Estimation of Healthy Bone Shape and Density Distribution from Partial Inputs for Implant Design

by

DP Kramer

Thesis presented in partial fulfilment of the requirements for the degree of Master of Engineering (Mechatronic) in the Faculty of Engineering at Stellenbosch University

Supervisor: Dr. J. Van der Merwe
Co-supervisor: Dr. M. Lüthi

April 2022

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: . . . . .April 2022. . . . .

i

Abstract

Estimation of Healthy Bone Shape and Density Distribution from Partial Inputs for Implant Design

D. Kramer

Department of Mechanical and Mechatronic Engineering,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Thesis: Master of Engineering (Mechatronic)

April 2022

When reconstructing segmental bone loss, segmentation and surface reconstruction require extensive specialist knowledge to be repeated for each new patient. This has proven to be time-consuming and cost-inefficient throughout literature and practice. Statistical modelling is widely used in biomedical fields for automated segmentation and is a viable alternative for reconstructing healthy bone anatomy in the absence of healthy contralateral geometry. Therefore, as part of this study, statistical models of shape and appearance were constructed from sample data based on femur and tibia data of the male and female South African population, and their application in automated segmentation, reconstruction and density estimation was investigated. The study uses a novel combination of an active shape and a mean appearance model to estimate missing bone geometry and density distribution from sparse inputs simulating segmental bone loss around the diaphyseal area. Estimations of diaphyseal resections were obtained by probabilistic fitting of the active shape model to sparse inputs consisting of proximal and distal bone data on computed tomography images. The resulting shape estimates of the diaphyseal resections were then used to map the mean appearance model to the patients' missing bone geometry, constructing density estimations. The models constructed reproduced the shape and density distribution of the population with an average error below 1.47 mm and a 90 % density fit. Resected bone surfaces were estimated with an average error below 1.64 mm, and density distributions were approximated above 84 % of the intensity of the original target images. These results fall within the acceptable tolerance limits of reconstructive surgery and appear promising for practical use in patient-specific implant design.

Uittreksel

Skatting van Gesonde Been Vorm and Digtheids Verspreiding van Gedeeltelike Insette vir Inplantaat Ontwerp

D. Kramer

Departement Meganiese en Megatroniese Ingenieurswese,
Universiteit van Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.

Tesis: Magister in Ingenieurswese (Megatronies)

April 2022

Tydens die herstel van segmentele beenverlies, vereis segmentering en oppervlak herkonstruksie uitgebreide spesialiskennis wat vir elke nuwe pasiënt herhaal moet word. Dit is bekend as 'n tydrowend en koste-ondoeltreffende aktiwiteite in literatuur sowel as in praktyk. Statistiese modellering word algemeen gebruik in biomediese velde vir geoutomatiseerde segmentering en is 'n lewensvatbare alternatief vir die herkonstruksie van gesonde beenanatomie in die afwesigheid van gesonde kontralaterale geometrie. Daarom, as deel van hierdie studie, is statistiese modelle van vorm en voorkoms geskep uit steekproefdata gebaseer op femur en tibia data van die manlike en vroulike Suid-Afrikaanse bevolking en hul toepassing in geoutomatiseerde segmentering, rekonstruksie en digtheidsskatting is ondersoek. Hierdie studie gebruik 'n nuwe kombinasie van 'n aktiewe vorm en 'n gemiddelde voorkomsmodel om ontbrekende beengeometrie en digtheidsverspreiding te skat vanaf gedeeltelike insette wat segmentele beenverlies rondom die diafiseale area simuleer. Beramings van diafise-reseksies is verkry deur waarskynlike passing van die aktiewe vorm model op gedeeltelike insette wat bestaan uit proksimale en distale been data op rekenaar tomografie beelde. Die gevolglike vormskattings van die diafise-reseksies is dan gebruik om die gemiddelde voorkomsmodel na die pasiënte se ontbrekende beengeometrie oor te dra en digtheidsskattings te konstrueer. Die modelle wat gekonstrueer is het die vorm en digtheidsverspreiding van die populasie weergegee met 'n gemiddelde fout onder 1.47 mm en 'n 90 % digtheidspassing. Gedeeltelike beenoppervlaktes is geskat met 'n gemiddelde fout onder 1.64 mm en digtheidsverspreidings was akkuraat tot meer as 84 % van die intensiteit van die oorspronklike teikenbeelde. Hierdie resultate val

binne die aanvaarbare toleransiegrense van herkonstruktiewe chirurgie en lyk belowend vir praktiese gebruik in pasiënt-spesifieke inplantaat ontwerp.

Acknowledgements

My sincerest appreciation to Dr. J. van der Merwe and Dr. M. Lüthi, for their guidance, advice and mentorship. Dr. J. van der Merwe, thank you for your diligence and expert guidance. Your input always carried much wisdom. Dr. M. Lüthi, your door was always open. Though we are continents apart, you were never more than a video call away when I needed your guidance.

To my friends at BERG, thank you for all the 5-minute coffee breaks and lamenting about the academic life. Rita and Vincent, in you, I have found lifelong friendship.

My parents, Braam and Lize, thank you for the love, sacrifice and financial support. Without you, this would not have been possible. Your trust and belief in me, drives me more than you will ever know. To my future in-laws, thank you for the laughter and comfort during the COVID-19 pandemic. The time spent with you will never be forgotten.

And finally, to my future wife, Ilicia. Thank you for all the patience, love, understanding and spell checking. You are my laughter, and without you, this time would have been much darker.

Dedications

To our Heavenly Father, who strengthens us each new day.

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations

AAM   Active appearance model

AD   Average distance

AIE   Average intensity error

AM   Appearance model

ASM   Active shape model

BMD   Bone mineral density

CT   Computed tomography

GP   Gaussian Processes

GPMM   Gaussian Process Morphable Model

GUI   Graphical User Interface

HD   Hausdorff distance

HU   Hounsfield units

ICP   Iterative closest point

IM   Intra-medullary

MCMC   Markov chain Monte Carlo

MH   Metropolis-Hastings

MRI   Magnetic resonance imaging

MAM   Metal additive manufacturing

PCA   Principal component analysis

PDM   Point Distribution Models

PPMA   Polymethyl methacrylate

PSD   Positive semi-definite

PSI   Patient-specific implant

SAM   Statistical appearance model

SPD   Symmetric and positive definite

SSAM   Statistical shape and appearance model

SSM   Statistical shape models

SVD   Singular value decomposition

Symbols

b	Scaling coefficient	[mm]
c	Number of variance modes	[−]
D	Observed distribution	[−]
$\overline{\mathbf{g}}$	Mean normalised texture vector	[HU]
$\mathbf{g}$	Normalised texture vector	[HU]
$\mathbf{g}_{im}$	Intensity vector	[HU]
GP	Gaussian Processes	[−]
$\mathbf{I}$	CT Image	[HU]
$\mathbf{k}$	A covariance function	[mm]
L	Length	[mm]
LM	Landmark coordinate	[mm]
n	Total number of training samples	[−]
N	Total number of landmarks or vertices	[−]
$\mathcal{N}$	Normal distribution	[−]
$p()$	Probability function	[−]
P	Intesity profile vector	[HU]
r	Rotation parameter	[rad]
R	Rotation axes	[−]
t	Translation parameter	[mm]
s	Scale of variance	[−]
$\mathbf{S}$	Covariance matrix	[mm]
$\mathbf{T}$	Transformation matrix	[mm]
u	Deformation vector	[mm]
$\mathbf{u}$	A set of deformation vectors	[mm]
$\mathcal{U}$	Uniform Distribution	[−]
x	Point coordinates	[mm]
$\mathcal{N}$	Normal Distribution	[−]

Greek letters

α	Shape coefficient	[−]
$\boldsymbol{\Gamma}$	Shape point set	[mm]
λ	Eigenvalue of covariance matrix	[mm]
μ	Mean distribution	[mm]
$\boldsymbol{\Phi}$	Eigenvectors of covariance matrix	[−]
σ	Range over which deformations are correlated	[−]
θ	Translation, rotation and shape coefficients of model	[mm, rad, −]

Subscripts

d	Dimensionality
g	Intensity voxel
i	Sample
k	Landmark
R	Reference
s	Mode of shape variation
T	Target

Chapter 1

Introduction

1.1 Background

Delayed treatment of severe skeletal trauma, infections, and malignancy can lead to segmental bone loss, which requires additional specialist healthcare in the form of surgical intervention for correction and effective treatment (Wang and Yeung, 2017; Wang *et al.*, 2016). Additional delays in treatment could result in further bone loss, increased morbidity, and possible amputation (Marais *et al.*, 2013).

Bone grafting is the most common method used to treat and correct segmented bone loss and is seen as the gold standard to which all segmented bone loss treatments are compared. Autografts use healthy bone tissue from the patient to treat the affected area. However, extracting the bone tissue introduces a second surgical site that may cause further complications, e.g. increased pain, blood loss, additional recovery time, possible infections etc. Allografts, or donor tissue, are also sometimes used, but this has a reduced success rate due to the patient's immune response and may also introduce viral transmission. It is also mostly not practical due to limited donor availability (Wang and Yeung, 2017; Mauffrey *et al.*, 2015). The use of grafting for lower limb salvage has also proven to be particularly challenging. It often results in 'non-union' and then requires multiple additional surgical procedures, extending the recovery and reconstruction process (Dekker *et al.*, 2018).

1.2 Motivation

Patient-specific implants (PSIs) offer a modern alternative to conventional treatment methods such as grafting and stock implants for segmented bone repair (Hawkins *et al.*, 2020; Siegmund *et al.*, 2019; Elledge *et al.*, 2019; Tetsworth *et al.*, 2017). The advantages of PSIs include reduced recovery time, improved patient outcomes, reduced surgical time, lower long-term costs, and no secondary morbidity (Hamid *et al.*, 2016; Wang *et al.*, 2016; So *et al.*, 2018). Additionally, PSIs for long-bone defects are mainly manufactured using ad-

ditive manufacturing such as 3D printing (Haglin *et al.*, 2016; Zhang *et al.*, 2019).

The PSI design process, however, requires extensive specialist knowledge and procedures, which essentially need to be repeated for each new case. This makes PSI design time-consuming and cost-inefficient when compared to current conventional treatment options (Haglin *et al.*, 2016; Tetsworth *et al.*, 2017). Traditionally, the geometry for PSI design is inferred from healthy contralateral bone anatomy derived from computed tomography (CT) images of the patient (Mauler *et al.*, 2017; Letta *et al.*, 2014). However, healthy contralateral bone anatomy is not always available due to high energy trauma or pre-existing conditions (Mauler *et al.*, 2017; Saleh *et al.*, 1999). Furthermore, significant differences between bilateral anatomy, especially in long bones, have been reported in literature (Schepers *et al.*, 2015; Auerbach and Ruff, 2006).

Statistical models can provide the turn-key solutions required to make PSIs financially and practically feasible. These models could allow automation of time-consuming activities such as the image segmentation, surface reconstruction and density estimation of missing bone anatomy, thereby reducing the overall cost, time and specialist knowledge required for PSI design (Zadpoor and Weinans, 2015; Tetsworth *et al.*, 2017; Heimann and Meinzer, 2009; Sarkalkan *et al.*, 2014; Mauler *et al.*, 2017). Statistical models can represent the normal variability of an anatomical shape and density distribution within a specific population due to the prior knowledge captured when trained.

Statistical shape models (SSMs) and statistical appearance models (SAMs) are trained on the variation of bone shape and density from segmentation and CT data for a statistical population (Lüthi *et al.*, 2017; Cootes *et al.*, 2004). Active shape models (ASMs) combines the shape and density information around the bone surface area from an SSM and SAM. When used in combination with optimization algorithms, the ASM can automate costly activities such as segmentation and estimation of bone geometry (Cootes *et al.*, 2004; Heimann and Meinzer, 2009; Sarkalkan *et al.*, 2014; Morel-Forster *et al.*, 2018). Additionally, active appearance models (AAMs) include the full SAM when a complete image of the entire density distribution from the population is required (Cootes *et al.*, 2004; Sarkalkan *et al.*, 2014). However, AAMs are computationally more expensive to build, process, store, and fit successfully than ASMs, making them inconvenient for practical design applications (Cootes *et al.*, 2004; Sarkalkan *et al.*, 2014).

To make PSIs financially and practically more feasible, a novel method for estimating healthy bone shape and density distribution using an ASM and mean appearance model (AM) is presented in this study. The technique combines ASMs with mean AMs to reconstruct simulated segmental bone lose pathologies within the diaphyseal area of the femoral and tibial long bones. The model combination enables relatively fast geometrical and reasonable density estimates for segmental bone repair, automating PSI design procedures such as segmentation and reconstruction.

1.3 Aim and Objectives

The aim of this study was to develop a method that automatically performs the segmentation and estimation of healthy bone shape and density distribution from partial inputs for segmental bone repair. To achieve this aim, the project objectives were as follows:

1. Construct and validate SSMs and mean AMs for a given bony population from CT images.

2. Fit the validated SSMs and mean AMs to a sparse dataset simulating segmental bones loss, solving the segmentation problem for the long bone structure.

3. Implement optimization and reconstruction algorithms that map healthy shape and density distribution data to partial inputs.

It is important to note that the partial inputs referred to throughout this study represent CT bone data from which the pathology or segmental bone was removed.

1.4 Scope

This project does not include actual PSI design. PSI design is covered only to establish the wider system of interest. The study is focused on the development and implementation of model construction software, the validation of the models constructed, and the estimation of healthy bone shape and density distribution from partial inputs.

Additionally, the fundamental mathematics and considerations of statistical modelling used throughout this thesis were outside the thesis's scope. However, the fundamentals are discussed in much detail within the appendices for the interested reader.

1.5 Overview

A detailed literature review is given in Chapter 2. This is followed by a brief study of statistical models and their mathematical theory in Chapter 3. Chapter 4 will discuss the actual model construction and validation results. Chapters 5 and 6 use the models for segmental bone repair by estimating healthy bone shape and density distribution. Final remarks are given in Chapter 7.

The appendices contain supplementary notes and data to further the understanding of the work represented in this thesis. Appendices A and B have additional notes on software implementation and the mathematical theory behind model construction, geometrical and density re-estimation, and fitting

procedures. Appendix C contains additional documentation such as interviews and the segmentation protocol. Due to size, the raw validation data and test results will be submitted to the main supervisor within the electronic database accompanying this thesis.

Chapter 2

Literature Review

This chapter contains the vital background research required to understand the study objectives set in the introduction. Sections 2.1 to 2.3 briefly discuss the medical and biological aspects of this project. After that, Section 2.4 will review current means of reconstruction for implant design in literature and industry. Additionally, a discussion on design constraints and tolerances for PSI design will also be given in Section 2.4. Chapter 3 will continue the literature review but with a focus placed on statistical modelling.

2.1 Anatomy

Understanding the anatomical features that define long bone structure is necessary to ensure statistical model construction and healthy bone estimation is done correctly. This will also aid in understanding the pathologies that afflict long bone structures and the treatment methods used to correct these pathologies.

2.1.1 Long Bone Anatomy

Long bone structures from the appendicular skeleton can be divided into three main parts called the diaphysis, metaphysis and epiphysis (Betts *et al.*, 2013; Visible, 2020). Figure 2.1 illustrates these divisions within long bone structures, in addition to anatomical attributes associated with long bone anatomy (Betts *et al.*, 2013).

The diaphyseal region consists of a tubular shaft that runs from the proximal to the distal end of a long bone structure. Diaphyseal walls are densely packed cortical bone and enclose a hollow region, the medullary cavity, filled with yellow marrow. The inner wall surrounding the medullary cavity has a membranous lining, called endosteum, responsible for the growth, repair and remodelling of surrounding bone. The metaphyseal regions are comprised of the widening diaphyseal tubular shaft and the epiphyseal growth plate, the leading site of longitudinal growth for long bones. When adulthood is reached
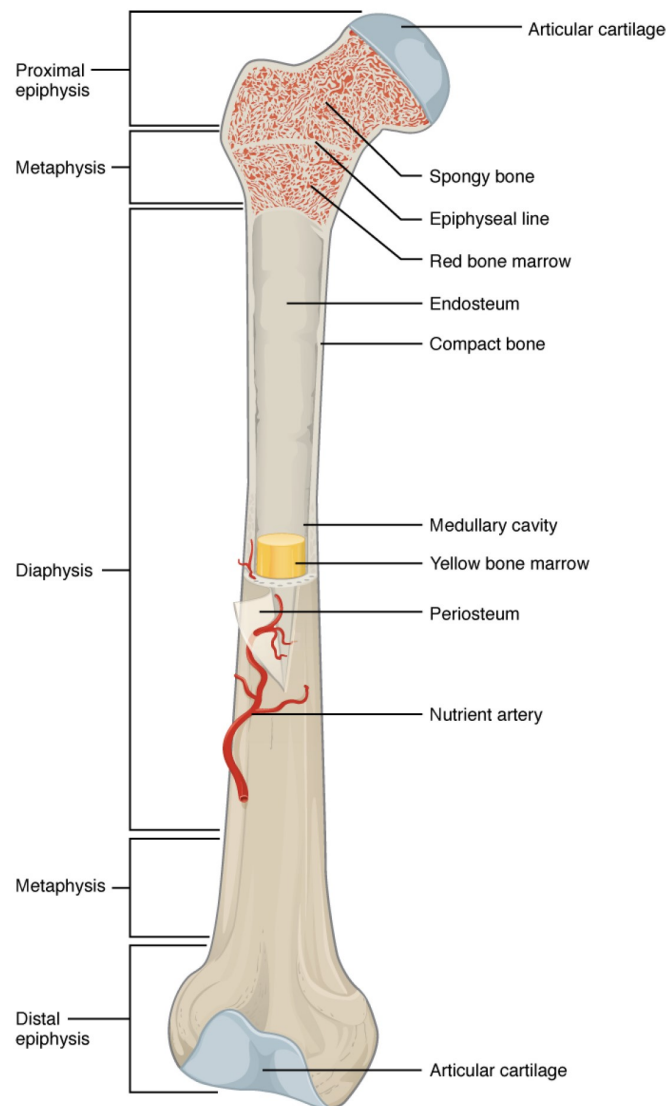
Figure 2.1: Long bone anatomy (Illustration: Betts *et al.* (2013, p. 221))

(18-21 years), a layer of transparent cartilage on the epiphyseal growth plate is replaced with osseous tissues forming the epiphyseal line (Betts *et al.*, 2013). The epiphyseal regions mainly consist of spongy, porous trabecular or cancellous bone with thinly layered sections of cortical bone. The porous trabecular bone is also mostly filled with red marrow (Betts *et al.*, 2013).

The periosteum, a fibrous membrane, forms on the outer surface of long bone structures and contains blood vessels, nerves, and lymphatic vessels that nourish the bone. Tendons and ligaments also attach to long bone structures through the periosteum. This membrane covers most of the long bone surface except where joints form on the epiphyseal regions; articular cartilage covers the joint area on these surfaces. The articular cartilage forms a thin layer that reduces friction and acts as a shock absorber (Betts *et al.*, 2013).
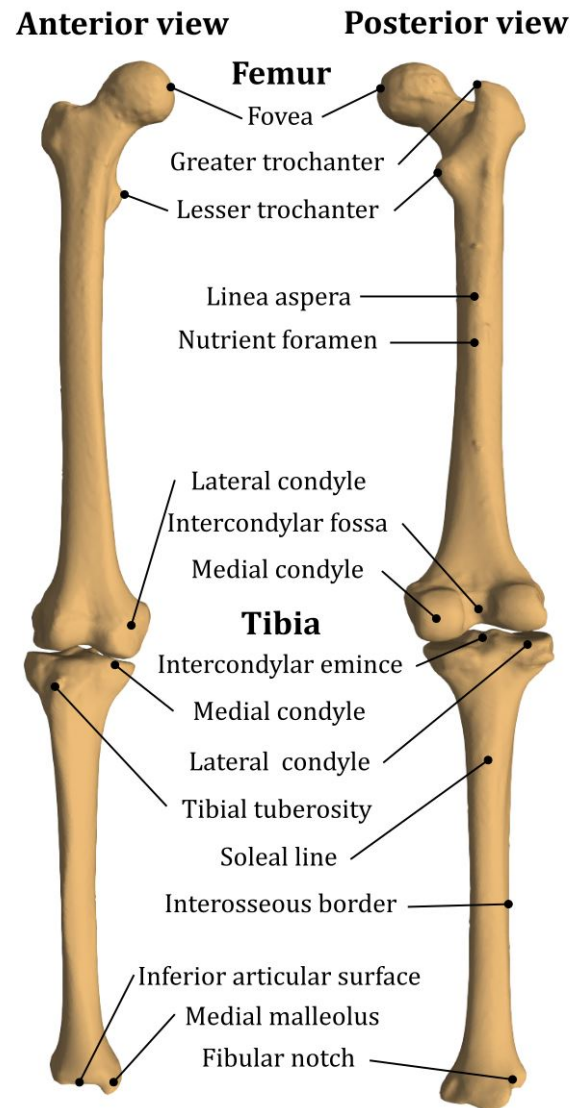
Figure 2.2: Femoral and tibial landmarks (Illustration: D. Kramer)

2.1.2 Anatomical Landmarks

The anatomical landmarks used throughout this study are listed in Tables 2.1 and 2.2, and illustrated in Figure 2.2 for both the femoral and tibial long bones. The definitions given are at times descriptive of an entire region; thus, the most protruding, sunken or central point in the area was used when identifying landmarks within samples.

2.2 Segmental Bone Loss

It is essential to know which areas and parts of long bone structures are affected by pathologies that cause segmental bone loss to automate segmental recon-

Table 2.1: Femoral landmark descriptions (Betts *et al.*, 2013)

Proximal Epiphysis/Metaphysis	
Fovea	A small oval-shaped dimple on the proximal end of the femur head. Forms part of the hip joint.
Greater trochanter	A quadrilateral elevation with a crest projecting upward from the lateral aspect of the femoral neck and shaft junction.
Lesser trochanter	A conical projection arising from the posteromedial surface of the neck-shaft with its angle directed medially.
Diaphysis	
Linea aspera	A longitudinally-oriented ridge on the posterior aspect of the femur, anchors several thigh muscles and functions as a buttress preventing the anterior bowing of the shaft.
Nutrient foramen	A canal through the cortex of the femur, where nutrient arteries and veins pass through.
Distal Epiphysis/Metaphysis	
Medial condyle	The lateral surface that creates the medial boundary of the intercondylar fossa. The epicondyle is the most notable point on the medial condyle.
Lateral condyle	Stouter then the medial condyle but less notable. Creates the lateral boundary of the intercondylar fossa. The epicondyle is the most notable point on the lateral condyle.
Intercondylar fossa	A deep notch, which divides the medial and lateral condyles posteriorly. Part of the knee joint.

struction. Understanding this will also give an estimate of the reconstruction size required.

Segmental bone loss is defined in literature as the smallest bone defect for a particular species that will not enter a state of spontaneous healing or regeneration without some form of intervention or correction (Wang and Yeung, 2017). Bone regions affected by segmental bone loss are generally subject to minimal or no re-vascularization and tissue differentiation, leading to density reabsorption, spontaneous bone fractures and non-union that drastically affect the patient's overall health and standard of life (Wang and Yeung, 2017; Strohmeyer, 2015).

For human adult patients, defects in bone regions greater than 20 mm in length or 50 % of the circumferential area are unlikely to spontaneously self regenerate (Tosounidis and Giannoudis, 2017; Mauffrey *et al.*, 2015). In an epidemiology study done by Keating *et al.* (2005) it was shown that of all the segmental bone loss injuries recorded, 68% was found in the tibial anatomical

Table 2.2: Tibial landmark descriptions (Betts *et al.*, 2013)

Proximal Epiphysis/Metaphysis	
Medial and lateral condyles	A greatly expanded area at the proximal end of the tibia with no epicondyles present. The top surfaces are flattened and smooth. Part of the knee joint.
Intercondylar emince	Situated between the medial and lateral condyles are the medial and lateral intercondylar tubercles, elevated crests that from part of the knee joint.
Tibial tuberosity	On the anterior side of the tibia, an elevated area near the proximal end, attaches a muscle-tendon associated with the patella.
Diaphysis	
Soleal line	A prominent ridge on the posterior aspect of the tibia which anchors several calf muscles. It begins below the base of the lateral condyle and continues down the medial and proximal third of the posterior tibia.
Interosseous border	Formed by a small ridge on the lateral side of the tibial shaft. An interosseous membrane connects to the ridge and unites the fibula and tibia bone.
Distal Epiphysis/Metaphysis	
Medial malleolus	A large bony expansion on the medial side of the distal tibia. Part of the ankle joint.
Inferior articular surface	A smooth, concave and quadrilateral surface articulates with the talus to form part of the ankle joint.
Fibular notch	A wide groove on the lateral side of the distal tibia. The distal tibiofibular joint is formed by the notch that articulates with the distal end of the fibula.

area and 22% in the femoral anatomical area, with a bone loss greater than 30 mm primarily encountered in the tibial region (Tosounidis and Giannoudis, 2017).

Segmental bone loss can usually be attributed to four reasons: a) high energy trauma, b) tumour resection, c) bacterial infection resection and d) revision surgery (Wang and Yeung, 2017; Wang *et al.*, 2016; Strohmeyer, 2015; Tosounidis and Giannoudis, 2017). A more detailed discussion of these reasons can be found in the subsections below.

2.2.1 High Energy Trauma

High energy trauma usually results from high kinetic energy transfer to bone structures within the human body. The energy transfer can be from traffic injuries, blast injuries, crushing injuries or falling from heights. The resulting fractures can be open or closed and generally result in high levels of tissue damage around the fracture site due to high energy transfer (Jovanović *et al.*, 2002).

High energy trauma resulting in open fracture or large amounts of bone splintering generally leads to significant bone and length loss. These fractures also have a 10 to 30 % chance of non-union and poor fracture healing depending on the treatment method used and the fracture location (Wang *et al.*, 2007).

With high energy segmental bone loss, the surrounding tissue damage and bone splintering directly affect re-vascularization and tissue differentiation. More precisely, tissue damage leads to the obstruction of bone volume increase through the recruitment of stem cell differentiation and bone retardation through chondrocyte apoptosis (processes active during blood clotting and inflammatory response stages). If left untreated, it usually results in non-union of the fractured bone site, seriously diminishing patients' standard of life and health (Wang and Yeung, 2017).

2.2.2 Tumour Resection

Tumours are formed when a particular cluster of tissue cells grow irregularly or do not die off during certain life cycle phases. Tumours are classified as benign (non-cancerous) or malignant (cancerous). Benign tumours can grow large but cannot transfer into surrounding tissues or body parts as malignant tumours can through blood and lymph systems. When malignant tumours spread in this way, it is referred to as metastasis (National Cancer Institute, 2018).

When looking at bone tissue, two kinds of tumours can form. The first is primary bone tumours that develop from bone tissue, and the second is metastatic tumours that form elsewhere and then spread to bone tissue. Benign primary bone tumours are more common than primary malignant tumours and are rarely life-threatening. Both compress healthy bone tissue, but benign tumours do not spread and destroy bone tissue as malignant tumours do (National Cancer Institute, 2018).

Table 2.3 lists common tumours that can afflict long bones, with Ewing's sarcoma and Osteosarcoma being the most common. These tumours can cause fractures, pain and hypercalcemia within long bone structures. Treatment methods vary between surgical resection, chemo therapy, radiation, cryosurgery and targeted therapy depending on the location, size and type of tumour (National Cancer Institute, 2018; Chen *et al.*, 2007).

Of these treatment methods, surgical resection has become the most common for malignant tumours in recent years as limb-salvage methods have de-

veloped. During surgery, the surgeon removes the entire tumour with clean margins (National Cancer Institute, 2018). Due to the significant tissue defects in long bone sarcoma, resection usually leads to high volumes of segmental bone loss (Chen *et al.*, 2007; Sewell *et al.*, 2011). In a study done by Chen *et al.* (2007) resection lengths due to sarcoma for the tibial bone ranged between 110 and 270 mm (Chen *et al.*, 2007; Sewell *et al.*, 2011).

Table 2.3: Long bone tumours (National Cancer Institute, 2018)

Tumour	Class	Initially Infected Tissue	Area
Osteosarcoma	Malignant	Osteoblasts in osteoid tissue	Metaphysis
Chondrosarcoma	Malignant	Cartilaginous tissue	Epiphysis
Ewing Sarcoma	Malignant	Bone and soft tissue	Diaphysis
Osteoclastoma (Giant cell tumour)	Benign or Malignat	Bone tissue	Epiphysis, Metaphysis,
Osteoblastoma	Benign	Cortical bone tissue	Diaphysis
Fibrous Dysplasia	Benign	Bone and marrow Tissue	Diaphysis, metaphysis
Myeloma	Malignant	Marrow tissue	Diaphysis
Osteoid Osteoma	Benign	Bone tissue	Diaphysis

2.2.3 Bacterial Infection

Osteomyelitis is a biofilm-based infection that can lead to necrosis and destruction of bone tissue while also inhibiting new bone growth. Osteomyelitis generally takes hold through bacterial inoculation in combination with either ischaemia, trauma or necrosis. It affects both adults and children and infects any bone structure. It is characterised by purulent secretion, necrotic bone, compromised soft tissue and fistulas (Marais *et al.*, 2013; Rao *et al.*, 2011).

Osteomyelitis can spread through haematogenous (blood) or contiguous (trauma) bacterial inoculation. Chronic osteomyelitis usually results from acute haematogenous osteomyelitis that forms necrotic bone segments, resulting in a source of ongoing infection if left untreated. Research shows that 10 to 30% of acute haematogenous osteomyelitis translates to a chronic infection. Contiguous bacterial inoculation generally results from open fractures. Depending on the injury and quality of the subsequent management, 3 to 50% of open fracture cases have been reported to translate to contiguous osteomyelitis (Marais *et al.*, 2013).

Generally, host cells establish a protective cellular layer that includes a defence mechanism and invading bacteria establishes a biofilm. The biofilm forms part of the growth pattern of the invading bacteria. It consists of interactive

communities capable of gene expression alterations, ensuring their survival. The formation of this layer-like aggregation occurs in five stages: 1) adhesion, 2) production of the extra-cellular matrix, 3) colonisation, 4) maturation, and 5) dispersion of bacteria (Marais *et al.*, 2013).

Even with the advancement in antibiotic therapy, treatment of osteomyelitis still proves challenging. Due to mechanical and osmotic difficulties, antibiotics struggle to penetrate biofilm. To date, chronic osteomyelitis, accompanied by necrotic bone, is most effectively treated through surgical debridement (Rao *et al.*, 2011). Depending on the stage of the infection, surgical treatment aims to be curative with wide clean resection margins, ensuring no tissue containing bacteria is left behind. Up to 220 mm defects have been recorded during limb salvage operations (Rao *et al.*, 2011).

2.2.4 Revision Surgery

Revision surgery is a term used to describe a wide variety of operations but generally refers to any surgery used to correct, replace or compensate for implant failure. Implant failure requiring revision include aseptic loosening, osteolysis around well-fixed implants, periprosthetic fractures, implant instability, infections like osteomyelitis, and implant fractures. Of these, aseptic loosening is the most common reason for revision surgery overall, with instability being the most common for early revision and osteolysis around well-fixed implants for late revision (Clohisy *et al.*, 2004).

Aseptic loosening, osteolysis around well-fixed implants and periprosthetic fractures generally occur due to stress shielding. Stress shielding is used to describe bone resorption and density reduction due to stress reduction within bone structures. Stress usually carried by the bone gets transferred to the implant through implant stiffness and fixation. This reaction fits perfectly into Wolff's law, which states that bone will model and remodel according to the loads being carried. Bone resorption and density reduction lead to a decrease in bone volume and strength, leading to spontaneous fractures and implant loosening (Sumner and Galante, 1992; Wang *et al.*, 2016).

The combination of stress shielding, bone resorption, revision and spontaneous fractures can lead to large volumes of segmental bone loss exceeding 80 mm (Tetsworth *et al.*, 2017). Revision surgery leads to increased patient morbidity and, sometimes, additional surgical sites (Wang and Yeung, 2017).

2.3 Segmental Bone Repair

To ensure the reconstruction methods used within this study are adequate, it is necessary to study and understand the current methods used for segmental bone repair. Historically, the standard treatment for long bone defects was amputation (Strohmeyer, 2015). It is only in the last several decades that limb-salvage techniques have developed to become as effective, if not more,

than amputation. If successful, limb salvage restores full limb functionality to the patient, increasing or maintaining the patient's quality of life (Chen *et al.*, 2007). However, limb salvage requires bone reconstruction through surgical intervention to ensure union between functional, healthy bone structures or segments (National Cancer Institute, 2018).

In the 1950s, Ilizarov devised many reconstructive techniques and devices, e.g. external fixation frames, that have become the standard for reconstruction surgeries in long bone defects, such as tibial deformities. Of these techniques devised, Ilizarov describe four fundamental principles for the reconstruction, correction and management of bone deformities: 1) precise amount of fracture support; 2) minimal surgical intervention; 3) immediate weight-bearing; and 4) mobilisation of joints (Saleh *et al.*, 1999). Following these principles, some standard and modern limb salvage methods were investigated in the subsections below.

