in figure 4.6. The global drift error of a landmark across the dataset is the average of drift errors of all its corresponding landmarks across the dataset. The global error for an observer is the average of those averages.

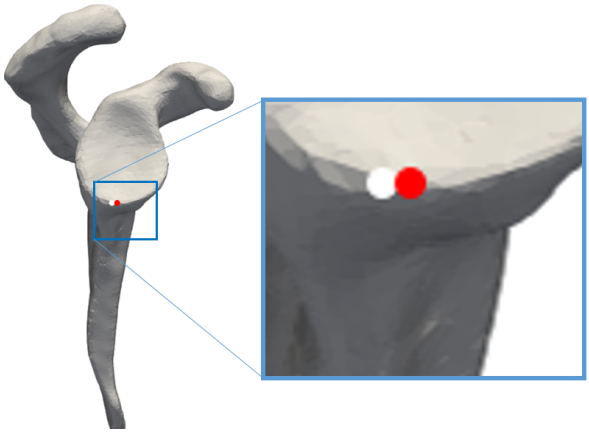


Figure 4.6 Selected landmark in white and the target point (original landmark) in red. The difference in location of these landmarks describes the drift resulting from the iterative registration process detailed above.

The drift error was computed for both datasets (humerus and scapula) and used the same scale from the precision assessment in Table 4.3. Precision here refers to the level of the drift magnitude (drift error). Table 4.5 shows the descriptive statistics of the drift magnitudes across the humerus and scapula dataset after the first registration step. Observe that the global drift magnitude average was 0.9 mm for the scapula and 0.97 mm for the humerus data. Notice, 68.75% of the total landmarks were selected with high precision for the scapula and 69.23% for humerus. None of the landmarks was imprecisely selected.

Table 4.5 Descriptive statistics of the point drift measurement

Mean and standard deviation of mean magnitudes of all landmarks.				
	Scapula		Humerus	
	mean	std	mean	std
average drift (mm) for the dataset	0.90	0.49	0.97	0.75
% of landmark selected associated to precisions				
	Scapula		Humerus	
% Highly precise (less than 1 mm)	68.75		69.23	
% Precise (1 to 1.5 mm)	12.5		23.08	
% Moderately precise (1 to 1.5 mm)	18.75		7.69	
% Imprecise (greater than 2 mm)	0.00		0.00	

The drift test shows that the homology was relatively preserved at the end of the iterative fitting.

4.6.5 Registration error for individual shapes

This project used the replacement method for establishing dense correspondence as outlined in section 2.4.3. In this approach, since each target was to be replaced by its corresponding fit after the registration, registration errors were computed to evaluate how close the fits were to the targets. Registration error is measured by computing the surface to surface distance between the fit and its corresponding target. The metric used was the Hausdorff distance (Hausdorff, 1918) between the fits after the second fitting step and their corresponding original surfaces. It measures how far two subsets of a metric space (here represented by mesh points) are from each other (Huttenlocher et al., 1993). The errors are shown for the humeri data in figure 4.7 and in figure 4.8 for the scapulae data. The results showed that the humerus data (36 samples) were fitted with a maximum Hausdorff distance of 0.91 mm with an average error (average over all samples) of 0.70 mm. The scapulae (36 samples) showed higher errors compared to the humerus data with a maximum Hausdorff distance of 1.71 mm and an average error of 1.28 mm. The larger errors for the scapulae registration can be attributed to the high variability of the scapula bone compared to the humerus (Mutsvangwa et al., 2015).

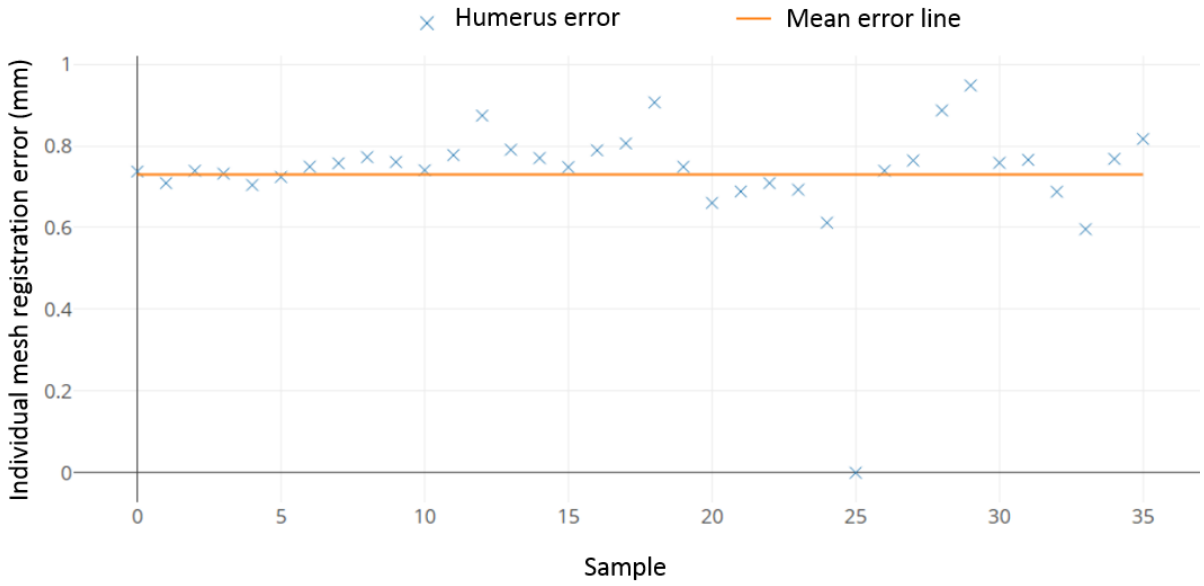


Figure 4.7 Registration errors (surface to surface Hausdorff distance) for the humerus.

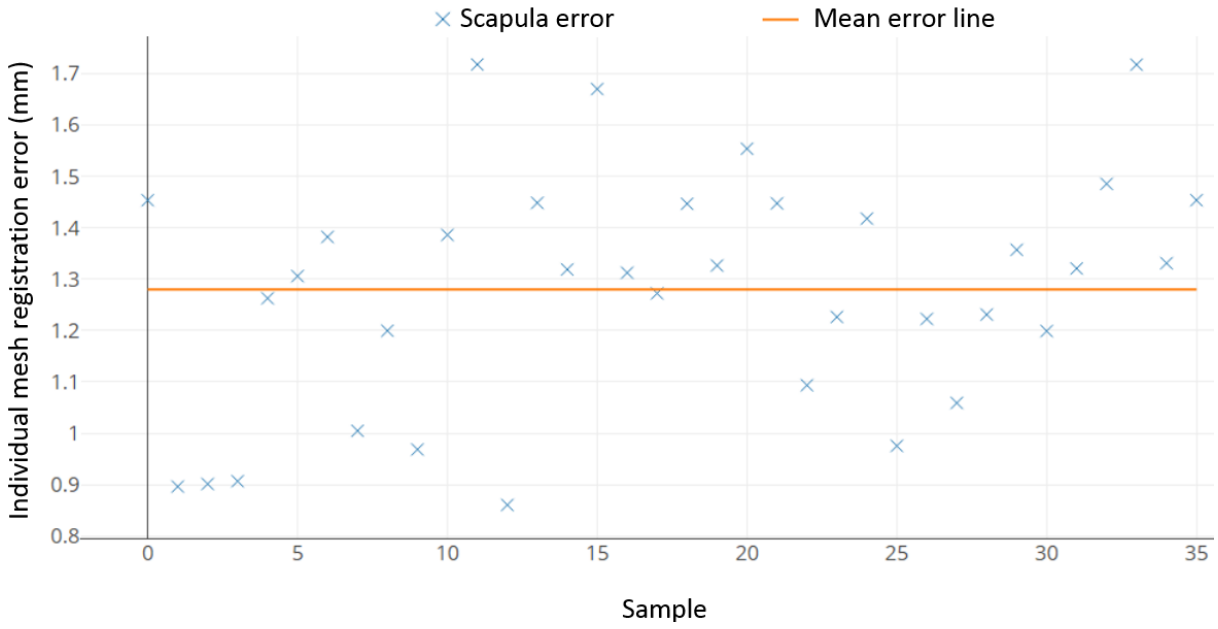


Figure 4.8 Registration errors (surface to surface Hausdorff distance) for the scapula.

The registration here was an improvement on the one in (Inyang et al., 2017) where both bones (humeri and scapulae for South African population) were registered with an average error of more than 1.5 mm. Given the low registration error, it was justifiable that each object was replaced by its fit for each respective bone.

After the validation of the registration, each dataset (scapula and humerus) now consisted of 3D surface meshes in dense correspondence. The General Procrustes analysis (GPA)

was applied to remove any differences that remained in the 3D dense correspondence objects obtained from the registration process to get form data. These differences are due to the translation and rotation and were removed by aligning all the objects around the mean shape.

4.7 Discussion

This chapter detailed a protocol to obtain dense correspondence across the sample of objects. First, the protocol for segmentation and 3D reconstruction was presented. Next, the landmark selection procedure and landmark quality assessments were done. Finally, a novel modified Gaussian fitting protocol was presented.

The fitting protocol developed allowed to minimize the Hausdorff distance between the "fit" meshes and the targets while conserving the homology of the features across the population. The results showed that the more the FFD model used for the fitting is similar to the target, the better the registration. However, the homology of the landmarks used to guide the registration still had to be preserved. The drift of the landmarks should be measured to assure that the two step registration process does not dramatically impact the overall homology.

The registration rounds can be repeated until drift has been eliminated or the Hausdorff distance between the transformed reference and the target is below a user defined threshold.

However,the vertices across the datasets may encode varying amounts of shape and form differences. Therefore, those encoding shape and form differences could be identified as investigated in the next chapter.

Chapter Five

Optimal representation of form and shape for morphological analysis

5.1 Introduction

Shape analysis seeks to study the relationship between the form/shape and function of biological structures (Barillot et al., 2005). The theoretical framework is drawn from geometric information wherein matrices are used to establish intrinsic distances between parametric shape coordinates (Peter and Rangarajan, 2009). Identification and selection of landmark features, usually done by human observation, may not be sufficient from a biological and geometrical point of view because such identification is subjective. Additionally, identified landmarks may not adequately describe some biological objects (, n.d.*a*). As detailed in the literature review, efforts have been made to automate the process of finding corresponding features in the biological objects resulting in objects which are in dense correspondence. However, while dense correspondence is desirable, in order to describe the biological objects more adequately, it comes at the cost of increased dimensionality further downstream in the analysis. In the situation where population form/shape differences are of interest, it may be necessary to use only densely corresponding features which encode the most differences between the populations. The added benefit of only using these features would be a reduction in dimensionality of the form/shape dataset. This section presents the development of an optimal representation of shape (ORS) method which identifies and selects features (in this case, vertices) that encode form and shape differences for a given population of biological shapes. The focus is on the sample of human humeri and scapulae shapes which are in correspondence as presented in the previous chapter.

5.2 Form and shape in GM

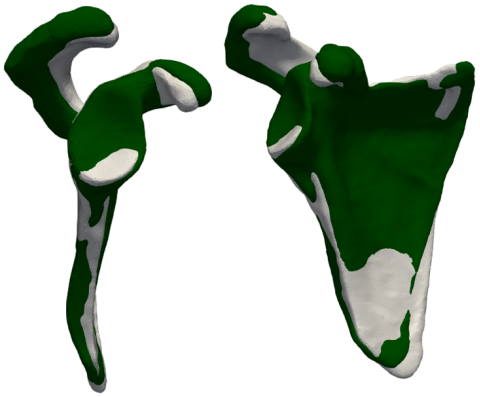
As defined in the literature, the term shape is the geometric property of an object that is independent of the object's overall scale (size), position, and orientation while form is shape including size (Kendall, 1977). Thus, instead of modeling differences between form, the size of the original objects can be removed, therefore, subsequent analysis becomes an analysis of shape; these two analyses can be significantly different (Vaclav and Pelikan, 2015). Objects with form belong to the form space and objects with shape belong to the shape space. Size can be removed using a quantity based on a single size measure such as centroid size (Kendall, 1977). The centroid size of a mesh $X = (x_1, y_1, z_1 \dots x_m, y_m, z_m)^T$

is defined as follows (Bookstein and Rohlf, 1990);

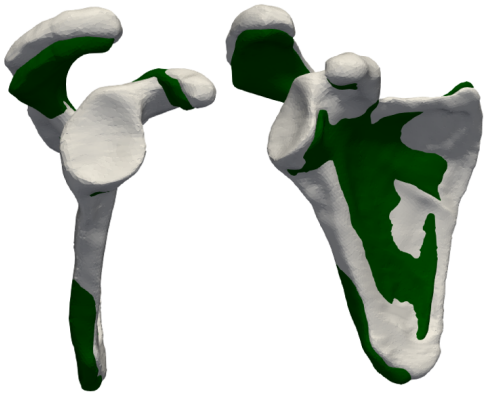
$$S(X) = \sqrt{\sum_{i=1}^{m}(x_i - \bar{x})^2 + (y_i - \bar{y})^2 + (z_i - \bar{z})^2}$$

where $\bar{x} = \sum_{i=1}^{m} x_i/m$, $\bar{y} = \sum_{i=1}^{m} y_i/m$, $\bar{z} = \sum_{i=1}^{m} z_i/m$ and m is the number of vertices.

Figure 5.1 Shows two super-imposed scapulae (white and green) in form space (figure 5.1a) and the same super-imposed scapulae in shape space (figure 5.1b). It can be observed that form differences and shape differences of the two objects are not the same.



(a) Two super-imposed scapulae in **form** space



(b) The same super-imposed scapulae in **shape** space.

Figure 5.1 Two scapulae (white and green) after GPA in form and shape spaces. a) Two scapulae after translation, rotation removed. b) Two scapulae after translation, rotation and size removed.

The ORS method is presented here using objects in form space but the method can be translated completely to analyse objects in shape space. For ease of reading, the terms **form** and **shape** are written in bold when referring to these terms defined above.

5.3 Description of ORS method for automated selection of landmarks

After establishing dense correspondence, each corresponding set of vertices, has an associated mean and standard deviation clustered around that mean. The vertices that encode **form** differences are identified by determining the standard deviation along each direction $(x, y, z - directions)$ of each vertex from this mean. Thereafter, the volume of the ellipsoid, defined by the standard deviations of the three directions about the mean, is computed.

To illustrate the process, let V_i^j be the j-vertex on the i-shape in the dataset and $\bar{V}^j = (\bar{x}^j, \bar{y}^j, \bar{y}^j)$ its correspondent on the reference shape, where j denotes the index

of the vertex in the triangular mesh. The set of the vertices across the dataset is defined as follows:

$$\mathbb{V} = \Big\{ V^j = \{V_i^j = (x_i, y_i, z_i), i = 1 \dots n\}, j = 1 \dots M \Big\} \tag{5.1}$$

where each element of this set is set of tuples, n is the number of the shape in the dataset and M is the number of vertices defining the triangular mesh. Each element $\{V_i^j = (x_i, y_i, z_i), i = 1 \dots n\} \in \mathbb{V}$ is a set of corresponding vertices across the dataset. For each element V^j, the standard deviation is computed in each direction: $\sigma_{x^j}, \sigma_{y^j}$ and σ_{z^j}; the standard deviation of the x-coordinates, y-coordinates and z-coordinates, respectivately. These are computed as follows:

$$\sigma_{x^j} = (\sum_{i=1}^{n}(x_i - \bar{x}^j)^2/n)^{\frac{1}{2}}, \quad \sigma_{y^j} = (\sum_{i=1}^{n}(y_i - \bar{y}^j)^2/n)^{\frac{1}{2}} \quad \text{and} \quad \sigma_{z^j} = (\sum_{i=1}^{n}(z_i - \bar{z}^j)^2/n)^{\frac{1}{2}} \tag{5.2}$$

The ellipsoid of the j^{th} vertex can then be defined as follows:

$$\mathbf{E}_j : \left\{ (x, y, z) \in \mathbb{R}^3 : \frac{x^2}{\sigma_{x^j}^2} + \frac{y^2}{\sigma_{y^j}^2} + \frac{z^2}{\sigma_{z^j}^2} = 1 \right\} \tag{5.3}$$

and its volume is defined as;

$$V_{\mathbf{E}j} = \frac{3}{4}\pi\sigma_{x^j}\sigma_{z^j}\sigma_{z^j} \tag{5.4}$$

Vertices with the volume $V_{\mathbf{E}j}$ above a fixed threshold may then be considered as **significant** (in terms of encoding **form** differences) and those with the volume $V_{\mathbf{E}j}$ below that threshold, considered as **insignificant**. For ease of reading, the terms **significant** and **insignificant** are written in bold when refereing to vertices (or areas) encoding **form** differences.

5.3.1 Identification of significant areas across South African and Swiss shoulder datasets using ORS method

The training dataset after registration had $24,882$ vertices for humeri and $24,180$ vertices for scapulae with the respective dimensionalities $d = 24,882 \times 3 = 74,646$ and $d = 24,180 \times 3 = 72,540$. This study aimed to measure how much morphological variation each vertex of the mesh encodes. The quantity of morphological variation encoded by a vertex was equivalent to the volume of its ellipsoid (equation 5.4). The ellipsoid volumes were computed for humeri and scapulae vertices, respectively. The logarithmic scale was used to allow a large range to be displayed without small values being compressed. In order to visualize the morphological variation encoded by each vertex, the mean shape was shown by associating to each vertex, its volume as a colour.

Figure 5.2 shows the morphological variation for the scapula bone. Vertices with high ellipsoid volume are found around the inferior angle, the medial border and the lateral border which are the areas in dark red. The acromion has vertices with higher volume (representing high variability in that region) compared to the glenoid and the coracoid process. This suggests that the glenoid of the scapula, with low volume ellipsoids (indicated by lighter red colour), encodes relatively low **form** differences in these data. In general, it can be observed that areas of high curvatures on the scapula vary the most

and therefore, are the most **significant** areas. One of the major findings is the anterior glenoid is shown to be the most varying area on the glenoid; this may be why it is often the site of shoulder dislocations as reported in a previous study (Saito et al., 2005).