2.3.1 Bone Grafting

Bone grafting utilises transplanted bone or other osteoconductive material to rebuild damaged bone tissue and stimulate bone regeneration. It has been used as the standard treatment method for segmental bone loss, bone defects, delayed union or non-union, and spinal fusions since the mid-1900s (Wang and Yeung, 2017; Mauffrey *et al.*, 2015; Masquelet and Begue, 2010). Osteoinductive material can roughly be divided into autografts, allografts and synthetic grafts. The osteoinductive material used during surgery is dependent on the defect site, defect size, the grafting material available, and the treatment methods used (Wang and Yeung, 2017; Mauffrey *et al.*, 2015). Autologous bone is seen as the gold standard in grafting material, stimulating bone regeneration through: a) osteoconduction, b) osteoinduction and c) ontogenesis. These three properties are usually not found in the same osteoconductive material (Wang and Yeung, 2017).

In addition to these properties, an optimal mechanical environment is also required to encourage and ensure early weight-bearing capability, secondary healing, bone regeneration and union (Tosounidis and Giannoudis, 2017; Mauffrey *et al.*, 2015). To achieve an optimal mechanical environment, internal or external fixation techniques are used to stabilise pathological or defective long bone structures, as shown in Figure 2.3. These techniques make use of external frames, intramedullary (IM) nails, screws, plates, meshes and cages usually made from a biocompatible material such as titanium (Saleh *et al.*, 1999). When defects exceed 200 mm, maintaining the limb axis becomes difficult externally, and intramedullary nails are used, as shown in Figure 2.3.b. Modern techniques and materials make early weight-bearing safe and desirable, with full weight-bearing capability recorded at 8-12 weeks after surgery (Mauffrey *et al.*, 2015; Saleh *et al.*, 1999).

Long bone defects greater than 50 mm often still leads to non-union due to
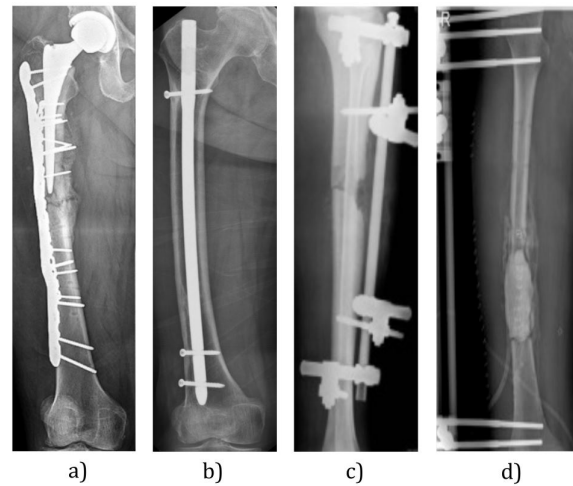
Figure 2.3: Bone fixation : a) Internal fixation for implant fracture repair on a femur (Prins *et al.*, 2018); b) Internal fixation through IM nail(Kim *et al.*, 2021); c) External fixation for grafting procedures on a tibia (Kostic *et al.*, 2019); d) External fixation for Masquelet PPMA spacer (Tong *et al.*, 2017)

graft resorption, even when using autografts (Mauffrey *et al.*, 2015; Masquelet and Begue, 2010). To combat this Masquelet and Begue (2010) describe an induced membrane technique that is performed in two stages. First, the defect is debrided, and a Polymethyl methacrylate (PMMA) cement spacer is used to fill the segmental defect, as shown in Figure 2.3.d. The bone defect is then stabilised with external or internal fixation. It then takes six to eight weeks for a foreign body membrane to form around the PMMA spacer. Second, the PMMA spacer is removed after the membrane is matured while the membrane is preserved and left intact. The membrane is then filled with autograft material. Weight-bearing capability is then dependent on the fixation method used (Mauffrey *et al.*, 2015; Masquelet and Begue, 2010). The technique has three main benefits:

- The induced membrane creates a boundary that limits graft resorption.

- The PMMA spacer allows for delayed bone grafting ensuring the proper grafting volume can be collected.

- The induced membrane is rich in growth factors, improving graft consolidation and union (Mauffrey *et al.*, 2015).

Masquelet and Begue (2010) reconstructed long bone defects between 50 to 240 mm in 35 patients with autograft compositions and the induced membrane, achieving a much lower graft resorption rate. Still, the induced membrane does not guarantee union and requires additional surgery for the spacer to be inserted and the membrane to mature, increasing patient morbidity in the short term.
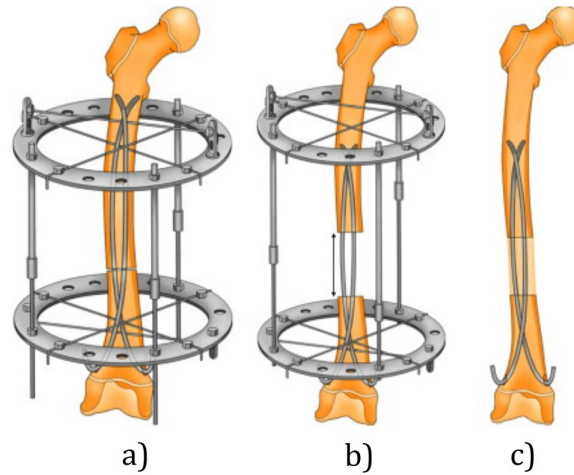
Figure 2.4: Distraction osteogenesis: a) A osteotomy is performed on the diaphyseal region, an Ilizarov frame is placed, and IM nails are introduced. b) The distal femur (transport segment) is then transported 1 mm per day using the frame structure till the defect length is recovered. c) The frame is removed after consolidation, the IM nails are left in place, and reconstruction is complete. (Illustration adapted from: Popkov *et al.* (2010))

2.3.2 Distraction Osteogenesis

Distraction osteogenesis requires the transport of a healthy bone segment through the length of the segmental bone defect (Mauffrey *et al.*, 2015). The bone transfer is controlled and maintained using external fixation or IM devices, e.g. Ilizarov frames, monolateral frames or IM lengthening nails. For long bone defects, an osteotomy is performed on healthy diaphyseal or metaphyseal bone regions to produce the bone segment that will be transported, as shown in Figure 2.4. The method of fixation and distraction is then applied, and the transfer phase can begin after surgery. The bone segment is transferred 1 mm per day for the length of the defect to allow bone tissue formation. Once the end of the defect is reached, the bone segment is compressed for several weeks until fracture callus forms, aiding consolidation (Mauffrey *et al.*, 2015).

The main advantages for distraction osteogenesis include weight-bearing during reconstruction, reliability, minimal soft tissue injury, and no constraints with regards to reconstruction length (Mauffrey *et al.*, 2015). The drawback is the time required for reconstruction. The consolidation phase is generally double the time of the transport phase. Thus, a 10 cm segment can take 300 days (ten to eleven months) to reconstruct, lowering the patient's morbidity for an extended period.

2.3.3 Scaffolds and Truss Cages

With the advancement of additive manufacturing, 3D printed scaffolds, truss cages and lattice structures have become a novel approach for managing seg-
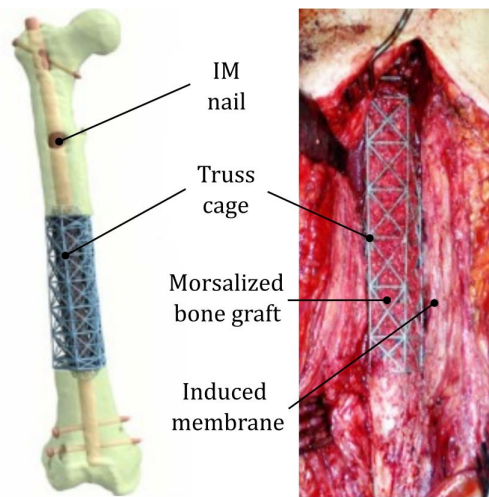
Figure 2.5: Femoral truss cage implant (Illustration adapted from: Tetsworth *et al.* (2017))

mental bone defects (Tetsworth *et al.*, 2017; Zhang *et al.*, 2019). Scaffolds, truss cages and lattice structures (grouped further as scaffolds) are printed using bio-compatible alloys or bio-substitutes, e.g. titanium alloys or bio-ceramics (Haglin *et al.*, 2016; Zhang *et al.*, 2019). Typically, scaffold design for PSIs requires the combined input of an orthopaedic surgeon, orthopaedic engineer and biomedical engineer. The design process typically starts with some form of medical imaging, such as CT or magnetic resonance imaging (MRI), followed by a semi-automated segmentation process that produces a 3D model. This model can then be used to plan and design the scaffold required for segmental bone reconstruction (van der Merwe *et al.*, 2018; Tetsworth *et al.*, 2017).

Scaffolds are used in combination with bone grafting and IM nails. It has also become practice to incorporate the Masquelet technique to minimise graft resorption and stimulate consolidation (Tetsworth *et al.*, 2017). An IM nail is inserted with the scaffold structure during surgery to stabilise the defect site and keep the scaffolding in place. The porous scaffold is then packed with morselised bone grafting material. If the Masquelet technique is used, the defect surface, scaffold and graft material is enclosed within the induced membrane, as shown in Figure 2.5 (Tetsworth *et al.*, 2017; Hamid *et al.*, 2016). Depending on the defect size and location, the patient can then be in full weight-bearing capacity 6 to 12 weeks after the surgery (Hamid *et al.*, 2016; Hollister, 2009).

Porous scaffolds have many advantages when compared to solid metal implants. Solid metal implants have higher stiffness levels than bone which can result in stress shielding. This could lead to resorption of the bone surrounding the implant, which in turn may cause implant failure and additional bone loss (Long and Rack, 1998; Sumner and Galante, 1992). Solid implants also

obstruct new bone growth and nutrient flow at the implant site where porous metal implants more closely match and support the natural properties of bone (Wang *et al.*, 2016; Tetsworth *et al.*, 2017; Ghavidelnia *et al.*, 2020). Porous implants not only carry the required anatomical load but also mitigate stress shielding by adjusting the porosity to reduce the equivalent stiffness (Hollister, 2009; Ghavidelnia *et al.*, 2020). Pore networks will also enable fluid transfer to facilitate bone regrowth and nutrient flow (Zhang *et al.*, 2013).

2.4 Reconstruction and Estimation for Implant Design

This section is focused on methods and tolerance limits used for surface reconstruction in industry and literature. First, surface reconstruction and the general implant design procedure will be investigated in Subsection 2.4.1. Second, the design constraints and tolerance limits used for PSIs will be discussed in Subsection 2.4.2. To gain a better understanding of how surface estimation for PSI design is achieved in practice, an interview with CranioTech (PTY) Ltd. (specialists in patient-specific solutions) was conducted (see Appendix C for interview documentation). The knowledge gained during the interview will be discussed in detail throughout this section.

2.4.1 Procedures for Surface reconstruction and Implant design

Figure 2.6 shows a generalised patient-specific implant design process. First, a medical image of the pathological bone region is taken with healthy contralateral anatomy if available. Manual segmentation on the medical image is then performed to generate 3D models of the surface geometry (Jun and Choi, 2010; CranioTech (PTY) Ltd., 2021). From the image and model, the surgeon decides what is required to surgically correct the pathology and stipulates initial design specifications for the implant (Jun and Choi, 2010; CranioTech (PTY) Ltd., 2021). The contralateral model, statistical model or technical expertise is then used to estimate the target shape required to reconstruct the pathological bone surface. The interface between current and reconstructed surface geometry, and the fixation method of the implant is then developed and integrated into the implant design. FEM analysis, stress, destruction or prototype testing is then done to assess the implant's load-bearing capacity and fixation to the required anatomical surface (CranioTech (PTY) Ltd., 2021). Though surgeon supervision is constant throughout the design process, surgeon approval is required to finalise the implant design and surgical plan. If approval is not given, the design, requirements, segmentation or medical image is altered and the process is repeated. With surgeon approval, the PSI is manufactured ac-
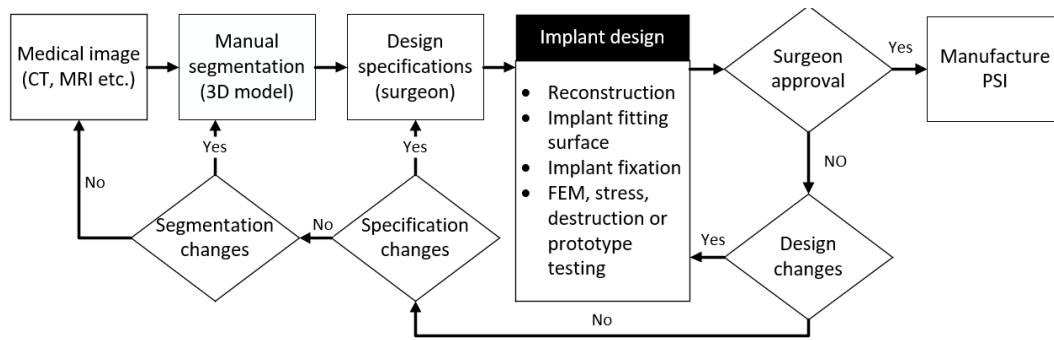
Figure 2.6: Patient-specific implant design procedure (Adapted from: Jun and Choi (2010))

cording to design specifications (Jun and Choi, 2010; CranioTech (PTY) Ltd., 2021).

As already stated, healthy contralateral bone anatomy is the main source used when reconstructing surface geometry, in both literature and practice (CranioTech (PTY) Ltd., 2021; Mauler *et al.*, 2017; Letta *et al.*, 2014). Using medical imaging and segmentation software, the geometrical information of both the pathological and healthy contralateral bone anatomy is extracted and stored in 3D models (CranioTech (PTY) Ltd., 2021; Jun and Choi, 2010). The healthy contralateral bone surface is then mirrored and superimposed on the pathological surface to estimate the surface geometry required for reconstruction and PSI design. In addition to this, specialist knowledge is required to integrate the remaining bone anatomy of the pathological site with the estimated surface geometry and establish the fixation location and methods (Haglin *et al.*, 2016; CranioTech (PTY) Ltd., 2021).

However, healthy contralateral bone anatomy is not always available due to high energy trauma or pre-existing conditions (Mauler *et al.*, 2017; Saleh *et al.*, 1999). Additionally, significant differences between bilateral anatomy, especially in long bones, have been reported in literature (Schepers *et al.*, 2015; Auerbach and Ruff, 2006). In such cases, statistical models or models of healthy bone anatomy are used to infer, based on prior knowledge, the surface geometry generally taken from healthy contralateral bone anatomy (Zadpoor and Weinans, 2015; Heimann and Meinzer, 2009). If no source of example geometry is available in the worst-case scenario, specialist knowledge and technical expertise alone are used to reconstruct the pathological surface geometry (Haglin *et al.*, 2016; CranioTech (PTY) Ltd., 2021).

2.4.2 Estimation Tolerances and Constraints

This study is focused on automated segmentation and reconstruction. Thus tolerances need to be established to evaluate the models constructed and reconstruction methods used. Unfortunately, due to the unique nature of each new case, setting tolerance limits for PSI design and surface reconstruction is

a complex subject. While literature does not state any well-established limits, pre-and post-operative reviews show successful outcomes for implant overhangs and variations in size ranging from 2 to 4 mm for knee and hip arthroplasty Abram *et al.* (2014); KR *et al.* (2020). In an interview with CranioTech (PTY) Ltd. (2021) it was revealed that anatomical and design constraints play a more significant role than actual tolerance limits.

In practice, constraints are unique to each patient and take priority during surface reconstruction, even when working with healthy contralateral anatomy (CranioTech (PTY) Ltd., 2021). Constraints include: a) medical imaging quality, b) bone stock quality and location, and c) fixation sites for implant attachment. The accuracy of the fitting surface and implant is highly dependent on the image quality used for segmentation. Inaccurate imaging could result in poor edge alignment, implant fitting and fixation stability. This could lead to implant instability and possible failure. Additionally, establishing areas for fixation is essential, and depending on the available healthy bone stock, the surface estimation is generally altered to accommodate implant fixation (CranioTech (PTY) Ltd., 2021).

Still, CranioTech (PTY) Ltd. (2021) stipulated that their surgeons are in general comfortable with 2 to 3 mm overhangs. For surface reconstruction, an average error of 5 to 10 mm is acceptable with a Hausdorff error of $\pm$ 10 mm at the centre of the diaphyseal area. A higher error in the centre of the diaphyseal area is accepted if it is required for IM-nail placement when working with truss cage implants. Again, the fixation method used takes priority over surface reconstruction.

Bone density considerations, for CranioTech (PTY) Ltd. (2021), are generally limited to the availability of healthy, stable bone stock for fixation. Implant strength is determined through stress and destructive testing. For femoral truss cages, the load caring capacity is designed to support ten times the weight of an average person. (CranioTech (PTY) Ltd., 2021). Thus, if the average weight was assumed as 80 kg, the lower limit designed for would-be 800 kg. CranioTech (PTY) Ltd. (2021) determined through stress and destructive testing that their titanium femoral truss cages, with varying truss diameters around $\pm$ 1.5 mm, could support a 2 ton load. In practice, truss cages are generally over-engineered to ensure the implant won't fracture or break apart within the patient (CranioTech (PTY) Ltd., 2021).

PSI design is generally more focused on surface reconstruction and implant stability (Zadpoor and Weinans, 2015; Mauler *et al.*, 2017). It is only in recent years that functionally graded lattice structures have been incorporated into PSI design procedures (Ghavidelnia *et al.*, 2020). Thus, accuracy limits for density estimation are not clear. To establish if estimated density distributions from this study are sufficient to generate a functionally graded lattice structure capable of supporting the load associated with the femur and tibial bones, additional work in future studies will be required.

Chapter 3

Study of Statistical Models

This chapter is focused on the basic theory required when approaching statistical modelling. Section 3.1 gives a theoretical overview of the main concepts necessary when constructing statistical models. Section 3.2 follows with a brief discussion on how statistical models are used and fitted to example data. Additionally, supplementary notes on statistical modelling and mathematical considerations can be found in Appendix B. Chapters 4 to 6 discuss the modelling considerations and implementations used throughout this thesis in greater detail.

3.1 Statistical Shape and Appearance Models

When applied to anatomical bone structures, SSMs and SAMs respectively describe the mean shape and density distribution along with the main modes of variation of a given population or shape family (Sarkalkan *et al.*, 2014). A population can, for example, refer to patients with a skeletal disease such as osteoporosis or healthy individuals with a shared ethnicity or gender (Sarkalkan *et al.*, 2014; Heimann and Meinzer, 2009). Anatomical shape is defined as the geometrical information that remains when location and rotational effects are filtered out (Luthi and Bouabene, 2020). Statistical models of bone are trained using a data set focused on a specific bone structure within a particular population. The focused data set provides the model with prior knowledge about the bone structure's main shape and density distribution. These models can then be fitted to an unseen bone structure within the same population to generate a statistical patient-specific three-dimensional approximation. The approximation can then be used in a variety of medical applications, including motion tracking, computer-aided surgeries or, in the scope of this project, estimating health bone anatomy for PSI design from partial inputs (Sarkalkan *et al.*, 2014; Mauler *et al.*, 2017). Estimating healthy bone anatomy gives way to a new form of treatment for skeletal bone diseases. For example, segmented bone loss can be remodelled specifically for PSI design, matching the patient's characteristics and physical traits without extensive specialist intervention.
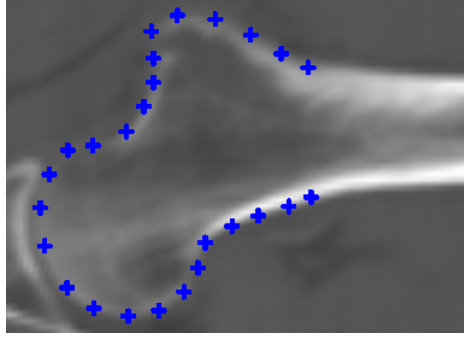
Figure 3.1: A discrete set of landmarks

Figure 3.2: A dense face-vertex surface mesh

3.1.1   Statistical Shape Models

Point distribution models (PDMs) are the most basic and important of all statistical model variations. PDMs represent object boundaries, i.e. shape, through a collection of point positions in a two- or three-dimensional plane. PDMs are trained using training objects or shapes, $\mathbf{\Gamma}_i$, of a specific bone structure within a statistical population. The size of the training set, $\{\mathbf{\Gamma}_1, ..., \mathbf{\Gamma}_n\}$, and its contents must represent all possible shape variations within the bone structure (Sarkalkan *et al.*, 2014; Lüthi *et al.*, 2017). Bone shapes are represented through a discrete set of landmark points; i.e $\mathbf{\Gamma}_i = \{x_{ik} | x_k \in \mathbb{R}^3, k = 1, ..., N\}$ where N denotes the number of landmark points Lüthi *et al.* (2017). Figure 3.1 illustrates a discrete set of landmarks describing a two-dimensional femur head derived from an image. These landmarks can also be represented through face-vertex surface meshes, containing not only a dense set of point coordinates but connectivity information describing the bony structure, as shown in Figure 3.2.

As described by Cootes *et al.* (2004) there are three main steps to follow when building an SSM. The first is to establish correspondence within the dataset. Establishing correspondence is important to ensure that the anatomical variance found in the data set is accurately interpreted when constructing the statistical model (Cootes *et al.*, 2004). Correspondence means that the kth landmark of shape $\mathbf{\Gamma}_i$ and $\mathbf{\Gamma}_j$, x_{jk} and x_{jk}, represent the same anatomical point on both instances (Cootes *et al.*, 2004; Luthi and Bouabene, 2020). Correspondence is established either through manual identification and assignment

of anatomical landmarks or automated registration algorithms such as the iterative closest point (ICP) or ASM algorithms (Heimann and Meinzer, 2009; Sarkalkan et al., 2014). A more detailed discussed on correspondence and how it may be established manually or automatically for three-dimensional shapes can be found in Subsection B.4.

The second step is to align the data set. After correspondence is established, the dataset can be aligned to eliminate all translational and rotational dependencies within the data set, leaving only dependencies that represent pure shape variation within the bone population being analysed (Heimann and Meinzer, 2009). Again, the alignment can be done manually through rigid alignment and the assignment of manually placed anatomical landmarks, or automated through alignment algorithms such as the generalised Procrustes alignment algorithm (Heimann and Meinzer, 2009; Sarkalkan et al., 2014). A detailed discussion on these methods can be found in Subsection B.5.

Finally, the shape variation from the dataset must be extracted using statistical analyses. Constructing the SSM then consists of extracting the mean shape and the main modes of variation within the bone population by assuming the shape variation can be modelled using a normal distribution, $s \sim \mathcal{N}(\overline{x}, S)$, where the mean shape, $\overline{x}$, and covariance matrix, $\mathbf{S}$, are calculated from the training data by:

$$\overline{x}_k = \frac{1}{n} \sum_{i=1}^{n} x_{ik} \tag{1}$$

$$\mathbf{S}_k = \frac{1}{n-1} \sum_{i=1}^{n} (x_{ik} - \overline{x}_k)(x_{ik} - \overline{x}_k)^T \tag{2}$$

The covariance matrix, $\mathbf{S}$, measures the variation in the training instances, x_{ik}, with regards to the the mean shape $\overline{x}_k$ (Cootes et al., 2004; Luthi and Bouabene, 2020).

The main modes of variation, $\mathbf{\Phi}_i$, present in the covariance matrix, S, can be estimated using principal component analysis (PCA), with $\mathbf{\Phi}_i$ and λ_i being the eigenvectors and eigenvalues of S. The principal modes of variation, $\mathbf{\Phi}_i$, are conventionally ordered in descending order with regards to the percentage of shape variation $\lambda_i \geq \lambda_{i+1}$. Ordering is done so that the first few modes of variation describe most of the possible shape distributions within the statistical population, be it a shape within or out of the data set (Cootes et al., 2004; Heimann and Meinzer, 2009). The bone shape of the entire population can be then described by:

$$x_k = \overline{x}_k + \sum_{s=1}^{c} b_s \mathbf{\Phi}_s \tag{3}$$

where c is the number of modes affecting the accuracy and the compactness of the resulting model and b_s is the contribution of $\mathbf{\Phi}_s$ in the model instance.

A higher number of modes results in a more accurate model, which is less compact and vice versa (Sarkalkan *et al.*, 2014). The most common approach for establishing the number of modes required is given by the threshold value r:

$$r = \frac{\sum_{s=1}^{c} \lambda_s}{\sum_{s=1}^{N-1} \lambda_s} \tag{4}$$

with r generally between 0.90 and 0.98 (Heimann and Meinzer, 2009). Modes beyond r are assumed to be noise, and including them may lead to overfitting.

3.1.2 Statistical Appearance Models

Just as SSMs, the SAM needs to be trained using a data set focused on a population. However, instead of using surface meshes or landmarks, the density distribution is obtained from the intensity values found in medical images. Thus, models of appearance can also be called models of intensity or density (Heimann and Meinzer, 2009). Where SSMs only contain the surface geometry of the bone shape, a SAM contains the density distribution of the bone shape and can thus infer the average bone density distribution within a population and its principal modes of variance (Sarkalkan *et al.*, 2014).

As described by Cootes *et al.* (2004) there are three main steps to follow when building a SAM. The first is to collect the intensity or pixel data from the data set. Through image warping, a consistent method can be devised to collect or capture the intensity values representing the bone density distribution. For this purpose, image warping is used to transform the images from one spatial configuration into another. Deriving the texture information from the training instances consists of warping the training set to fit a standard bone shape of the population which can be given by the mean bone shape of the SSM, $\overline{x}$. With correspondence already established during SSM construction, image warping then consists of a point set transformation $x_i' = f(x_i)$, where f is a continuous mapping function containing some form of transformation matrix. The warped images are often referred to as being shape normalized (Cootes *et al.*, 2004; Heimann and Meinzer, 2009; Sarkalkan *et al.*, 2014). In literature, two forms of the function f is commonly used, the piecewise affine and thin-plate spline (Cootes *et al.*, 2004). After shape normalization, the intensity information can then be sampled and stored into a texture vector $\mathbf{g}_{im}$.

The second step is to normalise the pixel and intensity data. Normalisation of intensity values is essential to obtain a SAM that will represent the statistical population as accurately as possible. Due to machine calibration, acquisition errors and imaging software, different intensity values in a data set can represent the same density value for various medical imaging devices (Cootes *et al.*, 2004; Sarkalkan *et al.*, 2014). In literature, intensity normal-

ization is often accomplished by scaling $\mathbf{g}_{im}$ with α while applying β as an additional offset (Sarkalkan *et al.*, 2014):

$$\mathbf{g} = \frac{\mathbf{g}_{im} - \beta\mathbf{1}}{\alpha} \tag{5}$$

where $\mathbf{1}$ is a unit vector. The α and β parameters are usually chosen such that all intensity values are transformed to a zero mean with a unit variance (Cootes *et al.*, 2004):

$$\alpha = \mathbf{g}_{im} \cdot \overline{\mathbf{g}}$$

$$\beta = \frac{\mathbf{g}_{im} \cdot \mathbf{1}}{N_{im}}$$

where N_{im} represents the length of vector $\mathbf{g}_{im}$. With $\overline{\mathbf{g}}$ being the mean normalized intensity vector, $\mathbf{g}$ (Equation 6)), determining α becomes an iterative process (Cootes *et al.*, 2004).

Finally, the SAM is constructed using the normalised intensity data. Constructing the SAM using the normalized intensity data follows the same procedures as constructing the SSM. The only difference is that instead of using the point position for shape representation the normalised intensity data $\mathbf{g}$ is used to extract the mean and modes of variance (Cootes *et al.*, 2004):

$$\overline{\mathbf{g}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{g}_i \tag{6}$$

$$\mathbf{S}_g = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{g}_i - \overline{\mathbf{g}})(\mathbf{g}_i - \overline{\mathbf{g}})^T \tag{7}$$

As in the SSM construction, the PCA method can be used to extract the main modes of variance $\mathbf{\Phi}_g$. In this way most of the possible density or intensity distributions within the statistical population can be described by:

$$\mathbf{g} = \overline{\mathbf{g}} + \sum_{g=1}^{c}b_g\mathbf{\Phi}_g \tag{8}$$

Similar to the SSM, the vector length of $\mathbf{b}_g$ could be determined by the value of c using Equation 4 (Cootes *et al.*, 2004; Sarkalkan *et al.*, 2014).

Furthermore, SSMs and SAMs can be combined into a compact model to simultaneously describe the shape and density distribution within the bone shape population. These models are often referred to as a combined statistical shape and appearance model (SSAM) (Sarkalkan *et al.*, 2014). The interested reader may refer to Cootes *et al.* (2004) for additional mathematical information.

Stellenbosch University https://scholar.sun.ac.za

3.2 Active Shape and Appearance Models

After establishing the SSM and SAM, a capable optimisation algorithm is still required to fit the model to a set of unseen data inputs. These inputs can be from two-dimensional scalar images such as CT or MRI scans, a partial surface mesh or a small group of landmark measurements. In either case, the optimisation algorithm is only focused on determining the parameter vectors $\mathbf{b}_s$ and $\mathbf{b}_g$, which best fit the models to the inputs (Sarkalkan *et al.*, 2014).

Establishing an optimisation algorithm requires consideration of certain factors. One such aspect forms part of the initialisation process, i.e. the starting point for the optimisation protocol. When the search procedure is initialised, rotational and translational dependencies must be minimised so that shape and density variation are the main dependencies between the model and the target data. A second factor relates to the cost function of the algorithm and its tendency to get stuck within a local minimum. Both these factors are amplified when working with medical images due to the large search space (Heimann and Meinzer, 2009). However, providing an initial pose, location and shape instance through user interaction can minimise these two factors. The image input and model can be aligned by manually placing corresponding landmarks, thereby minimising the possibility of getting trapped in local minima. The landmarks can also be used to calculate an initial posterior model, allowing for a more focused fitting procedure. Thus, $\mathbf{b}_s$ and $\mathbf{b}_g$ can be found in a way where the cost function, representing the quality of fit, is truly optimised (Heimann and Meinzer, 2009).

Optimisations algorithms used to successfully fit SSMs and AAMs to unseen data include:

- The Quasi-Newton method (Haslam *et al.*, 1994).

- The Nelder-Mead simplex approach (Tang and Ellis, 2005).

- The conjugate gradient algorithm (Lötjönen *et al.*, 2004).

- Gradient descent optimization (Nain *et al.*, 2007).

- Genetic algorithms (Hill and Taylor, 1992).

- Simulated annealing (Fleute *et al.*, 1999).

- The Marquardt-Levenberg algorithm (Okada *et al.*, 2008).

Though many different optimisation algorithms and methods can establish the shape and density parameters for automating registration and segmentation, the ASM and AAM have been well established within literature and industry (Heimann and Meinzer, 2009; Cootes *et al.*, 2004; Lüthi *et al.*, 2017). The algorithms and methods listed above have been shown to be slower and less efficient than ASM and AAM implementations (Heimann and Meinzer,

2009; Cootes *et al.*, 2001). For example, implementations of the Quasi-Newton method have been shown to be 40 times slower than ASM implementations (Haslam *et al.*, 1994). The ASM and AAM are generally used when fitting SSMs and SAMs to partial inputs. ASM traditionally only fits the SSM where AAM usually finds the best shape and density representation of the input data (Sarkalkan *et al.*, 2014).

ASMs are often used in industry to fit 3D models to a set of patient-specific two-dimensional scalar images. The three-dimensional modal can then estimate data specific to the patient, providing a patient-specific three-dimensional model at the cost of 2D imaging. If this were done using an AAM, it would both solve the segmentation problem and provide us with a patient-specific 3D model containing shape and density distribution. From this model, a PSI design can be established at a fraction of the time and cost usually required for designs reliant only on specialist knowledge (Heimann and Meinzer, 2009).

AAMs are generally very successful when matching a 2D model to a 2D image, e.g. facial recognition software. However, it is rarely used to segment 3D images due to the computational load and memory storage required to match and store the model, respectively. Even though methods for scaling down the intensity resolution exist, it diminishes the accuracy of the AAM. Additionally, the intensity representation of medical images also mostly lack the required resolution and quality to produce a robust model from a training set (Heimann and Meinzer, 2009).

Due to these problems, ASMs remain the most popular method for the segmentation of 3D medical images. The approach is straightforward to implement in software, providing a reasonably robust method at a fraction of the computation costs consistent with AAM search schemes. Many variants of the ASM have been developed over the years, trying to improve on outlier errors. Methods have also been identified to map a SAM to a segmented shape from the ASM. Though less robust than using AAMs, it remains computationally less intensive and easier to implement in software (Heimann and Meinzer, 2009).

Chapter 4

Model Construction

This chapter details the construction of the statistical models used for segmental bone repair in Chapters 5 and 6. Section 4.1 details data collection and processing, where Section 4.2 discusses the SSM, ASM and mean AM construction and developed software. Section 4.3 shows the results found when validating the models, and Section 4.4 discusses the models' validity.

4.1 Data Collection and Processing

The statistical models were constructed using retrospectively collected CT scans of the femur and tibia. As the study aimed to restore the non-pathological geometry of the femur and tibia, the training data consisted of healthy anatomy to avoid introducing a bias towards a pathology within the statistical models' behaviour. Furthermore, due to different anatomical traits between men and women, it is essential to analyse the statistical variability of the two groups separately (Nieves *et al.*, 2005).

Therefore, data collection was targeted at diagnostic CT scans of male and female South African patients between the ages of 18 and 45 with healthy bone anatomy, no prior radiologically identifiable disability, obesity, arthritis, or other musculoskeletal conditions that could affect bone mineral density. The target sample size for each model was 50 (n = 50) to reflect that of prior studies in literature (Heimann and Meinzer, 2009; Audenaert *et al.*, 2019). However, this was not possible for each training set due to limited availability.

Table 4.1 summarises all the essential information required to describe the statistical population of each training set. Data collection was done after ethical clearance was received from the Health Research Ethics Council (Ref: S20/07/171). Note that after data collection, either the left or right anatomy for the population was selected due to the available healthy bone samples in each data set.

After collection, the bone anatomy required from each CT scan was segmented using Mimics (Materialise NV, Belgium) software. During segmentation, the cortical and trabecular bone structure was isolated using thresholding

Table 4.1: Statistical population data

	Male femur	Male tibia	Female femur	Female tibia
Right or left anatomy	Right	Right	Left	Left
Sample size (n)	50	50	50	37
Age [years]	26.90 (± 6.08)	26.90 (± 6.08)	29.26 (± 7.24)	30.16 (± 7.63)
Slice thickness [mm]	1	1	1.06 (± 0.11)	1.06 (± 0.11)
Pixrel resloution [mm]	0.79 (± 0.11)	0.79 (± 0.11)	0.80 (± 0.14)	0.82 (± 0.12)

operations and manual editing, resulting in a mask of the intensity information of the femoral and tibial bone structures. Using the mask of each training sample, a face-vertex surface mesh was then constructed that defined the outer surface of each training sample. Appendix C contains the segmentation protocol used during data processing.

The surface meshes (or segmentations) produced consisted of ±75000 vertices each. Using *3-Matic* (Materialise NV, Belgium) software, the segmentations were reduced to 5 000 vertex points with a geometrical error lower than 0.1 mm. This produced samples that were computationally light and easy to work with during debugging but still gave good results during testing. After vertex reduction, the surface meshes were then used to programmatically crop and extract the voxel density distribution of each training sample from the CT images using the Scalismo (University of Basel, Switzerland) library.

4.2   Statistical Model Construction

The SSMs, ASMs, and mean AMs were constructed using the standard approaches proposed in the literature (Lüthi *et al.*, 2017; Morel-Forster *et al.*, 2018; Cootes *et al.*, 2001). Figure 4.1 gives an overview of the construction procedures with the data inputs and model outputs. First, the SSMs were constructed and validated to ensure the shape variation of the populations was captured sufficiently. The SSMs and cropped density distributions were then used to construct the ASMs and mean AMs. The ASMs, combined with an automated segmentation algorithm, could then produce automated face-vertex surface meshes from the training images. These automated surface meshes could then be compared to the manual segmentations for validation of the ASMs. To validate the mean AMs, healthy density estimates for training samples were produced by warping the mean AMs to the surface mesh boundaries of the training samples. The estimates were then compared to the original density distribution of each training sample. A more detailed discussion on the
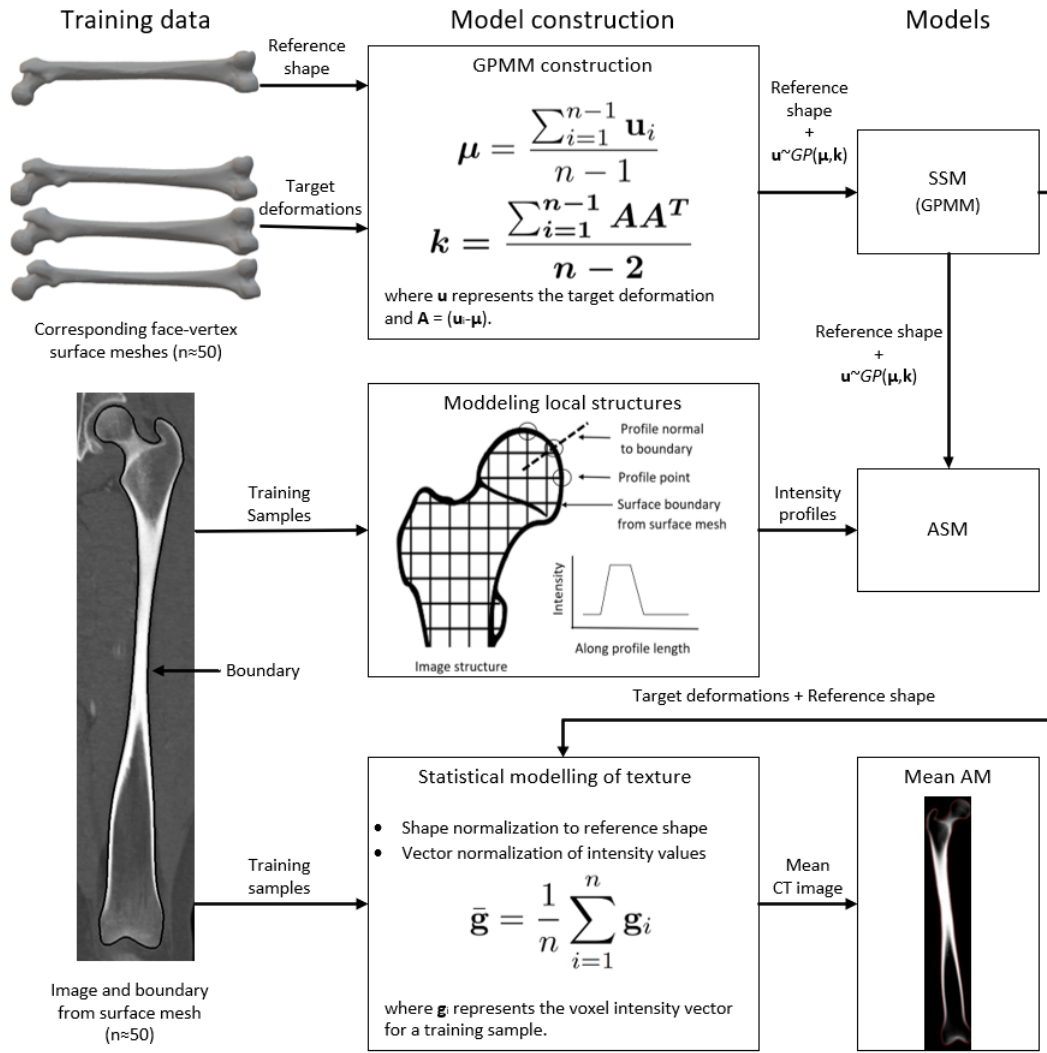
Figure 4.1: Model construction overview.

construction procedures for each model is given in the subsequent subsections. The detailed validation procedures and results will be given in Section 4.3.

All programming was done in Scala, a strong statically typed general-purpose programming language, using Scalismo library, a scalable image analysis and shape modelling library. Appendix A contains additional notes on software considerations that were not discussed in this chapter. Detailed notes on modelling and mathematical considerations, not fully addressed in this chapter, can be found in Appendix B. For the reader who is not familiar with Gaussian processes or PCA a review of Sections B.2 and B.6 is recommended.
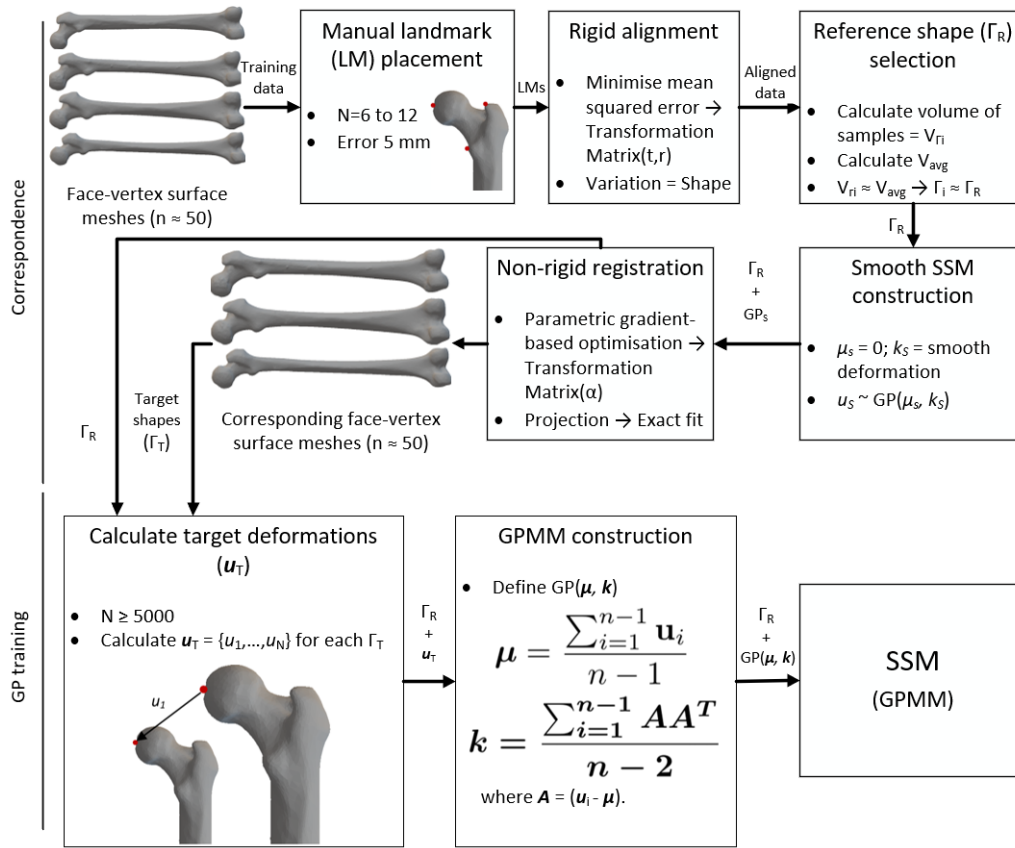
Figure 4.2: Statistical shape model construction.

4.2.1 Shape Model Construction

SSMs were constructed using the Gaussian Process Morphable Model (GPMM) construction methodology proposed by Lüthi *et al.* (2017). Where shape variation is modelled as deformations, $\{u_1, ..., u_N\}$, from a reference surface, $\Gamma_R \subset \mathbb{R}^3$, which is representative of the average shape of the target population (Lüthi *et al.*, 2017). In this study, the GPMM methodology can be simplified to two main activities: 1) establishing correspondence within the training data, and 2) storing shape variation within a Gaussian Process (GP). Figure 4.2 illustrates these two main activities and their underlying steps.

As already mentioned, correspondence means that the landmark of each shape in the training set represents the same anatomical points (Cootes *et al.*, 2004). Thus, for GPMMs all the training data must be in correspondence with Γ_R to ensure that the anatomical variance found in the data set is accurately interpreted when constructing the statistical model. Five steps were required to establish correspondence in the training set:

1. Manual landmark placement.

2. Rigid alignment.

3. Reference shape selection.

4. Initial SSM construction with smooth deformation kernels.

5. Non-rigid registration.

A set of landmarks (or points) ($N_{femur} = 12$; $N_{tibia} = 10$) similar to those seen in Figure 2.2, were manually placed on each training sample (or face-vertex surface mesh) of the particular training set used. It was assumed throughout the study that manual landmark placement had a radial uncertainty of 5 mm. To accommodate this error, a noise variable in the form of a GP was introduced during fitting operations to allow additional variance when calculating the posterior for registration, segmentation and reconstruction.

The manually placed landmarks were then used to align the training set rigidly by minimizing the mean squared error over the corresponding landmark points. Using the Scalismo library, a transformation matrix was calculated to rigidly align the training samples to the reference frame of the first training sample in the sequence. Using basic translation and rotation operations, the calculated transformation matrix for each sample was used to align the dataset to the first surface mesh. A detailed mathematical explanation of the alignment methods used can be found in Subsection B.5.

To construct a GPMM, a reference shape, Γ_R, close to the mean of the population has to be selected from the training set, $\Gamma_{1 \to n}$. In this study, it was assumed that the volume of a training sample, $V_{\Gamma_{1 \to n}}$, would be a good indicator for the population mean. The volume of each sample was estimated by multiplying the length, width and height of the bounding box around the surface mesh of the sample. The average volume, V_{avg}, was derived, and the training sample closest to the average volume, $V_{avg} \approx V_{\Gamma_i}$, was then selected to be the reference shape, $\Gamma_R \approx \Gamma_i$. Even though the volume estimate was not the actual sample volume, it was simple to compute. Visual inspection also showed that the selected Γ_R closely approximated the population's mean after the SSM was created.

During registration a shape instance, Γ, within the statistical population can be created by sampling $\hat{u} \sim GP(\mu, k)$ and warping the reference shape, $\Gamma = \{x + \hat{u}(x) | x \subset \Gamma_R\}$. If $\Gamma_R \approx \Gamma_{avg}$ a simple GP can be constructed by setting $\mu = 0$ and defining a scalar valued covariance matrix or kernel, $\mathbf{k}_{scalar}$, that models smooth deformations by:

$$\mathbf{k}_{scalar}(x_{i1}, x_{i2}) = e^{\frac{-\|x_{i1} - x_{i2}\|^2}{\sigma^2}} \tag{9}$$

where $\mathbf{k}_{scalar}$ is the covariance matrix and σ^2 defines the range over which the deformations are correlated, i.e. the smoothness. Thus, for example, for $\sigma \gg 10$, deformation fields will be much smoother over a specific region than for $\sigma = 10$. To translate $\mathbf{k}_{scalar}$ to the $x_{1 \to N}$ components of the vertex mesh, a matrix-valued kernel can be defined as

$$\mathbf{k}_{smooth}(x_{i1}, x_{i2}) = s \cdot \mathbf{I}_{3\times3}\mathbf{k}_{scalar}(x_{i1}, x_{i2}) \qquad (10)$$

where $\mathbf{I}_{3\times3}$ is a 3×3 identity matrix and $s \in \mathbb{R}^3$ determines the scale of the variance. Additionally, when modelling with GPs, simple kernels, such as the smooth deformation kernel in Equation 10, can be combined to define a new kernel that models different deformations over similar or different regions in the same covariance matrix (Lüthi *et al.*, 2017). In this study three scalar valued smooth deformation kernels with varying smoothness, σ, and constant scale, s, were combined to define $GP_{smooth}(\mu = 0, \mathbf{k}_{smooth^1} + \mathbf{k}_{smooth^2} + \mathbf{k}_{smooth^3})$. The GP could then be combined with Γ_R to define a smooth GPMM as an initial SSM. See Sections B.1 and B.7 for a detailed explanation of the mathematical considerations, and Table A.2 for the values of σ and s.

The smooth SSM could then be used in combination with parametric gradient based optimisation for non-rigid registration (Lüthi *et al.*, 2017; University of Basel, 2021). A non-rigid registration problem was established by defining four things:

1. A *'transformation space'* in which instances of the smooth SSM can be sampled.

2. A *'metric'* to measure the distance between the target surface, Γ_T, and Γ_R.

3. A *'regularizer'* which penalizes unlikely GP transformations.

4. An *'optimizer'* to ensure efficient registration.

Note that Γ_T is the sample shape, Γ_i, for which correspondence is being established. The defined registration problem could then be used to calculate the GP transformation, α, required to warp Γ_R to Γ_T. In the study, the registration problem was carried out for a sequence of iterations and only for a randomly selected number of vertices on Γ_R. Thus, after the defined number of iterations, the vertices not yet located on the target surface were projected normally to the closest point on Γ_T. This resulted in a reference shape where all the vertices were fitted to the surface of the target mesh, giving a corresponding face vertex surface mesh where a vertex on Γ_R closely relates to the same vertex on the new Γ_T. For a more detailed explanation of the registration algorithm, iteration sequence and sampled points used see Appendix A (University of Basel, 2021).

With correspondence established the shape variation within the statistical population can now be stored within a GP by training a GPMM using the $\Gamma_R \subseteq \Omega$ (where $\Omega \subset \mathbb{R}^3$) and the new target shapes. The vertices on Γ_R and a particular Γ_T can be used to calculate a sequence of deformation vectors, $\mathbf{u} = \{u_1, ..., u_N\}$. The deformations can be stored within a GP, $u_k \in GP(\mu_k, \mathbf{k}_k)$ with the mean deformation $\mu_k : \Omega \to \mathbb{R}^3$ and covariance matrix $\mathbf{k}_k : \Omega \times \Omega \to \mathbb{R}^3$. Similar to Equations 1 and 2, μ_k and $\mathbf{k}_k$, can be derived through:

$$\mu_k = \frac{1}{n-1} \sum_{i=1}^{n-1} u_{ik} \tag{11}$$

$$\mathbf{k}_k = \frac{1}{n-2} \sum_{i=1}^{n-1} (u_{ik} - \mu_k)(u_{ik} - \mu_k)^T \tag{12}$$

When working with surface meshes, the number of landmarks, N, equals the number of vertices used to describe the surface boundary in question. For the femur and tibia meshes used in this study, N is greater or equal to 5000 vertices. Thus, the covariance matrices derived during training are quite large, making it difficult to store and process. Fortunately, as with Equation 3, PCA can be used to represent the derived GP in terms of an orthogonal set of basis functions. Through low-rank approximation, the GP can be presented by:

$$\widetilde{u}(x) \sim \mu(x) + \sum_{r=i}^{r} \alpha_i \sqrt{\lambda_i} \Phi_i(x) \tag{13}$$

were α_i is the contribution of (λ_i, Φ_i), the eigenpairs. The GPMM can be stored and processed much easier in this form while sampling shape instances from the population can be done by simply adjusting α_i.

4.2.2 Active Shape Model Construction

The procedure for modelling local structures was used (Cootes *et al.*, 2004; Morel-Forster *et al.*, 2018) in combination with the GP Morphable Model construction methodology (Lüthi *et al.*, 2017) to construct the ASM. The GP of the SSM was combined with intensity profiles around the surface boundaries of each training sample. The intensity profiles were extracted from the cropped CT images using the corresponding face-vertex surface meshes to identify the surface boundaries. Figure 4.3 gives an illustration of the ASM construction procedure used during the study.

First, the reference shape was used to uniformly sample 250 vertices ($N_{profile} = 250$) on the surface mesh. Second, using the corresponding face vertex surface meshes and their accompanying images, intensity profiles were extracted on the same 250 corresponding points of each image. An intensity profile, as seen in Figure 4.3, consists of a centre point ($k^{th}_{profile}$ of $N_{profile}$) on the mesh boundary and ten sub-points, five on each side of the boundary line. Each sub-point was spaced 1 mm apart, and the profile line was set normal to the surface boundary. The intensity profile values on each image were sampled using the centre and sub-point coordinates. Third, using the sampled intensity profile values for each image in the training set, a GP of the intensity gradient was derived for each of the 250 vertices (University of Basel, 2021). The intensity gradient made the ASM more robust against intensity variation due to acquisition errors and machine settings. Finally, the profile data was
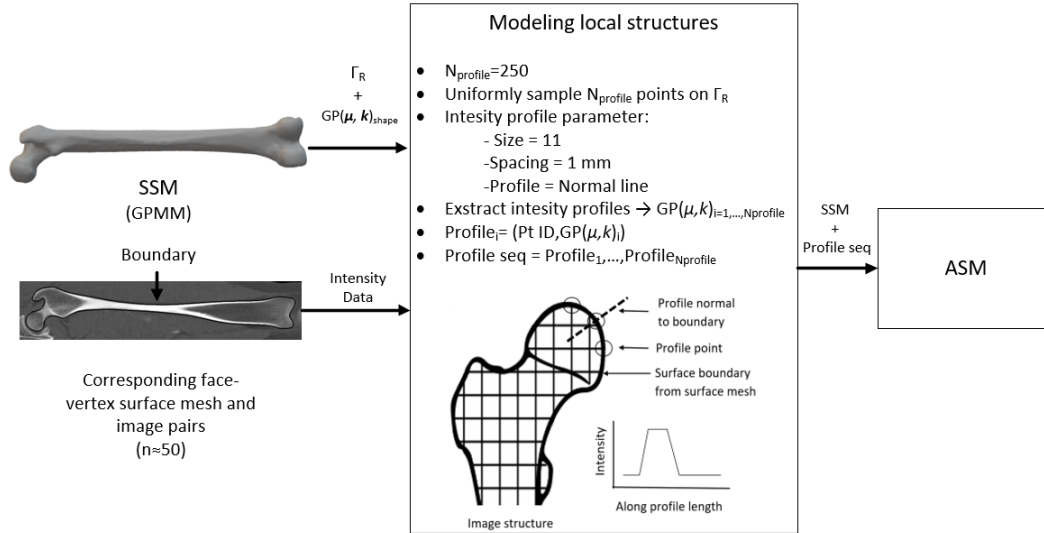
Figure 4.3: Active shape model construction.

stored in a sequence containing the profile GPs, and vertex ID where it was sampled. The profile sequence, Γ_R and SSM GP were then stored in the same data structure to create the ASM. See Table A.4 for more detail regarding the data structure.

A segmentation algorithm also had to be implemented in software to fit the ASM to unseen images. For this, Markov Chain Monte Carlo (MCMC) combined with the Metropolis-Hastings (MH) algorithm was used following the methodology set out by Morel-Forster *et al.* (2018). This method uses a probabilistic interpretation of the ASM as Bayesian inference. Thus, we assume a prior distribution on considered parameters such as shape and length. This method allows for the easy integration of user or specialist information during optimisation for a more accurate fit, it is simpler to implement, and is more robust against local minima than the traditional ASM fitting methodology (Morel-Forster *et al.*, 2018).

Following the segmentation procedures illustrated in Figure 4.4, fitting operations first require some user input through manual landmark placement. By manually placing corresponding points on the unseen image and Γ_R, initial parameters, θ_0, can be calculated to initialise the automated segmentation procedures. Parameters consist of the model's translation, rotation and shape coefficients, $\theta = (t, r, \alpha)$, with respect to the origin of the model's reference frame. Initial translation, t_0, and rotation, r_0, parameters can be derived by calculating the transformation matrix that rigidly aligns the ASM and the unseen image. The best initial GP coefficients, α_0, are assumed to be the mean shape deformation, Γ_{mean}, contained within $GP(\mu, k)$.

An initial posterior was calculated using the manually placed landmarks and MH algorithm to ensure the intensity profiles were fitted as accurately as possible. As seen in Figure 4.5, model parameters consisting of translation,
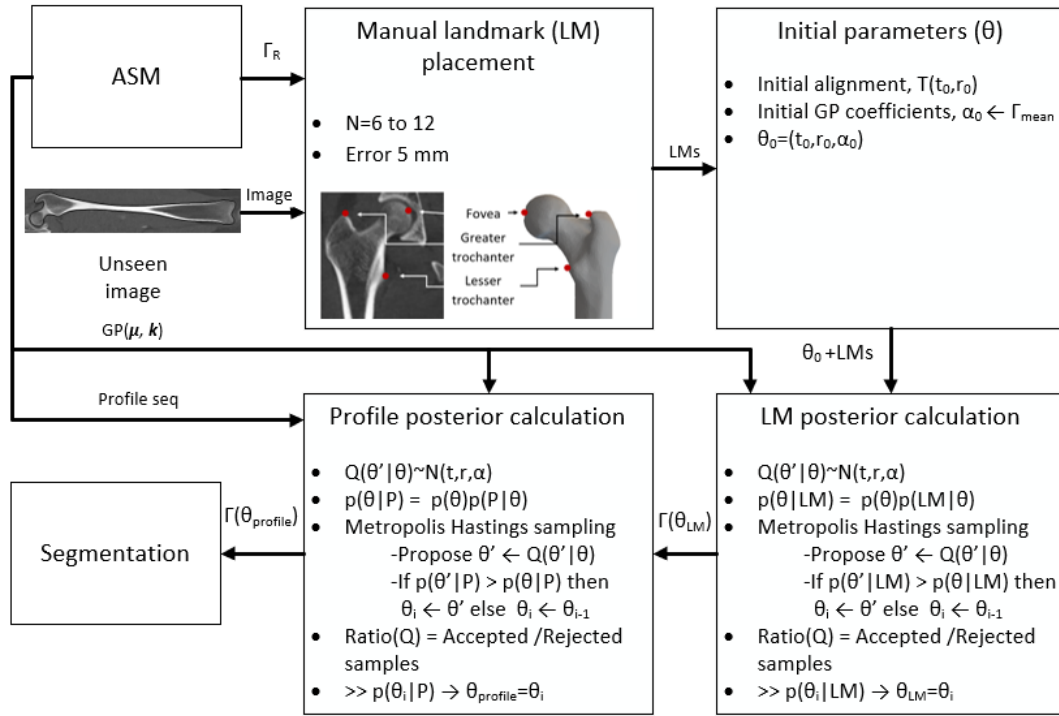
Figure 4.4: Segmentation procedures.

Figure 4.5: Metropolis Hastings sampling (Adapted from: Morel-Forster *et al.* (2018)).

rotation and GP coefficients are sampled normally from a user-defined normal distribution, $Q(\theta'|\theta) \sim \mathcal{N}(t, r, \alpha)$. The user must specify a normal distribution around zero for t, r and α at the beginning of the MH algorithm. The range of the distribution was classified as step sizes. If, for example, the step size for translation was 1 mm, then $t' = t + \mathcal{N}(-1, 1)$ for each new iteration.

Additionally, for each iteration, only one parameter (t, r or α) was sampled. The parameter selected for sampling in each iteration was selected randomly through a given probability, i.e. a random number, following a uniform distribution between zero and one, was generated and depending on the given probability, a parameter was selected and sampled. Assuming θ_0 was selected properly, shape variation would be the main concern during the fitting operation. Thus, adjustments to α would be favoured over adjustments of (t, r). In this study, it was assumed that 80 % of the samples from $Q(\theta'|\theta)$ would have to be α to gain a proper segmentation and (t, r) would both have a 10 % probability $(0.0 \geq p(t) \geq 0.1, 0.1 > p(r) \geq 0.2, 0.2 > p(\alpha) \geq 1.0)$. For example, if a new iteration began and the random value generated was 0.1 the t would be selected and sampled, $t' = t + \mathcal{N}(-1, 1)$. Thus (r, α) would remain constant for the iteration. In each iteration, the sampled step would be used to adjust the current θ, producing θ'. See Table A.3 and Morel-Forster et $al.$ (2018) for more detail.

The likelihood of the posterior for both the current θ and newly sampled θ' is then calculated using:

$$p(\theta|I) = p(\theta)p(D|\theta) \tag{14}$$

where I donates the observed image, $p(\theta)$ the likelihood given the observed distribution captured in the GP of the SSM, and $p(D|\theta)$ the likelihood given either the manually placed landmarks ($D = LM$) or ASM profiles ($D = P$). The likelihood $p(\theta)$ is determined by calculating the log-likelihood of the model parameters t, r and α. For t and r a uniformly distributed distribution was defined around the model origin ($t \sim \mathcal{U}(-1, 1)$ m, $r \sim \mathcal{U}(-\pi, \pi)$ rad). Thus, all transformations within this range would be equally likely. For α, the log-likelihood for the observed GP in the SSM was calculated. The three log-likelihoods were then combined through summation to produce $p(\theta)$. The likelihood $p(D|\theta)$ was calculated similarly using either a normal distribution around zero error for the manually placed landmarks or the GPs of the ASM profile sequence.

To calculate $p(LM|\theta)$, the corresponding manually placed landmark positions would be subtracted from each other to determine the root mean square error. Using a normal distribution around zero, errors closer to zero would have a higher log-likelihood. With each iteration, the positions of the landmarks would change, and the new log-likelihood would be calculated. For $p(P|\theta)$, using each profile's main vertex position, the intensity profiles of θ were sampled from the image and then compared to the ASM profiles by calculating the log-likelihood of the observations using the GPs' of the ASM profile sequence.

Additionally, the log-likelihoods of $p(\theta)$ and $p(D|\theta)$ can be combined through summation as before to produce $p(I|\theta)$.

Furthermore, if the θ' was more likely than θ, it was accepted, and the next iteration was started. Otherwise, θ' was rejected and θ was carried forward. The acceptance ratio, $Ratio(Q)$, was calculated by dividing the number of accepted θ' by the number of rejected θ', to demonstrate to the user the effectiveness of the proposed step sizes of (t, r, α). If the ratio was high ($Ratio(Q) > 0.5$), only local minima were sampled, and the step sizes are considered ineffective. If low ($Ratio(Q) < 0.25$), the step sizes were too large, and the samples are too far from the global minima. However, $Ratio(Q)$ is only for the user to analyse the suggested step sizes. The fitting operation will continue to run for the number of iterations defined at the start of the MH algorithm. For this study, 10 000 iterations were run for both the LM and profile posterior calculations. With code optimisation, 10 000 iterations took approximately 30 to 40 seconds.

Returning to Figure 4.4, the landmark posterior was then used for the profile fitting operation. Again, the MH algorithm was used, except the ASM profiles were used to calculate $p(D|\theta)$ instead of the manually placed landmarks. Using each profile's main vertex, the intensity profiles of θ and θ' were sampled from the image and then compared to the ASM profiles. The sample parameters, θ_i, with the highest log-likelihood $p(I|\theta)$ of all the iterations, was then used to set the model for the final segmentation.

4.2.3 Appearance Model Construction

The construction procedure used for the mean AM can be seen in Figure 4.6. Using linear interpolation, the deformations, $\mathbf{u}_t = \{u_1, ..., u_N\}$, calculated for each target shape during registration in Subsection 4.2.1, were extended to the entire CT image for each sample. This warped the voxel data of each CT image to the reference shape, Γ_R. With the sample images shape-normalised, the corresponding intensity vector, $\mathbf{g}_{im}$, from each training image could be extracted by calculating which voxels fall within the Γ_R boundary. Thus, the voxel data contained in $\mathbf{g}_{im}$ for each image included only bone intensity data of the bone structure studied.

To combat intensity variation from different machine imaging settings, the intensity data contained within $\mathbf{g}_{im}$ was normalized to the mean intensity vector, $\bar{\mathbf{g}}$ (Cootes *et al.*, 2004; Heimann and Meinzer, 2009). However, with $\bar{\mathbf{g}}$ of the sample population not yet known at this stage, the method discussed in Section 3.1.2 was used. Briefly summed up, Cootes *et al.* (2004) proposed the following process:

1. Assign one of the samples as the initial $\bar{\mathbf{g}}$.

2. Normalize the intensity values of all samples to $\bar{\mathbf{g}}$.

3. Calculate the new $\bar{\mathbf{g}}$ of the normalized samples using Equation 6.
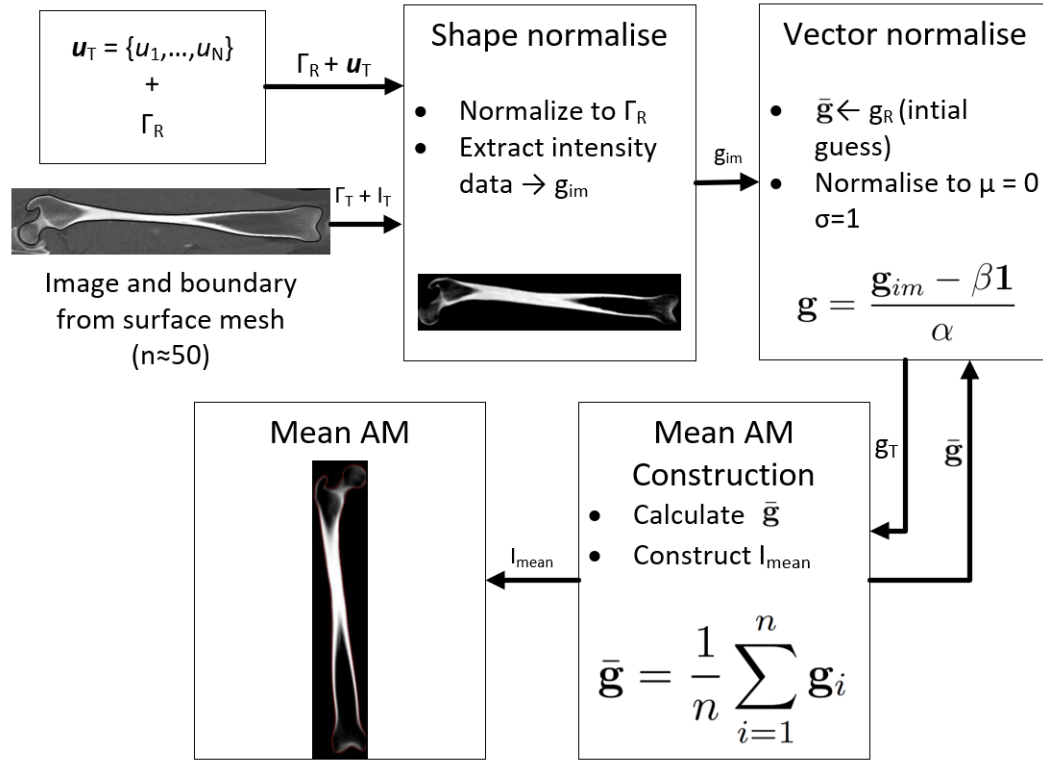
Figure 4.6: Mean appearance model construction.

4. Iterate until convergence.

In this study, the corresponding intensity vector, $\mathbf{g}_R$, of Γ_R was selected as the initial $\bar{\mathbf{g}}$. The vector zero mean unit standard deviation was selected to normalise the intensity values of all samples. Refer to Subsection 3.1.2 for more detail. The convergence criterion was the average difference between the elements of the old and new $\bar{\mathbf{g}}$ of each iteration. An average difference value of 0.1 yielded good results. After the final $\bar{\mathbf{g}}$ was calculated, it was used to generate a mean CT image.