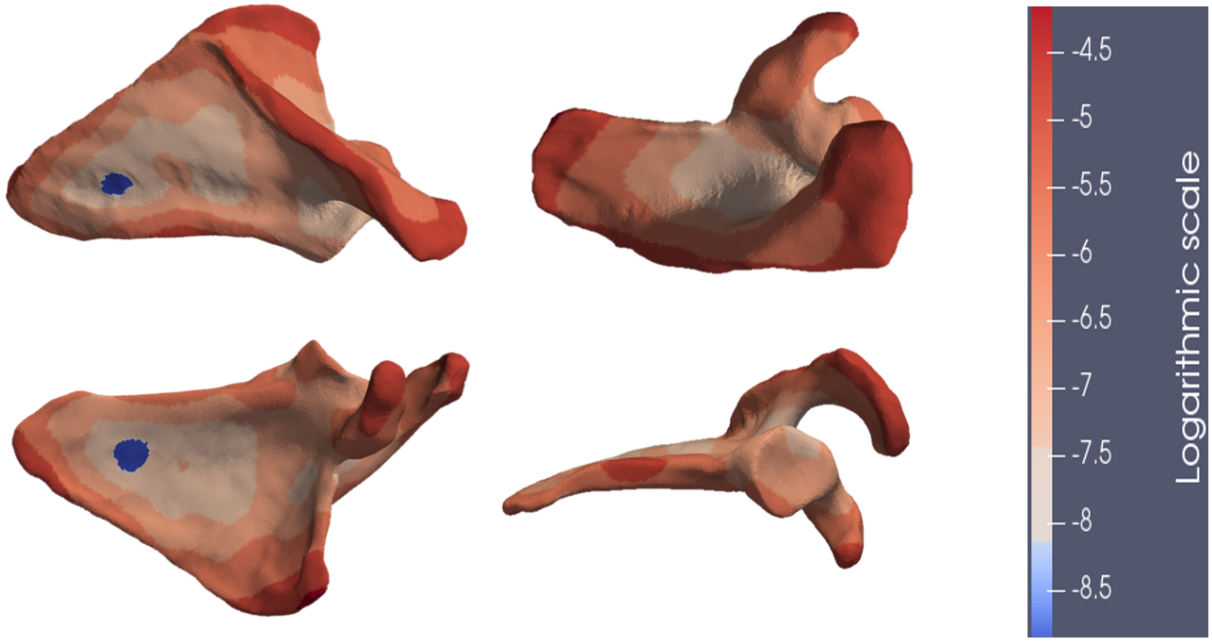


Figure 5.2 Vertices encoding scapula **form** differences. Areas in dark red vary the most across the sample of scapulae, while areas in blue vary the least.

For the humerus bone, the vertices with highest ellipsoid volume (areas in dark and light red) belong to the humeral head and distal humerus; more precisely around the greater tubercle and the epicondyles (figure 5.3). The shaft of the humerus, having vertices with lower volume, does not capture signal in terms of **form** differences. As observed in the scapula bone, areas with high curvature vary the most.

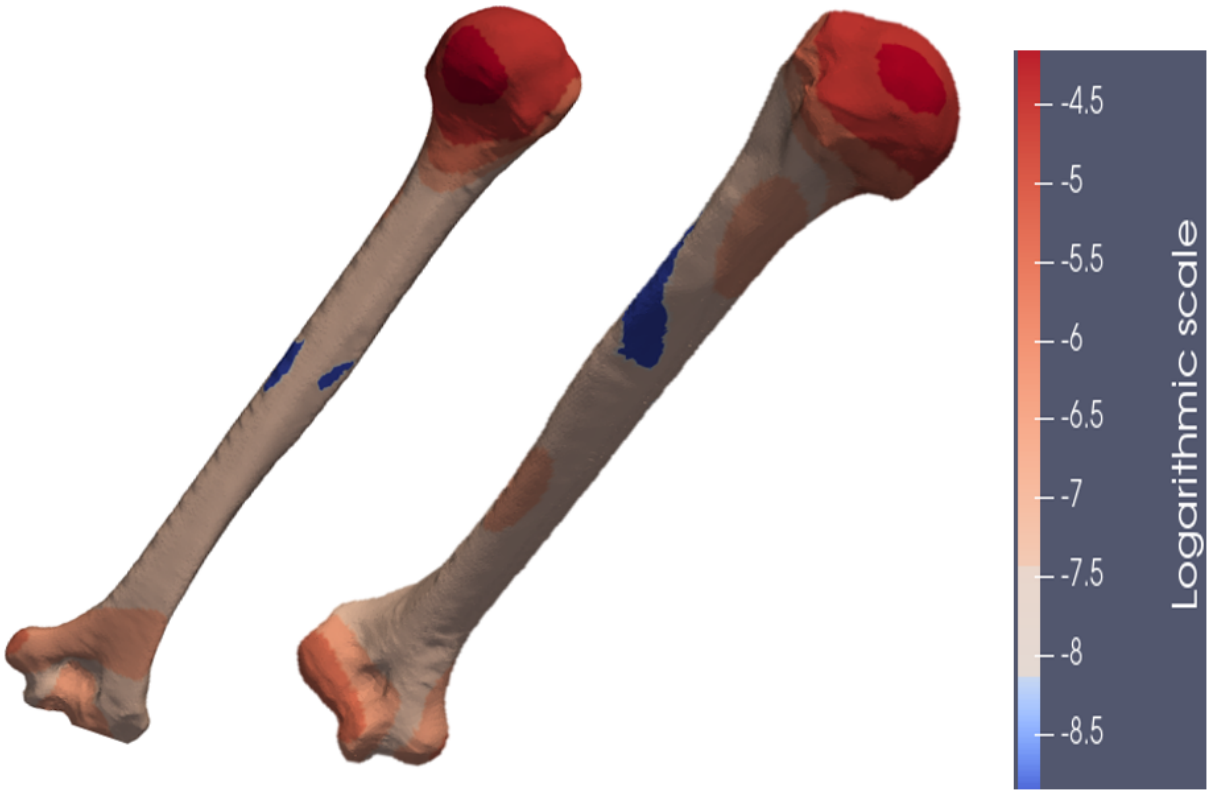


Figure 5.3 Vertices encoding humerus **form** differences variations. Areas in dark red are the areas that vary the most across the sample of humeri, while areas in blue vary the least.

5.4 Reduction of dimensionality in shape analysis using ORS method

This section shows how the dimensionality of the shape can be reduced using the volumes computed above. The aim here is to show that the covariance matrix of the data does not change if corresponding vertices close to their mean are removed across the dataset. In order to prove that, the coordinates of such "close to the mean" vertices are replaced with zeros in the data matrix and it is shown that the matrix still has the same covariance.

Assume that $d \times n$ matrix $\mathbf{X} = [X_1 \dots X_n]$ is a configuration matrix of data, where $\mathbf{X}_i = (V_i^1 \dots V_i^M)^T$, $i = 1 \dots n$ and $d = 3M$ as defined in the previous section. Let us assume that only vertices with a volume above a fixed threshold t are **significant**. We can denote by $\mathbf{X}_t$ (equation 5.5), the "submatrix" of $\mathbf{X}$ containing vertices with the volumes above t only. If all the vertices with the volume below t are replaced by zero then there is a new data matrix defined as:

$$\mathbf{X}_{new} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \\ x_{(p+1)1} & x_{(p+1)2} & \dots & x_{(p+1)n} \\ \dots & \dots & \dots & \dots \\ x_{d1} & x_{d2} & \dots & x_{dn} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{p\times n} \\ \mathbf{X}_t \end{bmatrix} \tag{5.5}$$

where, $\mathbf{X}_t = \begin{bmatrix} x_{(p+1)1} & x_{(p+1)2} & \dots & x_{(p+1)n} \\ \dots & \dots & \dots & \dots \\ x_{d1} & x_{d2} & \dots & x_{dn} \end{bmatrix}$ is a $(d-p)\times n$ matrix and

$\mathbf{0}_{p\times n} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}$, $p = 3k$,

with k denoting the number of **insignificant** vertices across the data. The matrix $\mathbf{X}_{new}$ is made by regrouping null rows and non null rows since the eigenvalues/eigenvectors of a matrix do not change by interchanging rows or columns. The sample covariance matrix of $\mathbf{X}_{new}$ and $\mathbf{X}_t$ are the same i.e $\mathbf{X}_{new}^T\mathbf{X}_{new} = \mathbf{X}_t^T\mathbf{X}_t$. Indeed,

$$\begin{aligned}
\mathbf{S} &= \mathbf{X}_{new}^T\mathbf{X}_{new} \\
&= \begin{bmatrix} 0 & \dots & 0 & \dots & x_{(p+1)1} \dots & x_{d1} \\ 0 & \dots & 0 & \dots & x_{(p+1)2} \dots & x_{(p+1)n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & x_{(p+1)n} \dots & x_{dn} \end{bmatrix} \begin{bmatrix} 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \\ x_{(p+1)1} & x_{(p+1)2} & \dots & x_{(p+1)n} \\ \dots & \dots & \dots & \dots \\ x_{d1} & X_{d2} & \dots & x_{dn} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=1}^{d-p-1} x_{(p+i)1}^2 & \dots & \sum_{i=1}^{d-p-1} x_{(p+i)1}x_{(p+i)n} \\ \sum_{i=1}^{d-p-1}\sum_{j=2}^{n} x_{(p+1)j}x_{(p+i)1} & \dots & \sum_{i=1}^{d-p-1}\sum_{j=2}^{n} x_{(p+1)j}x_{(p+i)n} \\ \dots & \dots & \dots \\ \sum_{i=1}^{d-p-1} x_{(p+i)n}x_{(p+i)1} & \dots & \sum_{i=1}^{d-p-1} x_{(p+i)n}^2 \end{bmatrix} \\
&= \mathbf{X}_t^T\mathbf{X}_t
\end{aligned}$$

The matrix $\mathbf{X}_t^T\mathbf{X}_t$ represents the new covariance but the dimensionality has been reduced from d to $d-p$.

This covariance matrix can be used to perform PCA with different thresholds in order to visualise the distribution variability of the data using different thresholds as investigated later in this thesis (see section 6.4.3).

5.4.1 Examples of variation with different significant thresholds

The tools described above were used on the same datasets in the previous experiment (see section 5.3.1). The maximum ellipsoid volumes were 10^{-3} mm^3 for the scapula and

69.1 mm^3 for the humerus. The minimum ellipsoid volumes were 10^{-6} mm^3 for the scapula and 0.5 mm^3 for the humerus. After applying the logarithmic scale, the corresponding values were from minimum -9.6 to the maximum -4.8 for the scapula; and from -0.6 to 4.2 for the humerus. Different **significant** thresholds were chosen from the maximum values with a step size of 1 to analyse the variation of the dimensionality of the shape for both bones. The thresholds were defined as an arithmetic sequence $\{t_k = t_{k-1} - 1, k = 0, \dots 5\}$, where t_0 denotes the maximum threshold. For each threshold t_k, the number of vertices (n_{inf_k}) with a volume below the threshold t_k were counted. The dimensionality of the shape after reduction with a selected threshold t_k was d_{t_k} and was computed as;

$$d_{t_k} = d - n_{inf_k} \times 3$$

This allows separation of a mesh into **significant** vertices (vertices with volume above the threshold t_k) and **insignificant** (vertices with the volume below the threshold t_k). Additionally, the corresponding percentage Per_{t_k} of **significant** vertices associated with different thresholds t_k also were computed as;

$$Per_{t_k} = (1 - n_{inf_k}/d) \times 100$$

Scapula example-Table 5.1 shows different thresholds (t_k) with their corresponding dimensionalities (d_{t_k}). It also shows the percentages (Per_{t_k}) of **significant** vertices for each threshold. Figure 5.4 shows the mean scapula surface separated into **significant** vertices (areas in red) and insignificant vertices (areas in blue) for each threshold. It was found that with high threshold t_0 (figure 5.4a), there are few **significant** vertices. When the threshold starts to decrease, the number of **significant** vertices increases. When the threshold reaches t_3, most of the areas are **significant** and with the threshold t_5 all the vertices in the mesh are **significant**. This can be of interest in decimation of scapula meshes that are in correspondence without loss of correspondence and **form** information.

Table 5.1 Scapula shape dimensionalities associated with different thresholds

Scapula						
k	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
t_k	$t_0 = -4.8$	$t_1 = -5.8$	$t_2 = -6.8$	$t_3 = -7.8$	$t_4 = -8.8$	$t_5 = -9.8$
n_{inf_k}	24177	20834	12972	5355	414	0
d_{t_k}	9	10038	33624	56475	71298	72540
Per_{t_k}	0.1%	13.8%	46.3%	77.8%	98.2%	100%

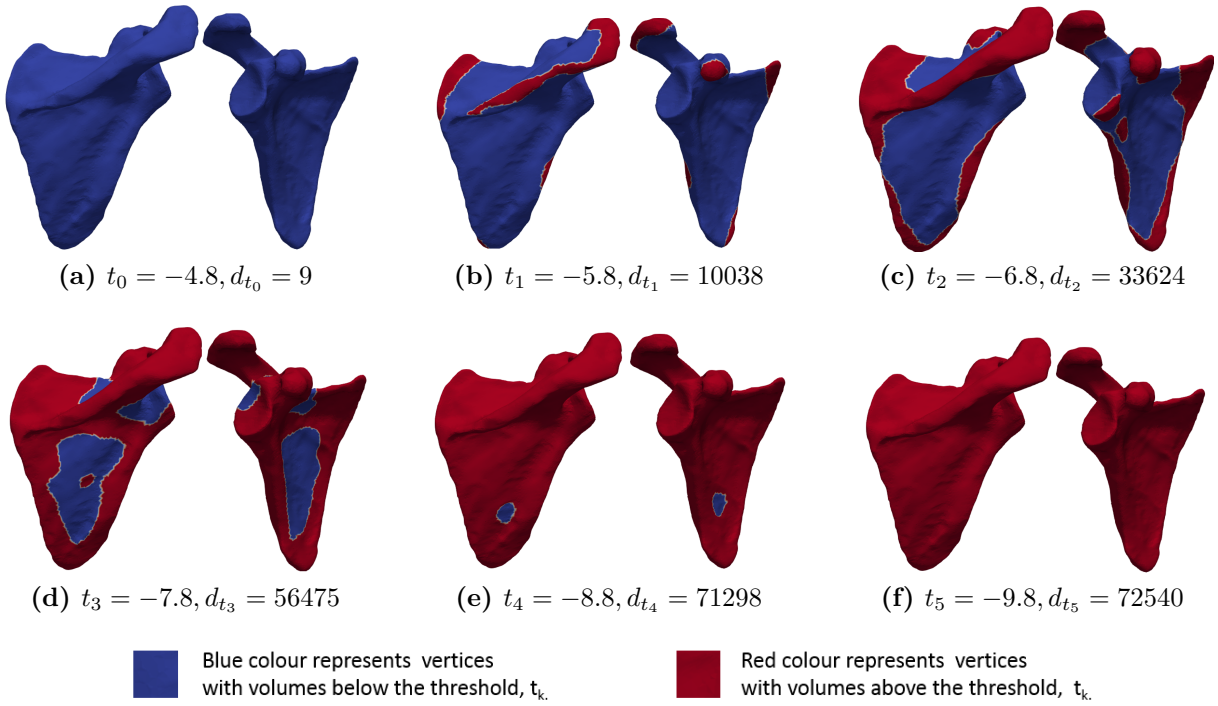
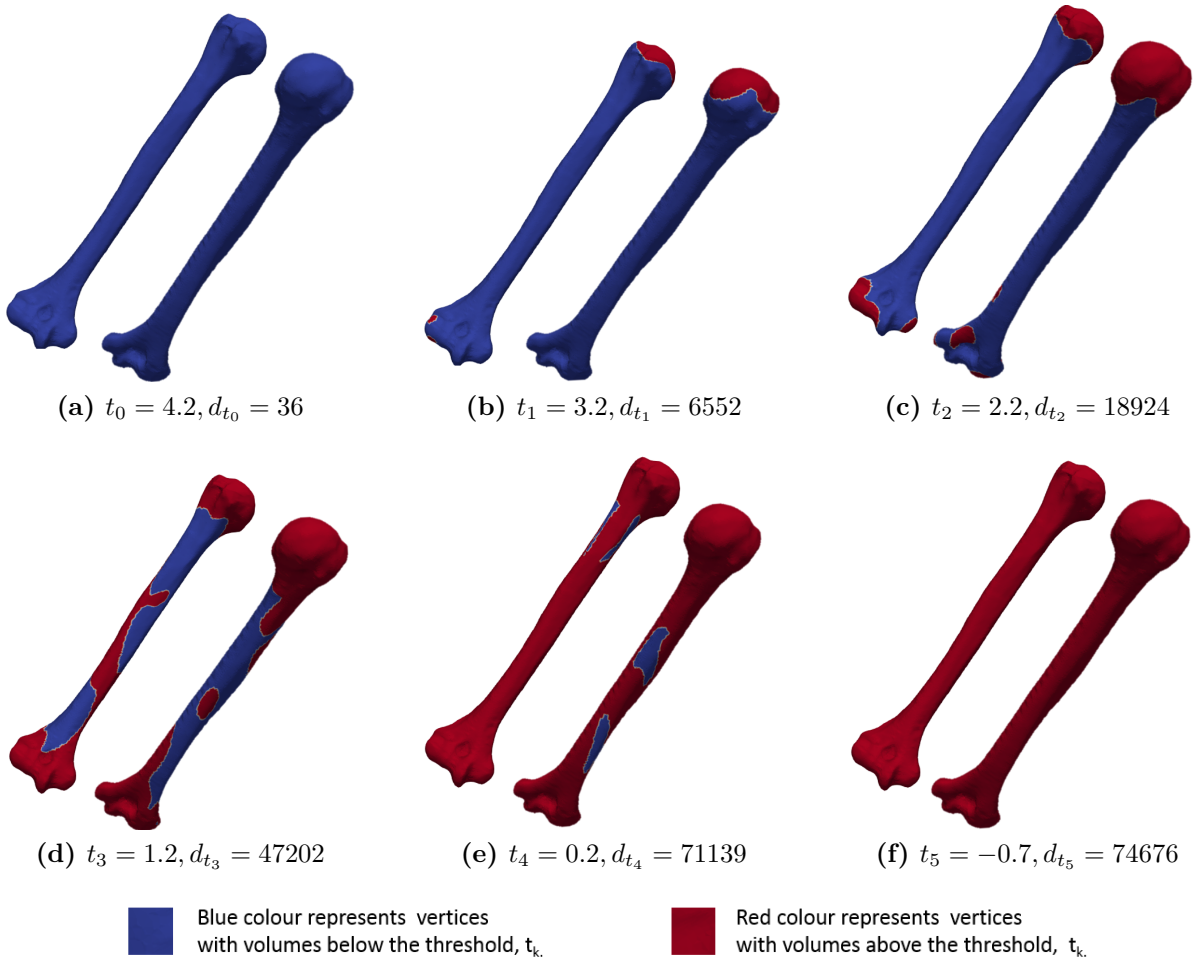


Figure 5.4 Variation of the **significant** vertices for selected thresholds: From a) to f) scapula surfaces showing **significant** vertices (red) and **insignificant** vertices (blue) for different thresholds from $t_0 = -4.8$ to $t_5 = -9.8$ respectively. **Significant** areas increase as the threshold decreases to the smallest threshold, t_5. The mean surface from all the scapulae is used as a template in the illustration above.

Humerus example-Table 5.2 shows different thresholds (t_k) with their corresponding dimensionalities (d_{t_k}). It also shows the percentages (Per_{t_k}) of **significant** vertices for each threshold. Figure 5.5 shows the mean humerus surface separated into significant vertices (areas in red) and insignificant vertices (areas in blue) for each threshold. When the threshold decreases from the largest value ($t_0 = 4.2$) to the smallest ($t_5 = -0.7$), the number of **significant** vertices increases. When the threshold reaches the smallest value, all the vertices in the meshes become **significant**. This can also be of interest in decimating the humerus mesh while preserving correspondence and without loss of **form** information.