4.3 Model Validation

Model generalisation was investigated through leave-one-out testing on ten samples to determine how well the SSM, ASM and mean AM represent members of the statistical population outside of the training data used during construction (Audenaert *et al.*, 2019). Partial models were constructed by removing a selected sample from the training data ($n' = n - 1$). The resulting models were then fitted to the corresponding unseen test sample. The shape and density fits were assessed by calculating the average distance error, Hausdorff distance error, R square (R^2), and the average intensity error (AIE). First, the average distance (AD) and Hausdorff (HD) errors were calculated

Table 4.2: Male femur model validation

Male femur	SSM			ASM			Mean AM	
Sample ID	AD [mm]	HD [mm]	HD Pos.	AD [mm]	HD [mm]	HD Pos.	R2 [%]	AIE [HU]
NS_401	0.57	3.28	Prox	1.25	7.00	Prox	96%	89
PB_126	0.80	4.59	Prox	1.66	8.07	Prox	92%	118
SD_235	0.56	3.43	Prox	1.47	8.53	Prox	92%	107
SM_517	0.58	3.30	Dist	1.26	8.82	Prox	95%	101
SN_234	0.66	4.01	Dist	1.11	5.51	Prox	92%	116
SZ_234	0.76	4.65	Prox	1.60	7.80	Dist	90%	132
TC_603	0.64	4.79	Prox	1.45	5.71	Dist	96%	84
TG_375	0.59	3.55	Prox	1.32	8.11	Prox	93%	112
TL_126	0.52	3.88	Dist	1.48	7.30	Prox	86%	148
TS_710	0.72	5.38	Dist	1.52	7.20	Dist	87%	141
Avg.	0.64	4.09		1.41	7.41		92%	115
Std.	0.09	0.69		0.16	1.05		3%	20

to evaluate the quality of the bone shape estimation. The shortest distance to the estimated surface was calculated for each vertex point on the unseen test surface. The average distance error was defined as the overall average of all distances, and the Hausdorff error was the maximum, minimum distance. Secondly, the density estimation was assessed using the R^2 metric and the average intensity error. Before calculating these metrics, the density estimation was shape normalised, and its intensity values vector normalised to that of the original density image. The normalised density estimation can be viewed in the same reference frame as the original CT image. This gives a visual representation of the estimations residuals, ensuring the correct interpretation of the R^2 value while aiding in understanding the location of density errors within the estimation. It is important to note that the CT intensity data were analysed in Hounsfield units (HU), and with the HU bone scale being quite large (226 to 3071 HU depending on the population and bone structure (Chougule *et al.*, 2018)), the R^2 value was also used to bring the average intensity error into perspective.

4.3.1 Shape Model Validation

Tables 4.2 to 4.5 show the average AD and HD errors for all the SSMs, ranging from 0.52 (± 0.08) mm to 0.64 (± 0.09) mm and 4.09 (± 0.69) mm to 3.34 (± 1.06) mm, respectively. Through visual inspection, the HD error for each test sample was found to occur at the proximal or distal end of both the femur and tibia models. In most cases, this was at the trochanteric fossa for the femur, and at the medial or lateral condyle, or tibial tuberosity for the tibia. Visual inspection also showed that the diaphyseal and metaphyseal areas of

Table 4.3: Male tibia model validation

Male tibia	SSM			ASM			Mean AM	
Sample ID	AD [mm]	HD [mm]	HD Pos.	AD [mm]	HD [mm]	HD Pos.	R2 [%]	AIE [HU]
NS_401	0.58	3.01	Prox	0.96	4.41	Dist	96%	95
PB_126	0.67	3.93	Prox	1.28	5.05	Dist	93%	124
SD_235	0.40	2.73	Dist	0.97	4.56	Prox	95%	97
SM_517	0.56	3.42	Prox	0.92	4.25	Prox	91%	131
SN_234	0.52	3.04	Prox	0.89	5.20	Prox	88%	141
SZ_234	0.74	4.75	Prox	1.69	6.86	Prox	91%	130
TC_603	0.54	2.92	Dist	0.99	5.35	Dist	94%	96
TG_375	0.58	3.02	Prox	1.22	6.23	Dist	91%	126
TL_126	0.63	3.49	Prox	0.92	4.09	Dist	90%	138
TS_710	0.69	3.51	Prox	1.25	4.77	Distal	89%	165
Avg.	0.59	3.38		1.11	5.08		92%	124
Std.	0.09	0.57		0.24	0.84		3%	21

Table 4.4: Female femur model validation

Female femur	SSM			ASM			Mean AM	
Sample ID	AD [mm]	HD [mm]	HD Pos.	AD [mm]	HD [mm]	HD Pos.	R2 [%]	AIE [HU]
SA_303	0.53	3.01	Prox	1.02	4.55	Prox	95%	83
SL_401	0.62	3.69	Dist	1.66	8.10	Prox	91%	105
SL_077	0.54	4.15	Prox	1.33	8.15	Prox	94%	95
SP_054	0.64	4.60	Prox	1.50	8.08	Prox	89%	145
SS_161	0.54	2.91	Dist	1.02	4.29	Prox	89%	111
SS_107	0.62	3.48	Prox	1.42	8.57	Prox	93%	96
TM_355	0.60	4.60	Prox	1.62	6.56	Dist	94%	86
VL_124	0.64	3.76	Prox	1.05	5.50	Prox	87%	119
VW_006	0.45	2.44	Prox	0.89	4.60	Dist	95%	78
WK_503	0.52	3.06	Prox	1.07	5.62	Prox	70%	175
Avg.	0.57	3.57		1.26	6.40		90%	109
Std.	0.06	0.69		0.27	1.61		7%	29

Table 4.5: Female tibia model validation

Female tibia	SSM			ASM			Mean AM	
Sample ID	AD [mm]	HD [mm]	HD Pos.	AD [mm]	HD [mm]	HD Pos.	R2 [%]	AIE [HU]
NB_320	0.43	2.36	Prox	0.78	4.70	Prox	95%	84
RC_631	0.48	2.55	Prox	0.89	4.46	Dist	95%	94
RS_700	0.66	5.57	Dist	0.97	6.74	Dist	90%	136
SA_303	0.44	2.82	Prox	0.75	4.22	Dist	95%	93
SL_401	0.48	2.70	Dist	0.90	4.06	Prox	90%	123
SL_077	0.61	3.87	Prox	1.46	7.44	Prox	92%	125
SP_054	0.53	3.68	Prox	0.98	5.76	Prox	91%	131
SS_161	0.48	2.96	Prox	0.75	3.86	Dist	92%	107
SS_107	0.64	4.76	Prox	0.95	6.05	Dist	93%	102
VW_006	0.49	2.16	Prox	1.06	4.88	Dist	95%	81
Avg.	0.52	3.34		0.95	5.22		93%	107
Std.	0.08	1.06		0.20	1.16		2%	19

the femur were fitted more accurately than the epiphyseal areas. The female tibia SSM performed the best of the four statistical populations tested, and the male femur SSM the worst.

4.3.2 Active Shape Model Validation

Tables 4.2 to 4.5 also shows the average AD and HD errors for all the ASMs, ranging from 0.95 ($\pm$ 0.20) mm to 1.42 ($\pm$ 0.16) mm and 5.08 ($\pm$ 0.84) mm to 7.41 ($\pm$ 1.05) mm, respectively. Again, the HD was inspected visually and observed to occur on the proximal or distal ends of both the femur and tibia models. This is most likely due to neighbouring bone geometry or protrusions at the joint areas within the CT image. The ASM profiles at these areas were fitted to the local minima of the bony protrusions close to the joints, not the femoral or tibial ends. The female tibia ASM performed the best for the four statistical populations tested and the male femur ASM the worst.

4.3.3 Appearance Model Validation

Figures 4.7 and 4.8 show the mean image and standard deviation from the training population for all four sample groups. From the figures, we can see that most of the variation for the groups occur within the medullary cavity and the proximal and distal ends. The diaphyseal and metaphyseal cortical bone were represented accurately.

When analysing the R^2 metric in Tables 4.2 to 4.5, it becomes apparent that the mean AM gave a reasonable estimation of the intensity distribution for the test samples, averaging above 90 % of the original density distribution.
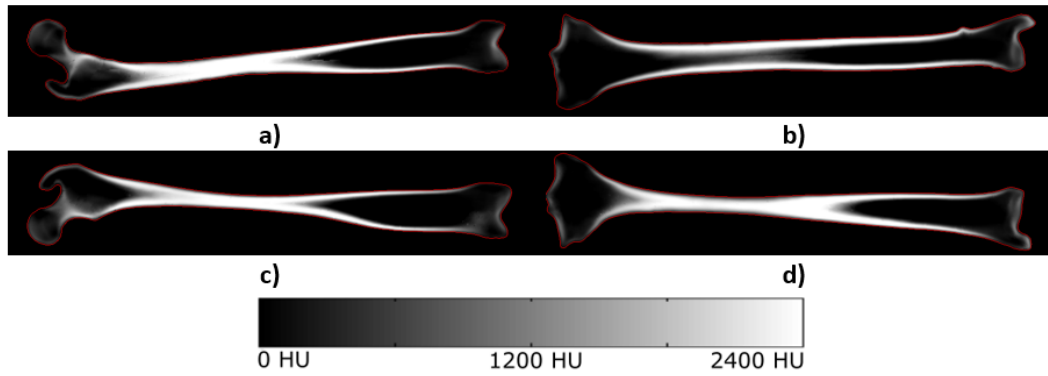
Figure 4.7: An image of the mean AM: a) Male femur, b) Male tibia, c) Female femur, and d) Female tibia. The images were viewed with a window width in Hounsfield Units (HU) ranging from 0 to 2400.
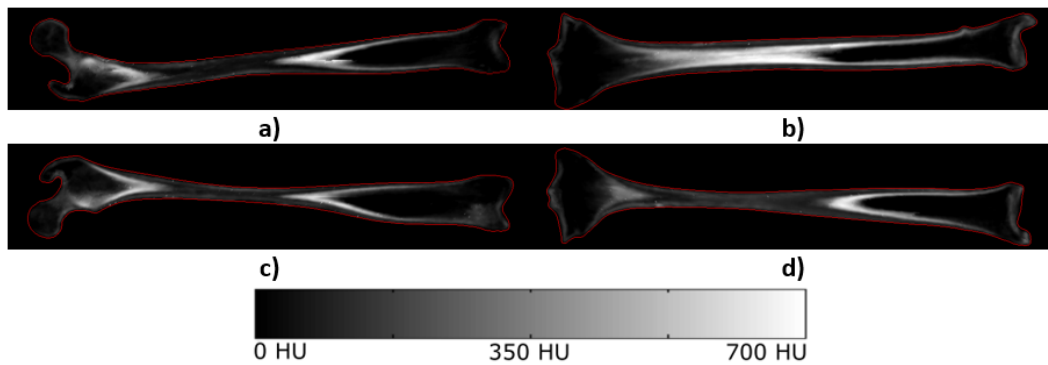
Figure 4.8: An image of the univariate standard deviation from the training population: a) Male femur, b) Male tibia, c) Female femur, and d) Female tibia. The images were viewed with a window width in Hounsfield Units (HU) ranging from 0 to 700.

Visual inspection of residuals throughout AM testing also revealed that the R^2 metric was most accurately calculated. Looking at the standard deviation for all the population groups in Figure 4.8, we find that most AIE lies within the cancellous bone regions.

4.4 Discussion

Audenaert *et al.* (2019) showed that SSMs constructed from 50 training femurs or tibias could have a model generalisation error below 1 mm. It is clear from Tables 4.2 to 4.5 that the full SSMs should have a generalization error below 0.7 mm. This showed that the constructed SSMs could describe members of the South African male and female population outside of the sample data used to train the models. Audenaert *et al.* (2019) also found that the generalisation error for their femur and tibia models did not go below 0.5 mm with a sample

size of 250. Though the generalisation error is also dependent on the statistical population, increasing the dataset would most likely have a limited effect on the generalisation ability of the SSM.

When analysing the results for the ASMs, it was found that the average AD and HD errors ranged from 0.95 ($\pm$ 0.20) mm to 1.42 ($\pm$ 0.16) mm and 5.08 ($\pm$ 0.84) mm to 7.41 ($\pm$ 1.05) mm, respectively. From Subsection 2.4.2 these are acceptable errors for reconstruction surgery as observed in literature and industry. More specifically, the AD error falls below the 2 to 4 mm overhangs reported from literature Abram *et al.* (2014); KR *et al.* (2020), and below the 5 to 10 mm tolerance from industry (CranioTech (PTY) Ltd., 2021). Though the HD error is below the 10 mm range, it lies within the proximal and distal epicondyle ends. Thus, the HD error based on diaphyseal resections and estimations, as measured in the following chapter, will be a better indicator of the models' performance with respect to this study's aim and objectives. Still, the segmentations were done automatically without any specialist knowledge and completed within a few minutes as opposed to typical manual processes which could take up to half an hour or longer. Additionally, the ASM profiles fitting to the bony protrusions close to the joints could be prevented through a better initial starting position, manual intervention during the fitting process, or an added MH parameter, θ. These considerations will be discussed further in Chapters 5 and 6.

In Figures 4.8.a and 4.8.b, it can be observed how the mean AM represents the intensity distribution of the sample sets. The largest standard deviations were found within the medullary cavity and epiphyseal ends, most likely due to the dynamic nature of cancellous bone tissue (Betts *et al.*, 2013). The relatively low standard deviation of the compact cortical bone, especially in the diaphyseal area, indicates an accurate representation by the mean image.

In literature, AMs of bone have been used for two applications. Firstly, to assess risk factors for fractures in long bone structures and joint anatomy related to osteoporosis (Sarkalkan *et al.*, 2014). Secondly, to calculate the stress and strain distribution, study bone tissue adaptation and simulate cycle loading on implants via patient-specific finite element (FE) models (Sarkalkan *et al.*, 2014; Poelert *et al.*, 2013). For both applications, an AM or FE model is populated by estimating the bone mineral density (BMD) of the patient from in vivo imaging using several direct methods (Poelert *et al.*, 2013; Duchemin *et al.*, 2008; Lekadir *et al.*, 2015). Lekadir *et al.* (2015) describe a novel method for estimating the trabecular micro-architecture from healthy CT images using a combined statistical shape and appearance model, trained from ex vivo micro-CT images. However, while direct methods are computationally less expensive, they do not incorporate prior knowledge of variation within a population to guide a statistically valid outcome. Secondly, while a statistical AM can accurately model the variation within a population, training requires a healthy high-resolution training database, and it is computationally much more expensive during data processing, model construction, and model fitting

(Cootes *et al.*, 2004; Sarkalkan *et al.*, 2014; Lekadir *et al.*, 2015).

Thus, the mean AMs used and validated during this study is a useful compromise. When estimating the density distribution, we get the speed of the direct methods with the additional statistical validity of a partial AM, as seen from the +90% average density fit throughout testing. Furthermore, the appearance estimate based on CT image intensity can be directly linked to BMD, which can be used to estimate the Young's modulus and ultimate strength of bone (Duchemin *et al.*, 2008). Thus, even with some error in the estimation, the density distribution captured within the mean AM could enable the design of patient-specific functionally graded lattice structures that closely resemble the mechanical properties of an individual's healthy bone structure to avoid stress shielding and stimulate new bone growth and nutrient flow at the implant site (Wang *et al.*, 2016).

This chapter presented a method in which a representative mean density image can be warped to a patient's image using ASMs. The results are promising, with accuracies within acceptable tolerances for an application aimed towards the surgical reconstruction of long bone defects. Chapters 5 and 6 will focus on estimating and reconstructing pathological geometry based on sparse inputs, typically seen in patients suffering from long bone defects.

Chapter 5

Estimating Missing Bone Shape and Density

In reference to the models and methods developed in Chapter 4, the ability to estimate missing bone anatomy was investigated within this chapter by reconstructing digitally simulated segmental bone loss in healthy femoral and tibial bones. This chapter is used to evaluate the estimated surface geometry and density distribution accuracy before introducing misalignment. This shows how well the prior knowledge captured within the statistical models can estimate missing anatomy.

5.1 Simulated Bone Loss

To simulate segmental bone loss on the test samples from Section 4.3, the centre of each sample's volume was calculated and used as a datum for clipping segments of the diaphyseal shaft, dividing the test sample into proximal and distal ends. The length of the cut segments ranged from 5 to 70 % in 5 % increments, based on the test sample's total length. After cutting, the sample's target image would only contain voxel data of the remaining proximal and

Table 5.1: Length of total bone structure and segmental bone loss segments

[mm]	Male femur	Male tibia	Female femur	Female tibia
Length	431.16 (±20.93)	390.03 (±28.36)	410.77 (±17.67)	366.06 (±18.17)
5%	21.56 (±1.05)	19.50 (±0.88)	20.57 (±0.88)	18.30 (±0.91)
35%	150.91 (±7.33)	136.51 (±9.92)	143.77 (±6.18)	128.12 (±6.36)
70%	301.81 (±14.6)	273.02 (±19.85)	287.54 (±12.37)	256.24 (±12.72)

45

distal ends. Table 5.1 gives a summary of the average length of the ten test samples used for each population group.

5.2 Estimation Method

Figure 5.1 gives a visual overview of the estimation process followed when re-estimating the diaphyseal segment cut from the test samples. Note that an understanding of Figures 4.3 and 4.4 is required to fully understand the estimation process explained below.

To ensure the ASM fits the global minima of the remaining proximal and distal ends, an appropriate initial alignment and shape instance was required (Cootes *et al.*, 2004; Morel-Forster *et al.*, 2018). To determine this alignment and shape instance, two sets of manually placed anatomical landmarks were needed. In this study, the landmarks used for alignment consisted of three proximal points and three distal points. Figure 2.2 illustrates the landmarks used for both the femoral and tibial bones. After the shaft segment was cut, the aforementioned corresponding landmarks were placed manually on the ASM and the partial CT image. Two landmarks were placed on the anterior and posterior of the model's shaft centre to guide the shaft shape profile during initialisation. The corresponding landmarks were again placed on the partial CT image. When calculating the initial shape instance, the posterior of the ASM was computed for the two sets of corresponding landmarks. To compute the posterior, a Metropolis-Hastings fitting was derived following the methodology set out in Subsection 4.2.2.

The resulting posterior and ASM profiles were then used to compute the most probable fit to the partial image. First, viable ASM profiles used during the fitting procedure had to be determined. Assuming $\Gamma(\theta_{LM})$ is a reasonable first estimate of the target boundary, the coordinates of the ASM profiles were used to sample intensity profiles from the partial image. If no voxel data was found for an image intensity profile, the ASM profile was assumed to be within the cut segment or out of the image frame. These intensity profiles would be disregarded during the final fitting procedure, as fitting profiles in the cut segment would lead to grouping of the profile points at the proximal and distal cut edges, resulting in a stretched model surface over the cut segment. The final ASM posterior was again calculated using the Metropolis-Hastings algorithm by evaluating the ASM profiles against the intensity profiles from the partial image with each new $\Gamma(\theta')$. This was done to establish correspondence between the ASM surface boundary and the partial image's proximal and distal femur surface. The mean surface of the resulting posterior was then used to infer the missing surface geometry of the partial image to estimate the healthy shape after simulated segmental bone loss. The accuracy of the approximation was then measured using the average distance error and Hausdorff error metrics.

Afterwards, the shape normalised mean AM was warped to fit the outer surface of the estimated diaphyseal segment via linear interpolation. The cor-
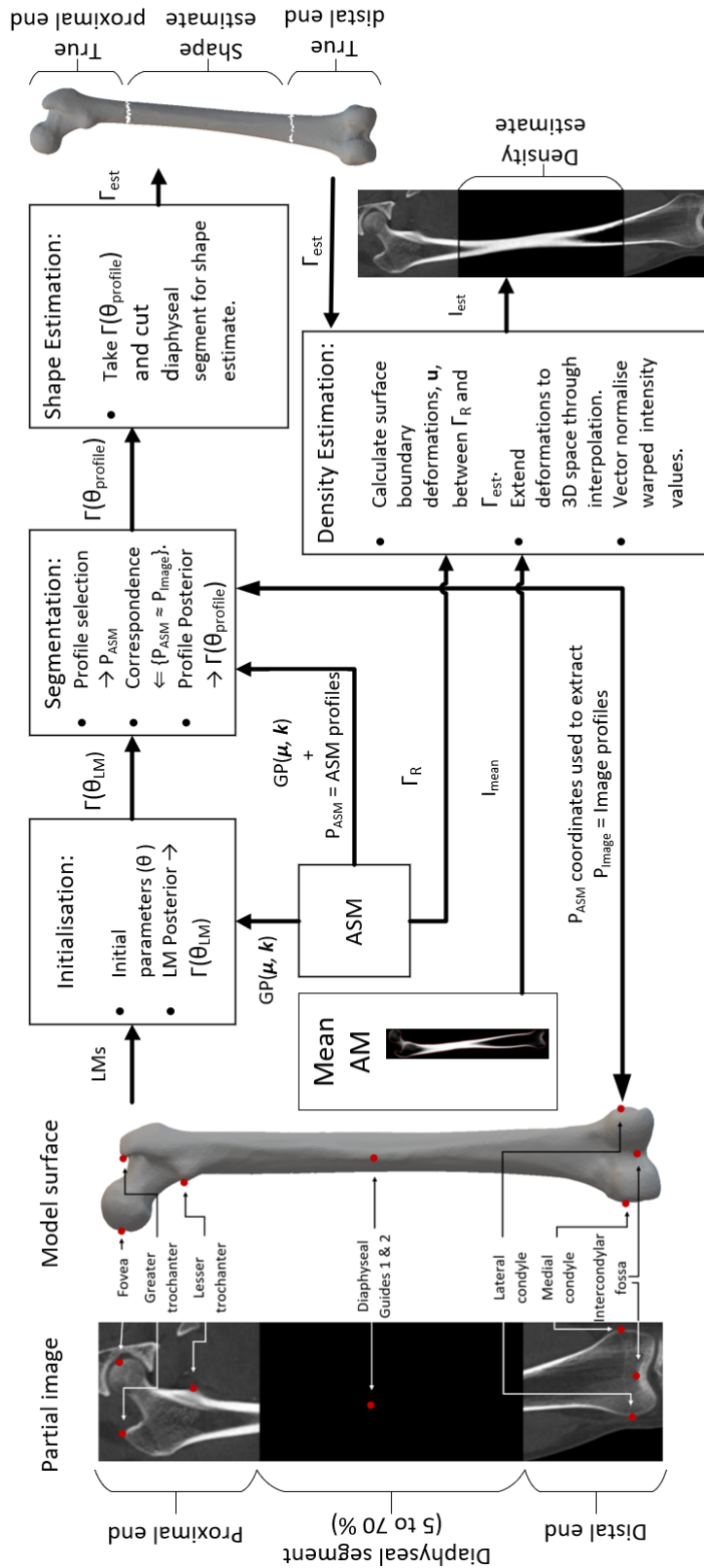
Figure 5.1: Shape en density estimation pipeline.

responding vertex coordinates on both the estimated mesh surface and the known AM surface boundary mesh was used to calculate a set of deformation fields, $\{u_1, ..., u_N\}$, for each point on the shape estimate. Through linear interpolation, the deformation fields were extended to the entire 3D space containing the mean AM. This warped the voxel data of the mean AM to produce an appearance estimate that was fitted to the shape estimate, thereby estimating a healthy density distribution based on the population mean. The accuracy of the density approximations was then measured using the R^2 and average intensity error metrics.

Additionally, to ensure a more accurate estimate, the intensity data of the mean AM was normalised to fit the remaining intensity data of the test sample. This was done to account for the anatomical variance of the test sample as well as to combat intensity variation from different machine imaging settings and algorithms (Cootes *et al.*, 2004; Heimann and Meinzer, 2009; Sarkalkan *et al.*, 2014). These procedures were iterated for each test sample of each population group, spanning the entire range of the length of the cut segments.

5.3   Estimation Results

The same leave-one-out test procedure and metrics, as during model validation, were used to measure the estimated bone shape and density distribution accuracy. The estimations were measured against the ground truths, i.e. the segments cut from the target shapes and images were stored and compared to the shape and density estimates. In this section, the results are shown through a series of graphs for each population group. Afterwards, figures showing the shape and density estimate of one test sample from each group are given and briefly discussed.

Figures 5.2.a and 5.2.b show the average and Hausdorff errors obtained for the male femur population. As the cut length increased, the accuracy of the shape estimation decreased, with the average distance error ranging from 0.71 ($\pm$ 0.21) mm to 2.36 ($\pm$ 0.53) mm for the 5 % and 70 % resections, respectively. Similarly, the average Hausdorff error ranged from 1.41 ($\pm$ 0.43) mm to 5.02 ($\pm$ 1.90) mm. Figures 5.2.c and 5.2.d show the R^2 and average intensity error of the density distribution obtained after warping the AM. There was a gradual decrease in R^2, ranging from 89 ($\pm$ 4) % for resections of 5 % of the total length to 87 ($\pm$ 6) % for resections of 70 % of the total length. The average intensity error was relatively constant regardless of the size of the resection, with an average error of approximately 180 HU.

Figures 5.3.a and 5.3.b show the average and Hausdorff errors obtained for the male tibia population. As the cut length increased, the accuracy of the shape estimation decreased, with the average distance error ranging from 0.82 ($\pm$ 0.31) mm to 1.47 ($\pm$ 0.32) mm for the 5 % and 70 % resections, respectively. Similarly, the average Hausdorff error ranged from 2.75 ($\pm$ 0.79) mm to 5.10 ($\pm$ 0.74) mm. Figure 5.3.c shows a gradual increase in R^2, ranging from
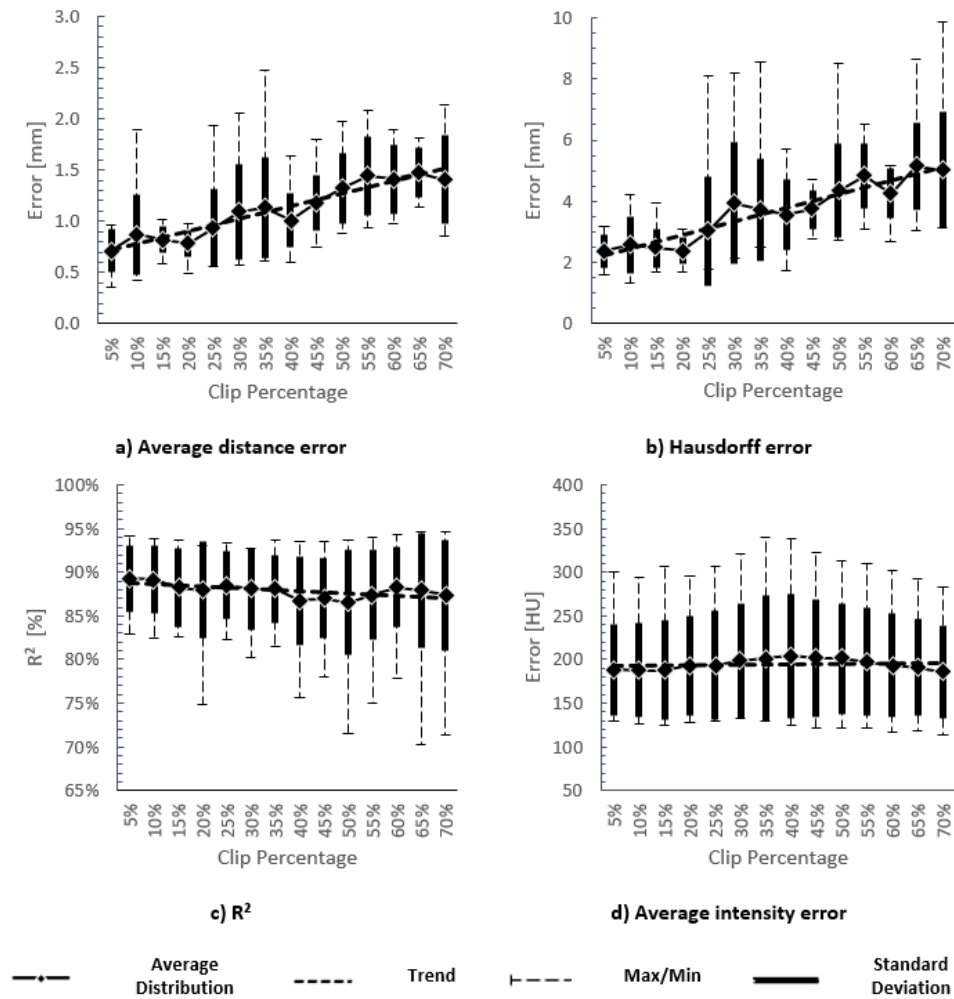
Figure 5.2: The combined statistical analysis of the test results for segmental bone repair of the male femur.

86 ($\pm$ 7) % for resections of 5 % of the total length to 90 ($\pm$ 4) % for resections of 70 % of the total length. The average intensity error, Figure 5.3.d, also showed an increase in accuracy, ranging from 226 ($\pm$ 70) HU for resections of 5 % of the total length to 180 ($\pm$ 44) HU for resections of 70 % of the total length.

Figures 5.4.a and 5.4.b show the average and Hausdorff errors obtained for the female femur population. As the cut length increased, the accuracy of the shape estimation decreased, with the average distance error ranging from 0.64 ($\pm$ 0.14) mm to 1.64 ($\pm$ 0.41) mm for the 5 % and 70 % resections, respectively. Similarly, the average Hausdorff error ranged from 2.00 ($\pm$ 0.65) mm to 4.49 ($\pm$ 0.87) mm. Figure 5.4.c shows that the average R^2 was relatively constant regardless of the size of the resection, with an average of approximately 89 ($\pm$ 6) %. The average intensity error, Figure 5.4.d, was also relatively constant, with an average error of approximately 160 HU.
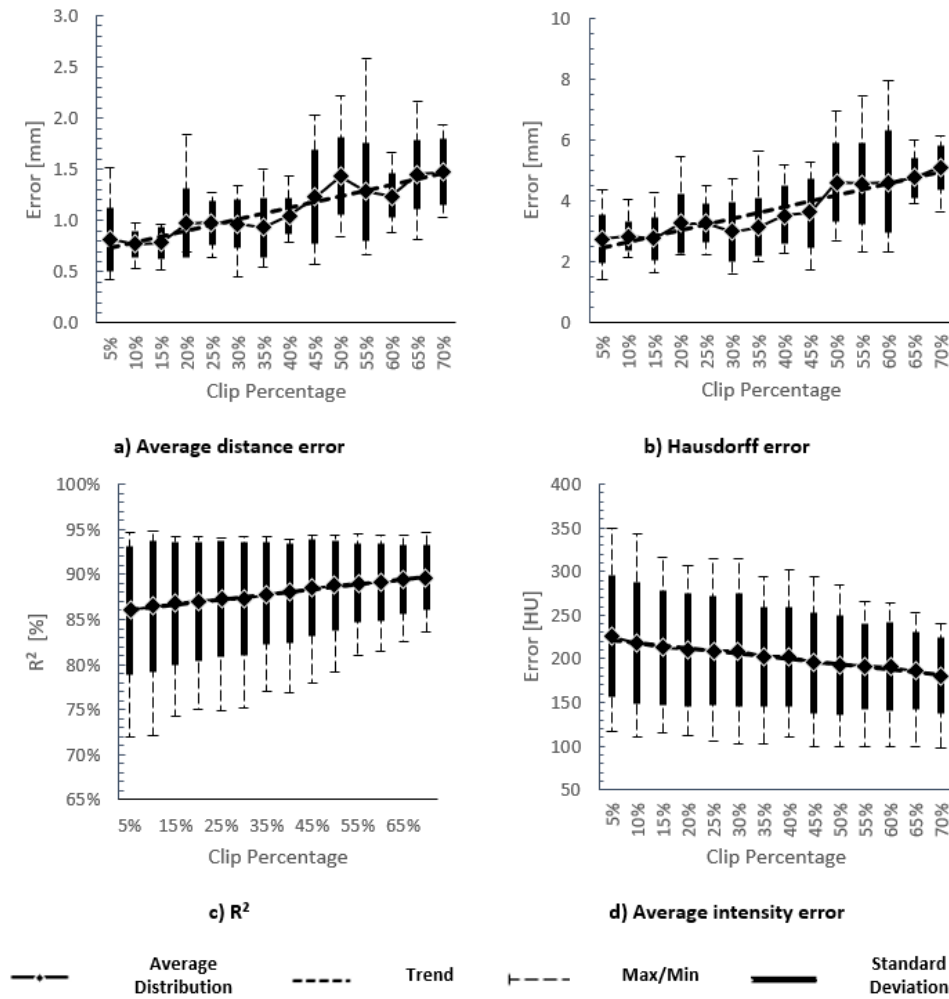
Figure 5.3: The combined statistical analysis of the test results for segmental bone repair of the male tibia.

Figures 5.5.a and 5.5.b show the average and Hausdorff errors obtained for the female tibia population. As the cut length increased, the accuracy of the shape estimation slightly decreased, with the average distance error ranging from 0.58 ($\pm$ 0.12) mm to 1.06 ($\pm$ 0.36) mm for the 5 % and 70 % resections, respectively. Similarly, the average Hausdorff error ranged from 1.61 ($\pm$ 0.41) mm to 3.60 ($\pm$ 1.03) mm. Figure 5.5.c shows that there was a gradual increase in R^2, ranging from 84 ($\pm$ 5) % for resections of 5 % of the total length to 89 ($\pm$ 2) % for resections of 70 % of the total length. The average intensity error, Figure 5.5.d, also showed an increase in accuracy, ranging from 208 ($\pm$ 25) HU for resections of 5 % of the total length to 159 ($\pm$ 20) HU for resections of 70 % of the total length.

Figures 5.6 to 5.9 show the shape of the estimated diaphyseal segments for one left-out sample of each population group. Visual inspection shows that as the size of the resected portion increases, the accuracy of the estimate
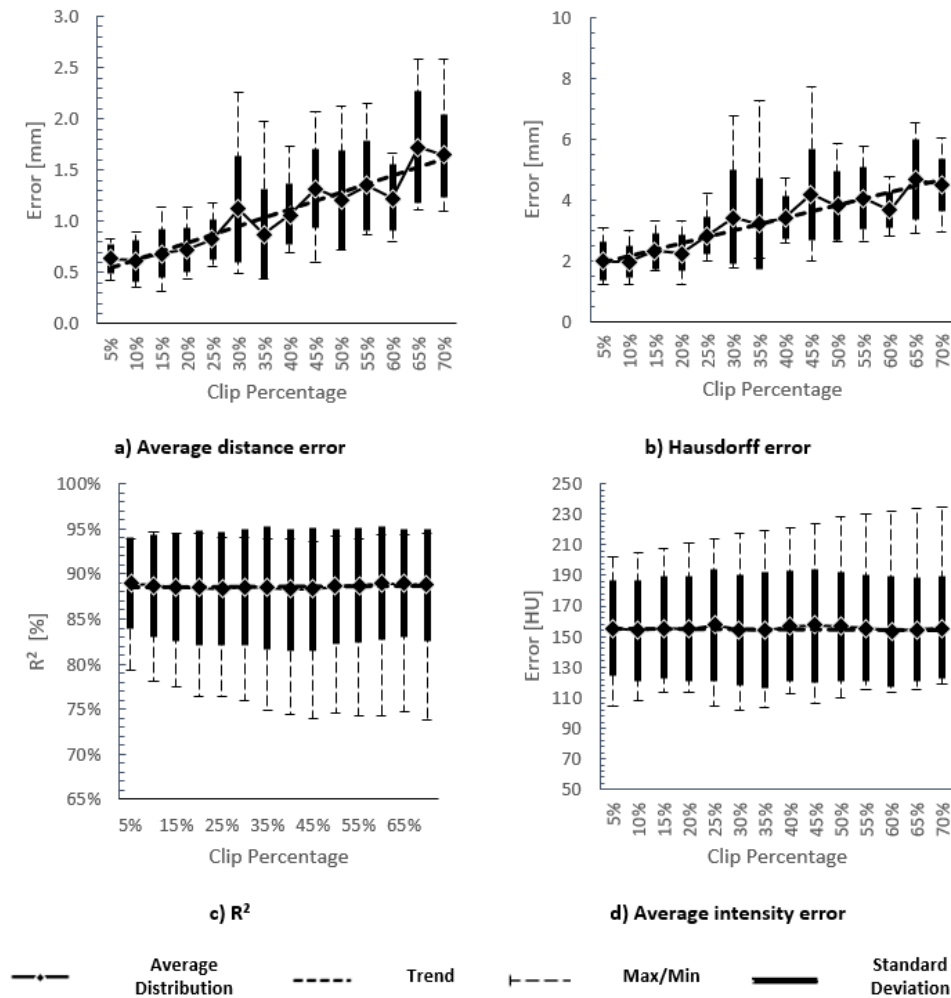
Figure 5.4: The combined statistical analysis of the test results for segmental bone repair of the female femur.

deteriorates. This is most noticeable at the centre of the resection, while sections towards the distal and proximal end more closely match the left-out shape's original geometry. Note that the edge alignment on the proximal end for the 70 % tibial resections are not as well estimated as the distal ends.

Figures 5.10 to 5.13 show axial and coronal slices of the original and estimated density distributions for one target sample of each population group. In general, the estimations follow the cortical wall thickness and profile, with slight variations in wall thickness. Furthermore, visual inspection reveals that the density estimations are more accurate towards the exterior of the shape profile. However, the estimations produce minor artefacts of cortical bone at the inner wall of the medullary cavity.
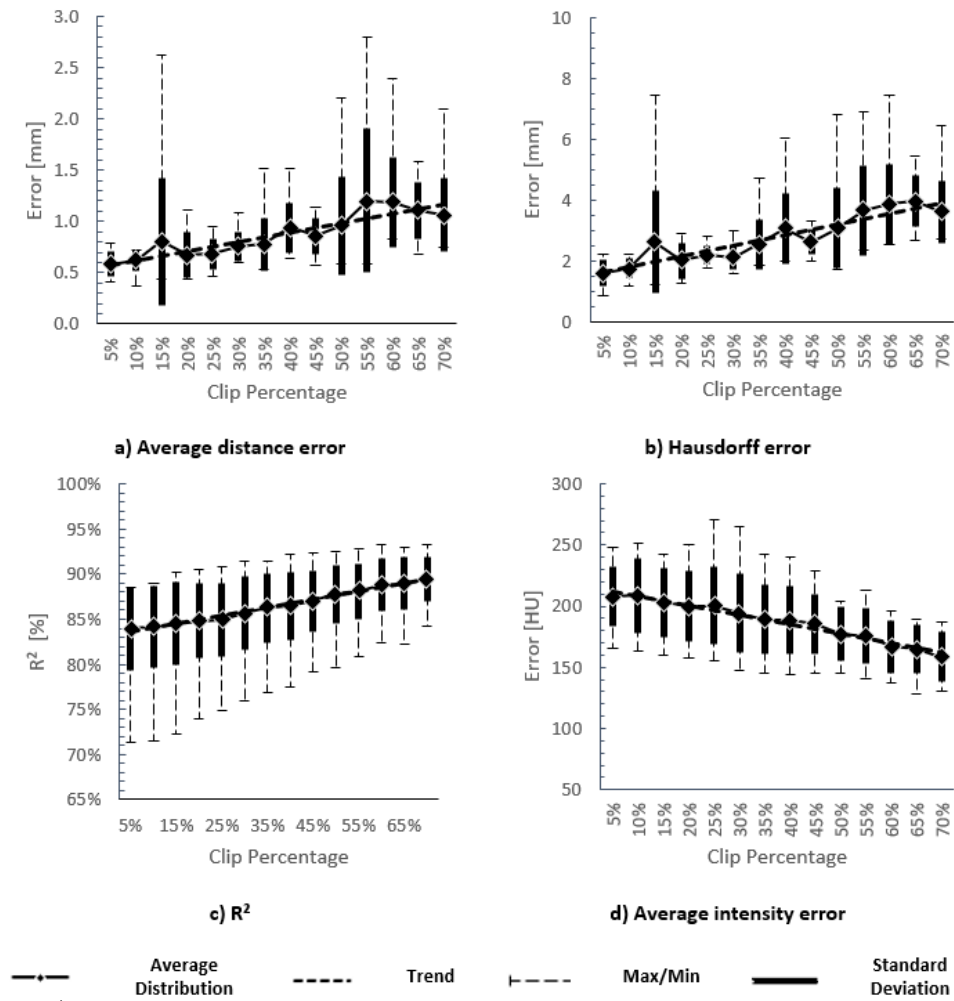
Figure 5.5: The combined statistical analysis of the test results for segmental bone repair of the female tibia.

5.4 Discussion

The combined ASM and mean AM used in this study could reconstruct and estimate the shape and density distribution of resected femoral or tibial diaphyseal segments with an average distance error ranging from 0.58 to 1.64 mm and a R^2 above 84 % of the intensity of the original target images. These results indicate that the method proposed in this study was able to estimate the original, healthy condition of long bones after segmental bone loss, suggesting that it is viable for automating PSI design procedures such as reconstruction and segmentation.

The ASM and automated segmentation algorithms produced diaphyseal estimates with a mean average error below 2.5 mm and a mean Hausdorff error below 5 mm during leave-one-out testing. While direct comparisons are difficult to make, pre-and post-operative reviews in literature show successful
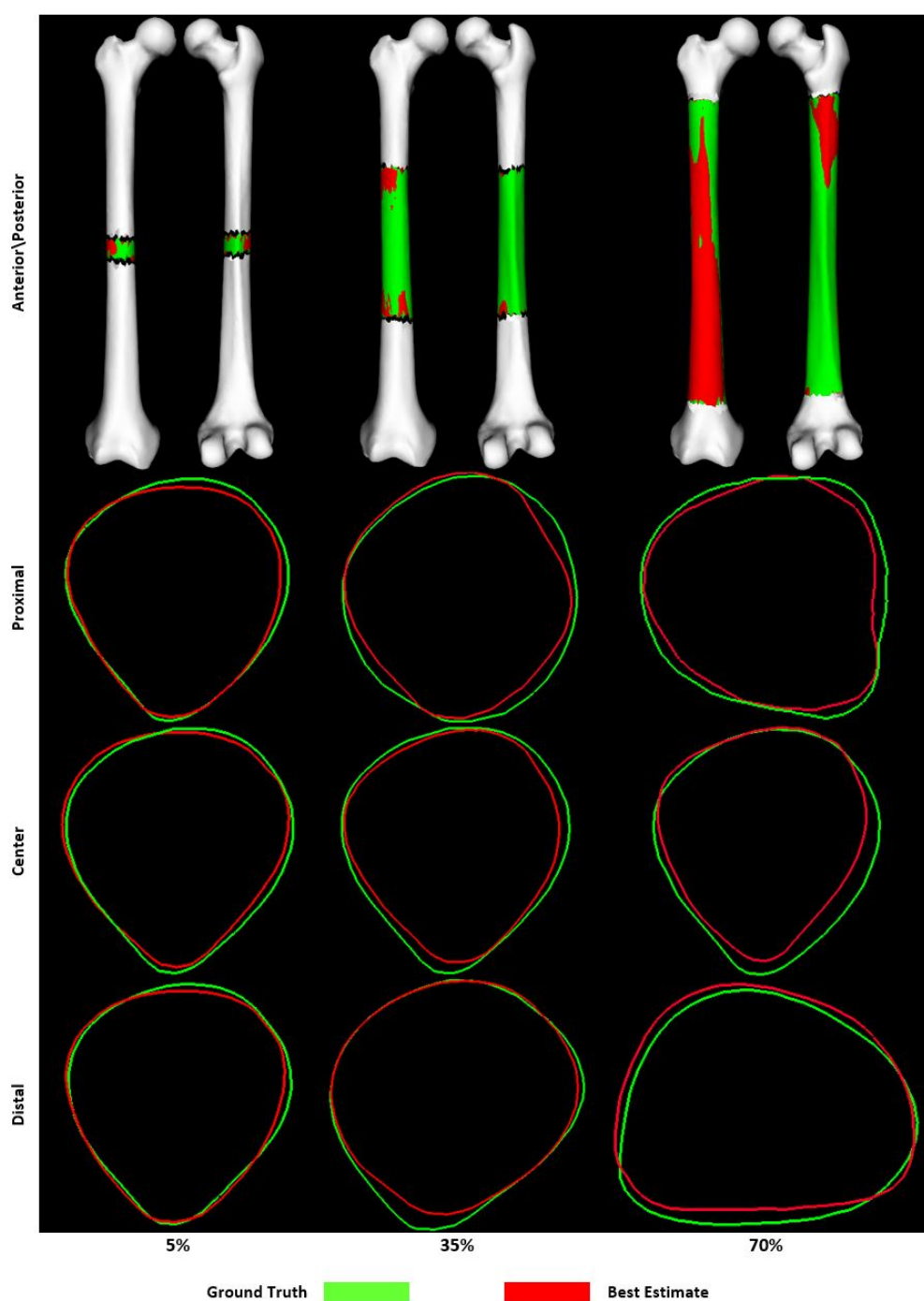
Figure 5.6: The shape estimation for sample NS-401 of the male femur population. The estimations had an average distance error and a Hausdorff error of 0.49 mm and 2.34 mm for the 5 % resection, 1.09 mm and 3.03 mm for the 35 % resection, and 1.11 mm and 3.13 mm for the 70 % resection (recorded length 458 mm).
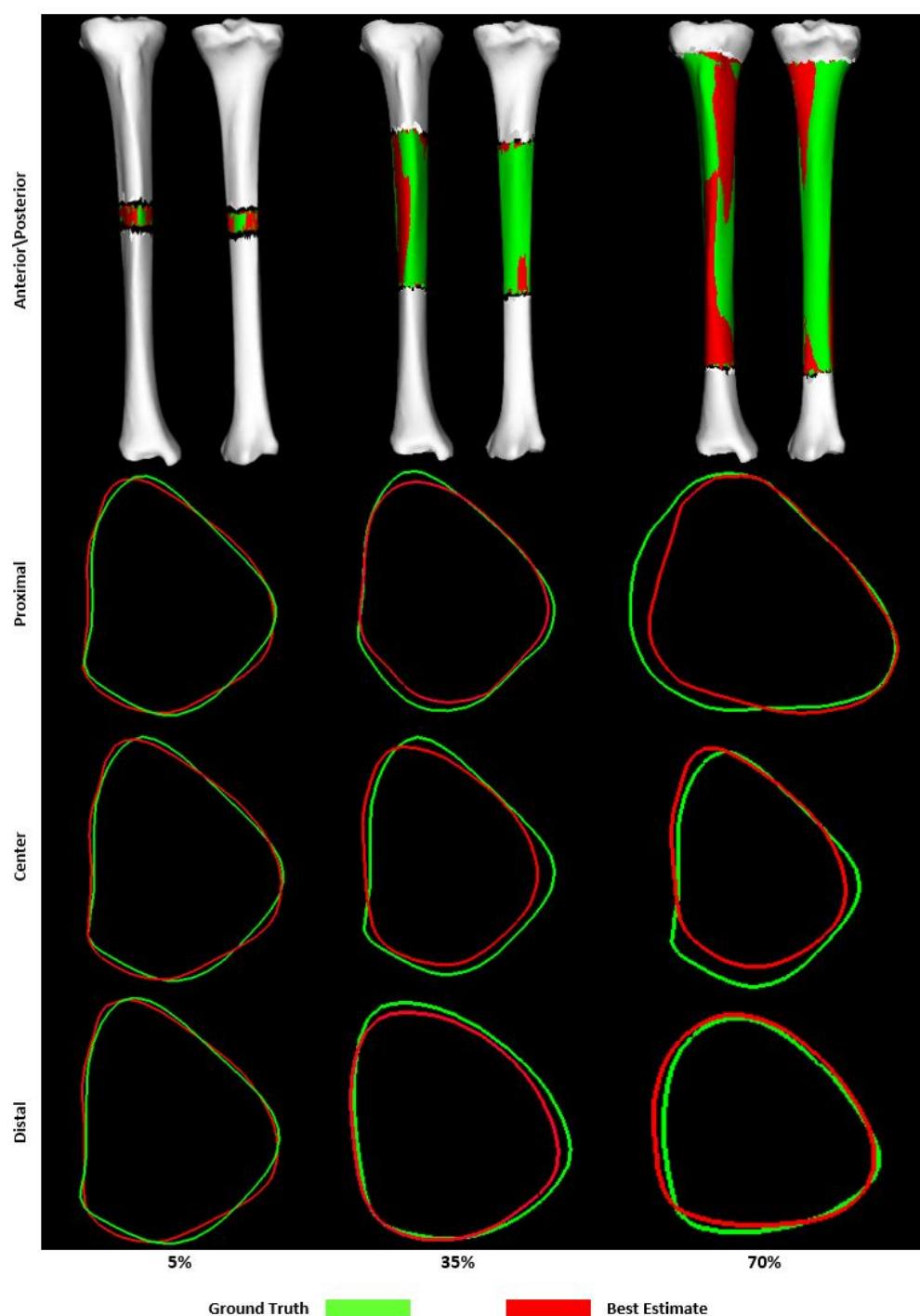
Figure 5.7: The shape estimation for sample NS-401 of the male tibia population. The estimations had an average distance error and a Hausdorff error of 0.79 mm and 2.19 mm for the 5 % resection, 1.00 mm and 2.69 mm for the 35 % resection, and 1.78 mm and 4.80 mm for the 70 % resection (recorded length 398 mm).
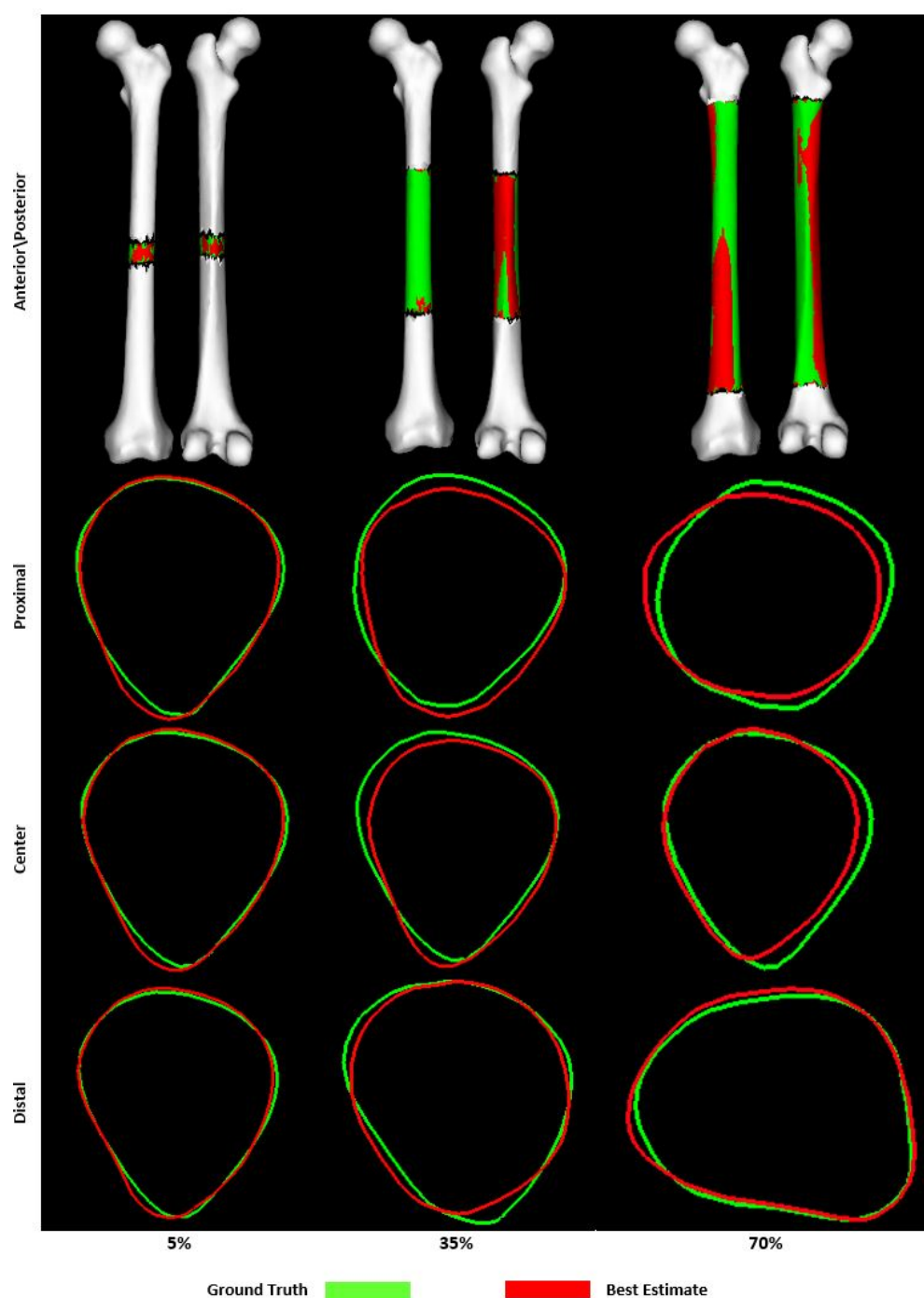
Figure 5.8: The shape estimation for sample SA-303 of the female femur population. The estimations had an average distance error and a Hausdorff error of 0.63 mm and 1.71 mm for the 5 % resection, 1.00 mm and 2.69 mm for the 35 % resection, and 1.78 mm and 4.80 mm for the 70 % resection (recorded length 398 mm).
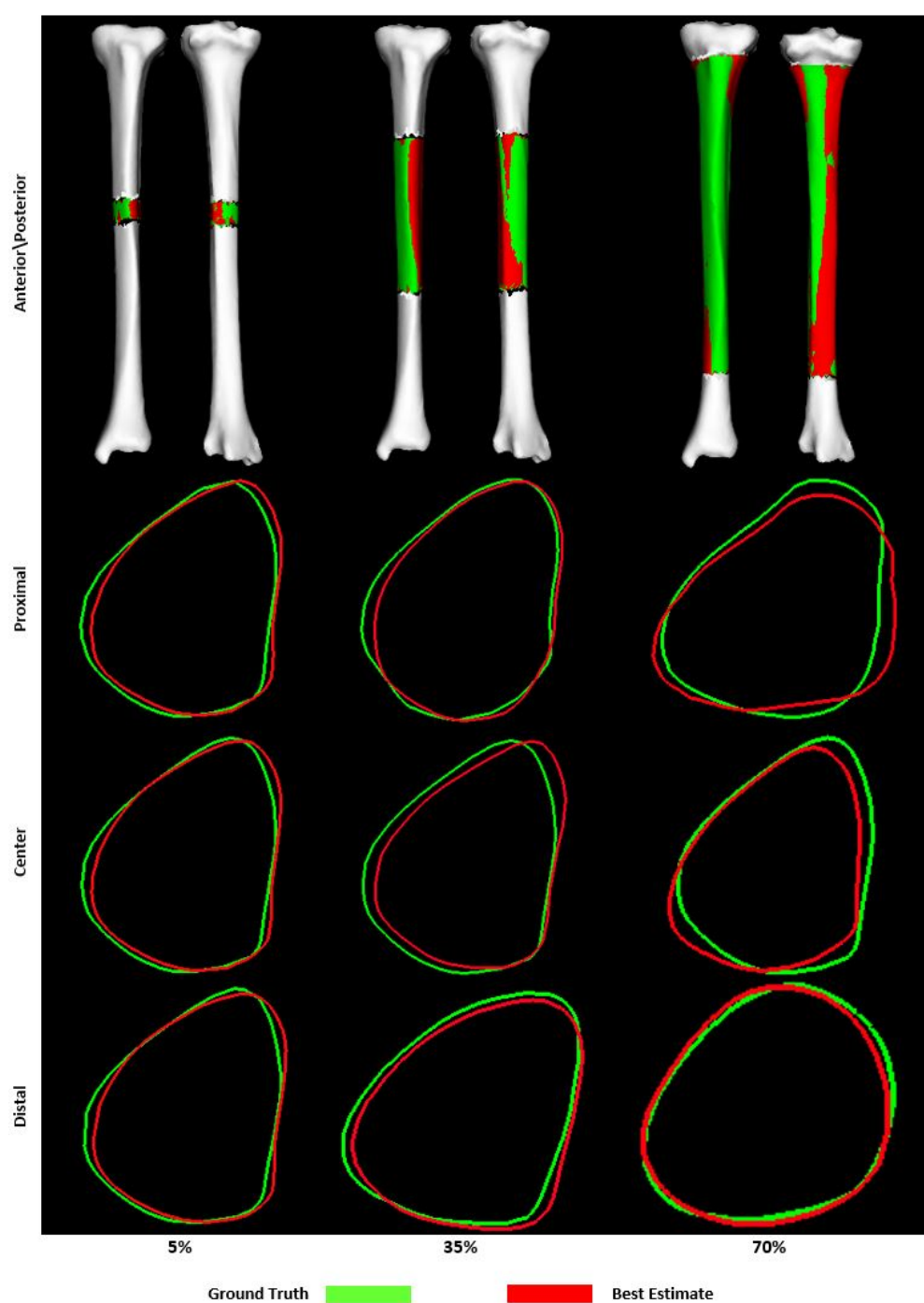
Figure 5.9: The shape estimation for sample NB-320 of the female tibia population. The estimations had an average distance error and a Hausdorff error of 0.41 mm and 0.88 mm for the 5 % resection, 0.71 mm and 2.44 mm for the 35 % resection, and 1.03 mm and 3.97 mm for the 70 % resection (recorded length 347 mm).
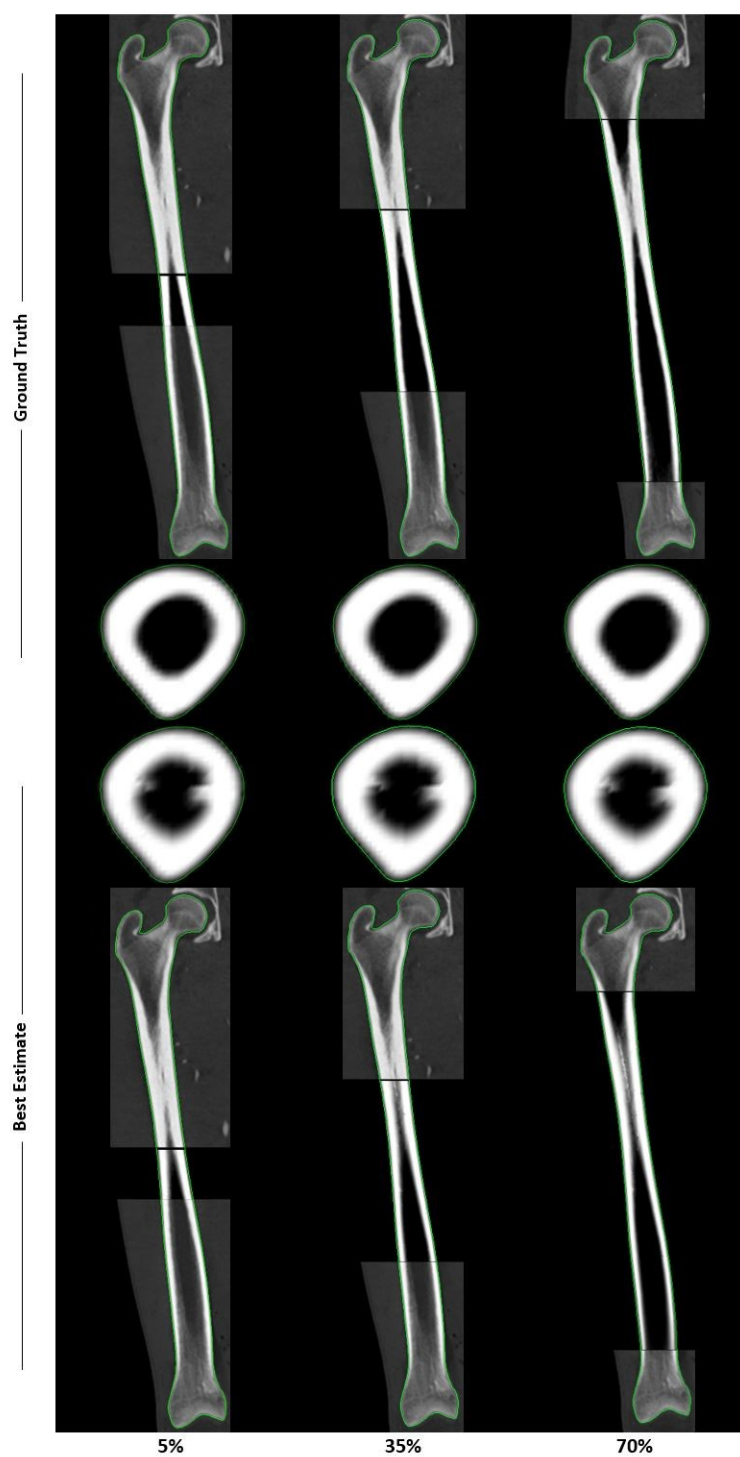
Figure 5.10: The appearance estimation for sample NS-401 of the male femur population. The estimations had a R square and average intensity error of 90 % and 169 HU for the 5 % resection, 91 % and 161 HU for the 35 % resection, and 92 % mm and 143 HU for the 70 % resection (recorded length 458 mm).
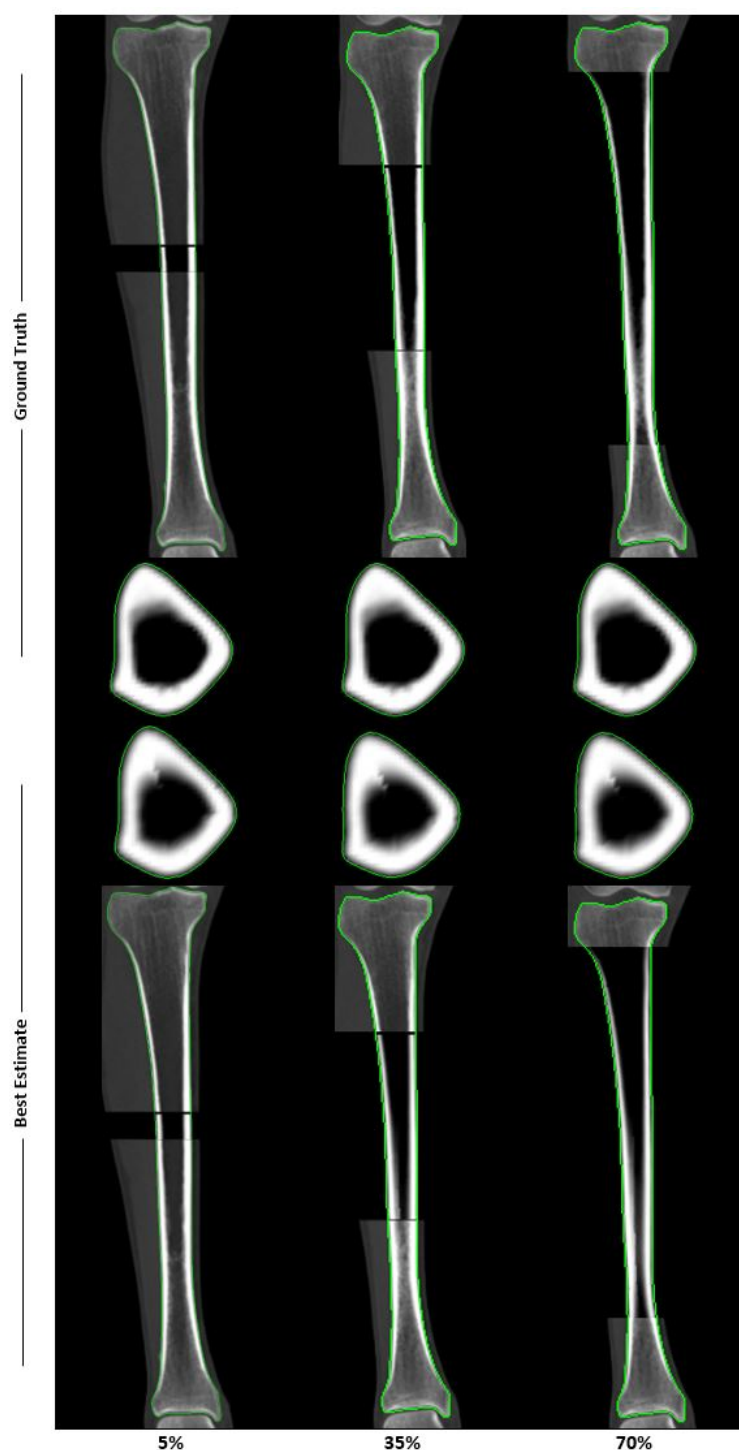
Figure 5.11: The appearance estimation for sample NS-401 of the male tibia population. The estimations had a R square and average intensity error of 93 % and 175 HU for the 5 % resection, 93 % and 168 HU for the 35 % resection, and 95 % mm and 135 HU for the 70 % resection (recorded length 398 mm).
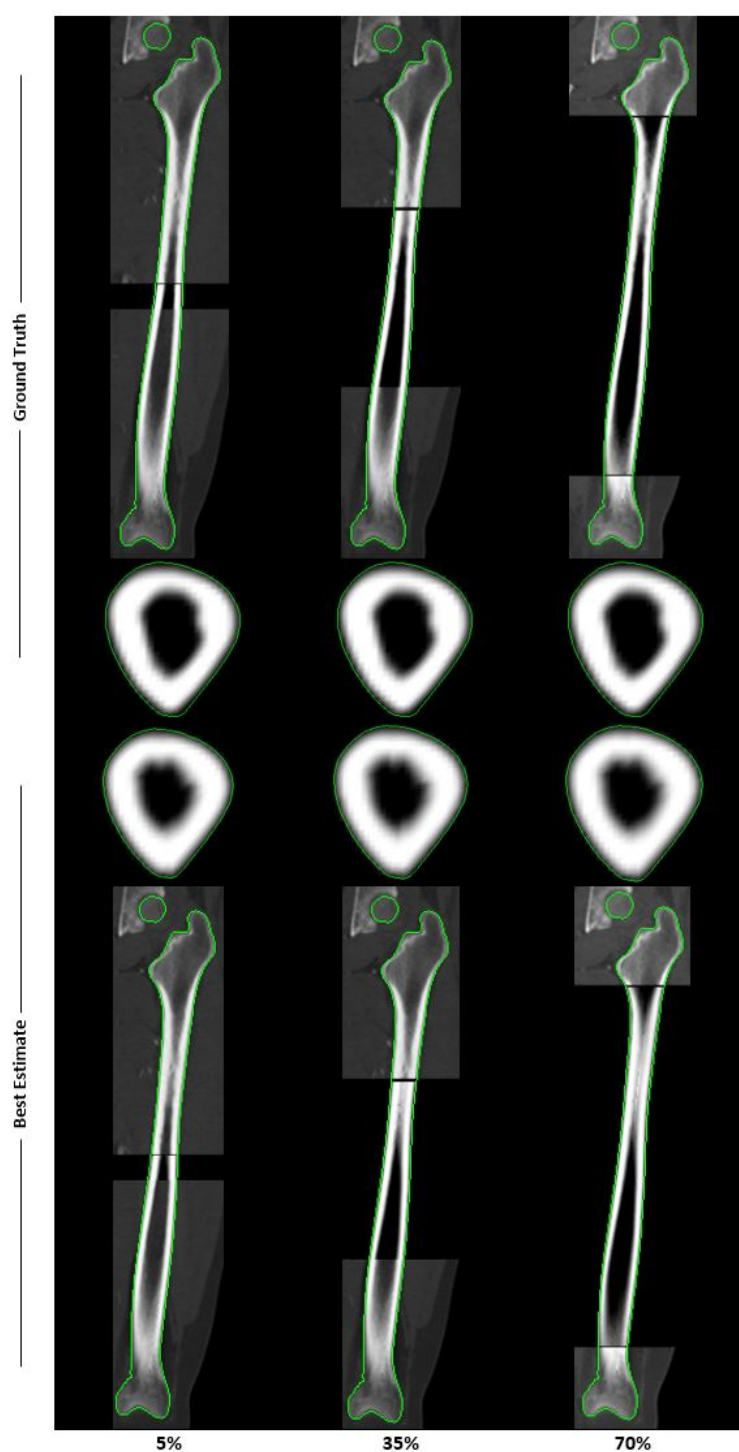
Figure 5.12: The appearance estimation for sample SA-303 of the female femur population. The estimations had a R square and average intensity error of 92 % and 160 HU for the 5 % resection, 91 % and 173 HU for the 35 % resection, and 92 % mm and 151 HU for the 70 % resection (recorded length 398 mm).
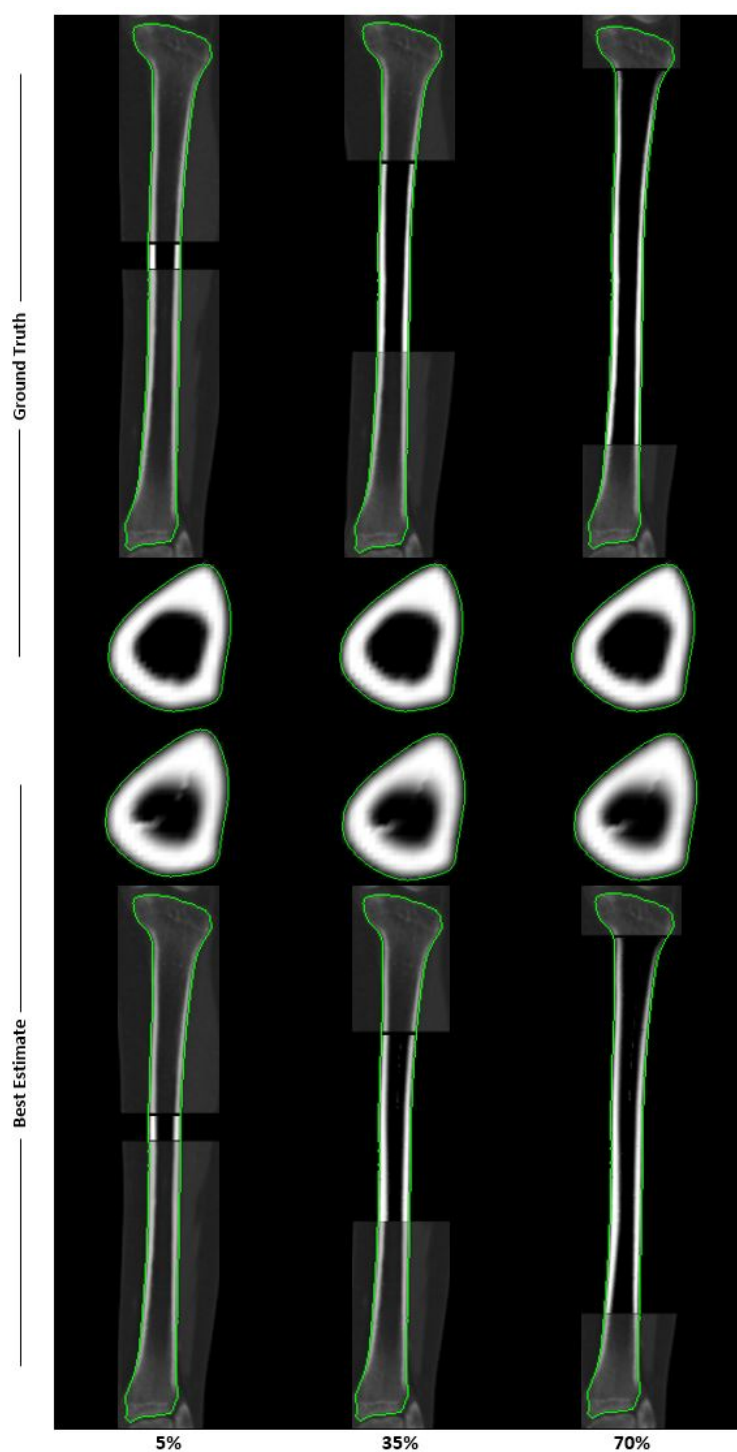
Figure 5.13: The appearance estimation for sample NB-320 of the female tibia population. The estimations had a R square and average intensity error of 85 % and 166 HU for the 5 % resection, 88 % and 147 HU for the 35 % resection, and 89 % mm and 135 HU for the 70 % resection (recorded length 347 mm).