Table 5.2 Humerus shape dimensionalities associated with different thresholds

Humerus						
k	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
t_k	$t_0 = 4.2$	$t_1 = 3.2$	$t_2 = 2.2$	$t_3 = 1.2$	$t_4 = 0.2$	$t_5 = -0.7$
n_{inf_k}	24880	22708	18584	9158	1179	0
d_{t_k}	36	6552	18924	47202	71139	74676
Per_{t_k}	0.01%	8.7%	25.3%	63.2%	95.3%	100%



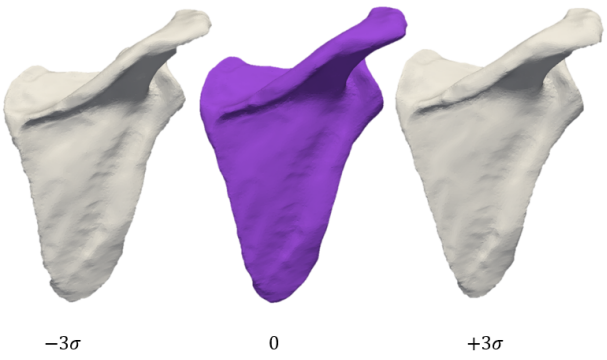
(a) $t_0 = 4.2, d_{t_0} = 36$ **(b)** $t_1 = 3.2, d_{t_1} = 6552$ **(c)** $t_2 = 2.2, d_{t_2} = 18924$

(d) $t_3 = 1.2, d_{t_3} = 47202$ **(e)** $t_4 = 0.2, d_{t_4} = 71139$ **(f)** $t_5 = -0.7, d_{t_5} = 74676$

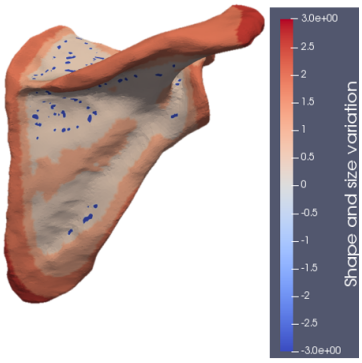
Figure 5.5 Variation of the significant vertices at selected thresholds: From *a*) to *f*) humerus surfaces showing **significant** vertices (red) and **insignificant** vertices (blue) for different thresholds from $t_0 = 4.2$ to $t_5 = -0.7$ respectively. **Significant** areas increase as the threshold decreases to smallest threshold $t_5 = -0.7$. The mean surface from all the humerus data is used as a template in the illustration above.

5.5 Validation of the form representation procedure with PCA

To validate the identification of the vertices (or areas) encoding **from** differences, in-correspondence **form** data were used to build a PCA model. The first PC of each bone family was used to evaluate the variation of the **form** since it explains the most variation. Figures 5.6*a* and 5.7*a* show the mean shape (middle) and two instances (left and right) with the largest variation ($\pm 3\sigma$). On figures 5.6*b*, and 5.7*b* the colour distribution shows **form** variations. The areas that vary the most are the same as those shown above in figures 5.2, and 5.3; this confirms that the method above performed as expected.

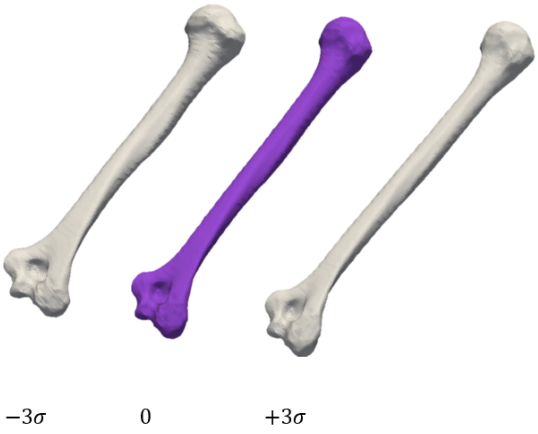


(a) Scapula instances from the PCA model

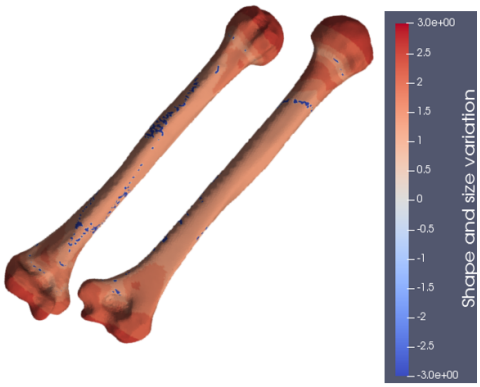


(b) Hausdorf surface to surface distance (mm)

Figure 5.6 a) Model mean **form** (in violet) and two shapes $\pm 3\sigma$ in the direction of the first PC (grey). b) Variation of surface distance between the scapula **form** with the largest variation (left and right on figure 5.6a). Dark red indicates high variation while blue indicates low variation.



(a) humerus instances from the PCA model



(b) Hausdorf surface to surface distance (mm)

Figure 5.7 a) Model mean shape (in violet) and two shapes $\pm 3\sigma$ in the direction of the first PC (grey). b) Variation of surface distance between the humeral **form** with the largest variation (left and right on figure 5.7a). Dark red indicates high variation while blue indicates low variation.

5.6 Discussion

This chapter detailed a method describing the automatic identification of the regions of the bones with the highest **form** variation across a population. This enables minimization of operator bias in determining landmarks for **form** representation because it enables the selection of user-determined vertices that encode **form** difference. It also enables the removal of subjectivity when selecting landmarks or areas encoding **form** differences since the method is repeatable for a given population of objects. For the shoulder joint,

the humeral head is shown to have high variation across the population while the glenoid does not vary significantly.

Several investigators have reported that 90-98% of shoulder dislocations occur at the anterior glenoid (Westin et al., 1995; Cutts et al., 2009). Saito et al. (2005) studied the glenoid defect location in the shoulders of 123 patients, unilaterally, with anterior dislocations of the shoulder and expressed the location of the bony defect as times on a clock face. The finding was that the bony defect of the glenoid was most frequently observed between 2:30 and 4:20 almost anterior to the glenoid as indicating in figure 5.8. However, our study found that the anterior glenoid is the most varying area of glenoid. Figure 5.8 shows that the area indicated by the previous study is included in the area with high variation (red) shown by our method. This may suggest that bony defects of the glenoid are almost always located at the anterior glenoid because of its high variation across the population in terms of **form**.

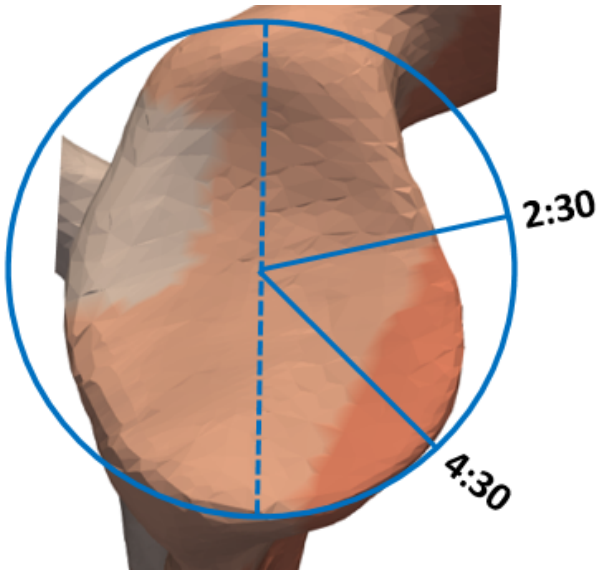


Figure 5.8 Comparison of the area where most shoulder dislocations occur (2:30 to 4:30) and the area of high variation (area in red) on glenoid

To the author's best knowledge, there is no method to decimate meshes that are in correspondence while preserving the correspondence and keeping vertices that capture the most morphological variation across the datasets. The method presented here could be of interest in decimation of meshes that are in correspondence. This could be done by selecting only vertices that encode low morphological variations while preserving correspondence across the dataset. One could consider that areas with vertices having low volumes do not encode enough morphological shape difference and therefore can be removed. This is only valid however when the analysis under consideration is focused on **form** or **shape** differences.

One of the most important applications of this method is dimensionality reduction when dealing with HDLSS. It can be used to reduce the dimensionality of the form/shape dataset when performing the PCA by considering the optimal form/shape representation.

This provides the ability to only consider a minimum number of landmarks/vertices that capture the most variability across the population. This last application of the method is demonstrated in the following sections. Again, the method above could be extended directly to **shape** analysis only as well.

Chapter Six

Application of principal component analysis methods in a HDLSS context

6.1 Introduction

Once the correspondences across the data are found, principal component analysis (PCA) can be performed where corresponding vertices are considered as landmarks (Vaclav and Pelikan, 2015). Principal component analysis is a multivariate technique that analyzes a data matrix in which several inter-correlated quantitative dependent variables are used to describe observations. It aims to extract the important information from the data; to represent it as a set of new orthogonal variables called principal components (PC), and to display the pattern of similarity of the observations and of the variables as points in maps (Abdi and Williams, 2010). Principal component analysis, however, may fail when using densely corresponding landmarks in geometric morphometrics (GM) (Ahn, 2006). This section presents a protocol to describe the **shape/form** variation within a population in the context of high dimensionality low sample size (HDLSS) in GM. As presented in the literature review, the noise reduction (NR) methodology has been developed in applied statistics to handle the HDLSS issue. This approach allows the estimation of the eigenvalues and eigenvectors of the sample covariance matrix so that it will no longer converge to the unit matrix when the dimensionality increases. In this chapter, the NR methodology was adopted for PCA. Before applying the method, however, data (humerus and scapula) were projected into tangent space (see section 6.2) in order to get the data in a linearised space which is a requirement for multivariate analysis (Stegmann and Gomez, 2002). Finally, a comparative study of **shape** and **form** differences was performed for the South African and Swiss shoulder datasets.

6.2 Projection of the data

Multivariate analysis requires data to belong to a linearized space (Stegmann and Gomez, 2002). This is because linear methods are well-behaved and are very well understood (Stegmann and Gomez, 2002). However, this stands in striking contrast to the curved surface of high dimensional space produced by the general Procrustes analysis (GPA). The aim of a tangent space projection is to modify the shape vectors to form a hyper plane, instead of a subpart of a hyper sphere (Stegmann and Gomez, 2002). The Euclidean distance in this plane (tangent space) can be employed as shape metric instead of

the true geodesic distance, i.e. at the hyper sphere surface. This is done by projecting all the data from the Procrustes analysis to the tangent plane at their mean (in our case the reference object). There are two common methods for tangent space projection namely stereographic and orthogonal projection. While stereographic is a smooth bijective conformal mapping, the orthogonal projection is a perspective projection (Rohlf, 1999, 2001). However, both methods are similar when shape variation is modest (Rohlf, 2001). The orthogonal projection was adopted here.

Using surface meshes as an example, the following derivation of tangent space projection is adopted from (Stegmann and Gomez, 2002). Let $X_i = (x_{i1}, y_{i1}, z_{i1} \ldots x_{im}, y_{im}, z_{im})^T \in \mathbb{S}s$ be the vector configuration of the i^{th} shape in the training dataset after the GPA and their mean shape configuration is $\bar{X} = (\bar{x}_1, \bar{y}_1, \bar{z}_1 \ldots \bar{x}_m, \bar{y}_m, \bar{z}_m)^T \in \mathbb{S}s$, where $\mathbb{S}s$ denotes the space of vectorized shape. Let us denote $\mathbb{T}s$, the tangent plane at the mean shape. Each element X_i of the training dataset is projected onto the tangent space as follows:

$$p : \mathbb{S}s \longrightarrow \mathbb{T}s \quad X_i :\longmapsto X_i^p = \frac{\| \bar{X} \|^2}{\bar{X}.X_i} X_i$$

where

Tangent space projection is illustrated in figure 6.1 where four samples are aligned around their mean by performing the GPA. The aligned data are projected onto the tangent plane at the mean. The tangent space is a linear space where the data belong to at the end of the process. The PCA can be then performed in this space as presented in the section below.

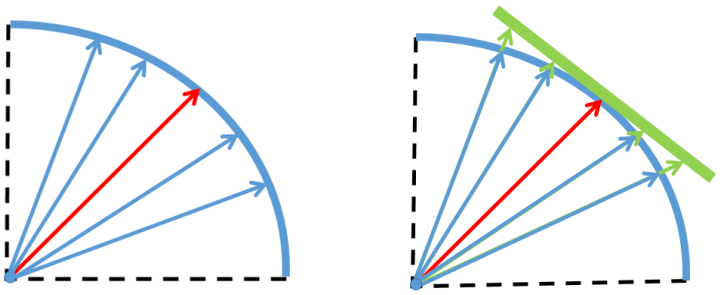


Figure 6.1 An illustration of surface mesh projection into tangent space. Left: A planar projection of four aligned surface meshes (blue) with the mean surface mesh (red). Right: Same as left with tangent space projection (shown in green). The blue arc is the hyper sphere (**shape** space) and the dashed thin line is its radius.

6.3 Traditional PCA

Traditional PCA denoted from now as PCA_{tr} is based on the sample covariance matrix of the data in a linear space. It is a way of determining the patterns in biological data and expressing the data in order to highlight the similarities and differences between them since patterns in the data can be hard to find in data of high dimension (Zhang et al., 2008).

The scatterplot of the principal component (PC) scores provides a visualisation of the underlying structure. This assists the researcher in making an inference from the biological data. However, PCA_{tr} is based on the covariance matrix that needs to be decomposed.

The covariance matrix, also known as dispersion matrix, is a matrix whose element in the (i, j) position is the covariance between the i^{th} and j^{th} elements of a vectorized shape. It is a square matrix that contains the variances and covariances associated with several variables. The eigen-decomposition of this matrix enables the obtaining of the eigenvalues, where the first eigenvalue explains the greatest variance in the data, and the second explains the second greatest variance and so on. This allows for the estimation of the variance of the original data. The PCs associated with the eigenvalues allows one to plot the data in low-dimensional space where the visualisation is compared to the original space. When the dimensionality of the shape is very large, it is challenging to handle with covariance matrix directly. A useful approach is to work with its dual which is the transpose of the covariance (Seshamani et al., 2011; Jung et al., 2009; Ahn, 2006). The dual allows a reduction of the dimensions of the covariance matrix, because the covariance matrix and its dual share the same non-zero eigenvectors. In the theoretical framework below, the data is assumed to be normally distributed. Gaussian data is preferable when performing PCA since this avoids pre-processing such as the exclusion of outliers, the use of non-parametric methods, the application of central limit theorem, etc (Han and Liu, 2013; Liu et al., 2008).

Assume that $d \times n$ matrix $\mathbf{X} = [X_1 \dots X_n]$ is a configuration matrix of a dataset, where $\mathbf{X}_i = (x_{i1} \dots x_{id})^T$, $i = 1 \dots n$ are independent and identically distributed as a d-dimensional multivariate distribution with a non-zero mean and the original covariance matrix Σ. The dimensionality of the data is d, the number of samples n and [,] denotes the horizontal concatenation of vectors. A PCA_{tr} computation based on eigen-decomposition of the full rank covariance matrix, assumes $d < n$. However, when considering PCA on dense meshes in correspondence, the dimensionality of the mesh becomes greater than the sample size ($d >> n$). In this scenario, the full rank covariance matrix can no longer be computable. An alternative algorithm has been used in GM to handle this issue (Vaclav and Pelikan, 2015), which leads to the computation of the $n \times n$ dual covariance matrix instead of computing the $d \times d$ covariance matrix itself (equation 6.1). These two matrices share the same non-zero eigenvalues but the dual covariance matrix has lower computational complexity. The sample covariance matrix $\mathbf{S}$ and its dual $\mathbf{S}_D$ are defined as;

$$\mathbf{S} = \frac{1}{n-1}(\mathbf{X} - \bar{\mathbf{X}})(\mathbf{X} - \bar{\mathbf{X}})^T \text{ and } \mathbf{S}_D = \frac{1}{n-1}(\mathbf{X} - \bar{\mathbf{X}})^T(\mathbf{X} - \bar{\mathbf{X}}), \tag{6.1}$$

with $\bar{\mathbf{X}} = [\bar{X} \dots \bar{X}]$ having d vector, $\bar{\mathbf{X}} = \frac{1}{n}\sum_{k=1}^{d} X_k$.
The eigen-decompostion of the dual covariance matrix is as follows:

$$\mathbf{S}_D = \sum_{i=1}^{n} \lambda_{tr_i} v_{tr_i} v_{tr_i}^T,$$

where $\lambda_{tr_1} \dots \lambda_{tr_n}$ are the eigenvalues and $v_{tr_1} \dots v_{tr_n}$ the eigenvectors.

The PCs are defined as follows:

$$PC_{tr_i} = v_{tr_i}\sqrt{n\lambda_{tr_i}}$$

6.3.1 Example analysis using PCA$_{tr}$

Figure 6.2 shows the scatterplot of first and second PC of 18 South African (blue) and 18 Swiss (red) scapulae **forms** using the PCA$_{tr}$ described above. Each mesh has 24180 vertices, the sample size has $n = 36$ and the dimensionality $d = 24180 \times 3 = 72540$. It can be observed that most of the samples are close to the first PC axis. Almost all the samples are distributed along the first PC axis while only two samples are far away from the first PC axis. Most of the South African samples are lying around the origin of the axis while most of the Swiss samples on the negative values of the first PC. One can consider this distribution and draw a biological conclusion from it, however, this distribution may be affected by the HDLSS phenomenon as detailed below. Thus, any conclusion drawn from this distribution may not represent the underlying biological phenomenon being investigated.

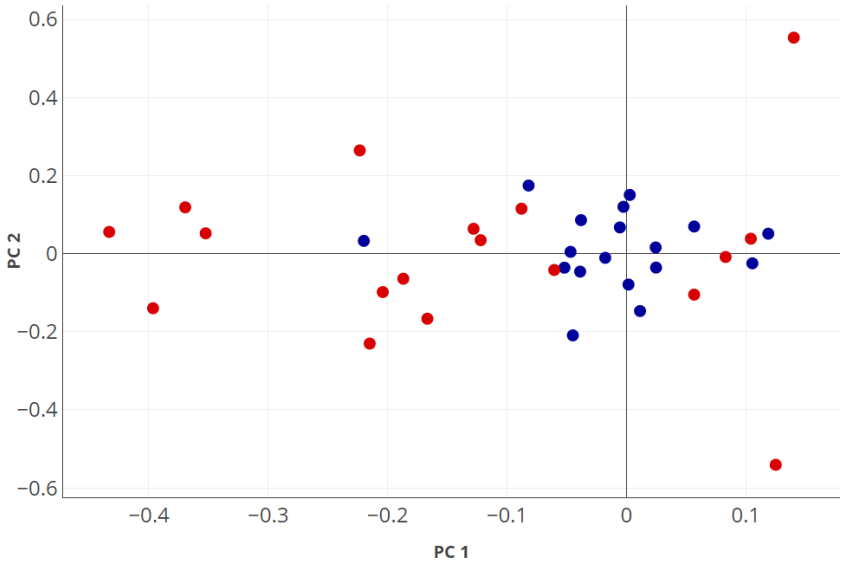


Figure 6.2 The scatterplot of first and second PCs using PCA$_{tr}$ of South African (blue) and Swiss (red) scapulae explaining 39.21% and 22.8% of the total variability, respectively.

6.3.2 Inconsistency of PCA$_{tr}$ in HDLSS context

As described above, HDLSS data may lead to the failure of the PCA$_{tr}$. When the dimensionality d increases and the sample size n is fixed, the $n \times n$ matrix $\mathbf{S}_D$ converges to the unit matrix i.e. it is almost certain that the covariance matrix will tend to the unit matrix (Ahn, 2006). Therefore the approximation of the data with the first few PCs can no longer provide important underlying structure in the data through visualization. This makes the eigenvalues λ_{tr_i} inconsistent in the sense that the eigenvalues may no longer converge to the eigenvalues of the original covariance matrix. This is summarized formally as follows;

$$\mathbf{S}_D \longrightarrow I_n, \text{ as } d \longrightarrow \infty,$$

where I_n is the $n \times n$ unit matrix.

Ishii et al. (2014) illustrated how this inconsistency of sample eigenvalue convergence can affect the distribution of the data in the PCA_{tr} as shown in figure 2.6 in the literature review. From that illustration, it can be observed that when the dimensionality tends to infinity, and the sample size is fixed, any biological conclusion drawn from this data is for the spherically distributed case and not for the original data.

The NR methodology proposed by Yata and Aoshima (2012) enables us to find the eigenvectors that converge to the original eigenvalues (λ_{o_i}) when the dimensionality increases and the sample size is fixed. The method is detailed below.

6.4 Resolving eigenvalue inconsistency with NR

Principal component analysis using NR (PCA_{nr}) aims to estimate the eigenvalues/eigenvectors pairs of the original covariance matrix Σ of the data (Yata and Aoshima, 2012). This is done by removing the "noise" (which prevents the eigenvalue estimators from converging to the original one when the dimensionality increases) from eigenvalues of the sample covariance matrix. The PCA_{nr} is then a way to represent biological data in a low-dimensional space using PC coordinates defined by its estimated eigenvectors. The scatterplots can then be visualized by plotting PCs and some underlying structure can be provided through the visualization of the **shape/form** variability. Original patterns, similarities and differences between biological data are then also provided and not information that does not reflect the original data.

Considering the eigen-decomposition of the covariance matrix $\mathbf{S}_D$ above, the NR eigenvalue estimators are defined as:

$$\lambda_{nr_j} = \lambda_{tr_j} - \frac{tr(\mathbf{S}_D) - \sum_{i=1}^{j} \lambda_{tr_i}}{n-j}, \text{ for } j = 1 \ldots n-1. \tag{6.2}$$

where the sequence $(\hat{\lambda}_j)_j$ is positive. Assume that $\mathbf{X}$ is Guassian, the NR eigenvalues above converge to original eigenvalues as d tends to infinity, n fixed; under the conditions in (Yata and Aoshima, 2012) for a fixed $m \geq 1$. This is formulated as follows:

$$\frac{\lambda_{nr_j}}{\lambda_{o_j}} = 1 + o_p(1), \text{ for } j = 1 \ldots m. \tag{6.3}$$

The jth PC is estimated as follows:

$$PC_{nr_j} = v_{tr_j}\sqrt{n\lambda_{nr_j}}. \tag{6.4}$$

6.4.1 Example analysis using PCA_{nr}

To illustrate the difference between PCA_{nr} and PCA_{tr}, the same data used to demonstrate the PCA_{tr} (see section 6.3) were used. Figure 6.3 shows the scatterplot of first and second PCs using NR methodology. It can be observed that the amount of variance explained by

the first PC is slightly larger than the one using PCA_{tr}. South African samples (blue) are not lying around the first PC axis, contrary to PCA_{tr} (figure 6.3), where all South African samples were lying around the first PC axis. All Swiss samples (red) are lying around the first PC axis, contrary to the case in the PCA_{tr} where some of them were far away from the first PC axis. It can be observed that this distribution in PCA_{nr} is different from the one in PCA_{tr} space (see figure 6.2). Deciding on which PCA is more representative of the underlying biological data relies on theoretical consideration (as illustrated above) and through validation as presented in the next section.

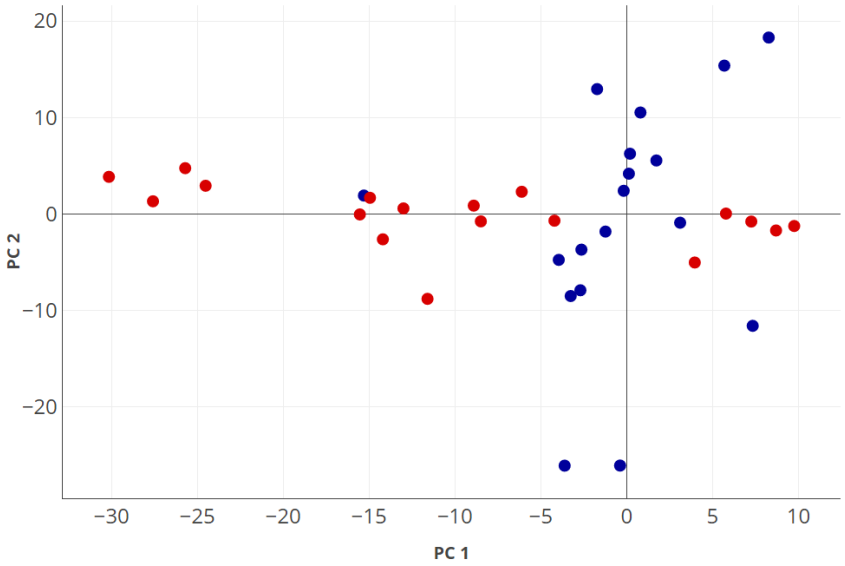


Figure 6.3 The scatterplot of first and second PC using PCA_{nr} of South African (blue) and Swiss (red) scapulae. The first and the second PC explain 39.8% and 22.17% of the total variability respectively.

6.4.2 Validation of PCA quality

For PCA results, the observations are considered to be the population of interest and the interpretations are based on these specific observations. The PCA is a descriptive method and its importance is indicated by the amount of the variance of the data (configuration matrix $\mathbf{X}$) explained by a component. The quality of the PCA is evaluated by the compacity of the first k components to be used to reconstruct the data (configuration matrix $\mathbf{X}$). The first k components are used to compute an estimator $\mathbf{X}^k$ of $\mathbf{X}$ while the matrix $\mathbf{X}$ is better estimated when more and more PCs are used; a PCA is better than another one when it uses fewer PCs to perfectly reconstruct the data (Abdi and Williams, 2010).

Let us consider $V = [v, \ldots, v]$ where v are the eigenvectors of the covariance matrix and D an $n \times n$ matrix where diagonal elements are the eigenvalues λ and the off-diagonal elements are zeros. By the single value decomposition (SVD), the estimator of $\mathbf{X}$ for the first k componnents is $\mathbf{X}^{[k]} = U^{[k]}(\sqrt{D})^{[k]}V^{[k]T}$, where $U^{[k]} = \mathbf{X}V^{[k]}\sqrt{D}^{-1[k]}$. Therefore $\mathbf{X}^{[k]} = \mathbf{X}V^{[k]}V^{[k]T}$.

The quality of the PCA is evaluated by measuring the similarity between $\mathbf{X}$ and its

reconstitution $\mathbf{X}^{[k]}$ with the first k PCs. The error for the first k components $er^{[k]}$ (Abdi and Williams, 2010) is computed as follows.

$$er^{[k]} = \| \mathbf{X} - \mathbf{X}^{[k]} \|^2 = trace\{\mathbf{X} - \mathbf{X}^{[k]}\},$$

where$\| \ . \ \|$ is the l_2-norm. The error $er^{[k]}$ measures how accurate the first k components carry the patterns of the biological data. Small values of this error imply better estimation of the original data and increases the chance of reflecting the true underlying structure regarding the similarities and differences between data.

Figure 6.4 shows the data reconstruction error using the same data used above for the illustration of the PCA_{tr} and PCA_{nr}. It can be seen that the data was better reconstructed using the PCA_{nr} (blue) than the PCA_{tr} (orange) when using the first five components. This shows agreement with the result presented in Yata and Aoshima (2012) where it was clearly shown the PCA_{nr} performs better than the PCA_{tr} for HDLSS scenarios. However, in both cases, the data were perfectly reconstructed with approximately the first fifteen components as shown in figure 6.4.

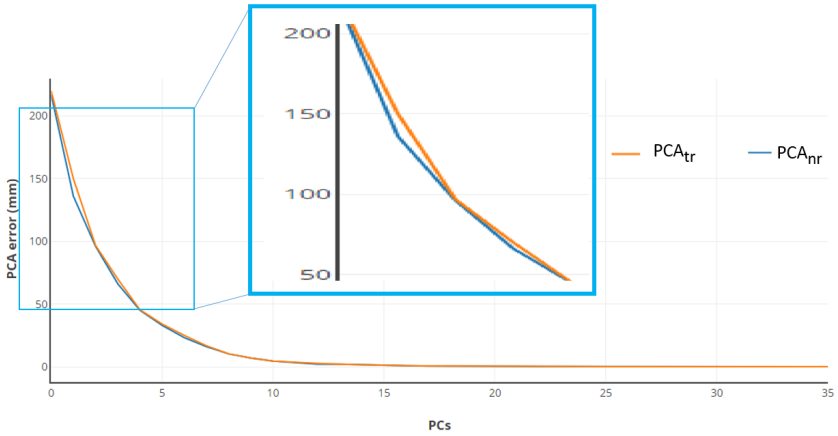


Figure 6.4 The reconstruction error ($er^{[k]}$) associated with the first k components used for PCA_{nr} (blue) and PCA_{tr} (orange) for South African and Swiss scapula data.

.

6.4.3 Combining OSR and NR: Comparison of the data distribution in low-dimensional space

In this section, a combination of OSR and NR is presented to illustrate the utility of both and show how our data behave in low-dimensional space. Although, the ORS method developed in section 5.4.1 could be applied to any PCA, this example illustration was applied to the PCA_{nr} described above. A comparative analysis of data distribution in PCA_{nr} space with different thresholds was performed. Our data were normally distributed since the data were registered using the Gaussian process (GP) model which is based on multivariate normally distributed parameters. Data were first projected onto a linear space

as described in section 6.2. Thereafter, the rest of the study was based on the same experiment in section 5.4.1, where different thresholds and corresponding dimensionalities are shown in table 5.1 for the scapula and table 5.2 for the humerus. However, in this experiment the primary purpose was to visualize the effect of thresholds (or dimensionalities) on the data distribution in the PCA_{nr} space and evaluate how the number of PCs that explains the total variation of data varies with thresholds. This can be useful in deciding which vertices/landmarks and how many of them are necessary to represent biological data in PCA space without loss of information. In order to achieve that, the covariance matrix was computed as detailed in section 5.4.1 for each threshold and the PCA_{nr} was performed. The cumulative variance (eigenvalue) of the PCA_{nr} was computed for each threshold. The first and the second PCs (which are the ones explaining the most variability of the data) were plotted.

Example with scapula data

Figure 6.5 shows the cumulative variance of the PCA_{nr} for each threshold. It was found that when the threshold decreases (or dimensionality increases), then the number of the PCs that explains the total variability of the data increases. The amount of variance explained by the first PC decreases as the dimensionality increases. From this, it can be concluded that the larger the dimensionality, the larger **form** (or **shape**) information captured.

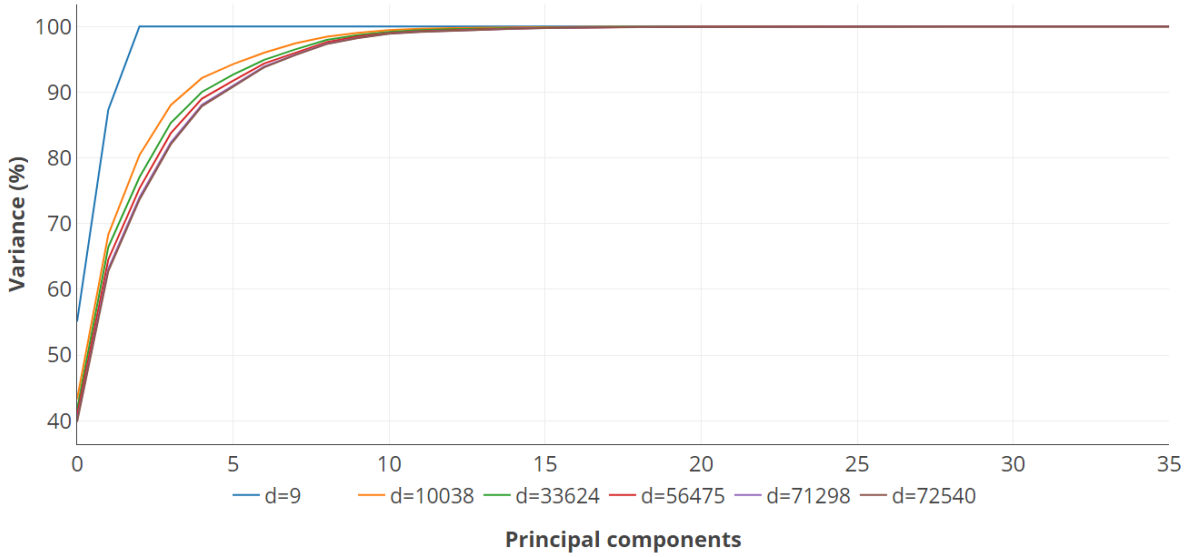
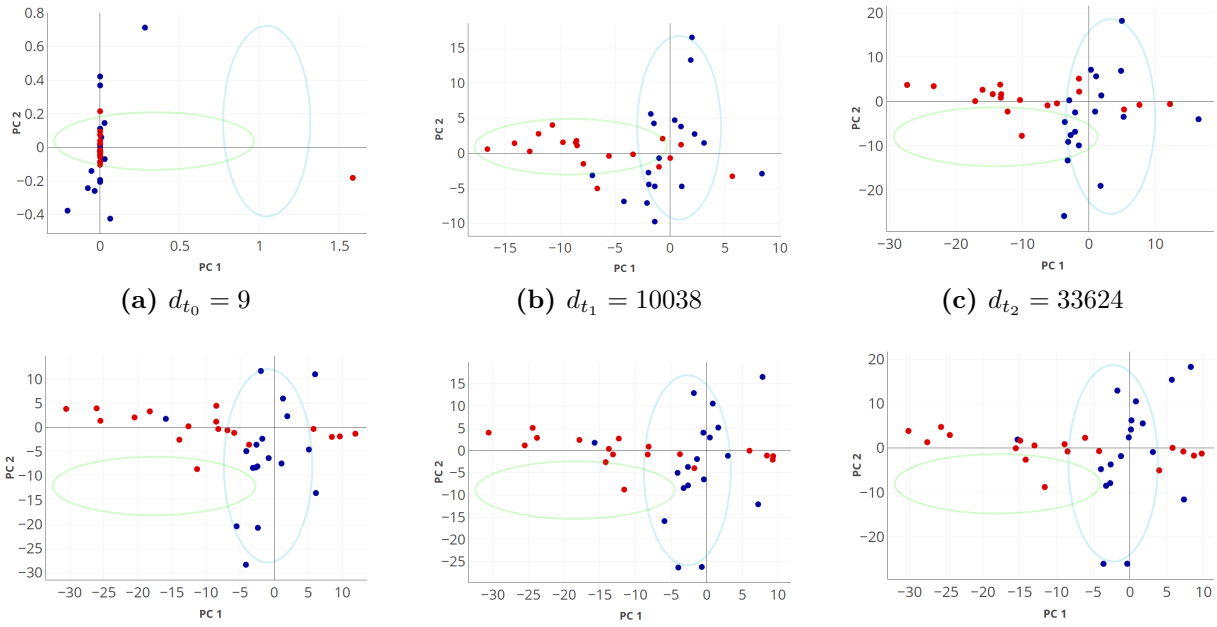


Figure 6.5 Cumulative variances of the data with different dimensionalities using PCA_{nr}. From the dimensionality $d = 9$ to $d = 72540$, the respective coloured plots show the variation of the amount of variance carried by the PCs. The blue plot shows the cummulative variance for minimum dimensionality and the dark purple plot for the maximum dimensionality.

Figure 6.6 shows the scatterplots of the PCA_{nr} with the variation of the thresholds (or dimensionalities). The changing of the data distribution in the PCA_{nr} space when the dimensionality increases can be visualized. In order to track the change in the PCA_{nr} space, two ellipsoids (blue and green) were plotted arbitrary as shown in figure 6.6. The

ellipsoids do not change (in terms of position and size) when the dimensionality increases. This allows for the visualization of the change of the data distribution in **form** space. It was found that when the dimensionality increases from the smaller value ($d_{t_0} = 9$) to the largest value ($d_{t_5} = 72540$), the distribution of data changes considerably. Observe the distribution of the data for smaller dimensionality; almost all the samples are lying on the second PC axis and there is very little clustering of the specimens for each population. When the dimensionality starts to increase, the samples start moving away from the axis and begin to cluster for each population. The distribution becomes relatively similar from the dimensionality d_{t_3} to the largest dimensionality d_{t_5} and it is much easier to distinguish South African data (blue) from the Swiss data. This could be due to the increase of **significant** vertices capturing similarities and differences between data considered for the analysis. This experiment shows how the dimensionality can affect the distribution of data in the PCA$_{nr}$ space. It can be concluded that when the dimensionality increases, the data converges to a stable position in the space. This may help to decide how many **significant** vertices are needed when applying PCA. This result lies with the cumulative variance shown above (figure 6.5).



(a) $d_{t_0} = 9$ **(b)** $d_{t_1} = 10038$ **(c)** $d_{t_2} = 33624$

(d) $dt_3 = 56475$ **(e)** $d_{t_4} = 71298$ **(f)** $d_{t_5} = 72540$

Figure 6.6 Scatterplots of the PCA$_{nr}$ with the variation of the thresholds. From a) to f) scatterplots show the distribution of the South African (blue) and Swiss scapula data (red) for dimensionality from 9 to 72540, respectively. The distribution of the data converges to stabilty as the dimensionality increases.

Example with humerus data

Figure 6.7 shows the cumulative variance of the PCA$_{nr}$ for the humerus data for each threshold. It can be observed that as the dimensionality increases, the amount of variation explained by the first PC decreases. The number of the PCs that explain the total

variability of the data increases with the dimensionality, but it is low compared to the scapula. This may be due to the fact that there was a lot of variation in terms of scale in the humerus data, therefore very few PCs explain the total variability. As previously discussed for the scapula, it can be concluded that the larger the dimensionality, the larger the number of signals capturing similarities and differences between the data.