outcomes for implant overhangs and variations in size ranging from 2 to 4 mm for knee and hip arthroplasty (Abram *et al.*, 2014; KR *et al.*, 2020). Based on the results from Figures 5.2.a to 5.5.a, the average distance errors were below 2.5 mm regardless of the size of the resected segment. Considering Figures 5.2.b to 5.5.b, the average Hausdorff errors were below 5 mm in all cases, although some cases had maximum errors of close to 10 mm. Nevertheless, it is essential to note that the Hausdorff errors occurred close to the centre of the resected shaft in all test cases. Furthermore, the highest Hausdorff error was less than 4 % of the total estimated length when 70% of the diaphysis was resected. With the femur and tibia being the longest bones in the body and their length being the main contributor to shape variance (Zhang *et al.*, 2016), the effect of the Hausdorff error measured in the radial direction close to the centre of the shaft becomes comparatively negligible.

The proximal and distal edge alignment in Figures 5.6 and 5.9 exhibit minor misalignment for estimates of resections 30 % and less. The edge alignment deteriorates from the 35 % increment as the registration and segmentation algorithms have fewer data points from which to derive a reasonable estimate. Still, due to the statistical data captured within the validated SSM, the surface geometry at the edge can be considered anatomically valid (Mauler *et al.*, 2017) and that, within reason, overhangs are a general occurrence when working with long bone implants (Abram *et al.*, 2014; KR *et al.*, 2020). Even when working with healthy contralateral bones to infer surface geometry for PSIs, adjustments are still required due to the bilateral asymmetry between left and right long bones (Mauler *et al.*, 2017; Auerbach and Ruff, 2006). However, it may be possible to improve the edge alignment of the estimation by altering the ASM used during this study. The ASM was constructed using the method of the local structure set out by Cootes et al. (Cootes *et al.*, 2004), where intensity profiles were sampled during construction. The positions of these profiles were chosen randomly over the model's surface, using a uniformly distributed algorithm. Suppose instead that the profiles were assigned to have a radial-grid configuration over the length of the shaft. In that case, the intensity profiles might have fit better to the remaining edge of the healthy proximal and distal end, resulting in better edge alignment.

Analysing graphs c and d in Figures 5.2 to 5.5, we see that the mean AM gives a similar result for the entire cutting range. The model shows an average fit quality above 84 % and an average intensity error below 226 HU. The R square metric measured how well the fitted mean AM explained the actual density distribution of the target sample; the higher the value, the better the fit was between the estimated density distribution and the ground truth. For the average intensity error, which could be an over or underestimation, it is essential to note that the target cortical bone density within all four data sets could be as high as 3200 HU. From these results, the mean AM can be considered to produce density estimations that closely resemble the original healthy density distribution.

Within Figures 5.10 to 5.13 we find artefacts of cortical bone on the medullary cavity wall. These artefacts were a product of the warping procedures used and not knowing the shape boundary of the medullary cavity. Still, we see that the estimations follow the cortical wall thickness and profile accurately when looking at the coronal slices. Some slices show a slightly larger wall thickness but maintain a similar wall profile. It may be possible to correct the cortical artefact formation by including the inner wall geometry of the medullary cavity in the face-vertex surface meshes when training the statistical GP for the SSM and ASM. This would then establish a clear boundary when warping the voxel data. Nevertheless, prior knowledge of only the outer surface geometry yielded a good density distribution profile against the outer surface.

Additionally, in Figures 5.3 and 5.5 we see the density estimate improve slightly as the cut segment increases. From Figures 5.11 and 5.13 we also see that the cortical bone artefacts were less prominent in the proximal and distal ends of the tibial estimations as the cut segment increases. Thus, by increasing the medullary cavity with each cut, the voxel data increases and the error due to the cortical bone artefacts average out. In Figures 5.2 to 5.5 we also find, for some cuts, that the maximum value falls below the standard deviation. This is due to several observations and outliers that fall below the calculated average. Thus, even though the standard deviation of the observations shows a statistical possibility for the cut size, no observations past the maximum were recorded.

This chapter partially simulated segmental bone loss by removing diaphyseal segments and leaving the proximal and distal ends in their original positions. Segmental bone loss could be the result of bacterial infection, pending resection, revision surgery, or high energy trauma (Wang *et al.*, 2016; Wang and Yeung, 2017). With bacterial infection and pending resection, there is a good chance that the remaining healthy bone anatomy will be available in the original position. However, with revision surgery and high energy trauma, it may be necessary to realign the remaining healthy anatomy so that the shape and density estimate will be accurate. The alignment and re-estimation of simulated segmental bone loss due to high energy trauma will be investigated in Chapter 6.

Chapter 6

Aligning Bone Segments

This chapter investigates a basic method for aligning the proximal and distal ends of long bones that have been misaligned due to segmental bone loss. As in the previous chapter, reconstructing bone anatomy requires proper alignment between the remaining proximal and distal ends. As already stated, this is not always the case when segmental bone loss results from high energy trauma or revision surgery. A basic method for aligning the epiphyseal ends was developed and tested within this study to investigate shortcomings and consider future work.

6.1 Simulated Misaligned Bone Loss

The proximal and distal ends were programmatically misaligned to simulate segmental bone loss as a result of high energy trauma or revision surgery. Using the same simulated segmental bone loss procedure set out in Section 5.1, the distal end of the image was translated and rotated in a random direction with a distance and angle ranging from 0 to 50 mm and 0 to 180 degrees, respectively. A simple transformation matrix was constructed for translation by randomly assigning values to the (x, y, z) coordinates from the given range. The same was done for the (R_x, R_y, R_z) rotation axes around the zero coordinate point of the samples 3D space. Combined, a rigid transformation matrix was formed and applied to each voxel coordinate on the distal image. This random transformation was calculated for each sample after simulating segmental bone loss.

6.2 Alignment Method

To align the proximal and distal ends, knowledge of the bone structure before pathology is required. This knowledge is gained through the statistical data captured within the SSMs and ASMs constructed throughout this study. To align the proximal and distal images, full segmentations of both ends are
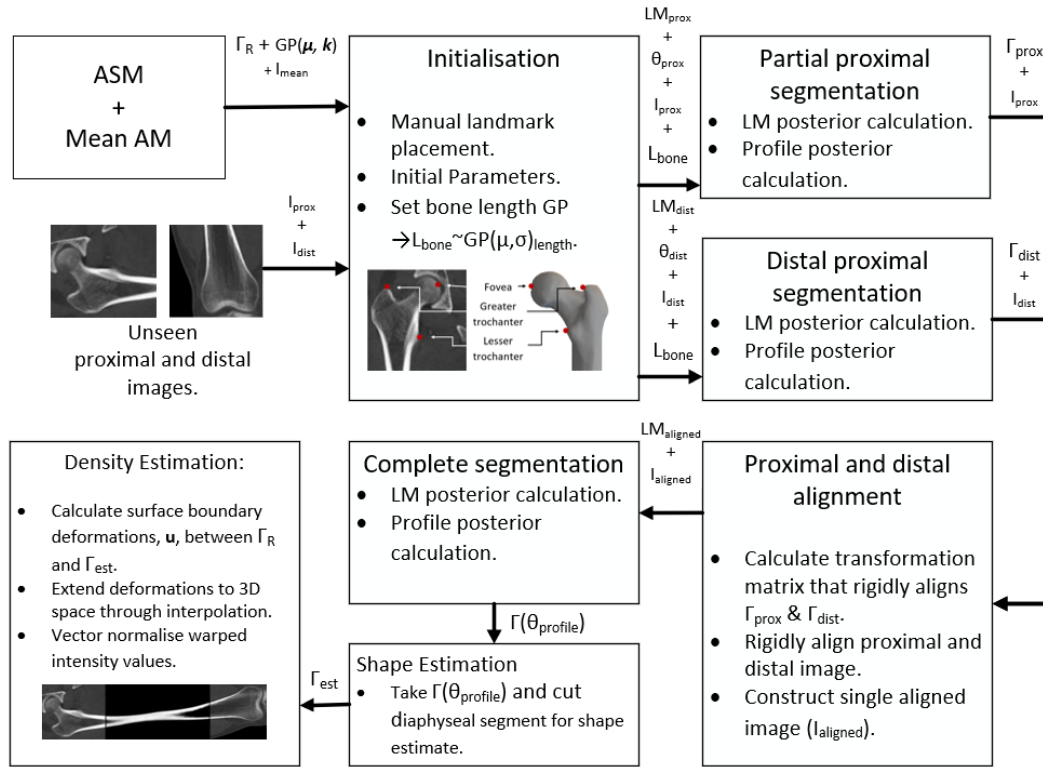
Figure 6.1: Shape en density estimation pipeline for misaligned proximal and distal ends.

calculated by fitting the ASM to each partial misaligned image. The segmentations can then be rigidly aligned and a single partial image constructed. A final segmentation of both ends can then be calculated, and the partial image reconstructed as in Section 5.2. Figure 6.1 gives an overview of the alignment and reconstruction procedure used.

First, the input data is constructed by simulating segmental bone loss and misaligning the distal end of each sample. The proximal and distal images, I_{prox} and I_{dist}, are then used in combination with the ASM to derive the initial parameters for each end. Additionally, the user is required to input an assumed total length for the sample being tested. With the ends misaligned, an added parameter L_{bone}, is introduced along with θ. When fitting each end, L_{bone} aids in ensuring the segmentations are not too long or short, as this would compromise proximal and distal end alignment. In practice, the assumed total length can be derived from contralateral anatomy, the patient's total length or surgical expertise. For this study, the assumed length was measured from the CT image before simulating misalignment. However, to accommodate error generally made in practice when aligning and reconstructing segmental bone, L_{bone} did not just consist of a single length but a prior distribution where the user given length is μ_{length} and the error is θ_{lenth}, $L_{bone} \sim \mathcal{N}(\mu_{length}, \theta_{lenth})$. In this study, a 5 mm error was assumed.

To segment the proximal and distal image, the MH algorithm was used as in Subsection 4.2.2. First, the LM posterior was calculated, then the intensity posterior using only the landmarks and intensity profiles over the end being fitted. During the fitting procedure, the likelihood of both $p(\theta)$ and $p(L_{bone})$ was evaluated when determining the likelihood of the sampled distribution (see Subsection 4.2.2 and Figure 4.5). At the end of the fitting procedure, full-length segmentations for the proximal, Γ_{prox}, and distal, Γ_{dist}, ends were produced.

By using the corresponding points on Γ_{prox} and Γ_{dist} a transformation matrix that minimises the root mean square between the vertices could be calculated and used to align I_{prox} and I_{dist}. A single image, $I_{aligned}$, was then constructed by combining I_{prox} and the aligned I_{dist}. Using L_{bone} during proximal and distal segmentation ensures the ends are far enough apart to accurately represent the segmental bone loss segment cut when constructing the partial image. Additionally, the landmarks manually placed during initialisation are aligned using the transformation matrix and combined to a corresponding LM set. Using $I_{aligned}$ a final segmentation, $\Gamma(\theta_{profile})$, is made and the missing shape and density, Γ_{est} and I_{est}, is estimated and reconstructed as in Section 5.2.

6.3 Results

Again, the same leave-one-out test process and metrics were used to measure the estimated bone shape and density distribution accuracy after alignment. The estimation accuracy was measured against the ground truth, i.e. the segments cut from the target shapes and images before misalignment were stored and compared to the shape and density estimates. The test results are shown below through a series of graphs and figures for each population group.

Figures 6.2.a and 6.2.b show the average and Hausdorff errors obtained for the male femur population. As the cut length increased, the accuracy of the shape estimation decreased, with the average distance error ranging from 1.27 ($\pm$ 0.52) mm to 3.04 ($\pm$ 0.70) mm for the 5 % and 70 % resections, respectively. Similarly, the average Hausdorff error ranged from 3.14 ($\pm$ 1.06) mm to 10.10 ($\pm$ 2.29) mm. Figure 6.2.c shows that the average R^2 was relatively constant regardless of the size of the resection, with an average of approximately 87 %. The average intensity error, Figure 6.2.d, was also relatively constant, with an average error of approximately 190 HU.

Figures 6.3.a and 6.3.b show the average and Hausdorff errors obtained for the male tibia population. As the cut length increased, the accuracy of the shape estimation decreased, with the average distance error ranging from 1.55 ($\pm$ 0.87) mm to 2.91 ($\pm$ 0.94) mm for the 5 % and 70 % resections, respectively. Similarly, the average Hausdorff error ranged from 3.84 ($\pm$ 1.61) mm to 12.14 ($\pm$ 5.49) mm. Figure 6.3.c shows a gradual increase in R^2, ranging from 86 ($\pm$ 7) % for resections of 5 % of the total length to 90 ($\pm$ 3) % for
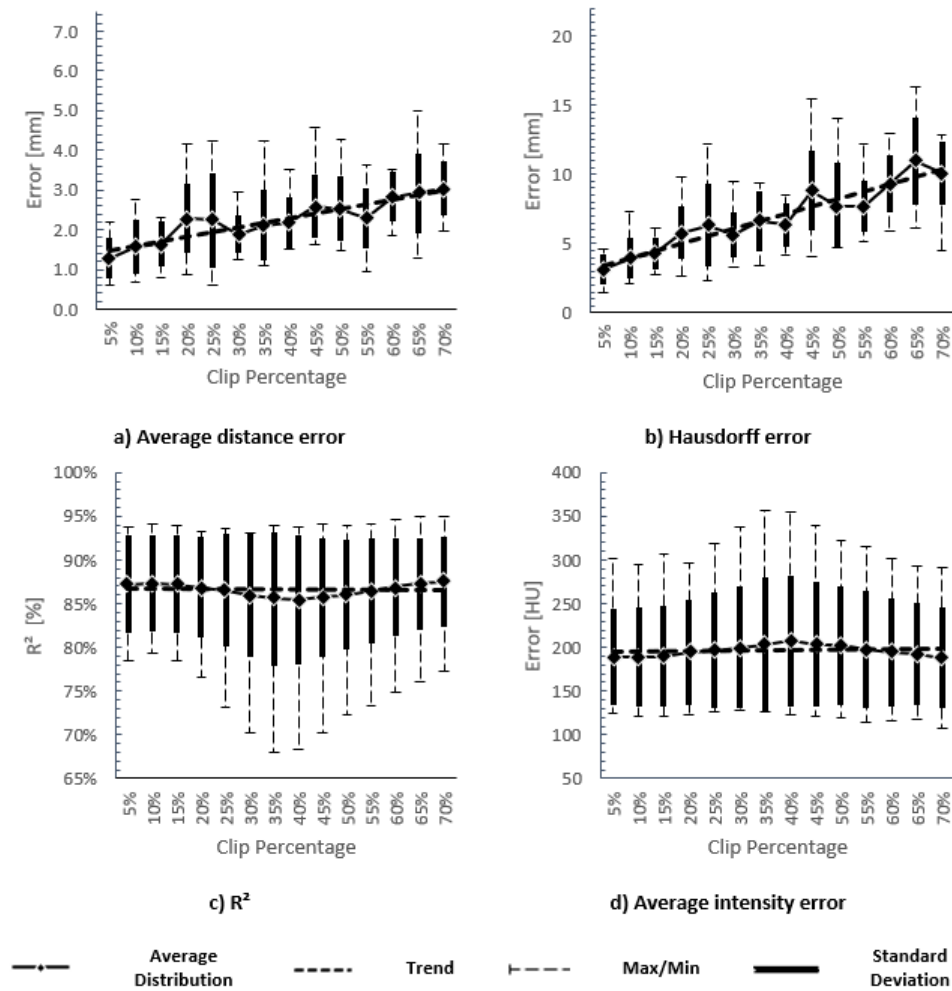
Figure 6.2: The combined statistical analysis of the test results for segmental bone repair of misaligned male femur samples.

resections of 70 % of the total length. The average intensity error, Figure 6.3.d, also showed an increase in accuracy, ranging from 227 ($\pm$ 68) HU for resections of 5 % of the total length to 180 ($\pm$ 41.39) HU for resections of 70 % of the total length.

Figures 6.4.a and 6.4.b show the average and Hausdorff errors obtained for the female femur population. As the cut length increased, the accuracy of the shape estimation decreased, with the average distance error ranging from 1.66 ($\pm$ 0.87) mm to 2.87 ($\pm$ 1.50) mm for the 5 % and 70 % resections, respectively. Similarly, the average Hausdorff error ranged from 3.42 ($\pm$ 2.18) mm to 11.02 ($\pm$ 3.83) mm. Figure 6.4.c shows that the average R^2 was relatively constant regardless of the size of the resection, with an average of approximately 89 ($\pm$ 6) %. The average intensity error, Figure 6.4.d, was also relatively constant, with an average error of approximately 160 HU.

Figures 6.5.a and 6.5.b show the average and Hausdorff errors obtained
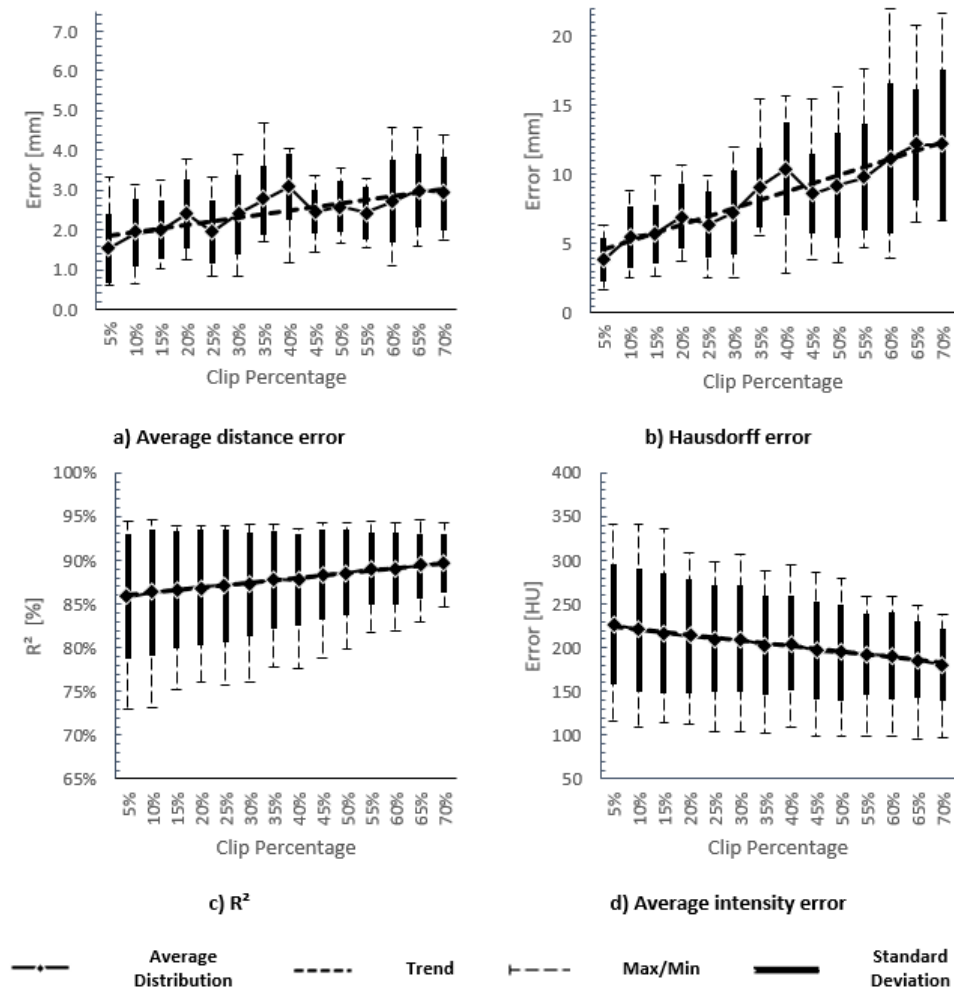
Figure 6.3: The combined statistical analysis of the test results for segmental bone repair of misaligned male tibia samples.

for the female tibia population. As the cut length increased, the accuracy of the shape estimation slightly decreased, with the average distance error ranging from 1.35 (± 0.74) mm to 2.41 (± 0.85) mm for the 5 % and 70 % resections, respectively. Similarly, the average Hausdorff error ranged from 3.78 (± 1.66) mm to 9.12 (± 3.32) mm. Figure 6.5.c shows that there was a gradual increase in R^2, ranging from 85 (± 5) % for resections of 5 % of the total length to 90 (± 3) % for resections of 70 % of the total length. The average intensity error, Figure 5.5.d, also showed an increase in accuracy, ranging from 204 (± 25) HU for resections of 5 % of the total length to 158 (± 20) HU for resections of 70 % of the total length.

Figures 6.6 to 6.9 show the shape of the estimated diaphyseal segments for one misaligned left-out sample of each population group. Visual inspection shows that as the size of the resected portion increases, the accuracy of the estimate deteriorates as well as the alignment quality. The added error due
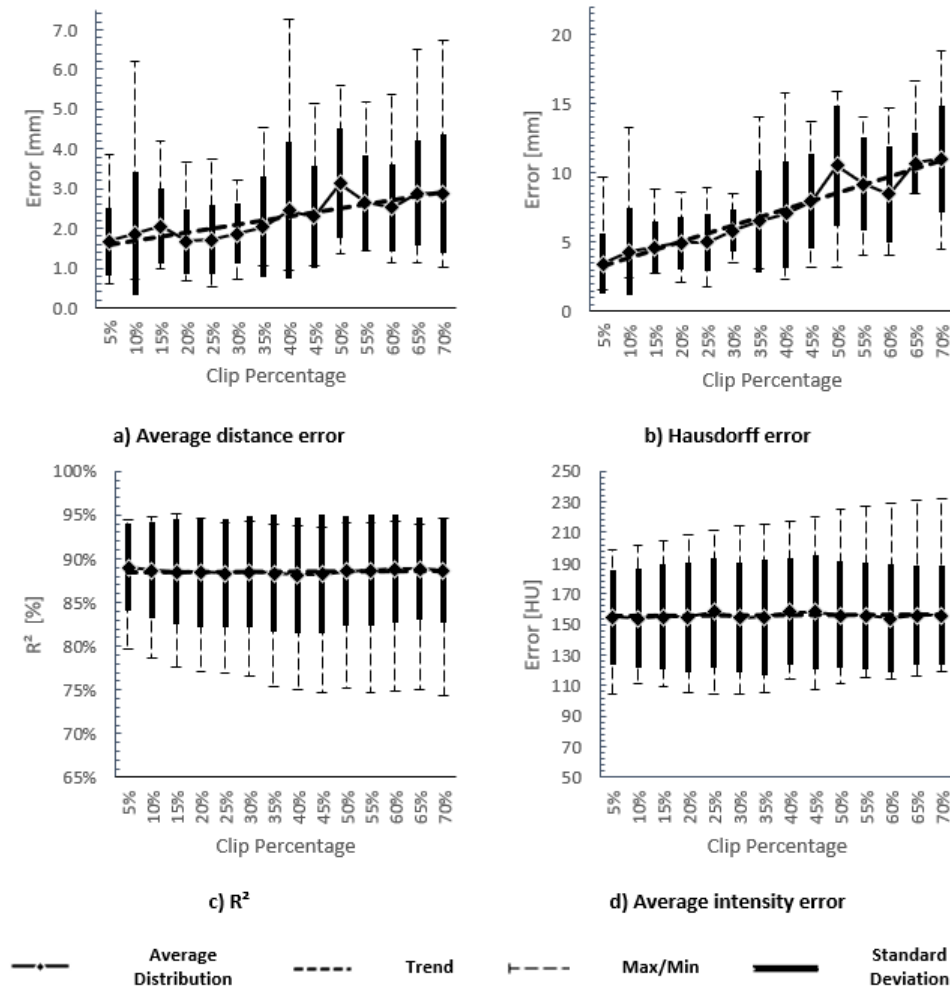
Figure 6.4: The combined statistical analysis of the test results for segmental bone repair of misaligned female femur samples.

to alignment is clear from the distal ends of the 70 % resections. Still, shape estimations below the 35 % resection remain accurate for the proximal and centre segments, where the accuracy of distal estimation varies throughout testing. Due to the similarities in the density estimation, visual examples are not given in this chapter. Figures 5.10 to 5.13 are very similar to the estimations made during misalignment testing.

6.4 Discussion

When comparing Figures 6.2 to 6.5 to the results from Section 5.3 we find that the shape estimates follow the same trend but with error almost two or three times the size. However, the density estimation results are similar for both the aligned and misaligned reconstruction. When measuring the density
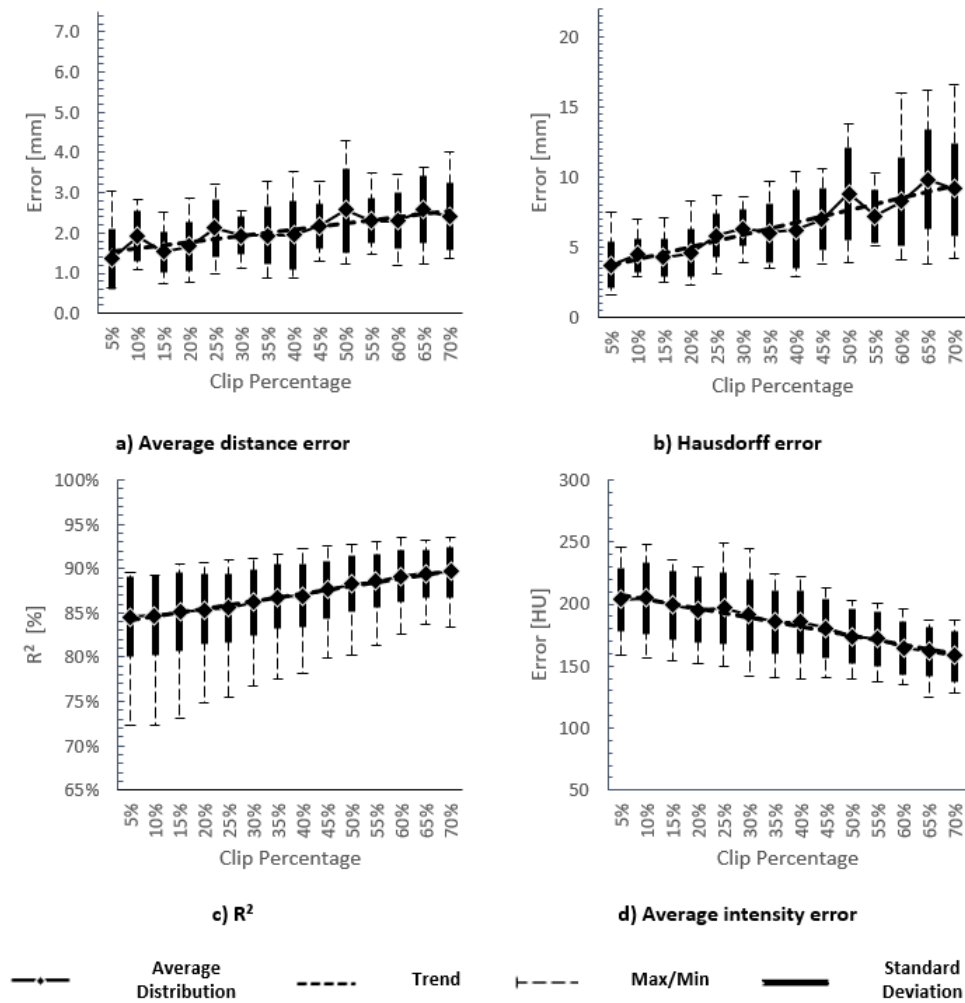
Figure 6.5: The combined statistical analysis of the test results for segmental bone repair of misaligned female tibia.

estimation accuracy, the voxel data is shape normalised to ensure the density estimation accuracy is calculated correctly. Thus, with the intensity data of each voxel of the mean AM remaining constant, similar results are to be expected. Therefore, the performance of the alignment method developed are best evaluated by comparing the shape estimations from the previous chapter to the results found in this chapter.