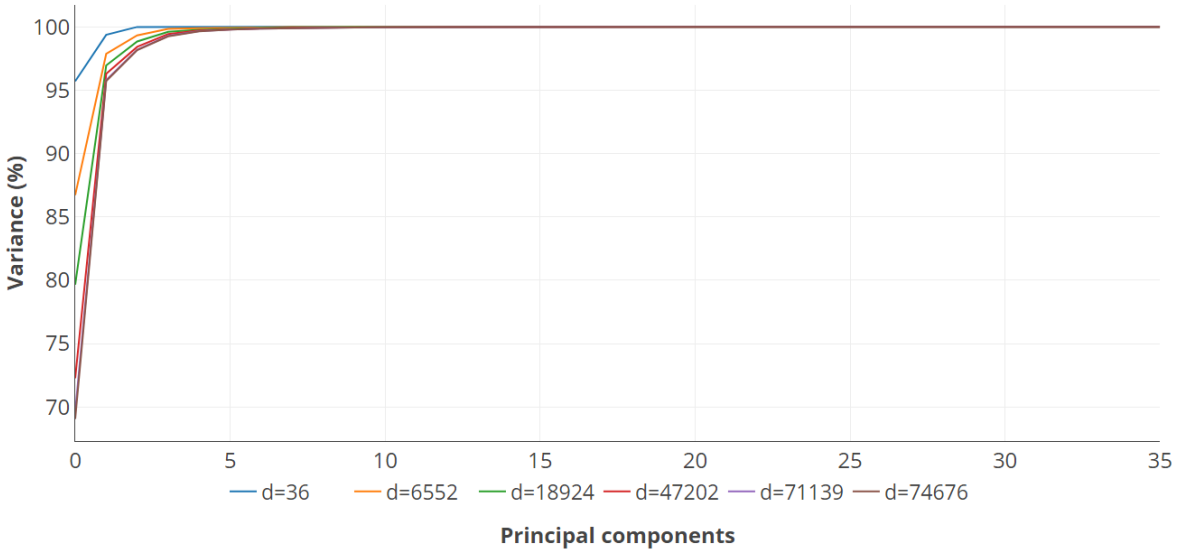
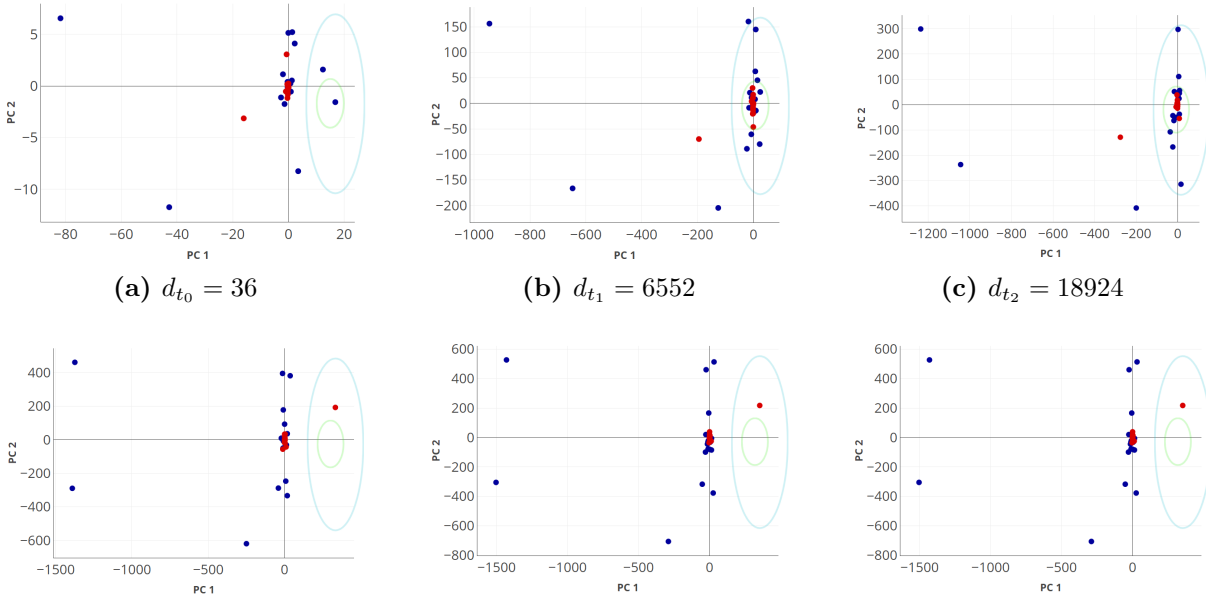


Figure 6.7 Cumulative variances of the humerus data with different dimensionalities using PCA_{nr}. From the dimensionality $d = 36$ to $d = 74676$, the respective coloured plots show the variation of the amount of the variance carrying by the PCs. The blue plot shows the cummulative variance for minimum dimensionality and the dark purple plot for the maximum dimensionality.

Figure 6.6 shows the scatterplots of the PCA_{nr} with the variation of the thresholds (or dimensionalities) for the humerus data. The changing of the data distribution in the PCA_{nr} space when the dimensionality increases can be visualized. As above for the scapula, two ellipsoids (blue and green) were plotted arbitrary as shown in figure 6.8 to track the changing of the distribution. It was found that when the dimensionality increases from the smaller value ($d_{t_0} = 36$) to the bigger value ($d_{t_5} = 74676$), the distribution of data does not change considerably. Observe the distribution of the data for the smaller dimensionality; almost all the samples are lying on the second PC axis and there is little population clustering. When the dimensionality starts to increase, still no clusters form. The distribution becomes relatively similar from the dimensionality d_{t_3} to the largest dimensionality. Contrary to the distribution of scapula above, it is very hard to distinguish South African (red) and Swiss (red) in terms of position in low-dimensional space. This could be due to the high scale/size related variation of the humerus data which represents a high amount of total variation of the data. This experiment shows how the dimensionality can affect the distribution of data in PCA_{nr} space. For this data, it can be concluded that when the dimensionality increases, the distribution of data is relatively stable.



(a) $d_{t_0} = 36$ **(b)** $d_{t_1} = 6552$ **(c)** $d_{t_2} = 18924$

(d) $d_{t_3} = 47202$ **(e)** $d_{t_4} = 71139$ **(f)** $d_{t_5} = 74646$

Figure 6.8 Scatterplots of the PCA$_{nr}$ with the variation of the thresholds for the humerus data. From a) to f) scatterplots showing the distribution of the South African (blue) and Swiss scapula data (red) for dimensionality from 36 to 74676 respectively. The distribution of the data is relatively stable as the dimensionality increases distribution to the biggest dimensionality (f).

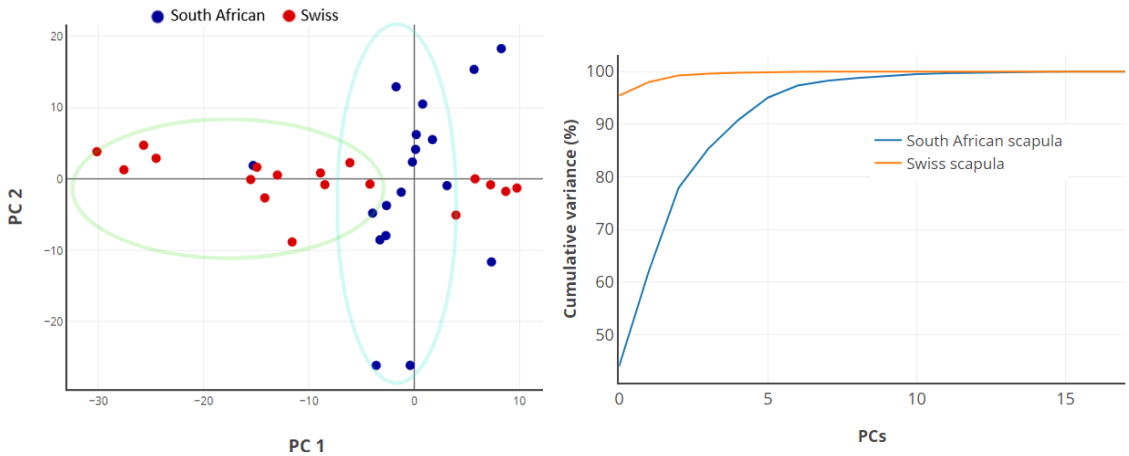
6.5 Comparative study of the South African and Swiss shoulder datasets using PCA$_{nr}$

This study presents the use of the PCA$_{nr}$ for population comparisons. To the author's best knowledge, no mesh-based assessment PCA has been presented in the literature to evaluate shoulder **shape**/**form** differences. In the PCA$_{NR}$ application above using different thresholds, there was no clear comparison of both populations since it was not the aim of that experiment. The goal of this experiment is to draw a comparative mesh **shape** and **form** variability analysis between the South African and Swiss shoulder (humerus and scapula) datasets. The potential outcome of this analysis is to clearly establish a mesh-based difference between South African and Swiss shoulder shapes in low-dimensional space. The analysis was performed for **shape** and **form**, separately, and both were compared. In order to analyse, the data with **form**, the size was not removed after. To consider **shape** information only, the size was removed by scaling each mesh with its centroid size. The PCA$_{nr}$ was first performed to a combined dataset that includes both populations (South African and Swiss) in order to visualise both populations in the same low-dimensional space. Secondly, the PCA$_{nr}$ was applied to each dataset, separately (South African; Swiss), in order to compare the cumulative variance of both populations. Each step above was done for each bone (scapula and humerus). Only the first two PCs were considered which are the ones capturing the most variability. Two ellipsoids were plotted for study case to form clusters for each population in order to distinguish them. The aim

was to draw the ellipsoid that forms the largest cluster for each population. In each case, the blue ellipsoid represents South African data and the green ellipsoid represents the Swiss data.

Case study 1: Comparison of South African and Swiss scapulae datasets

Form comparison-Figure 6.9a shows a scatterplot of the data scores in the first two PCs showing projection of the **form** space. The green ellipsoid in the plot shows approximately 72% of the Swiss population and the blue one approximately 83% of the South African population. Only one South African specimen overlays with the Swiss one. A South African specimen belongs to the blue confidence ellipse with approximately 80% certainty and a Swiss specimen belongs to the green ellipse with approximately 92% certainty. This difference between the populations is confirmed by the cumulative variance in figure 6.9b. The first PC explains approximately 95.4% of the total variance for the Swiss specimens while it only explains approximately 44.1% of the total variance. This clearly shows that South African and Swiss scapula specimens are highly distinguishable in **form** space.

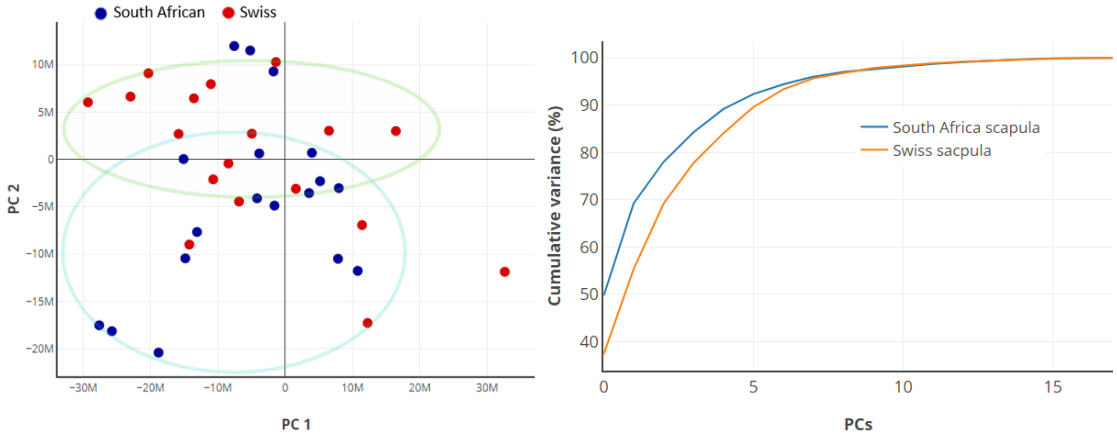


(a) (b)

Figure 6.9 Comparison of South African and Swiss scapula in **form** space using PCA_{nr}. *a*) Scatterplot of the two first PCs in **form** space (Swiss (red) and South African data (blue)); *b*) Cumulative variation of the PCA_{nr} of the populations considering **form** (South African (blue) and Swiss (orange)).

Shape comparison-Figure 6.10a shows a scatterplot of the data scores in the first two PCs showing projection of the **shape** space where the difference in scapula **shape** is the most distinguishable between two populations. The green ellipsoid in the plot shows approximately 72% of the Swiss population and the blue one approximately 83% of the total South African population. A South African specimen belongs to the blue ellipsoid with approximately 68% certainty and Swiss specimen belongs to the green ellipse with approximately 61% certainty. This means a specimen of a population could be correctly classified in the **shape** space with approximately 60% probability. It is difficult to distinguish between both populations and this could be confirmed by the differences between the cumulative variance of the populations in figure 6.10b. The first PC explains approximately 37.3% of the total variance for the Swiss data (orange plot) while it explains

approximately 49.6% for the South African data (blue). Comparing the distribution of the populations in **shape** space with the one in **form** space, it is easier to distinguish between the populations in **form** space than in **shape** space. The populations are more quantitatively different in **shape** than **form** space.

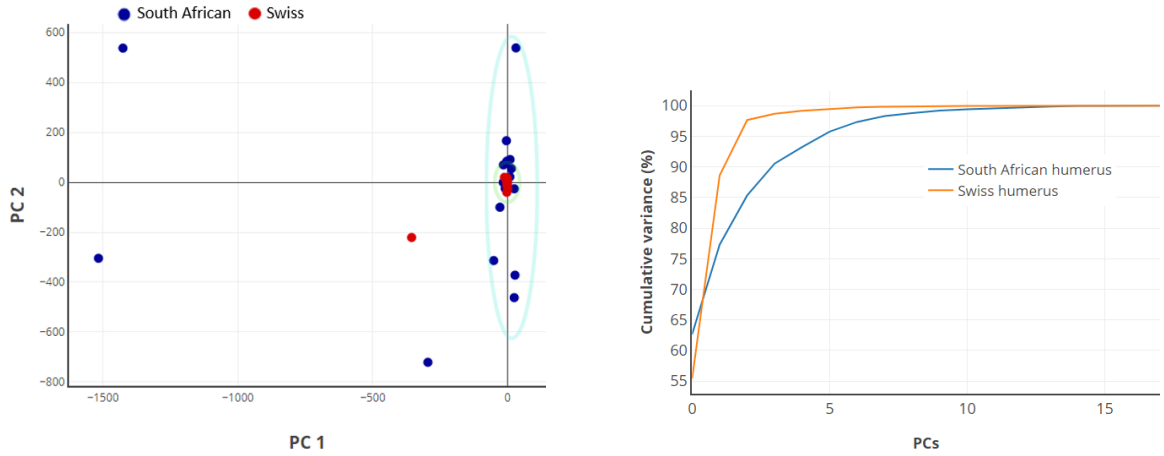


(a) **(b)**

Figure 6.10 Comparison of South African and Swiss scapula in **shape** space using PCA_{nr}. *a*) Scatterplot of the two first PCs in the **shape** space (Swiss (red) and South African data (blue)); *a*) Cumulative variation of the PCA_{nr} of both population considering **shape** only (South African (blue) and Swiss (orange)).

Case study 2: Comparison of South African and Swiss humerus

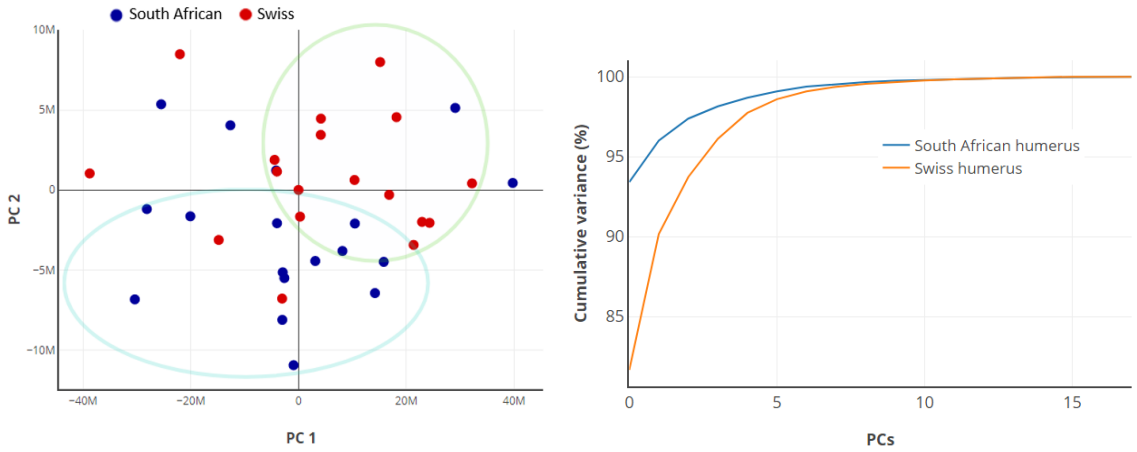
Form comparison-Figure 6.11a shows a scatterplot of the data scores in the first two PCs showing projection of the **form** space for the humerus. It was found that approximately 95% of Swiss specimen were clustered around each other and 61% of the South African were clustered around each other. The blue ellipsoid in the plot represents approximately 88% of data for both populations. The green ellipsoid approximates 95% of the Swiss data. The amount of variance explained by the first PC for both population data was different, 62.6% for South African and 55.5% for Swiss (figure 6.11b). Therefore, the humerus population are difficult to distinguish in **form** space.



(a) **(b)**

Figure 6.11 Comparison of South African and Swiss humerus in **shape** space using PCA_{nr}. *a*) Scatterplot of the two first PCs in **form** space (Swiss (red) and South African data (blue)); *a*) Cumulative variation of the PCA_{nr} of both population considering **form** (South African (blue) and Swiss (orange)).

Shape comparison-Figure 6.12a shows a scatterplot of the humerus data in the first two PCs showing projection of the **shape** space. The green ellipsoid in the plot shows approximately 77.7% of the Swiss population and the blue one approximately 72% of the total South African. A South African specimen belongs to the blue ellipsoid with approximately 81% certainty and a Swiss specimen belongs to the green ellipse with approximately 73.6%. This means a specimen of a population could be correctly identified in the **shape** space with at least 73% of probability. Figure 6.12b shows the difference between the variance of both populations. It was found that the first PC of the South African humerus data explains 93.4% of the total variability while it explains 81% for the Swiss data. This suggests that the populations are distinguishable in **shape** space. Comparing the distribution in **shape** with the one in the **form** space in figure 6.12a, individual specimens are further from each other in **shape** space than in **form** space.



(a) (b)

Figure 6.12 Comparison of South African and Swiss humerus in **shape** space using PCA_{nr}; *a*) Scatterplot of the two first PCs in **shape** space (Swiss (red) and South African data (blue)); (a) Cumulative variation of the PCA_{nr} of both population considering **shape** only (South African (blue) and Swiss (orange)).

6.6 Discussion

Mesh-based PCA presented in Vaclav and Pelikan (2015) did not address the HDLSS issue. In this chapter, a PCA protocol using the NR method in the context of HDLSS for GM was presented. First, the PCA_{tr} was presented and its failure was addressed using the NR methodology which led to the PCA_{nr}. Next, the ORS method introduced in the previous chapter was used to evaluate the distribution change in the PCA_{nr} space for different dimensionality. Finally, PCA_{nr} was applied in a comparative study between South African and Swiss shoulder **form** and **shape** difference analysis.

The distribution of data in the PCA_{tr} and PCA_{nr} space was shown to be different. The distribution in the PCA_{tr} maybe affected by the issue of HDLSS as shown in Yata and Aoshima (2012). The PCA_{nr} also performed better in reconstructing the data than the PCA_{nr}, although marginally. This protocol opens a new way of performing mesh based GM analysis while managing the HDLSS phenomenon that was until recently not addressed in this field (Vaclav and Pelikan, 2015). However, there was no ground truth distribution to measure the accuracy of the PCA_{nr}. For further work, it will be important to design an experiment where the original distribution of the data is known and both PCAs (PCA_{nr} and PCA_{tr}) are used to approximate the original distribution. This will further validate the PCA_{nr} from an experimental point of view.

The second part of this chapter illustrated the utilisation of the ORS method developed in PCA_{nr} using different thresholds. It was found that considering more **significant** vertices in the mesh based PCA_{nr} enables a better estimation of the original distribution of the data in low-dimensional space. It was also found that removing vertices that are **insignificant** does not affect the estimation of the original distribution of the data. The scatterplot of first two PCs in the PCA_{nr} space for the scapula and the humerus data

showed that the Swiss is distinguishable from the South African samples by visualisation when considering almost the whole surface as a **significant** area. In traditional sparse correspondence analysis, this difference would not be visible. This highlights the importance of dense correspondence in GM. However, it is important to note that no outlier detection was performed before the PCA. There may be some outliers affecting the distribution of the data in low-dimensional space. Further study could aim to investigate this as it could obscure the effect that one wishes to observe in the PCA space.

The last part of this chapter detailed the mesh-based variability analysis between South African and Swiss shoulder (scapula and humerus) data. It was found that any mesh-based variability analysis should be very specific either in terms of **form** or in terms of **shape** for both bones. The South African and Swiss shoulder datasets were more distinguishable between them in **form** rather than in **shape** space. However, humerus specimens were less distinguishable between the two populations in **form** than in the **shape** space. Statistical analysis to assess whether differences between the populations are significant, is the topic of the next chapter.

Chapter Seven

Statistical analysis for population comparisons

7.1 Introduction

One of the principal tasks of modern statistical learning is to make meaningful inference based on variables. The multivariate analysis involves simultaneous analysis of multiple independent variables with multiple dependent variables using matrix algebra (Cohen et al., 2013). Data sets with more variables than observations (HDLSS data) are now common in many fields, especially in medical imaging (Mazumder, 2012). The multivariate analysis of these types of data has become a serious challenge to the statistical community. An important and challenging research direction in this context is in building an efficient protocol for statistical inference (Mazumder, 2012). As presented in the literature review, some methods have been developed to address this issue in the field of applied statistics. However, these methods need to be applied in medical imaging fields especially in GM. This section is devoted to the test of significant mean differences using the "two-sample tests for HDLSS data under a strongly spiked eigenvalue model" developed by Ishii et al. (2017). This enables quantitative conclusions on biological differences between the two groups of surface meshes. The classification is finally performed using the discriminant analysis (DA), which allows identification of an unknown biological structure by classifying it in one of the existing family of biological structures (Venables and Ripley, 2002). This analysis focuses on different groupings of the same humerus and scapula data on which the PCA_{NR} was applied in chapter 6.

7.2 Multivariate analysis of variance (MANOVA) test

The multivariate analysis of variance (MANOVA) aims to study the main effects and the interactions among the independent variables, the importance and the strength of association between dependent variables (French et al., 2008). This analysis is useful in experimental situations where independent variables are manipulated. In the context of surface mesh data, the variables are vertex coordinates. The MANOVA assumes that there are linear relationships among all pairs of variables (Stegmann and Gomez, 2002) i.e the space in which the analysis is conducted is linear, which is the case for tangent space onto which our data were projected in the previous chapter (section 6.2).

7.2.1 Test for significant mean shape/form differences

The first goal of the MANOVA is to derive the hypothesis tests and find the appropriate statistical test. Assume two hypotheses: H_0 (null hypothesis) and H_1(non-null hypothesis), as defined below. The MANOVA is a powerful method to test in a probabilistic way whether the null hypothesis has to be rejected or not (Everitt and Skrondal, 2002). However, the popular Hotelling's $\mathbf{T}^2$ test fails when dealing with HDLSS data (Chi and Muller, 2013*b*). Hotelling's $\mathbf{T}^2$ test is not defined when $d > n$ since the sample covariance matrix does not have an inverse. This makes power statistics in GM analysis collapse when dealing with HDLSS data. The NR methodology allows the definition of an appropriate statistic that could be used for **F**-test in the HDLSS context (Ishii et al., 2017). The **F**-test coupled with NR methodology avoids the inversion of the covariance matrix which is not possible to compute using traditional methods.

The following derivation of statistical test is adopted from Ishii et al. (2017). Assume two populations π_i defined by the $d \times n_i$ matrix, $\mathbf{X}_i = [X_{1(i)} \dots X_{n_i(i)}], i = 1, 2$; where $X_{k(i)} = (x_{k1(i)} \dots x_{kd(i)})^T$, $k = 1 \dots n_i$ are independent and identically distributed as a d-dimensional multivariate distribution with mean $\boldsymbol{\mu}_i$; covariance matrix $\boldsymbol{\Sigma}_i$ and $\mathbf{X}_i, i = 1, 2$ are the configuration vectors of shape landmarks; d is the dimensionality of the data; n the number of samples and [,] is the horizontal concatenation of vectors. There is no assumption between the covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ (Ishii et al., 2017). Consider the following two-sample mean test,

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \qquad vs \qquad H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$$

Let's define the following MANOVA statistic

$$T_n = \| \bar{X}_1 - \bar{X}_2 \|^2 - \sum_{i=1}^{2} tr(\mathbf{S}_{Di})/n_i$$

where $\mathbf{S}_{Di} = \sum_{k=1}^{n_i} (X_{j(i)} - \bar{X}_i)^T (X_{j(i)} - \bar{X}_i)$ is the sample covariance matrix of the population π_i and $\bar{X}_i = \sum_{k=1}^{n_i} X_{k(i)}$ its sample mean. An approximation of the distribution (F_0) of the data can be defined as follows:

$$F_0 = \frac{n_1 n_2}{n_1 + n_2} \times \frac{T_n + (\sum_{i=1}^{2} \lambda_{nr1(i)}) n_i^{-1}}{\lambda_{nr_1(n)}}$$

where $\lambda_{1(n)} = (n_1 + n_2 - 2)^{-1} \sum_{i=1}^{2} (n_i - 1) \lambda_{nr_{1(i)}}$ and $\lambda_{nr_{1(i)}}$ is the NR eigenvalues estimator of $\mathbf{S}_{Di}$. The significant difference between π_1 and π_2 is tested as follows for a given upper percentage $\alpha \in (0, \frac{1}{2})$:

$$\text{rejecting } H_0 \iff F_0 > F_{1,\nu(\alpha)} \tag{7.1}$$

where $F_{1,\nu(\alpha)}$ is the F-test with 1 and $\nu = n_1 + n_2 - 2$ degrees of freedom, α is the maximum p-value (significance level) for which the null hypothesis is rejected (Ishii et al., 2017). The significance level α corresponds to the value of the test statistic $F_{1,\nu(\alpha)}$ and the p-value hypothesis test result corresponds to the test statistic F_0.

The null hypothesis H_0 above is either rejected or not rejected. Consequently, before undertaking the experiment, two possible things can happen. Either a difference between

groups is found when one does not actually exist; i.e. the probability of making the wrong decision when the null hypothesis is true. This is called a Type I error (α). Likewise, it is possible that the test will not be able to identify a difference that does exist. This type of wrong outcome is called a Type II error and generally denoted β. This ambiguity is solved by computing the following power and sample determination.

7.2.2 Sample size and power determination

Once the statistical test method is derived, it remains important to evaluate the credibility of the outcome of the test. This credibility is measured by the power which refers (equation 7.3) to the probability that the test will find a significant difference when such a difference actually exists. It is generally accepted that power should be greater or equal to a certain fixed value p; that is, one should have a $p \times 100\%$ or greater chance of finding a statistically significant difference when it exists. The power statistic is directly correlated with the sample size (Hauer, 2004). Increasing the sample size can also give greater power to detect differences. Therefore is it important to know the sample size required to have a specific power.

To derive the sample size required for a given power, the following method in Aoshima and Yata (2011) is adopted. Let $\delta = \|\mu_1 - \mu_2\|^2$ be the square norm difference between the two population means. The objective is to find a required sample size for the given values of the type I error α and the power no less than $1-\beta$ when δ is greater than a fixed value Δ, where β is the type II error and $(\alpha, \beta) \in (0, 1/2)$. It is of interest to know whether the sample size $n_i, i = 1, 2$ is enough for the given type II error (β) and significant level (α). If the sample size required is less or equal to the actual sample sizes then one can conclude that the statistically significant difference found by the test actually exists with a probability of $1 - \beta$, which is the power.

The following is the formulation of the sample size required given the hypothesis above. Let $k \geq 4$, θ_{i*} $(i = 1, 2)$ be a positive lower bound of $\sqrt{tr(\Sigma_i^2)}$ such that $\theta_{i*}/\sqrt{tr(\Sigma_i^2)} \in (0, 1)$, as $d \longrightarrow \infty$ and $\sigma_* = \min_{i=1,2}(\sqrt{\theta_{i*}} \sum_{j=1}^{2} \sqrt{\theta_{j*}})$. (n.d.$b$) defined an integer m as:

$$m = \max\left\{k, \left[\frac{(z_\alpha + z_\beta)\sqrt{2}\sigma_*}{\Delta}\right] + 1\right\}, \tag{7.2}$$

and estimated the total sample size (N_i) for each population π_i as :

$$N_i = \max\left\{m, \left[\frac{(z_\alpha + z_\beta)\sqrt{2}}{\Delta} tr(\mathbf{S}_{im_1}\mathbf{S}_{im_2})^{1/4} \sum_{k=1}^{2} tr(\mathbf{S}_{km_1}\mathbf{S}_{km_2})^{1/4}\right] + 1\right\} \tag{7.3}$$

where $[x]$ denotes the largest integer less than x, $\mathbf{S}_{im_1} = \sum_{k=1}^{m_1}(X_{j(i)} - \bar{X}_{m_1(i)})(X_{j(i)} - \bar{X}_{m_1(i)})^T/(m_1 - 1)$ is the sample covariance matrix of the first block the data and $\mathbf{S}_{im_2} = \sum_{k=m_1+1}^{m}(X_{j(i)} - \bar{X}_{m_2(i)})(X_{j(i)} - \bar{X}_{m_2(i)})^T/(m_2 - 1)$ the sample covariance matrix of the second block of the data as introduced in Yata and Aoshima (2010), with $m_1 = [m/2] + 1$ and $m_2 = m - m_1$. The critical value z_α is the upper bound point of $N(0, 1)$.

In equation 7.2, m should be less than n_i (The actual size of the population π_i). The objective is to find the required sample size N_i only if it is greater than the actual sample size

n_i, otherwise $N_i = n_i$. Taking this into consideration, m can be taken as $m = \min\{n_1, n_2\}$. The value of δ is approximated by $\delta = \| \bar{X}_1 - \bar{X}_2 \|^2$ and Δ could be approximated by $\Delta = \delta - \epsilon$. In order to get the minimum sample size required, Δ should be maximal, therefore ϵ should be the smallest positive number. Once the sample required to get a given power is known, it can then be said if the statistical test result is conclusive or not.

7.3 Mesh-based MANOVA test between South African and Swiss shoulder data: Comparative analysis of form and of shape

A recent study reported that there were no significant differences between South African and Swiss glenoid (Dey et al., 2017). Another study has shown that there are no bilateral differences in the South African shoulder (Inyang et al., 2017). Those studies did not provide any information related to shape and did not provide any quantitative mesh-based comparison. The aims of study presented here were to evaluate the quantitative mesh based differences between the South African and Swiss scapula and humerus both in **shape** and in **form** as well as bilateral (left and right) shoulder bone differences in the South African population. In addition to this, the surface to surface difference was computed between both means for each group and compared by computing the Hausdorff distances between them.

The datasets (humerus and scapula) were divided into the two groups (South Africa (SA) and Swiss (SW)) and subgroups (left and right South African) as described in table 7.1.

Table 7.1 Dataset partitions for each bone (scapula and humerus).

	Group		Subgroup	
	SA	SW	Right	Left
Group size	18	18	9	9

Tests were made using the rejection region approach (Anderson and Hauck, 1983), which allows the finding of the significant level that must be kept at or below 5% (cut-off for significance). The cut-off of the power statistics was set at 80% which gives a reasonable probability (Cohen, 1992) to detect an effect when there is actually one. To put it another way, the study had no more than $\beta = 20\%$ probability of making type II error since $power = 1 - \beta$.

Case study 1: Tests for significant differences between SA and SW scapulae

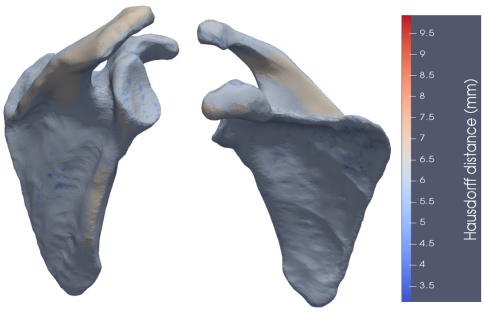
In terms of form- Table 7.2 shows the outcome of the MANOVA **form** difference test for the scapula data. The test failed to find a significant difference between the SA and Swiss scapulae with a p-value greater than 0.05 when considering **form**. This result was confirmed by the surface to surface differences between the mean of SA and SW data as shown in figure 7.1a. The failure of finding a significant difference between SA and SW

datasets corroborates the previous result in Dey et al. (2017), which reported no significant difference between the SA and SW glenoid. However, this could also be attributed to the small size sample; there were not enough samples to get a statistical power at 80%. A further study could be focused on performing the same analysis but with the required size sample as indicated in tables 7.2.

In terms of shape-Table 7.3 shows the outcome of the MANOVA **shape** difference test for the scapulae data. There were a significant difference between SA and SW scapulae at 5% significant level when considering **shape**. This result could also be confirmed by the surface to surface distance differences between the mean of SA and SW scapula as shown in figure 7.1b. In figure 7.1b many areas including the acromion, the coracoid and the glenoid showed a high variation confirming the difference between the scapula means in **shape**; contrary to when considering **form**. This outcome was significant with a statistical power greater than 80% since the sample size was large enough for this test. A previous study (Dey et al., 2017) only focused on **form** differences and found that there were no significant differences between the two population glenoids. However, the results here show that there is actually a significant difference between SA and SW scapulae, at least for these data in terms of **shape**. However, it should be noted that the datasets were not screened for ancestral homogeneity. Future research may perform a screening test for the homogeneity of the populations and redo the analysis.

Table 7.2 F-test for the significant mean **form** and **shape** differences between the SA and SW scapulae data at $\alpha = 5\%$ significant level with a power greater than 80%.

	Form		**Shape**	
Groups	SA	SW	SW	SA
Group size	18	18	18	18
p-value	> 0.05 (fail to reject H_0)		< 0.05 (reject H_0)	
Sample size required (N_1, N_2)	$(18, 23)$		$(18, 18)$	



(a) Population **form** differences.

(b) Population **shape** differences.

Figure 7.1 South African and Swiss scapulae mean differences using surface to surface distance. a) In terms of **form**; b) In terms of **shape**. Red color indicates areas of high differences while blue color indicates areas of lower difference.

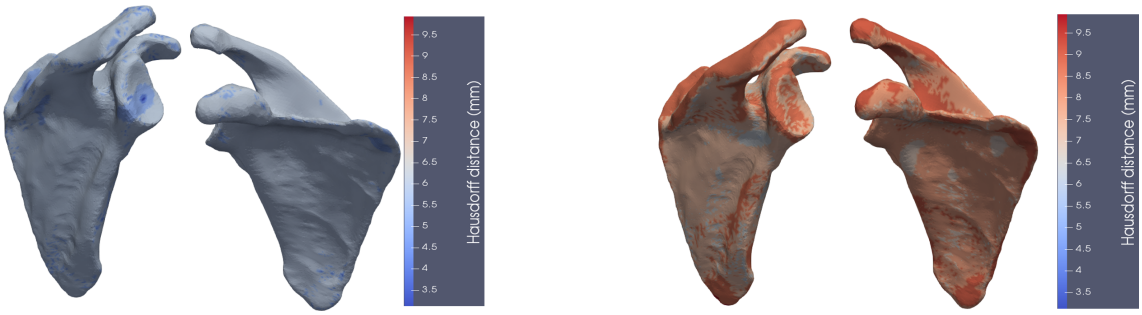
Case study 2: Test for significant differences between left and right scapulae in the SA data

In terms of form-Table 7.2 shows the outcome of the MANOVA test for bilateral **form** differences in the SA scapulae. The test failed to find significant difference with a p-value greater than 0.05. This result was confirmed by the surface to surface differences between the mean of SA left and right scapulae as shown in figure 7.2a. The results here corroborate a previous report (Inyang et al., 2017) which found that there are no bilateral differences between the SA left and right scapula bones. However, this lack of significant difference may be attributed to the small size of the sample. Performing the same analysis but with a large enough sample size (see table 7.2) may lead to different results.

In terms of shape-Tables 7.3 shows the outcome of the MANOVA test for bilateral **shape** difference in the SA scapula. It was found that there was a significant difference between left and right scapula (p-value less than the significant level 0.05). This result was also confirmed by the surface to surface distance differences between the mean of SA left and right scapulae as shown in figure 7.2b. In figure 7.2b many areas including the acromion, the coracoid and the glenoid show relatively higher variation confirming the difference in **shape** compared to **form**. This result is significant with a statistical power greater than 80% since the sample size was large enough. Previous work (Inyang et al., 2017) only focused on **form** differences and concluded that there were no bilateral differences for the same population comparisons. However, the results here show that there are in fact significant bilateral differences for SA scapulae in terms of **shape**; albeit for the heterogeneous data used.

Table 7.3 F-test for the significant mean **form** and **shape** differences between left and right scapulae data at $\alpha = 5\%$ significant level with a power greater than 80%.

	Form		**Shape**	
Groups	Right	Left	Right	Left
Group size	9	9	9	9
p-value	> 0.05 (fail to reject H_0)		< 0.05 (reject H_0)	
Sample size required (N_1, N_2)	$(19, 9)$		$(9, 9)$	



(a) Bilateral **form** differences. **(b)** Bilateral **shape** differences.

Figure 7.2 Bilateral scapulae means differences using surface to surface distance. a) In terms of **form**; b) In terms of **shape**. Red color indicates areas of high differences while blue color indicates areas of lower difference.

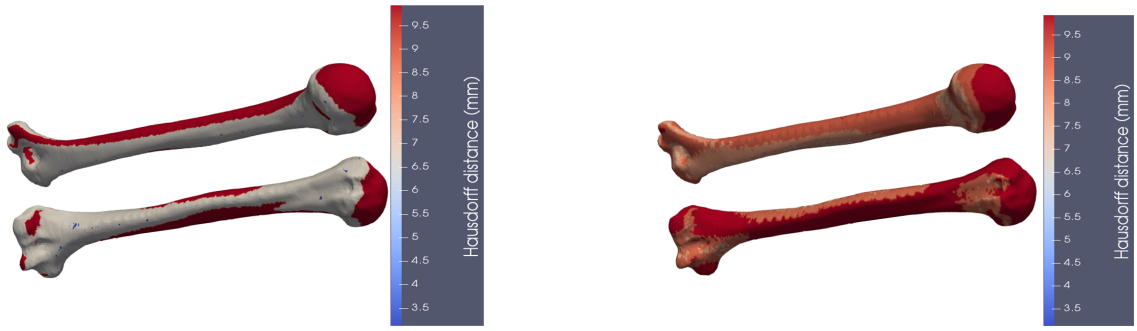
Case study 3: Tests for significant difference between SA and SW humeri

In terms of form- Table 7.4 shows the outcome of the MANOVA **form** difference test between SA and SW humerus datasets. There is a **form** difference at 5% significant level between SA and SW humeri. Figures 7.3a shows the surface to surface **form** differences between the SA and SW means where it can be observed that the humeral head and part of the humeral shaft show large differences. These results also confirm previous findings (Inyang et al., 2017). The differences are significant with a statistical power of at least 80%. However, there was no screening for population homogeneity.