For resections below 35 % of the samples' total length, enough anatomy is available to calculate proper full segmentations from only the proximal or distal end. That is to say that enough of the diaphyseal shaft is available to accurately estimate the opposing end's distance, angle and size from the prior knowledge captured in the SSM. However, for resections 35 % and above, the available anatomy renders statistically viable full segmentations, but with angle and size estimations not similar to the test samples' opposing ends. When
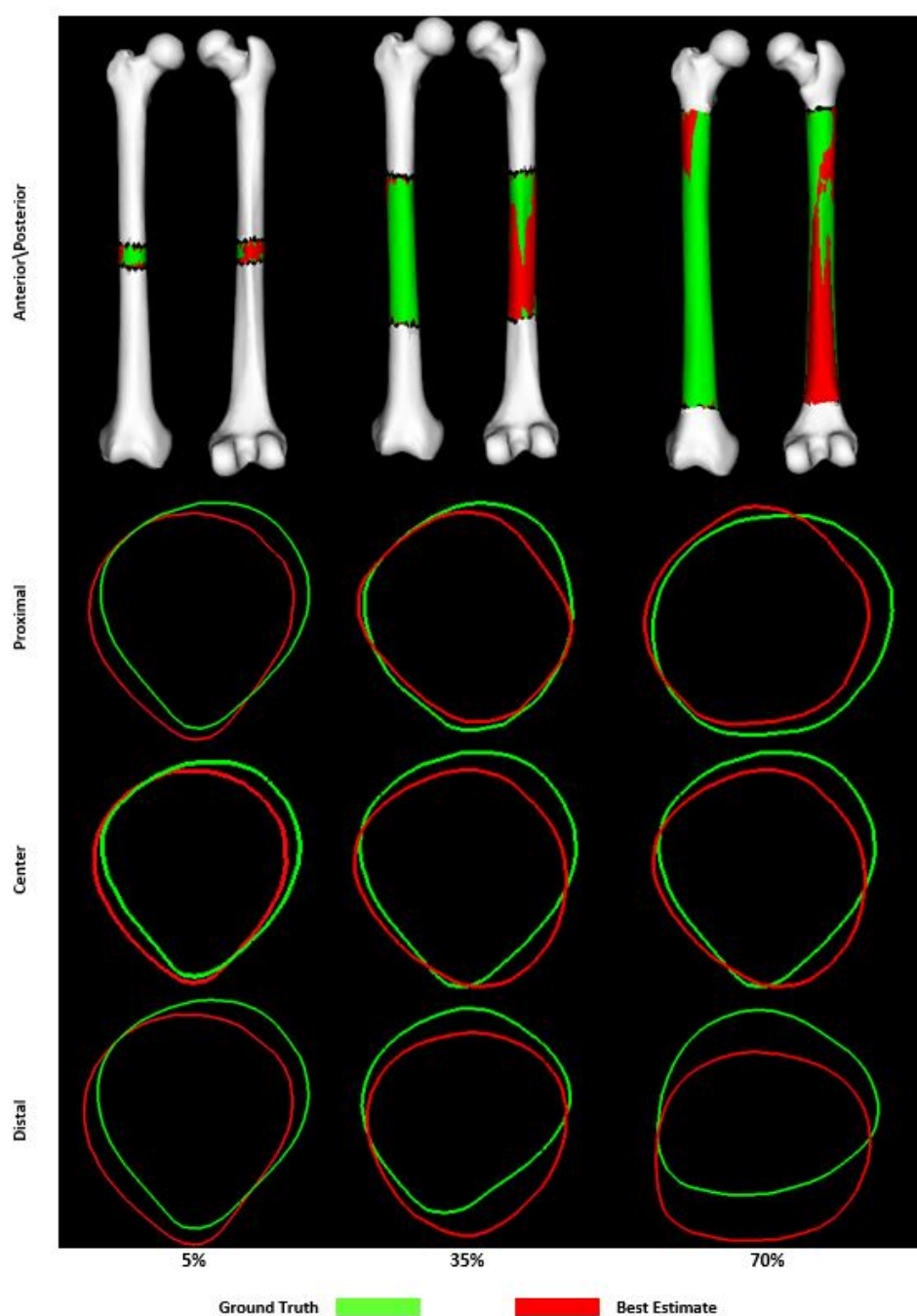
Figure 6.6: The shape estimation for misaligned sample NS-401 of the male femur population. The estimations had an average distance error and a Hausdorff error of 1.11 mm and 2.75 mm for the 5 % resection, 3.01 mm and 9.32 mm for the 35 % resection, and 3.56 mm and 7.90 mm for the 70 % resection (recorded length 458 mm).
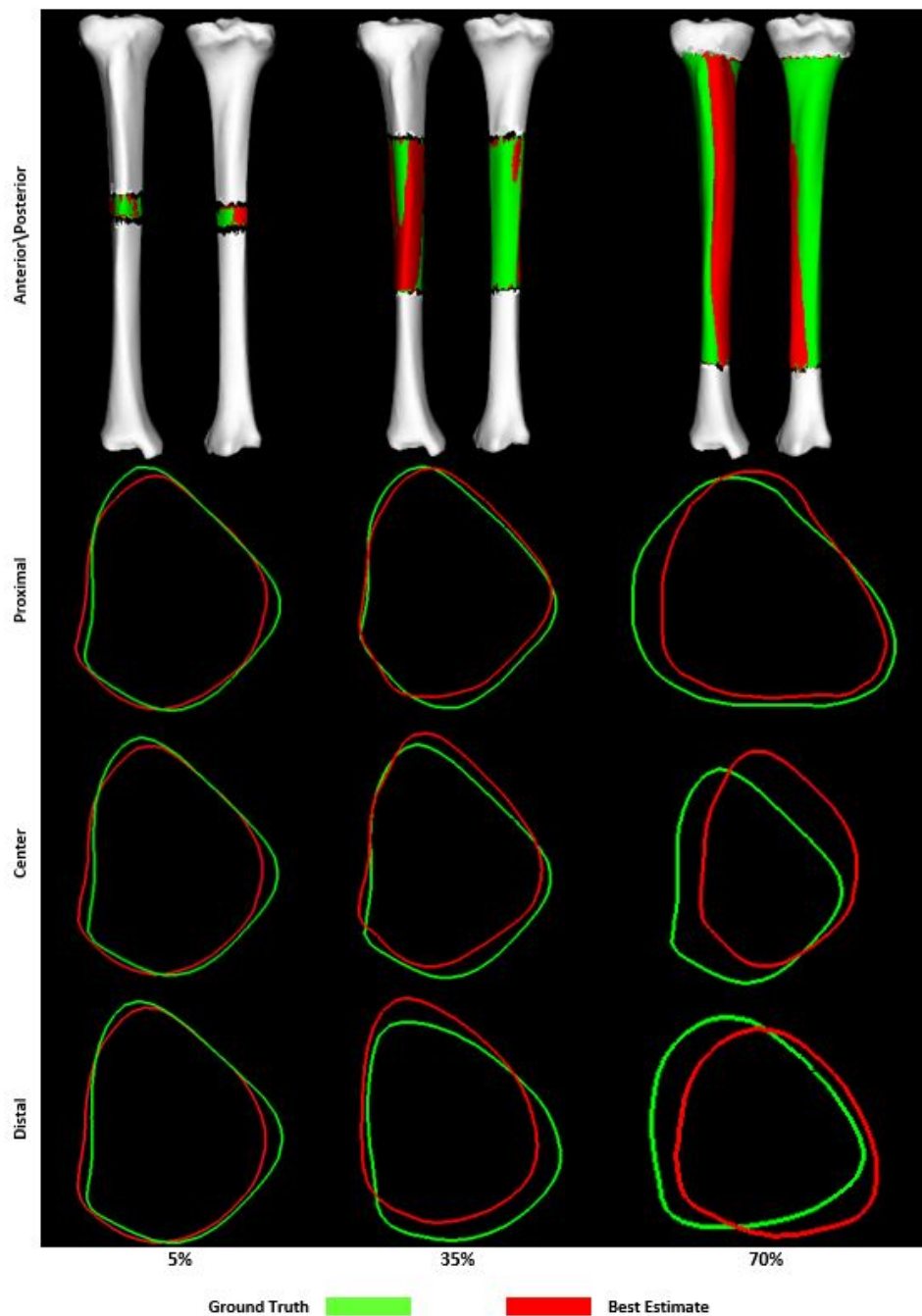
Figure 6.7: The shape estimation for misaligned sample NS-401 of the male tibia population. The estimations had an average distance error and a Hausdorff error of 0.94 mm and 2.78 mm for the 5 % resection, 2.32 mm and 5.61 mm for the 35 % resection, and 2.43 mm and 6.83 mm for the 70 % resection (recorded length 398 mm).
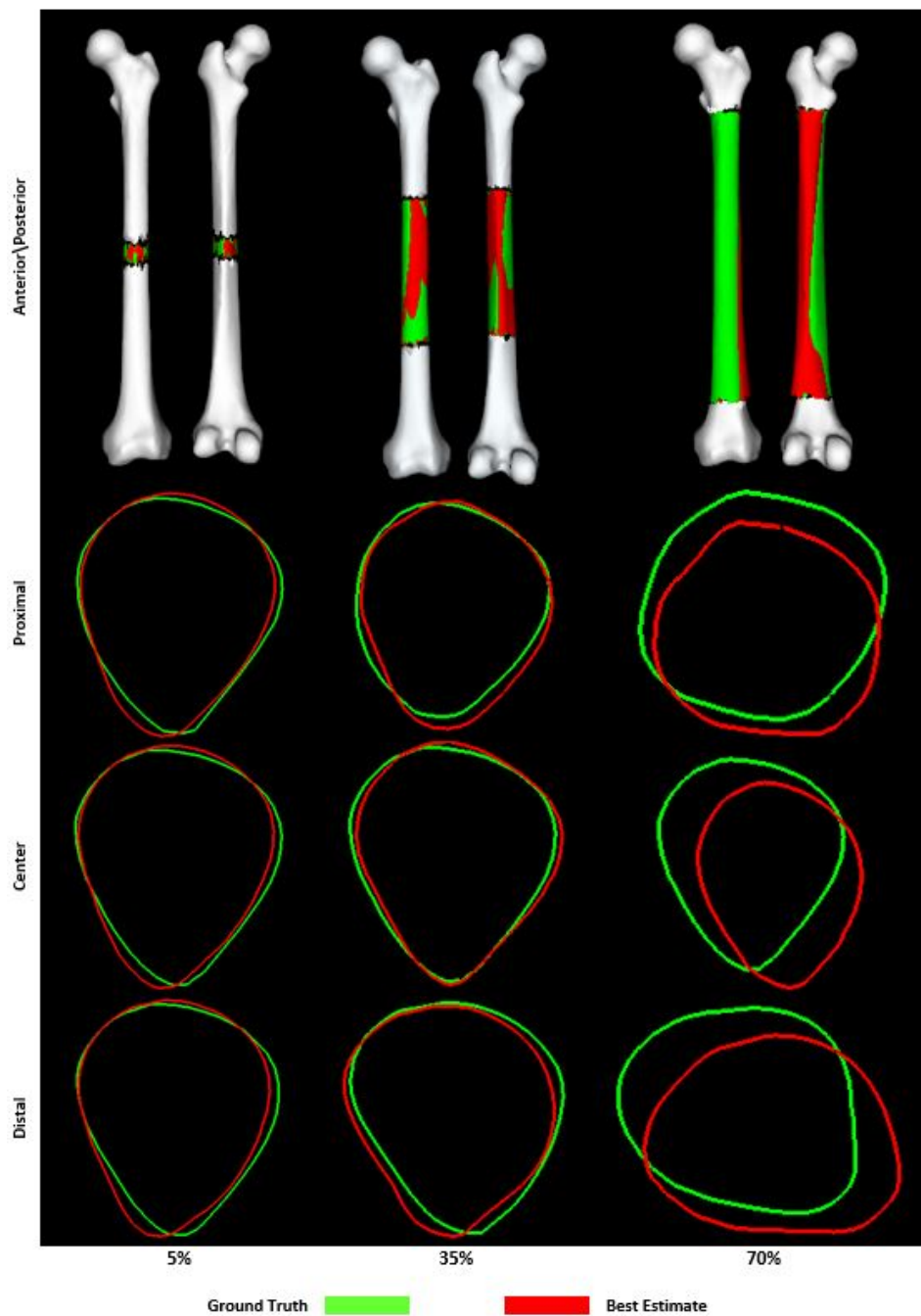
Figure 6.8: The shape estimation for misaligned sample SA-303 of the female femur population. The estimations had an average distance error and a Hausdorff error of 0.61 mm and 1.57 mm for the 5 % resection, 1.48 mm and 4.12 mm for the 35 % resection, and 3.63 mm and 14.20 mm for the 70 % resection (recorded length 398 mm).
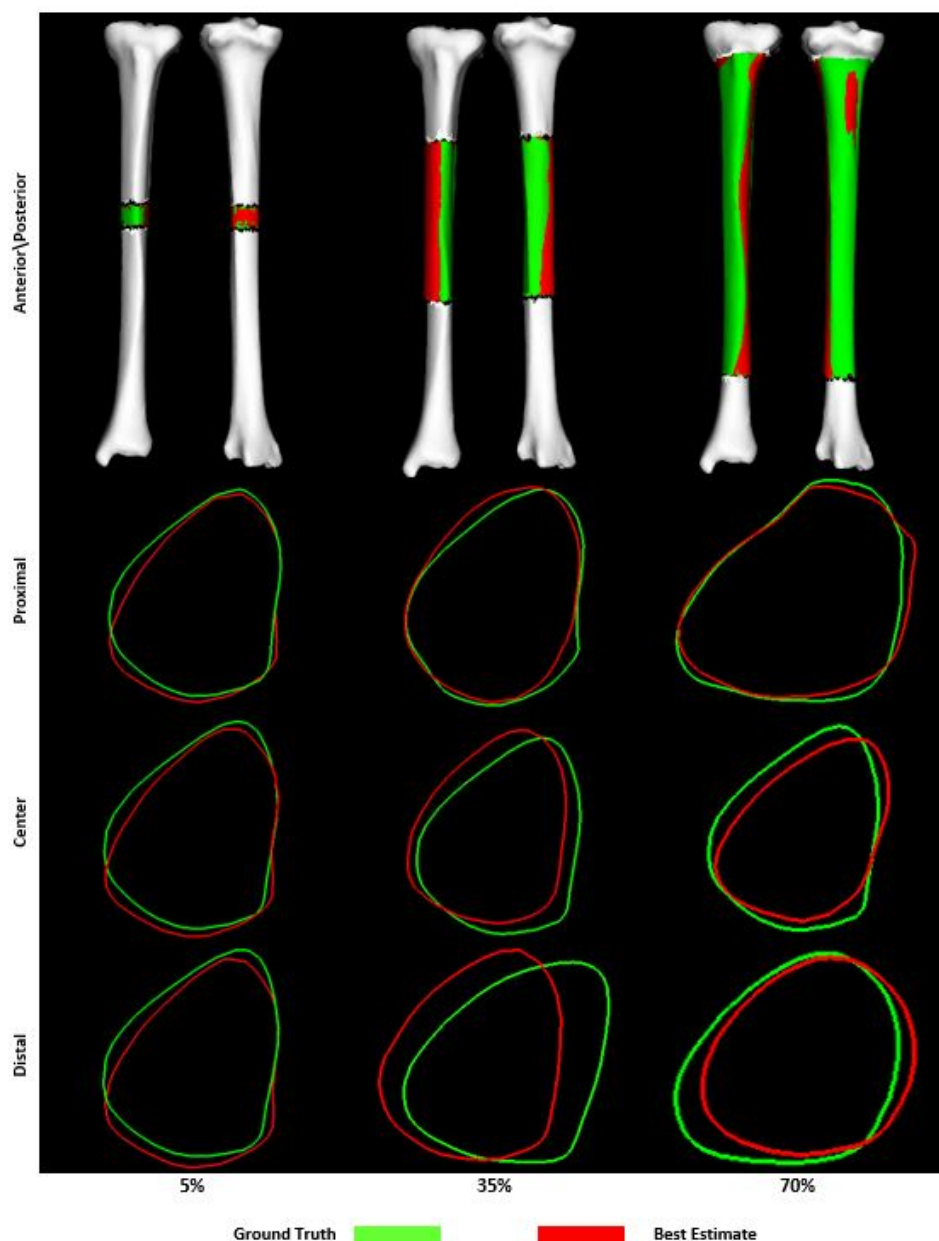
Figure 6.9: The shape estimation for misaligned sample NB-320 of the female tibia population. The estimations had an average distance error and a Hausdorff error of 0.88 mm and 3.00 mm for the 5 % resection, 1.29 mm and 3.97 mm for the 35 % resection, and 2.00 mm and 9.12 mm for the 70 % resection (recorded length 347 mm).

comparing the shape estimations to the results from the previous chapter, we find that the estimations for resections 35 % and below still render average distance errors within the 2 to 4 mm tolerance limit generally seen in long bone stock implants (Abram *et al.*, 2014; KR *et al.*, 2020). When resections are larger than 35 % the Hausdorff error greatly increases, this points to local regions with significant surface estimation errors, mostly found at the distal end.

To correct this, possible future work could include an alteration to the pipeline suggested in Figure 6.1. By segmenting the proximal and distal end individually, data on the opposite end's angle and size are not considered. By solving both segmentations in a single MH fitting operation, the full data range can be utilised. An MH iteration would be able to adjust the translation, t, and rotation, r, of both the proximal and distal image while sampling from the model's shape distribution, α. Thus, the shape boundary data contained within I_{prox} and I_{dist} can be utilised to align and derive the missing surface geometry within a single segmentation, giving a more complete picture of the available healthy bone anatomy. This would reduce three separate segmentations to one, saving on time, and most likely render more accurate estimations as there are more data points to fit the ASM to. With the MH functions developed during this study (Appendix A) minor alterations can be made to the function's input and pipeline structure to make this possible in future studies.

The alignment algorithm used during this chapter was able to correct misaligned segmental bone to some degree. Although additional work is required to improve the method's results, it proved to be a step in the right direction for developing an alignment method for automating not only reconstruction but alignment using the prior knowledge captured in the constructed SSMs and ASMs.

Chapter 7

Conclusions

7.1 Introduction

Segmental bone loss can be treated with PSIs, which offer advantages such as reduced recovery time, improved patient outcomes, reduced surgical time, lower long-term costs, and no secondary morbidity (Siegmund *et al.*, 2019; Tetsworth *et al.*, 2017; Wang *et al.*, 2016). PSI design, however, requires extensive specialist knowledge and procedures, which essentially need to be repeated for each new case, making PSI design time-consuming and cost-inefficient when compared to current conventional treatment options such as autografting (Haglin *et al.*, 2016; Tetsworth *et al.*, 2017). Of these time-consuming activities, automating image segmentation, surface reconstruction, and density estimation of missing bone anatomy would reduce the overall cost, time and specialist knowledge required for PSI design (Zadpoor and Weinans, 2015; Tetsworth *et al.*, 2017; Heimann and Meinzer, 2009; Sarkalkan *et al.*, 2014; Mauler *et al.*, 2017).

The study aimed to develop a method that automatically performs the segmentation and estimation of healthy bone shape and density distribution from partial inputs for segmental bone repair. To achieve this aim, the project objectives were as follows:

1. Construct and validate SSMs and mean AMs for a given bony population from CT images.

2. Fit the validated SSMs and mean AMs to a sparse dataset simulating segmental bones lose, solving the segmentation problem for the long bone structure.

3. Implement optimisation and reconstruction algorithms that map healthy shape and density distribution data to partial inputs.

7.2 Main Findings and Contributions

To make PSIs financially and practically more feasible, a novel method for estimating healthy bone shape and density distribution using an ASM and mean AM was presented in this study. The technique combined ASMs with mean AMs to reconstruct simulated segmental bone lose pathologies within the diaphyseal area of the femoral and tibial long bones for a male and a female South African population. The model combination enabled relatively fast geometrical and reasonable density estimates (60 to 90 seconds and 50 to 60 seconds, respectively) for segmental bone repair, automating PSI design procedures such as segmentation and surface reconstruction.

During model validation, leave-one-out testing revealed that the SSMs could represent the statistical population with average distance and Hausdorff errors below 0.65 mm and 4.10 mm, respectively. When validating the ASM, the maximum average length and Hausdorff errors increased to 1.47 mm and 7.41 mm. For the mean AMs, density distributions were approximated within $\pm$ 90 % of the intensity of the test samples, with average intensity errors around $\pm$ 124 HU.

The combined ASMs and mean AMs used in this study could reconstruct and estimate the shape and density distribution of resected femoral or tibial diaphyseal segments with an average distance error ranging from 0.58 to 1.64 mm and a R^2 above 84 % of the intensity of the original target images for the 5 to 70 % resections. When aligning and reconstructing misaligned proximal and distal segments, the accuracy of surface reconstruction declined but remained constant for density estimation. The shape and density distribution of resected and misaligned femoral or tibial diaphyseal segments were re-estimated with an average distance error ranging from 1.35 to 3.04 mm and a R^2 around 84 % of the intensity of the original target images for the 5 to 70 % resections.

Through literature and an interview with Craniotech (PTY) Ltd. (see C), acceptable design and reconstruction methods and tolerances were investigated. Traditionally, surface reconstruction uses healthy contralateral bone anatomy to infer the missing surface geometry (Mauler *et al.*, 2017; Letta *et al.*, 2014). However, healthy contralateral bone anatomy is not always available due to high energy trauma or pre-existing conditions (Mauler *et al.*, 2017; Saleh *et al.*, 1999). In cases such as these, missing geometry is estimated through SSMs or specialist knowledge (Schepers *et al.*, 2015; Zadpoor and Weinans, 2015; Heimann and Meinzer, 2009). The literature revealed an acceptable 2 to 4 mm over or under-hang error when working with hip and knee stock implants (Abram *et al.*, 2014; KR *et al.*, 2020).

CranioTech (PTY) Ltd. (2021) stated that healthy contralateral anatomy, if available, is the primary resource used when reconstructing segmented bone. When designing PSIs, the main concern is the anatomical constraints posed by each unique patient and not a specific set of design tolerances. Though

there are curtain *rules of thumb* when reconstructing surface geometry, design tolerances are generally established through patient-specific constraints and surgeon expertise. During the interview, however, it became clear that surgeons are usually comfortable with surface reconstruction errors between 5 to 10 mm, and edge alignment errors between 2 to 3 mm. When considering both the lost and available bone in the PSI design, the main area of concern is viable bone stock for fixation screws, not the implant strength. A PSI is generally over-designed to handle a much greater load than it will ever experience within the human body. For the femoral truss cages designed by CranioTech (PTY) Ltd. (2021), the truss diameter generally used is approximately 1.5 mm and is capable of supporting a 2 ton load.

When looking at the shape model validation and surface reconstruction results from Chapters 4 and 5 we find that the surface reconstruction and estimation methods used within this study are viable for automating PSI design procedures such as surface reconstruction and segmentation. Though density estimations seem reasonable, additional studies will have to be done in future work to relate the HU density estimations to bone strength for the South African population. This relation would allow for FEM analysis, which would reveal if the density estimations made in this study are capable of the same load-carrying capacity as the original bone sample.

The results from Chapter 6 revealed, however, that the alignment method used still requires some additional work. The edge alignment for the 35 % to 70 % resections, specifically on the distal end, was much greater than the 2 to 4 mm (over or under-hang) or 2 to 3 mm error from literature or industry.

In summary, the author implemented and combined existing and well-established methods to meet the aim and objectives of this study. Existing functionality was acquired through Scalismo, a Scalable Image Analysis and Shape Modelling library. Table 7.1 contains a list of existing functionality compared to the list of contributions by the author.

7.3 Recommendations

As the author's understanding of the material and methods grew throughout the study, certain assumptions, implementations and methods were used that could be improved upon if the study was repeated. This section aims to highlight these recommended improvements for future studies. These recommendations have already been discussed in previous chapters in greater detail. Thus, they are summarised within this section for completeness.

When constructing the SSM and ASM, the face vertex surface mesh segmentations used for training were reduced to 5 000 vertex points. Using *3-Matic* (Materialise NV, Belgium) software, the segmentations from CranioTech, of ±75000 vertices each, were reduced to 5 000 vertex points with a geometrical error lower than 0.1 mm. This produced SSMs and ASMs that were computationally light and easy to work with during debugging but still

Table 7.1: List of contributions

Existing Functionality	Author Contributions
SSM and ASM construction functionality.	The processing of an extensive data set for viable samples.
Vertex mesh software toolset.	The drafting of a segmentation protocol that could be used to professionally segment all four data sets.
Medical imaging software toolset.	Establishing correspondence within all four data sets.
Statistical modelling user interface.	Functionality to construct mean AMs.
MCMC and parametric optimisation toolkit.	Validated statistical models for all 4 data sets.
Statistcal modelling data structures and classes.	A tailored optimisation function to fit the validated models and estimate the missing bone shape and density distribution.
	Promising results showing how the prior capture within the statistical models can be used as an alternative to the specialist knowledge required for conventional reconstruction procedures.

gave good results during testing. The author would recommend increasing the reduction scale 10 000 points. With software debugging complete, this would produce higher resolution models, mitigating the possible effects of *pixelation* within the automated segmentation procedures.

The proximal and distal edge alignment in Chapter 5 could be improved by altering the intensity profile construction of the ASMs. The ASMs' main intensity profile points were sampled randomly over the SSMs' surface, using a uniformly distributed algorithm during construction. Assigning the profiles in a radial-grid configuration over the length of the shaft could better fit the intensity profiles to the remaining edge of the healthy proximal and distal end, resulting in better edge alignment.

Artefacts of cortical bone on the medullary cavity wall were found when warping the Mean AM for density estimation in Chapter 5. These artefacts were a product of the 3D interpolation and warping procedures used and not knowing the shape boundary of the medullary cavity. It may be possible to correct the cortical artefact formation by including the inner wall geometry

of the medullary cavity in the face-vertex surface meshes when training the statistical GP for the SSM and ASM. The mask and segmentation data used to segment the training samples were stored within their separate Mimics files. Thus, adding the medullar cavity surface would require recalculating the mesh surface from the available mask data without wrapping the outer bone structure. The added medullar cavity surface would then establish a clear boundary when warping the voxel data. Additionally, the interpolation method used to shape-normalise or warp the voxel data of the training images and the mean AM was linear interpolation. The author would recommend changing the interpolation method to a B-spline or Thin Plate Spline interpolation method, which is better suited for image warping (Cootes *et al.*, 2004).

When aligning proximal and distal ends for surface reconstruction, alterations to the current alignment method in Chapter 6 are required. The full data range can be utilised by solving both the proximal and distal segmentations in a single MH fitting operation. An MH iteration would be able to adjust the translation, t, and rotation, r, of both the proximal and distal image while sampling from the model's shape distribution, α. Thus, the shape boundary data contained within I_{prox} and I_{dist} can be utilised to align and derive the missing surface geometry within a single segmentation, giving a more comprehensive picture of the available healthy bone anatomy. This would most likely produce more accurate estimations as there are more data points to fit the ASM profiles.

Lastly, within the study, the vertex and profile data of the ASMs and the corresponding voxel data of the mean AM were stored separately. Tetrahedral meshes, already operational within the Scalismo library (University of Basel, 2021), can combine both data sets within a single model or data structure. Using a combined tetrahedral model would most likely decrease computational power and benefit shape normalising the mean AM as the ASM is fitted to image data.

7.4   Future Work

With segmentation, surface reconstruction and density estimation automated, the next steps in making PSIs more feasible and financially more cost-effective would be to automate lattice or truss generation. The density estimate can be directly linked to bone mineral density and thus Young's modulus and ultimate bone strength (Duchemin *et al.*, 2008). The density estimate could then be used to design a functionally graded lattice structure, closely resembling the mechanical properties of healthy bone structures. This would mitigate stress shielding, stimulate new bone growth, and allow nutrient flow at the implant site (Wang *et al.*, 2016; Tetsworth *et al.*, 2017; Long and Rack, 1998; Sumner and Galante, 1992; Zhang *et al.*, 2019).

Future studies would then need to first calculate the relationship between the intensity values in the density estimate and the bone mineral density for

the statistical population in question (Duchemin *et al.*, 2008). Once complete, software will have to be developed that takes the density and shape estimates of this study as input to derive a functionally graded lattice structure as output. Additional considerations will also have to be made for the fixation method used. From literature, we find that surgeons who use lattices or truss cages for segmental bone repair generally prefer an IM-nail for internal fixation (Hollister, 2009; Ghavidelnia *et al.*, 2020). The IM-nail keeps the truss cage stable and in position while the bone graft calcifies. It also allows for early weight-bearing, further improving new bone growth and so patient morbidity.

7.5 Conclusion

SSMs, ASMs, and mean AMs were constructed and validated throughout this study for four bony populations (male femur, male tibia, female femur and female tibia). The validated models and automated segmentation algorithms were used to successfully reconstruct missing bone anatomy that simulated segmental bone loss. These results indicate that the methods proposed in this study successfully estimated the original, healthy condition of long bones after segmental bone loss, suggesting that it is viable for automating PSI design procedures such as reconstruction and segmentation. This achieved the aims and objectives set out at the beginning of this study.

Appendices

Appendix A

Supplementary Notes on Software Structure and Development

This appendix contains notes on the model construction, automated segmentation, estimation, reconstruction and alignment software developed and implemented. The code will be submitted to the main supervisor within an electronic record of all the work done throughout this study. This appendix will briefly describe the packages and the software structure and give a record of the user-supplied inputs. The code itself is well commented throughout to aid in future work.

All programming was done in Scala, a strong statically typed general-purpose programming language, using Scalismo library, a scalable image analysis and shape modelling library (University of Basel, 2021). Within Scala, *Objects* were defined within *Packages*. Each *Object* contained code, functions and data structures for each main activity in Chapters 4 to 6. Table A.1 will illustrate the package structure and briefly describe each of the developed objects. Figure A.1 will illustrate the sequence of object implementations required for Model construction and shape and density estimation. Table A.2 and A.4 gives a record of the user-defined inputs and the outputs of each object. Note that the inputs given are constant within each package. Finally, Table A.5 will give a brief description of the functions written in the *Modelling Function* object. Each function's input and output is indicated within the code comments.