In terms of shape- Table 7.4 show the outcome of the MANOVA **shape** difference test for the humerus dataset. It was found that there is a significant **shape** difference between SA and SW humeri (p-value less than 0.05). Figures 7.3b shows the surface to surface mean **shape** differences of the SA and SW. It can be observed that the humeral head and the humeral shaft exhibit large difference between the two population mean shapes. However, this result is not significant because the difference could be attributed to the small sample size i.e. the type II error of the test was greater than the cut-off of 0.2.

Table 7.4 F-test for the significant mean **form** and **shape** differences between SA and SW humeri data at $\alpha = 5\%$ significant level with a power greater than 80%.

	Form		**Shape**	
Groups	SA	SW	SA	SW
Group size	18	18	18	18
p-value	< 0.05 (H_0 rejected)		< 0.05 (reject H_0)	
Sample size required (N_1, N_2)	$(18, 18)$		$(61, 45)$	



(a) Population **form** differences. **(b)** Population **shape** differences.

Figure 7.3 South African and SW humeri mean differences using surface to surface distance. a) In terms of **form**; b) In terms of **shape**. Red color indicates areas of high differences while blue color indicates areas of lower difference.

.

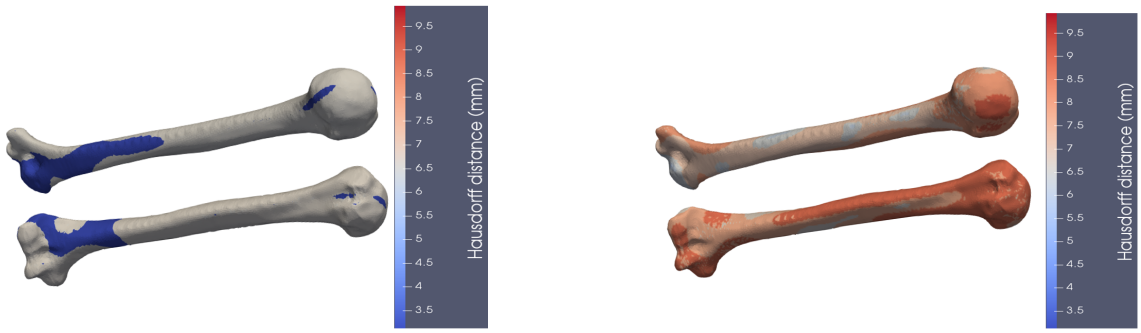
Case study 4: Test for significant differences between left and right SA humerus in the SA data

In terms of form- Table 7.5 shows the outcome of the MANOVA **form** test for bilateral **form** difference in the SA humerus means. The test failed to find significant difference with a p-value greater than 0.05 and it can be observed in figure 7.4a that almost all area of the surface are in grey meaning there is relatively small difference. The results here corroborate a previous report (Inyang et al., 2017). However, the failure in finding significant differences may be attributed to the sample size, which was too small to attain 80% power.

In terms of shape- When considering **shape**, the MANOVA test between left and right SA humerus means failed to find significant difference. However, it can be observed in figure 7.4b that there is difference between both means. This lack in finding significant difference could be attributed to the small sample size.

Table 7.5 F-test for the significant mean **form** and **shape** difference between left and right SA humeri data at $\alpha = 5\%$ significant level with a power greater than 80%.

	Form		**Shape**	
Groups	Right	Left	Right	Left
Group size	9	9	9	9
p-value	> 0.05 (fail to reject H_0)		> 0.05 (fail to reject H_0)	
Sample size required (N_1, N_2)	$(52, 46)$		$(50, 54,)$	



(a) Bilateral **form** differences. **(b)** Bilateral **shape** differences.

Figure 7.4 Bilateral humeri means differences using surface to surface distance. a) In terms of **form**; b) In terms of **shape**. Red color indicates areas of high differences while blue color indicates areas of lower difference.

7.4 Discriminant analysis: Classifying into two population groups

Linear discriminant analysis (DA) of two classes is defined as a classification problem. The goal is to find a good predictor for a class for any sample given an observation (Venables and Ripley, 2002). This analysis is the inverse of the MANOVA in the sense that the independent variable are predictors and the dependent variable are the groups. The DA automatically determines some optimal combination of variables so that the first function provides the most overall discrimination between groups, the second provides the second most and so on (Poulsen and French, 2008). Moreover, the functions will be independent or orthogonal; this is to prevent the overlapping of the classification between groups. The first function picks up the most variation; the second function picks up the greatest part of the residual unexplained variation (Poulsen and French, 2008). The discriminant function is defined below for the classification of two populations.

Assume the two populations π_i defined above (section 7.2) and two conditional probabilty density functions $p(X_k|y=0)$ and $p(X_k|y=1), k=1\ldots n_1+n_1$, X_k is an object in the training set, $\{y=0\}$ and $\{y=1\}$ are two classes. The discriminant function f_i is the minimum value to assign an object X_k to the class i,where class 1 is $\{y=0\}$ and class 2 is $\{y=1\}$ (Venables and Ripley, 2002). This function is defined as:

$$f_i = (X_k - \bar{X}_i)^T \mathbf{S}_{Di}^{-1}(X_k - \bar{X}_i) + ln(det(\mathbf{S}_{Di})),\ k=1\ldots n_1+n_2 \tag{7.4}$$

where $det(\mathbf{S}_{Di})$ denotes the determinant of $\mathbf{S}_{Di}$ and $\mathbf{S}_{Di}^{-1}$ is the Moore-Penrose pseudoinverse of $\mathbf{S}_{Di}$ i.e $\mathbf{S}_{Di}\mathbf{S}_{Di}^{-1}\mathbf{S}_{Di} = \mathbf{S}_{Di}$ (Ben-Israel and Greville, 2003). The object X_k is classified as an element of the class i if f_i is the minimum of the set $\{f_i, i=1,2\}$.

7.5 Example of mesh-based classification using DA

The DA method detailed above was used to classify the same datasets as in the MANOVA tests. The analysis also focused on the same group and subgroup of data (see table 7.1)

used for the mean difference tests above which considered both **form** and **shape**. In this study, however, the aim was to perform a mesh-based classification in both **form** and **shape**, and then compare the results. Discriminant functions were computed for each sample in order to estimate its class (1 or 2). A sample was correctly classified in the class $i(i = 1, 2)$ if its discriminant function for that class was maximal. The specimens were classified using leave-one-out cross-validation. Leave-one-out cross-validation removes a specimen from the training dataset and uses the classifier computed from the remaining training dataset to classify it. This tests how the DA can correctly or incorrectly classify an unknown biological structure into existing groups of biological structures.

a) Scapula classification

SA and SW scapulae -Table 7.6 shows the results for population classification. Swiss samples were classified correctly (100%) while 100% of the SA samples were incorrectly classified when considering **form**. When considering **shape**, 16.67% of SW samples were classified correctly while 66.67% of SA smaples were correctly classified. The classifications were poorer when considering **form** compared to **shape**; this could be attributed to the failure of finding a significant difference when considering **form** while tests found significant differences when considering **shape**.

Table 7.6 Classification of the SA and SW scapulae data with DA

	Form		**Shape**	
	SA samples	SW samples	SA samples	SW samples
Correctly classified	00%	100%	66.6%	16.6%
Incorrectly classified	100%	00%	33.3%	83.3%

Left and right SA scapula-Table 7.7 shows the results of the bilateral classification of the SA humerus. Left-sided samples were all correctly classified and the right samples were all incorrectly classified when considering **form**. With regard to **shape**, 22.22% of the left samples were correctly classified while the right samples were all correctly classified.

Table 7.7 Classification of the left and right SA scapulae data with DA

	Form		**Shape**	
	Right	Left	Right	Left
Correctly classified	00%	100%	100.0%	22.2%
Incorrectly classified	100%	00%	00.0%	77.7%

b) Humerus classification

Left and right SA humeri-Table 7.8 shows the results for the population classification. SW samples were classified correctly (100%) while 55.55% of the SA samples were classified correctly when considering **form**. With regard to **shape**, 83.33% of the SW samples were classified correctly while 77.78% of the SA samples were classified correctly. It could be

observed that at least 55.55% of the SW and SA samples were correctly classified in both cases (**from** and **shape**); this shows the good performance of the DA when there is significant difference between two class means.

Table 7.8 Classification of the SA and SW humeri data with DA

	Form		**Shape**	
	SA samples	SW samples	SA samples	SW samples
Correctly classified	55.6%	100.0%	77.7%	83.3%
Incorrectly classified	44.4%	00.0%	22.2%	16.7%

Left and right SA humeri-Table 7.9 shows the results of the bilateral classification. More than 55.55% of the right-sided samples were classified correctly while 100.0% of the left-sided were incorrectly classified when considering **form**. With regard to **shape**, 100% of the left-sided samples were correctly classified while 100.0% of the right-sided samples were incorrectly classified.

Table 7.9 Classification of the left and right SA humeri data with DA

	Form		**Shape**	
	Right	Left	Right	Left
Correctly classified	55.6%	00.0%	00.0%	100.0%
Incorrectly classified	44.4%	100.0%	100.0%	00.0%

7.6 Discussion and computational challenges

-Discussion
In this chapter, a statistical method for difference in mean **forms**/**shapes** and classification in HDLSS context have been detailed.

In the first section, a method for mean differences tests and power analysis was presented. It uses a common mathematical concept with the previous chapters of this thesis. This algorithm was applied to the training datasets (scapula and humerus bones from SA and SW populations). It was found that the significant mean **shape** and **form** difference results of the data differ. Only the difference between the SA and SW humerus means was shown to be significant in both **shape** and **form**. Most of the group and subgroup means tests were shown to be significantly different at 5% level when considering **shape** while they show insufficient evidence at the 5% significance level to support the claim that there is significant difference when considering **form**. The sample size required was small to get a power statistic greater or equal to 80% which is the cut-off give an acceptable type II error for MANOVA test. However, the significant difference between left and right SA scapulae when considering **shape** is significant. The failure of finding significant bilateral **form** difference in the SA shoulder bones corroborates a previous report (Inyang et al., 2017), which confirms that there is no bilateral difference. From a prosthesis design

perspective it is important to know a priori whether or not the design is based on **form** or on **shape** since a **shape** analysis result cannot be interpreted as **form** analysis.

In the second section, the DA exploired classification of the training datasets into different groups and subgroups. It was found that samples are more frequently correctly classified when considering **shape** than when considering **form**. At least 66.66% of the SA samples were correctly classified in **shape** while all of them were incorrectly classified in **form** for the scapula. With regard to the SA humerus samples, 77.73% and 66.67 were correctly classified in **shape** and in **form**, respectively.

-Computational challenges
To complete the study above, there were computational challenges. Mesh-based quantitative analysis has a price in terms of time-consumption, computer memory cost and finding the right linear algebra algorithmic libraries. In this project, this complexity came from the fact that the number of vertices defining the meshes was relatively large and the multivariate analysis method above requires the covariance matrices of data (see equations 7.3). The covariance matrix size increases with the dimensionality of the shape i.e. the number of the vertices that define the mesh. For the above experiments, the size of the covariance matrix was $size_{data} = 5,576,504,976$ elements for the humerus and $size_{data} = 5,262,051,600$ elements for the scapula. The elements of the covariance matrix were stored as type double which is represented by 64 bits. The memory required for the computation of the covariance matrix was thus:

$$Memory = size_{data} \times 8 \times 10^{-9}\ GB$$

From this formula, the minimum memory required for the covariance matrix computation is 44.61 GB for humerus and 42.1 GB for scapula. The computer available in the author's research group only had 32 GB of RAM. This led the author to perform the computations requires in this chapter using the University of Cape Town's ICTS High-Performance Computing (HPC) facility (http://hpc.uct.ac.za).

Chapter Eight

General discussion and conclusion

8.1 Summary

The main goal of this research was to develop GM for 3D surface meshes in dense correspondence for population **shape** and **form** comparisons. It was hypothesized that using dense correspondence points for GM would facilitate more insight into **form**/**shape** differences between and within different population. This is relevant to the field of bialogical anthropology; surgery; diagnosis and prosthesis design.

A review of existing literature found that very little literature existed on the GM of the shoulder bone in general, in particular for the South African (SA) and Swiss (SW) populations. The experiments of this project used healthly humerus and scapula bones from SA and Swiss populations obtained from fresh cadavers. Until recently, the common approach of quantitatively analyzing various **shapes**/**forms** has been based on landmarks and semi-landmarks. It has been shown that these approaches may be subject to limitations regarding the description of the shapes using few points (landmarks and semi-landmarks), resulting in different descriptions of shapes. Little research has considered mesh-based GM, which uses all the vertices in the mesh as landmarks for the quantitative analysis. Vaclav and Pelikan (2015) brought the concept of mesh-based quantitative analysis of **shape** and **form**. However, the failure of the PCA and Hotelling's T^2 test was not addressed in their work when dealing with HDLSS which is a common situation in the medical image analysis. Additionally, correspondence remains a big challenge when considering the mesh-based approach.

A modified Gaussian process fitting protocol was developed to establish correspondence using the Gaussian fitting method (section 4.6). This protocol improved the registration of the training datasets compared a similar one presented in Inyang et al. (2017). However, all the vertices in the mesh do not encode equitably morphological variation, which led to the notion of optimal shape representation (ORS). An ORS method was developed, which allowed identification of vertices that encode the most **shape**/**form** differences. The amount of morphological variation carried by a vertex was defined by the volume of the ellipsoid defined by the standard deviations (in x, y and z-direction) of all its corresponding vertices across the datasets. It was found that the humeral head has high variation across the population while the glenoid does not vary significantly. The anterior glenoid was found to be the most varying area of glenoid which may be why it is often the area where 90-98% of shoulder dislocations occur as reported in (Saito et al., 2005). The ORS method was validated using a PCA model by comparing the morphological change of the first PC with the one showed by the ORS method.

The failure of the traditional PCA (PCA_{tr}) when handling a HDLSS data was addressed through the noise reduction (NR) method which allowed to resolve the inconsistency of the eigenvalues by estimating the original eigenvalue of the data matrix with the NR one. The PCA_{tr} was compared with the one using NR method (PCA_{nr}) using scapula data and the PCA_{nr} was shown to perform better in the HDLSS context. The novel ORS method combined with the PCA_{nr} were applied to the training datasets using different thresholds which allowed to visualize the effect of thresholds on the distribution of the data in a low-dimensional space. It was found that when the dimensionality increases, the distribution of the data is relatively stable i.e. the data converges to a stable position in the space.

Mesh-based **shape** and **form** differences across the training datasets were evaluated separately and then compared. This **shape** and **form** comparison was done for the PCA_{nr} and quantitative statistical analysis using F-test coupled with the NR method. The distribution of the data in the low-dimensional space when considering **form** was different to one when considering **shape**. Specimens were farther away from each other in **form** space than in **shape** space using PCA_{nr}. These **shape** and **form** differences were also observed in the significant mean difference tests where most of the specimen group means were significantly different in terms of **shape** while there were not when considering **form**. A significant result here is the difference between left and right SA scapulae when considering **shape**. Discriminant analysis (DA) developed for specimens classification shown to perform well only when there is significant difference between class means of the specimens in question.

8.2 Limitations and future work

Population comparisons require the training sample to be representative. The training dataset sizes were not large enough to be representative for SA and SW populations and datasets were not screened for ancestral homogeneity. It would be beneficial to screen for ancestral homogeneity for SA and SW datasets; acquire the sample size required and then reapply this pipeline for GM analysis.

Apart from improving the sample size and the ancestral homogeneity of the data, one could investigate the effect of the type of the mesh (isotropic or anisotropic) on the mesh-based morphometric analysis framework. It could be also important to investigate the decimation of the meshes using ORS method developed while preserving correspondence and without loss of **shape**/**form** information. Furthermore, it may be of interest to investigate methods for outlier detection that may have obscured the distribution of the data in the PCA space. One could consider using a synthetic or consensus shape as the template for registration that is sensitive to outliers and not the mean shape as was done in this study.

The sample size and power determination with Ishii's two-sample test may not be accurate since it only considers the strongly spiked eigenvalue (SSE) model. It would be of interest to investigate data transformation techniques (Aoshima and Yata, 2018) that transform the SSE model to the non-SSE model, before applying Ishii's two-sample test for sample size and power determination.

The computations for the MANOVA tests and DA were very challenging due to a large number of vertices in the meshes. It would be interesting for future work to fix as small as possible the number of vertices in the mesh when reconstruction the 3D surface mesh after segmentation. This leads to investigate the minimum number of vertices required to represent a specific bone mesh. For the classification, the inverse used still depends on the traditional PCA, possibly affecting the DA. Using an inverse based on the noise reduction technique or other inverse types maybe potential alternatives for addressing the HDLSS effect. One could use a biased corrected support vector machine (SVM) under asymptotic properties as reported in (Nakayama et al., 2017). Another topic is the comparison of the method developed with the one using semi-landmarks. One could consider developing GM method for a large number of semi-landmarks and compare it with the one with dense surface correspondence developed.

8.3 Outcome of the research project

It could be concluded in general that GM analysis for 3D dense surface correspondence has potential to facilitate more insight into **form/shape** differences between and within different groups (populations). However, it should be specified which analysis in terms of **form** or in terms of **shape** is being performed since the two analyses have shown different results.

8.4 Project outputs

-Conference proceedings

- Inyang, A. O., Fouefack, J.R, Borotikar, B., Burdin, V., Roche, S., Sivarasu, S. and Mutsvangwa, T. An anthropometric evaluation of the bilateral shoulder: A measure of handedness, in 2017 26^{th} Congress of the International Society of Biomechanics 2017 (ISB 2017), Brisbane, Australia.

- Fouefack, J.R, Inyang, A. O., Douglas, T. S and Mutsvangwa, T. 2017, Optimal representation of shape for morphological analysis, in 2017 South African Institution of Mechanical Engineering (SAIMechE) Postgraduate Conference, Stellenbosh, Western Cape, South Africa.

- Fouefack, J.R, Inyang, A. O., Douglas, T. S, Borotikar, B., Burdin, V. and Mutsvangwa, T. Addressing high dimensionality and low sample size data for mesh-based morphometric analysis: Shoulder form and shape analyses and implications to biomechanics modelling, submitted to 2017 8th World Congress of Biomechanics, Dublin, Ireland.

-Presentations

- Fouefack, J.R, Inyang, A. O., Douglas, T. S and Mutsvangwa, T. 2017, Minimizing operator bias in determining shape representation, in 2017 Postgraduate research day, Faculty of Health Sciences, University of Cape Town, South Africa.

- Fouefack, J.R, Inyang, A. O., Douglas, T. S and Mutsvangwa, T. 2017, Optimal representation of shape for morphological analysis, in 2017 Health Care Technology Conference, Cape Town, South Africa.

-Data and Coding

- Segmented and categorized South African and Swiss shoulder data are available in the author's research group for any future study.

- A well-commented reusable code was also developed for future use.

References

n.d.*a*.

n.d.*b*.

Abdi, H. and Williams, L. J. 2010, 'Principal component analysis', *Wiley interdisciplinary reviews: computational statistics* **2**(4), 433–459.

Adams, D. C., Rohlf, F. J. and Slice, D. E. 2013, 'A field comes of age: geometric morphometrics in the 21st century', *Hystrix* **24**(1), 7–14.

Ahn, J. 2006, 'High dimension low sample size data analysis', *PhD thesis, University of North Carolina at Chapel Hill. Library* .

Amini, A. R., Laurencin, C. T. and Nukavarapu, S. P. 2012, 'Bone tissue engineering: recent advances and challenges', *Critical Reviews in Biomedical Engineering* **40**(5).

Anderson, S. and Hauck, W. W. 1983, 'A new procedure for testing equivalence in comparative bioavailability and other clinical trials', *Communications in Statistics-Theory and Methods* **12**(23), 2663–2692.

Ansari, N. and Delp, E. J. 1990, 'Partial shape recognition: A landmark-based approach', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(5), 470–483.

Aoshima, M. and Yata, K. 2011, 'Two-stage procedures for high-dimensional data', *Sequential analysis* **30**(4), 356–399.

Aoshima, M. and Yata, K. 2018, 'Two-sample tests for high-dimension, strongly spiked eigenvalue models', *Statistica Sinica* **28**, 43–62.

Bai, Z. and Saranadasa, H. 1996, 'Effect of high dimension: by an example of a two sample problem', *Statistica Sinica* pp. 311–329.

Barillot, C., Haynor, D. and Hellier, P. 2005, 'Medical image computing and computer-assisted intervention', *Medical Image Analysis* **9**(5), 411–412.

Bejdová, Š., Krajíček, V., Peterka, M., Trefnỳ, P. and Velemínská, J. 2012, 'Variability in palatal shape and size in patients with bilateral complete cleft lip and palate assessed using dense surface model construction and 3d geometric morphometrics', *Journal of Cranio-Maxillofacial Surgery* **40**(3), 201–208.

Ben-Israel, A. and Greville, T. N. 2003, *Generalized inverses: theory and applications*, Vol. 15, Springer Science & Business Media.

Besl, P. J. and McKay, N. D. 1992, Method for registration of 3-d shapes, *in* 'Robotics-DL tentative', International Society for Optics and Photonics, pp. 586–606.

Bigoni, L., Velem, J. and B, J. 2010, 'Three-dimensional geometric morphometric analysis of cranio-facial sexual dimorphism in a central european sample of known sex', *HOMO-Journal of Comparative Human Biology* **61**(1), 16–32.

Boileau, P. and Walch, G. 1997, 'The three-dimensional geometry of the proximal humerus', *Bone &Joint Journal* **79**(5), 857–865.

Bookstein, F. L. 1997, *Morphometric tools for landmark data: geometry and biology*, Cambridge University Press.

Bookstein, F. L. and Rohlf, F. J. 1990, *Proceedings of the Michigan Morphometrics Workshop*, University of Michigan Museum of Zoology.

Brombin, C., Mo, G., Zotti, A., Giurisato, M., Salmaso, L. and Cozzi, B. 2009, 'A landmark analysis-based approach to age and sex classification of the skull of the mediterranean monk seal (monachus monachus)(hermann, 1779)', *Anatomia, histologia, embryologia* **38**(5), 382–386.

Brombin, C. and Salmaso, L. 2009, 'Multi-aspect permutation tests in shape analysis with small sample size', *Computational Statistics & Data Analysis* **53**(12), 3921–3931.

Chen, S. X., Qin, Y.-L. et al. 2010, 'A two-sample test for high-dimensional data with applications to gene-set testing', *The Annals of Statistics* **38**(2), 808–835.

Chhabra, N., Prakash, S. and Mishra, B. 2015, 'An anatomical study of glenoid cavity: Its importance in shoulder prosthesis', *Internationa Journal of Anatomy and Research* **3**(3), 1419–1424.

Chi, Y.-Y. and Muller, K. E. 2013*a*, 'Two-step hypothesis testing when the number of variables exceeds the sample size', *Communications in Statistics-Simulation and Computation* **42**(5), 1113–1125.

Chi, Y.-Y. and Muller, K. E. 2013*b*, 'Two-step hypothesis testing when the number of variables exceeds the sample size', *Communications in Statistics-Simulation and Computation* **42**(5), 1113–1125.

Cohen, J. 1992, 'A power primer.', *Psychological bulletin* **112**(1), 155.

Cohen, J., Cohen, P., West, S. G. and Aiken, L. S. 2013, *Applied multiple regression/-correlation analysis for the behavioral sciences*, Routledge.

Cutts, S., Prempeh, M. and Drew, S. 2009, 'Anterior shoulder dislocation', *The Annals of The Royal College of Surgeons of England* **91**(1), 2–7.

Dey, R., Inyang, W., Mutsvangwa, T., Charilaou, J., Roche, S. and Sivarasu, S. 2017, 'A comparative shoulder morphometric study towards a novel metric:peak points', *Bone Joint J* **99**(SUPP 1), 24–24.

Dryden, I. L. and Mardia, K. V. 1998, *Statistical shape analysis*, Vol. 4, J. Wiley Chichester.

Duchon, J. 1977, *Splines minimizing rotation-invariant semi-norms in Sobolev spaces*, Springer.

Egger, B., Kaufmann, D., Schönborn, S., Roth, V. and Vetter, T. 2016, Copula eigenfaces with attributes: Semiparametric principal component analysis for a combined color, shape and attribute model, *in* 'International Joint Conference on Computer Vision, Imaging and Computer Graphics', Springer, pp. 95–112.

Everitt, B. and Skrondal, A. 2002, *The Cambridge dictionary of statistics*, Vol. 106, Cambridge University Press Cambridge.

Fakhry, N., Puymerail, L., Michel, J., Santini, L., Lebreton-Chakour, C., Robert, D., Giovanni, A., Adalian, P. and Dessi, P. 2013, 'Analysis of hyoid bone using 3d geometric morphometrics: an anatomical study and discussion of potential clinical implications', *Dysphagia* **28**(3), 435–445.

Fisher, R. A. 1936, 'The use of multiple measurements in taxonomic problems', *Annals of eugenics* **7**(2), 179–188.

Franklin, D., Freedman, L., Milne, N. and Oxnard, C. 2006, 'A geometric morphometric study of sexual dimorphism in the crania of indigenous southern africans.', *South African journal of science* **102**.

French, A., Macedo, M., Poulsen, J., Waterson, T. and Yu, A. 2008, 'Multivariate analysis of variance (manova)', *San Francisco State University* .

Gerig, T., Shahim, K., Reyes, M., Vetter, T. and Lüthi, M. 2014, Spatially varying registration using gaussian processes, *in* 'International Conference on Medical Image Computing and Computer-Assisted Intervention', Springer, pp. 413–420.

Good, C. R. and MacGillivray, J. D. 2005, 'Traumatic shoulder dislocation in the adolescent athlete: advances in surgical treatment', *Current opinion in pediatrics* **17**(1), 25–29.

Green, D. J., Serrins, J. D., Seitelman, B., Martiny, A. R. and Gunz, P. 2015*a*, 'Geometric morphometrics of hominoid infraspinous fossa shape', *The Anatomical Record* **298**(1), 180–194.

Green, D. J., Serrins, J. D., Seitelman, B., Martiny, A. R. and Gunz, P. 2015*b*, 'Geometric morphometrics of hominoid infraspinous fossa shape', *The Anatomical Record* **298**(1), 180–194.

Green, D. J., Spiewak, T. A., Seitelman, B. and Gunz, P. 2016, 'Scapular shape of extant hominoids and the african ape/modern human last common ancestor', *Journal of human evolution* **94**, 1–12.

Gunz, P. and Mitteroecker, P. 2013, 'Semilandmarks: a method for quantifying curves and surfaces', *Hystrix, the Italian Journal of Mammalogy* **24**(1), 103–109.

Guo, J., Mei, X. and Tang, K. 2013, 'Automatic landmark annotation and dense correspondence registration for 3d human facial images', *BMC bioinformatics* **14**(1), 232.

Han, F. and Liu, H. 2013, Principal component analysis on non-gaussian dependent data, *in* 'International Conference on Machine Learning', pp. 240–248.

Hauer, E. 2004, 'The harm done by tests of significance', *Accident Analysis & Prevention* **36**(3), 495–500.

Hausdorff, F. 1918, 'Dimension und äußeres maß', *Mathematische Annalen* **79**(1), 157–179.

Hennessy, R. J. and Stringer, C. B. 2002, 'Geometric morphometric study of the regional variation of modern human craniofacial form', *American journal of physical anthropology* **117**(1), 37–48.

Hopkins, J. 2016, 'Orthopaedic Surgery', `http://www.hopkinsmedicine.org/orthopaedic-surgery/specialty-areas/shoulder/treatments-procedures/total-shoulder-replacements.htm`. [Online; accessed 17-November-2016].

Hotelling, H. 1931, 'The generalization of student's ratio', *The Annals of Mathematical Statistics* **2**(3), 360–378.

Huttenlocher, D. P., Klanderman, G. A. and Rucklidge, W. J. 1993, 'Comparing images using the hausdorff distance', *IEEE Transactions on pattern analysis and machine intelligence* **15**(9), 850–863.

Hutton, T. J., Buxton, B. F., Hammond, P. and Potts, H. W. 2003, 'Estimating average growth trajectories in shape-space using kernel smoothing', *IEEE transactions on medical imaging* **22**(6), 747–753.

Hutton, T. J., Buxton, B. and Hammond, P. 2001, Dense surface point distribution models of the human face, *in* 'Mathematical Methods in Biomedical Image Analysis, 2001. MMBIA 2001. IEEE Workshop on', IEEE, pp. 153–160.

Inyang, A. O., Fouefack, J.-R., Sivarasu, S., Roche, S., Borotikar, B., Burdin, V. and Mutsvangwa, T. 2017, Assessment of 3d morphological characteristics of the shoulder bones using statistical shape modeling: Prospective application to handedness, *in* 'Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE', IEEE, pp. 1629–1632.

Ishii, A., Yata, K. and Aoshima, M. 2014, 'Asymptotic distribution of the largest eigenvalue via geometric representations of high-dimension, low-sample-size data', *Sri Lankan Journal of Applied Statistics* **5**(4).

Ishii, A. et al. 2017, 'A two-sample test for high-dimension, low-sample-size data under the strongly spiked eigenvalue model', *Hiroshima Mathematical Journal* **47**(3), 273–288.

Jung, S., Marron, J. S. et al. 2009, 'Pca consistency in high dimension, low sample size context', *The Annals of Statistics* **37**(6B), 4104–4130.

Kendall, D. G. 1977, 'The diffusion of shape', *Advances in applied probability* **9**(3), 428–430.

Key, T. and Funk, L. 2013, 'The use of geometric morphometrics as a new method to analyse glenoid bone loss after shoulder dislocation', *World Journal of Medical Education and Research* **3**(1).

Kimmerle, E. H., Ross, A. and Slice, D. 2008, 'Sexual dimorphism in america: geometric morphometric analysis of the craniofacial region', *Journal of forensic sciences* **53**(1), 54–57.

Klingenberg, C. P. 2013, 'Visualizations in geometric morphometrics: how to read and how to make graphs showing shape changes', *Hystrix, the Italian Journal of Mammalogy* **24**(1), 15–24.

Koester, M. C., George, M. S. and Kuhn, J. E. 2005, 'Shoulder impingement syndrome', *The American journal of medicine* **118**(5), 452–455.

Krøner, K., Lind, T. and Jensen, J. 1989, 'The epidemiology of shoulder dislocations', *Archives of orthopaedic and trauma surgery* **108**(5), 288–290.

Kybic, J. and Unser, M. 2003, 'Fast parametric elastic image registration', *IEEE transactions on image processing* **12**(11), 1427–1442.

Laga, H., Kurtek, S., Srivastava, A. and Miklavcic, S. J. 2014, 'Landmark-free statistical analysis of the shape of plant leaves', *Journal of theoretical biology* **363**, 41–52.

Lee, S.-M., Abbott, A. L., Clark, N. A. and Araman, P. A. 2003, Spline curve matching with sparse knot sets: applications to deformable shape detection and recognition, *in* 'Industrial Electronics Society, 2003. IECON'03. The 29th Annual Conference of the IEEE', Vol. 2, IEEE, pp. 1808–1813.

Liu, X., Xie, L., Kruger, U., Littler, T. and Wang, S. 2008, 'Statistical-based monitoring of multivariate non-gaussian systems', *AIChE journal* **54**(9), 2379–2391.

Lüthi, M., Gerig, T., Jud, C. and Vetter, T. 2017, 'Gaussian process morphable models', *IEEE Transactions on Pattern Analysis and Machine Intelligence* .

Lüthi, M., Jud, C. and Vetter, T. 2013, A unified approach to shape model fitting and non-rigid registration, *in* 'International Workshop on Machine Learning in Medical Imaging', Springer, pp. 66–73.

Mardia, K. and Dryden, I. 1989, 'The statistical analysis of shape data', *Biometrika* pp. 271–281.

Mayya, M., Poltaretskyi, S., Hamitouche, C. and Chaoui, J. 2013, Scapula statistical shape model construction based on watershed segmentation and elastic registration, *in* '2013 IEEE 10th International Symposium on Biomedical Imaging', IEEE, pp. 101–104.

Mazumder, R. 2012, Topics in Sparse Multivariate Statistics, PhD thesis, Stanford University.

Morrison, D. F. 1998, *Multivariate analysis, overview*, Wiley Online Library.

Mutsvangwa, T., Burdin, V., Schwartz, C. and Roux, C. 2015, 'An automated statistical shape model developmental pipeline: application to the human scapula and humerus', *IEEE Transactions on Biomedical Engineering* **62**(4), 1098–1107.

Mutsvangwa, T. E., Veeraragoo, M. and Douglas, T. S. 2011, 'Precision assessment of stereo-photogrammetrically derived facial landmarks in infants', *Annals of Anatomy-Anatomischer Anzeiger* **193**(2), 100–105.

Myronenko, A. and Song, X. 2010, 'Point set registration: Coherent point drift', *IEEE transactions on pattern analysis and machine intelligence* **32**(12), 2262–2275.

Myronenko, A., Song, X. and Carreira-Perpinán, M. A. 2006, Non-rigid point set registration: Coherent point drift, *in* 'Advances in Neural Information Processing Systems', pp. 1009–1016.

Nakayama, Y., Yata, K. and Aoshima, M. 2017, 'Support vector machine and its bias correction in high-dimension, low-sample-size settings', *Journal of Statistical Planning and Inference* **191**, 88–100.

Perez, S. I., Bernal, V. and Gonzalez, P. N. 2006, 'Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation', *Journal of anatomy* **208**(6), 769–784.

Peter, A. M. and Rangarajan, A. 2009, 'Information geometry for landmark shape analysis: Unifying shape representation and deformation', *IEEE transactions on pattern analysis and machine intelligence* **31**(2), 337–350.

Pham, D. L., Xu, C. and Prince, J. L. 2000, 'Current methods in medical image segmentation', *Annual review of biomedical engineering* **2**(1), 315–337.

Pontil, M. and Verri, A. 1998, 'Support vector machines for 3d object recognition', *IEEE transactions on pattern analysis and machine intelligence* **20**(6), 637–646.

Poulsen, J. and French, A. 2008, 'Discriminant function analysis', *San Francisco State University: San Francisco* .

Püschel, T. A. and Sellers, W. I. 2016, 'Standing on the shoulders of apes: Analyzing the form and function of the hominoid scapula using geometric morphometrics and finite element analysis', *American journal of physical anthropology* **159**(2), 325–341.

Rohlf, F. J. 1999, 'Shape statistics: Procrustes superimpositions and tangent spaces', *Journal of Classification* **16**(2), 197–223.

Rohlf, F. J. 2001, 'Comparative methods for the analysis of continuous variables: geometric interpretations', *Evolution* **55**(11), 2143–2160.

Saito, H., Itoi, E., Sugaya, H., Minagawa, H., Yamamoto, N. and Tuoheti, Y. 2005, 'Location of the glenoid defect in shoulders with recurrent anterior dislocation', *The American journal of sports medicine* **33**(6), 889–893.

Schlecht, S. H., Bigelow, E. M. and Jepsen, K. J. 2015, 'How does bone strength compare across sex, site, and ethnicity?', *Clinical Orthopaedics and Related Research®* **473**(8), 2540–2547.

Sederberg, T. W. and Parry, S. R. 1986, 'Free-form deformation of solid geometric models', *ACM SIGGRAPH computer graphics* **20**(4), 151–160.

Seshamani, S., Chintalapani, G. and Taylor, R. 2011, Alternative statistical methods for bone atlas modelling, *in* 'SPIE Medical Imaging', International Society for Optics and Photonics, pp. 79642O–79642O.

Shahrani, A. I. 2012, 3D geometric morphometric analysis of tooth shape in hypodontia, PhD thesis, Newcastle University.

Sheets, H. D., Covino, K. M., Panasiewicz, J. M. and Morris, S. R. 2006, 'Comparison of geometric morphometric outline methods in the discrimination of age-related differences in feather shape', *Frontiers in zoology* **3**(1), 1.

Srivastava, M. S. 2006, 'Minimum distance classification rules for high dimensional data', *Journal of multivariate analysis* **97**(9), 2057–2070.

Srivastava, M. S. and Kubokawa, T. 2007, 'Comparison of discrimination methods for high dimensional data', *Journal of the Japan Statistical Society* **37**(1), 123–134.

Stegmann, M. B. and Gomez, D. D. 2002, 'A brief introduction to statistical shape analysis', *Informatics and mathematical modelling, Technical University of Denmark, DTU* **15**(11).

Vaclav, K. and Pelikan, J. 2015, 'Correspondence problem in geometrics morphometric tasks.', *University of prague* .

Venables, W. N. and Ripley, B. D. 2002, *Random and mixed effects*, Springer.

Webster, M. and Sheets, H. D. 2010, 'A practical introduction to landmark-based geometric morphometrics', *Quantitative Methods in Paleobiology* **16**, 168–188.

Wei, S., Lee, C., Wichers, L. and Marron, J. 2016, 'Direction-projection-permutation for high-dimensional hypothesis tests', *Journal of Computational and Graphical Statistics* **25**(2), 549–569.

Westin, C. D., Gill, E. A., Noyes, M. E. and Hubbard, M. 1995, 'Anterior shoulder dislocation: a simple and rapid method for reduction', *The American Journal of Sports Medicine* **23**(3), 369–371.

Yata, K. and Aoshima, M. 2010, 'Effective pca for high-dimension, low-sample-size data with singular value decomposition of cross data matrix', *Journal of multivariate analysis* **101**(9), 2060–2077.

Yata, K. and Aoshima, M. 2012, 'Effective pca for high-dimension, low-sample-size data with noise reduction via geometric representations', *Journal of multivariate analysis* **105**(1), 193–215.

Young, N. M., Capellini, T. D., Roach, N. T. and Alemseged, Z. 2015, 'Fossil hominin shoulders support an african ape-like last common ancestor of humans and chimpanzees', *Proceedings of the National Academy of Sciences* **112**(38), 11829–11834.

Zelditch, M. L., Swiderski, D. L. and Sheets, H. D. 2012, *Geometric morphometrics for biologists: a primer*, Academic Press.

Zhang, J., Zhang, X.-d. and Ha, S.-w. 2008, A novel approach using pca and svm for face detection, *in* 'Natural Computation, 2008. ICNC'08. Fourth International Conference on', Vol. 3, IEEE, pp. 29–33.

Zhou, C., Zwilling, C. E., Calhoun, V. D. and Wang, M. Y. 2014, 'Efficient blockwise permutation tests preserving exchangeability', *International journal of statistics in medical research* **3**(2), 145.