Table A.1: Package and Object Structure

Pakage	Object	Description	Chapter
Statistical Model Generation		Contains all objects required for model construction.	4
	Landmarking Tool	Used to place and store manual landmarks on all samples.	4.2
	Referencing Tool	Used to identify the reference shape of each bony population (shape closest to mean volume).	4.2
	Reslice Tool	Used to reslice and crop sample images to 1 mm slice thickness around the bone structure in question (lowers computational intensity).	4.1
	SSAM Construction	Used to construct the GPMM and mean AM for each bony population.	4.2
	SSM Validation	Calculates metrics for validation (leave-one-out testing) of the GPMM for each bony population.	4.3
	SAM Validation	Calculates metrics for validation (leave-one-out testing) of the mean AM for each bony population.	4.3
	ASM Construction	Constructs and samples profile sequence from all samples of each bony population and combines with GPMM.	4.2
	ASM Validation	Calculates metrics for validation (leave-one-out testing) of the ASM for each bony population.	4.3
	Modelling Functions	Contains all the functions written for both the 'Statistical Model Generation' and 'Statistical Model Fitting' packages.	4 to 6
Statistical Model Fitting		Contains all objects required for simulating segmental bone loss, misaligning proximal and distal ends, and estimating the shape and density distribution.	5 to 6
	ASM Fiting Partial Images	Simulates and repairs segmental bone loss for shape and density estimation. Proximal and distal ends remain aligned. Leave-one-out testing and metric calculations are performed simultaneously.	5
	ASM Fiting Partial Misaligned Images	Simulates and repairs segmental bone loss for shape and density estimation. Proximal and distal ends are misaligned. Leave-one-out testing and metric calculations are performed simultaneously.	6

Table A.2: Model construction inputs

Input	Value	Description
Directory	data/Male/ Femur/	Directory path where training data of bony the population is stored.
ModelName	Male_Femur	Used to label models after construction.
TestingRange	40 to 50	Samples used for leave-one-out testing.
Noise	5 mm	A noise variable to account for human error when placing landmarks.
KernelScale	80 mm	Used define smooth deformation kernel for establishing correspondence (Variance scale).
KernelSigma1	100	Large deformations.
KernelSigma2	60	Medium deformations.
KernelSigma3	30	Small deformations.
RegWeight	$e^{-1}, e^{-2}, e^{-4}, e^{-6}$	Penalty weights for unlikely transformations during registration sequence (Parametric registration, regularisation).
RegIterations	20, 30, 40, 50	Max number of iterations for each registration sequence.
RegSamplePts	1000, 1000, 2000, 4000	Number of points sampled during each registration sequance.
InterDegree	1	Degree 0 is nearest neighbour, degree 1 is linear interpolation, and degree 3 scalar images with not many artefacts.
LowDomVal	0	Used when resampling images. Lowest HU value, water = 0 HU.
SAMiter	20	Max number of iterations when constructing stable mean AM.
SAMtol	1 HU	When the average difference between the current and previous mean AM is smaller than 'SAMTolerance', stability is assumed.
ProNr	250	Number of profiles for ASM.
ProSize	11	Number of points per profile.
ProSpacing	1 mm	Distance between each point.
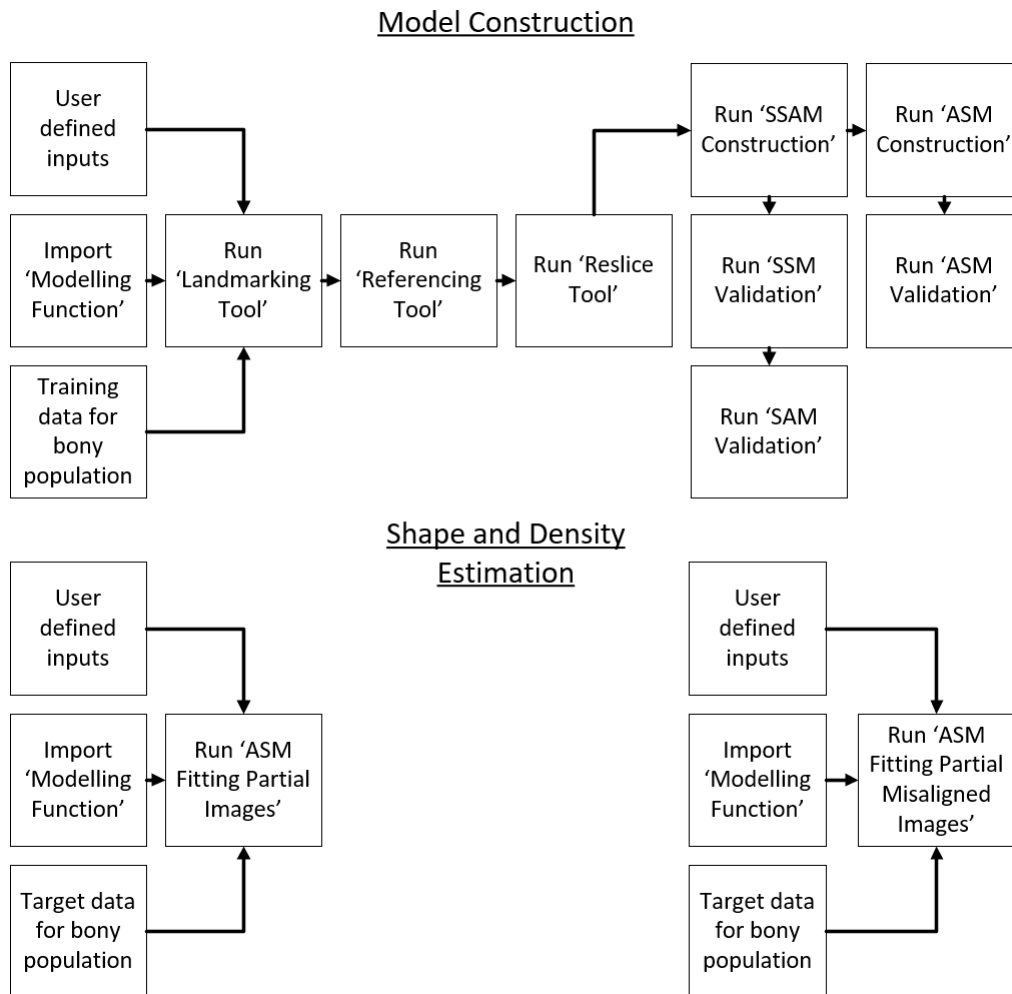
Figure A.1: Object implementation sequence

Table A.3: Shape and density estimation inputs

Input	Value	Description
Directory	data/Male/ Femur/	Directory where training data of bony the population is stored.
ModelName	Male_Femur	Used to label models after construction.
TestingRange	40 to 50	Samples used for leave-one-out testing.
Noise	5 mm	A noise variable to account for human error when placing landmarks.
Tstep	1	MH mean stepping range for translational adjustments (t).
Rstep	0.01	MH mean stepping range for rotational adjustments (r).
LSstep	0.5	MH mean stepping range for large shape adjustments (α).
SSstep	0.1	MH mean stepping range for small shape adjustments (α).
Burn-Counter	1000	MH algorithm requires a burn-in phase due to slight misalignment after initialisation. Thus, the first 1000 samples are run but not evaluated or stored.
Samp-Counter	11000	Maximuim number samples during MH fitting operation.
Clip-Range	5 to 70 % steps of 5 %	Range of segments cut when simualting segmental bone loss.
Toffset-Range	0 to 50 mm	Range for translation when misaligning distal end.
Roffset-Range	0 to 90 degrees	Range for rotation when misaligning distal end.

Table A.4: Object outputs

Model Construction		
Landmarking Tool	'.json'	Coordinates for manually placed landmarks of the proximal and distal end for each training sample.
Referencing Tool	'.txt'	The sample ID of the shape closest to the mean bone volume of the population.
Reslice Tool	'.vtk'	The resliced and croped sample images.
SSAM Construction	'.h5' '.vtk'	The full and partial SSMs and mean AMs for leave on out testing.
ASM Construction	'.h5'	The full and partial ASM for leave-one-out testing.
SSM Validation	'.txt'	Metric results.
SAM Valdiation	'.txt'	Metric results.
ASM Validation	'.txt'	Metric results.
Shape and Density Estimation		
ASM Fitting Partial Images	'.txt'	Metric results. Note that all .stl and .vtk file of each estimations would be to large to store for each poppulation. Thus, only the metric results were stored after reconstruction.
ASM Fitting Partial Misaligned Images	'.txt'	Metric results. Note that all .stl and .vtk file of each estimations would be to large to store for each poppulation. Thus, only the metric results were stored after reconstruction.

Table A.5: Function descriptions

Function	Description
FileComp()	Evaluate if loaded .stl IDs and .vtk IDs match.
ParametricReg()	Used to fit generic GPMM to training data for correspondence.
MCMCReg()	Used for automated segmentation. Contains MH algorithm and implmentation.
IsInside()	Determines if an image voxel is within the corresponding vertex mesh.
ExstractBony()	Removes all voxel data not within the corresponding vertex mesh.
IntensityNorm()	Normalises the intensity data of an image to a corresponding image.
VectorNorm()	Normalises the data of one vector to a corresponding vector through zero mean unit standard deviation.
BuildSSM()	Extracts deformation vectors from corresponding images and constructs a GPMM using Scalismo.
SmoothSSM()	Constructs a GPMM using a reference mesh and user-defined smooth deformation kernel. Used for registration and correspondence.
BuildMeanAM()	Uses shape and intensity normalised image data to construct a mean image.
UnivariateSAM()	Uses mean AM and normalised training images to calculate the univariate standard deviation of the sample population.
WarpTransform()	Uses a target and original vertex mesh to calculate a transformation matrix that would transform the original point set to the target surface.
CenterOfMesh()	Calculates the centre of volume for a vertex mesh.
ModelCoef()	Calculates the shape coefficients, (α), that best describes a shape instance corresponding to the model.
ShapeNorm()	Uses a corresponding target and original vertex mesh to calculate deformation fields and warp the original image to the target shape through linear interpolation.
Project()	Projects vertex points of an original mesh, normally towards the surface of the target mesh.
LMPlacement()	Used to sample manually placed landmarks coordinates from corresponding vertex meshes and images.
DistanceMeshLMs()	Calculates the distance between two landmarks placed on a target vertex mesh.
DistanceImagesLMs()	Calculates the distance between two landmarks placed on a target image.
AverageImage()	Calculates the deformation fields between an estimated and target vertex mesh. Then displays the error visually.
HausdorffImage()	Calculates the Hausdorff deformation field between an estimated and target vertex mesh. Then displays the error visually.

Appendix B

Supplementary Notes on Modelling and Mathematical Considerations

This appendix provides a detailed overview of the concepts and mathematical theories required when constructing SSMs using Scalismo. It aims to expand on the theory and mathematical consideration used in Chapters 3 to 6. Most of the theory discussed in this appendix is a summary of the knowledge gained through a statistical modelling course set out by Luthi and Bouabene (2020).

B.1   Basic Statistical Concepts

Shape is defined as the geometrical information which remains when location, scale and rotational effects are filtered out from an observed object set. This geometrical information then describes a shape family. Anatomical shapes, however, do not quite fit this description due to the role scale plays in anatomical shape variance, e.g. the femur of a child would not match that of a grown man. Thus, to accommodate this characteristic of anatomical shapes, the definition of anatomical shape is the geometrical information that remains when location and rotational effects are filtered out from an anatomical object Luthi and Bouabene (2020).

This definition becomes important when constructing SSMs and ASMs in software. Statistical models are trained using training data; therefore, when looking at the definition of anatomical shape, it is essential to ensure the entire data set is correctly aligned to one image frame before extracting any geometrical information.

Assuming we were to align the data to a single reference frame and extract the geometrical information, we still require some way to store and model the observed shape variance. In software, the biggest assumption made is that the shape variations within the shape family can be modelled by a normal distribution (Luthi and Bouabene, 2020):

$$p(x) = \mathcal{N}(\mu, \sigma^2) \tag{15}$$

Where μ is the mean, i.e. location of the distribution and σ is the shape variance within the family. $p(x)$ is an indication of how likely a given observation is.

It is also assumed that the normal distribution is unimodal, symmetric and centred around the mean. $p(x)$ values far from the mean quickly become unlikely. In shape modelling, we are not interested in modelling a single random variable. Instead, we always have a set of random variables that we want to model together. Thus, we use a multivariate normal distribution (Luthi and Bouabene, 2020):

$$\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{16}$$

Let $\mathbf{x}$ be a set of jointly normally distributed random variables with $\boldsymbol{\mu}$ defined as:

$$\boldsymbol{\mu} = (\mu_1...\mu_N)^T \tag{17}$$

containing the mean vector with the location of the distribution. μ_k can easily be calculated using Equation 1 if $mu_k = \overline{x}_k$. Using Equation 2, $\boldsymbol{\Sigma}$ in this appendix is defined by:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{S}_{11} & ... & \mathbf{S}_{1N} \\ | & ... & | \\ \mathbf{S}_{N1} & ... & \mathbf{S}_{NN} \end{bmatrix} \tag{18}$$

When the geometrical information has been extracted from the data set, an analysis of which shapes are more likely than others will be required. In other words, what shapes can be described by the observed geometrical information and which are so unlikely that they probably don't belong to the shape family? This probability can be analysed by using the density function:

$$\begin{aligned} p(\mathbf{x}) &= \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ &= (\frac{1}{(2\pi)^{N/2}\sqrt{det(\boldsymbol{\Sigma})}})exp(-\frac{1}{2}[(\mathbf{x} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})]) \\ &= (Normalization)exp(-\frac{1}{2}[\text{Mahalanobis Distance}]) \end{aligned} \tag{19}$$

where N is the total number of variables or landmark points. Equation 19 primarily consists of a normalisation factor and Mahalanobis Distance, which respectively ensures the density integrates to one and takes the shape of the distribution into account, through the presence of $\boldsymbol{\Sigma}^{-1}$.

When given a full joint distribution, we are mostly interested in a subset of the variables when constructing an SSM. It is then logical to ask :

1. What is the distribution of x_1 given no knowledge of $x_2 \rightarrow x_k$?

2. What is the distribution of x_1 given we have observed the positional values of $x_2 \rightarrow x_k$?

Respectvily these two questions are catagorised as the Marginal and Conditional distribution of **x**. For practical reasons the amount of random variables will be limited to x_1 and x_2 for further discusion. Transforming Equation 16:

$$p(x_1, x_2) = \mathcal{N}(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}) \tag{20}$$

where the corelation of the two random variables can be calculated by $\rho = \frac{\sigma_{12}}{\sigma_1 \sigma_2}$

As can be seen from Equation 20, $\mathbf{\Sigma}$ needs to be symmetric and positive definite (SPD). The variance stored within the matrix along the diagonal line determines how spread out the values are in each direction. The covariance within the upper and lower triangle of the matrix determines how the random variables, $x_1 \rightarrow x_N$, correlate according to the observed data, i.e. how much the points within the model change together. This correlation can easily be calculated for these two variables by:

$$\rho = \frac{\sigma_{12}}{\sigma_1 \sigma_2}$$

If the correlation between the two random variables is zero, the variables don't influence each other, and they can change their values entirely and independently. A positive correlation will show similar large or small variable values, where a negative correlation will have a significant value if the other is small and visa versa (Luthi and Bouabene, 2020).

Following Equation 20 the density functions of the marginal and conditional distribution can respectively be expressed as $p(x_1)$ and $p(x_1|x_2 = \tilde{x}_2)$. The marginal and conditional distribution can thus be represented in a multivariate normal distribution form as seen below:

$$p(\mathbf{x}_1) = \mathcal{N}(\mu_{x_1}, \mathbf{\Sigma}_{x_1 x_1}) \tag{21}$$

$$p(x_1|x_2 = \tilde{x}_2) = \mathcal{N}(\overline{\mu}, \overline{\mathbf{\Sigma}}) \tag{22}$$

where through linear algebra:

$$\overline{\mu} = \mu_{x_1} + \mathbf{\Sigma}_{x_1 x_2} \mathbf{\Sigma}_{x_2 x_2}^{-1} (\tilde{x}_2 - \mu_{x_2})$$

$$\overline{\mathbf{\Sigma}} = \mathbf{\Sigma}_{x_1 x_1} - \mathbf{\Sigma}_{x_1 x_2} \mathbf{\Sigma}_{x_2 x_2}^{-1} \mathbf{\Sigma}_{x_2 x_1}$$

This shows that both marginal and conditional distributions are normal distributions (Luthi and Bouabene, 2020).

The last distribution to take into account is a degenerate multivariate normal distribution. When constructing an SSM based on anatomical shapes, we usually have many more landmark points than examples in the data set. Using Equation 2 to estimate the covariance matrix results in a matrix that is only positive semi-definite. While it is still possible to define a valid multivariate normal distribution, we can no longer derive a density function to assess the probability of a particular shape recurring. This will be addressed within Section B.6.

In summary, assuming that shape variations can be modelled using a normal distribution, it is sensible to think that a mean shape exists around which all plausible shapes of the family can cluster. It is then equally likely that shapes of the same anatomical family can be smaller or larger than this mean, but we are highly unlikely to ever observe much larger or smaller shapes than the mean. When compared to bone shapes, we can see that these are reasonable assumptions when we restrict the statistical data used to train the model to a specific population based on age, gender and ethnicity (Luthi and Bouabene, 2020).

B.2   The Gaussian Process

Where PDMs model different shape distributions of a given shape family directly, in Scalismo, a different approach is taken. SSMs and ASMs are established by modelling shape changes as deformations from a given shape, i.e. how the points move between the shapes (Luthi and Bouabene, 2020).

Modelling shape deformations using Scalismo one assigns a point set describing the reference shape, $\boldsymbol{\Gamma}_R$, from a data set:

$$\boldsymbol{\Gamma}_R = \{\mathbf{x}^1 | \mathbf{x}^1 \epsilon \mathbb{R}^d\}$$

All other shapes from the data set then function as traget shapes, $\boldsymbol{\Gamma}_T$:

$$\boldsymbol{\Gamma}_{T^{2 \to n}} = \{\mathbf{x}^{2 \to n} | \mathbf{x}^{2 \to n} \epsilon \mathbb{R}^d\}$$

The target shapes are then seen as a set of deformations from the reference shape, and thus the variables x_k can be found on all the shapes of the data set. This means a deformation vector $\mathbf{u}(x_k)$ can be found, containing the deformation of x_k from the reference shape in all the target shapes:

$$\mathbf{u} : \boldsymbol{\Gamma}_R \to \mathbb{R}^d$$

For a fixed reference shape we can reformulate the problem of shape modelling as the problem of modelling shape deformations. Thus, we need to specify the deformation $\mathbf{u}(\mathbf{x})$ for the point set $\mathbf{x}$ (Luthi and Bouabene, 2020).

Assuming the deformation follows a multivariate normal distribution we can derive and state the following:

$$\sim \mathcal{N}(\overline{\mathbf{u}}(\mathbf{x}), \boldsymbol{\Sigma}(\mathbf{x}))$$

1. The mean deformation, $\overline{\mathbf{u}}(\mathbf{x})$, moves the applicable landmark points (x_k) to the position where it will be found in an average shape according to the dataset.

2. The covariance matrix, $\boldsymbol{\Sigma}(\mathbf{x})$, defines how much the position of the landmark point can vary

3. The model derived above can be generalised to any number of points.

However, generalising the model above to any number of points can be quite cumbersome. Therefore, assume we have a mean function available which assigns for every landmark point in the reference a mean deformation vector stating where it will move on average:

$$\vec{\boldsymbol{\mu}} : \boldsymbol{\Gamma}_R \to \mathbb{R}^d$$

and a covariance function which defines the covariance between the landmark points

$$\mathbf{k} : \boldsymbol{\Gamma}_R \times \boldsymbol{\Gamma}_R \to \mathbb{R}^{d \times d}$$

Using these two functions we can now compute for any finite set of reference points $\boldsymbol{\Gamma}_R$, the corresponding normal distribution:

$$\mathbf{u} \sim \mathcal{N}(\vec{\boldsymbol{\mu}}, \mathbf{K}) \tag{23}$$

with

$$\vec{\boldsymbol{\mu}} = (\boldsymbol{\mu}(\mathbf{x}))_{\mathbf{x} \epsilon \boldsymbol{\Gamma}_R}$$

$$\mathbf{K} = (\mathbf{k}(\mathbf{x}, \mathbf{x}'))_{\mathbf{x}, \mathbf{x}' \epsilon \boldsymbol{\Gamma}_R}$$

The model defined in Equation 23 is known as a Gaussian Process and is defined by:

$$\mathbf{u} \sim GP(\boldsymbol{\mu}, \mathbf{k}) \tag{24}$$

Where

$$\boldsymbol{\mu} : \boldsymbol{\Gamma}_R \to \Re^d$$

$$\mathbf{k} : \boldsymbol{\Gamma}_R \times \boldsymbol{\Gamma}_R \to \Re^{d \times d}$$

A GP is a more general model of the multivariate normal distribution and works for an infinite number of points. It allows us to model a distribution over a function without choosing a discretisation upfront. Gaussian Processes are commonly used in machine learning and statistics, providing us with a wide variety of tools already well established in software (Luthi and Bouabene, 2020).

B.3 Building a SSM From Example Data

Using the GP model described above and a data set focusing on a specific bone shape, an SSM can be constructed by following six steps as set out by Luthi and Bouabene (2020):

1. Define a reference shape,$\mathbf{\Gamma}_R = \{\mathbf{x}|\mathbf{x}\epsilon\mathbb{R}^d\}$, from one of the shapes in the data set.

2. Align and establish correspondence between the reference shape, $\mathbf{\Gamma}_R$ and the remaining target shapes, $\mathbf{\Gamma}_{T^{2\to N}} = \{\mathbf{x}^{2\to n}|\mathbf{x}^{2\to n}\epsilon\mathbb{R}^d\}$.

3. Assuming correspondence has been established, find the deformations $\mathbf{u}^i = \{u_1, ..., u_N\}$ from the target shapes ,$\mathbf{\Gamma}_{T_{1\to n-1}}$.

4. Learn shape variability through the estimation of $\boldsymbol{\mu}$ (Equation 11) and $\mathbf{k}$ (Equation 12).

5. Construct $\mathbf{u} \sim GP(\boldsymbol{\mu}, \mathbf{k})$.

6. The final shape model can be described by:

$$\mathbf{\Gamma} = \{\mathbf{x} + \mathbf{u}(\mathbf{x})|\mathbf{x}\epsilon\mathbf{\Gamma}_R\} \tag{25}$$

with $\mathbf{u}(\mathbf{x})$ being the probabilistic entity for some $\mathbf{u} : \mathbf{\Gamma}_R \to \mathbb{R}^d$

These steps are the basic framework for establishing an SSM using landmark deformations. However, a list of intermediary steps discussed below needs to be considered before steps two to six can be performed.

B.4 The Registration Problem

When constructing an SSM or fitting a model to a 2D scalar image or 3D mesh, it is essential to establish correspondence between the reference shape Γ_R, and the target shape, Γ_T, so that all discretised points describing the contour line, define the same landmark or position in both instances. This ensures the calculated deformation field $\mathbf{u}$ truly represents the shape variation within the statistical population. Establishing correspondence, however, is difficult due

to the translation and rotational effects and acquisition calibration when the scalar images were taken or segmented to the required surface mesh for training (Luthi and Bouabene, 2020).

The first thing required for establishing correspondence is the alignment of the refernce and target shapes, Γ_R and Γ_T. This mitigates the effects of rotation and translation, ensuring the model represents the actual shape variation within the shape family. One method for the alignment of shapes is the generalised Procrustes analysis found in Section B.5.

Now assuming the Γ_R and Γ_T shapes are already aligned, we will need to find the deformation between the two shapes. This is easy to do if we know the discretised points, describing the contour in each shape, are in correspondence, which is not the case in most practical applications, as already discussed.

To find correspondece we first need to build a posterior model (Section B.8) using the reference shape, Γ_R, and a smooth defomation Gaussian kernel (Section B.7) so defining a GP, $\mathbf{u} \sim GP(\boldsymbol{\mu}, \mathbf{k}_{smooth})$.

We will then be able to use the posterior model to establish correspondence by fitting the reference shape model to the entire set of target shapes $\Gamma_{T_{1 \to n-1}}$. Many optimisation algorithms can be used for this purpose. One of the simplest is the Iterative Closest Point (ICP) algorithm, which is well established within the shape modelling community (Luthi and Bouabene, 2020). It is a classical optimisation algorithm used for minimising the distance between two point sets and can be described for model fitting as follows:

1. Find the closest points between the Γ_T and Γ_R point sets.

2. Perform GP regression using deformations as noisy observations (Subsection B.8).

3. Set $\Gamma_R = \Gamma_{regresion}$.

4. Iterate till convergence criteria are met.

In practice, convergence criteria are characterised by assuming correspondence after a set number of iterations. The number of iterations used is usually verified visually (Luthi and Bouabene, 2020).

At the end of the ICP cycle, the point set should reasonably fit the target shape, resulting in correspondence with the posterior model's reference shape. If we were to then run all the target shapes $\Gamma_{T_{1 \to n}}$ through the ICP cycle, we should have a new data set containing the corresponding $1 \to n-1$ point sets. These point sets should then contain all the variance from the original target shapes while corresponding to each other and the reference shape. This new data set can then be used in Section B.3 to construct an SSM.

B.5 Aligning Shapes

In practice, a collected data set is very rarely in correspondence or even in the same reference frame (Luthi and Bouabene, 2020). Thus, to ensure that the deformation fields measured between the reference and target shapes are pure shape changes, we require methods to align and establish correspondence between the shapes in the data set.

The translational and rotational effects between two shapes described by individual point sets can be minimised through Procrustes alignment. This method states that for two shapes $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ the correct pose can be established by minimising:

$$(t^*, R^*) = argmin \sum_{k=1}^{N} ||x_{k,1} - R(x_{k,2} - t)||^2 \tag{26}$$

where $t \ \epsilon \ \mathbb{R}^d$, $R \ \epsilon \ \mathbb{R}^{d \times d}$, $R^T R = I$ and $det(R) = 1$. By obtaining the translation vector t^* and rotation matrix R^* we can superimpose the two shapes to fit the same reference frame. This method provides a closed-form solution for both 2D and 3D shapes and can be efficiently computed using a singular value decomposition (SVD) (Luthi and Bouabene, 2020).

When applying this process to a data set of more than two-point sets, $\mathbf{\Gamma}_1, ..., \mathbf{\Gamma}_N$, we can use an iterative process called Gengeneralised Procrustes alignment:

1. Assign a reference shape from the data set, e.g. $\mathbf{\Gamma}_R = \mathbf{\Gamma}_1$.

2. Align all shapes, $\mathbf{\Gamma}_2, ..., \mathbf{\Gamma}_N$ individually to $\mathbf{\Gamma}_R$ by optimizing Equation 26 and then superimposing the target shape.

3. Compute the mean shape $\mathbf{\Gamma}_{\boldsymbol{\mu}}$ from the aligned data set using Equation 11.

4. Assign $\mathbf{\Gamma}_R = \mathbf{\Gamma}_{\boldsymbol{\mu}}$ and iterate from step 2.

The process is repeated until the mean shape stops showing any noticeable change with each additional iteration, a feature which is usually examined through visual inspection.(Luthi and Bouabene, 2020)

Another term for aligning shapes in this manner is rigid-alignment, which is defined as a composition of translation and rotation. This sort of alignment comes into play when we need to align a data set in which correspondence has not yet been established. The task at hand is then to retrieve the transformation, which best aligns the target shapes, $\mathbf{\Gamma}_{T_{2 \to n-1}}$, to the reference shape, $\mathbf{\Gamma}_R$. It is essential to normalise the pose so that the target shapes, $\mathbf{\Gamma}_{T_{2 \to N}}$, represent transformed shapes dependent solely on shape variance. The easiest way to retrieve this transformation in practice is by manually predefining some landmarks using a Graphical User Interface (GUI) framework on all the shapes in

the data set. Using the Procrustes analysis above, we can then retrieve the best transformation and align the shapes (Luthi and Bouabene, 2020).

It is also worth mentioning that the ICP method can be adapted to automatically rigidly align a data set:

1. Find the closest points between Γ_T and Γ_R point sets.

2. Estimate transformation based on these corresponding points by using Equation 26.

3. Transform the reference using the transformation estimated above

4. Iterate till convergence criteria are met

B.6   Principal Component Analysis

The problem with GPs are that even for a finite discrete multivariate normal distribution, the resulting covariance matrix can be too large to computational process and store on computers. For example a 3D face mesh consisting of ± 50000 vertices ,a standard resolution, will produce a covariance matrix of the size $(50000 \cdot 3) \times (50000 \cdot 3)$. A solution to this problem is the low-rank approximation of the Karhunen-Loéve expansion (Luthi and Bouabene, 2020):

$$\mathbf{u} = \boldsymbol{\mu} + \sum_{i=1}^{r} \alpha_i \sqrt{\lambda_i} \Phi_i \tag{27}$$

with Φ_i as a eigenfunction with associated eigenvalue λ_i and $\alpha_i \sim \mathcal{N}(0,1)$ (Luthi and Bouabene, 2020). Looking at Equation 27, its similarities to equations found in the previous two sections should be evident. This subsection aims to better explain Principal Component Analysis (PCA).

The main idea of the low-rank approximation of the Karhunen-Loéve expansion is to Represent the GP using only the first r components. Giving us a finite, parametric representation of the GP from which the deformation $\mathbf{u}$ could be determined by coefficients $\alpha = \{\alpha_1, ..., \alpha_r\}$. This expansion also allows us to calculate the probability of a given shape deformation using only the α coefficients:

$$p(\mathbf{u}) = p(\alpha) = \prod_{i=1}^{r} 1 \frac{1}{\sqrt{2\pi}} exp(\frac{-\alpha_i^2}{2}) \tag{28}$$

This expansion also holds some beneficial computational aspects. Φ_i and λ_i can easily be computed from the covariance matrix using conventional numerical methods, which are well established within software and practice. Computing elements are very efficient for models learned from examples and general models. The entire covariance model is never fully computed; only estimations must be made according to the desired accuracy. And lastly, if the rank of r

is low, the representation is efficient, and the expansion can represent shapes with millions of landmark points (Luthi and Bouabene, 2020).

PCA is a popular method in shape modelling and closely relates to the Karhunen-Loéve expansion as described above. PCA is used for a GP defined on a discrete domain. The covariance matrix from this discrete domain can be estimated and stored as a compact covariance function for exploring the shape variations of a shape family (Luthi and Bouabene, 2020).

A more detailed explanation can be given by making two assumptions. Firstly, we have a set of discrete deformation fields $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$, and secondly, by using Equation 29, we can estimate the covariance matrix. Thus, the rank of Σ is at most n, the number of target shape deformations.

$$\Sigma = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{u}_i - \boldsymbol{\mu})(\mathbf{u}_i - \boldsymbol{\mu})^T := \frac{1}{n} \mathbf{X} \mathbf{X}^T \tag{29}$$

These assumptions produce two advantages:

1. By performing a SVD we are able to compute the decomposition for a much smaller data matrix $\mathbf{X} = (\mathbf{u}^i - \boldsymbol{\mu})$ more efficiently.

2. Σ has only n non-zero eigenvalues, thus $r = n$ and the expansion reduces to:

$$\mathbf{u} = \boldsymbol{\mu} + \sum_{i=1}^{n} \alpha_i \sqrt{\lambda_i} \Phi_i \tag{30}$$

where $\alpha_i \sim \mathcal{N}(0, 1)$. From Equation 30 we can then conclude that any deformation $\mathbf{u}$ can be accurately specified using a coefficient $\alpha \, \epsilon \, \mathbb{R}e^n$ (Luthi and Bouabene, 2020).

Additionally, PCA can also generate shape instances using the principal components or eigenmodes, Φ_i. The first few eigenmodes represent the main modes of variation and contain the highest variance directions within the data set. This represents most of the variation present within the population using the rank n eigenmodes. To analyse the particular variation presented by a specific eigenmode, j, we set the coefficient $\alpha^j = v$ and $\alpha_{i \neq j} = 0$ we can then visualise this effect through:

$$\boldsymbol{u} = \boldsymbol{\mu} + v\sqrt{\lambda_i}\Phi_i \tag{31}$$

v can than be changed in accordance with the 3σ standard deviation, thus $v \epsilon \{-3, 3\}$ (Luthi and Bouabene, 2020).

B.7   Kernel Construction

The parameters used to define Gaussian Processes can be derived using many different techniques and inputs. Using example data to determine the mean

shape and covariance matrix of a shape population ensures that the GP will only represent shapes within that shape population. This raises the question of how many samples are required to describe a statistical population and what can be done if a sufficient set of data samples are out of reach?

The first question is difficult to answer when it comes to anatomical shape variation. To the best of our knowledge, there has been no accurate method determined in literature to estimate the required sample size. In literature, this question is mainly answered by validating the SSM through one of two methods. The first method consists of a 'leave-one-out cycle' where the SSM is constructed by leaving a sample out of the data set with each cycle. The mean shape is then analysed to see if the particular sample left out causes drastic changes. The second method consists of fitting the SSM to a set of unseen data samples from the same statistical population. Therefore, if the mean does not change and the SSM can be fitted to a series of unseen data, it is logical to assume the data set is large enough (Heimann and Meinzer, 2009).

The second question can be answered through kernel construction, i.e. covariance function manipulation. As established $\mathbf{k} : X \times X \to \mathbb{R}^{d \times d}$ defines how deformation, $\mathbf{u}$, can deviate from the mean. Kernel construction can make $\mathbf{k}$ more robust or flexible so that we can include variances not seen within a data set.

To illustrate the power of kernel construction or manipulation, we first discuss the possibility of defining a GP with only one point set available to describe a statistical population:

1. If we have but one shape, the mean function becomes $\mathbf{0}$. The reference shape then illustrates the average shape found in the population.

2. It is then logical to asume $\boldsymbol{\mu}(\mathbf{x}) = \mathbf{0}$,where $\mathbf{0}$ is a vector the size of $\boldsymbol{\mu}$.

3. The covariance function or kernel defines the deformation fields' characteristics, and we assume the deformation fields describing the population are smooth.

4. Mathematically we know the covariance function $\mathbf{k}(\mathbf{x}, \mathbf{x}')$ should be a symmetric, positive semi-definite (PSD) kernel. And we know $\mathbf{k}$ is positive semi-definite if it gives rise to a positive semi-definite kernel matrix $\mathbf{K} \to \Re^{d \times d}$ which satisfies $v^T \mathbf{K} v \geq 0$ for all vectores $v \ \epsilon \ \Re^d$.

5. Then lastly, with $\mathbf{K}_{ij} = \mathbf{k}(\mathbf{x}_i, \mathbf{x}_j)$, $i, j = 1, ..., k$ for any choice of k and $\mathbf{x} = (x_1, ..., x_k)$ we are ensured that the GP defines a valid $N(\boldsymbol{\mu}, \mathbf{K})$ for any discritization (Luthi and Bouabene, 2020).

As discussed in Section B.1 we know the kernel shows how two function values $\mathbf{x}$ and $\mathbf{x}'$ are correlated. Thus, multiple components are available when modelling vector fields to describe the correlation, and a matrix-valued kernel is required. But if only real functions had to be modelled, it would be enough

to get one scalar value to describe the correlation. From this discussion, we define a classical scalar-valued Gaussian kernel used everywhere in literature:

$$\mathbf{k}(\mathbf{x}, \mathbf{x}') = s \times exp(-\frac{\parallel \mathbf{x} - \mathbf{x}' \parallel^2}{\sigma^2}) \tag{32}$$

where s relates to the scale of the deformation field variance and σ characterises the smoothness of the correlation between the two function values $\mathbf{x}$ and $\mathbf{x}'$. From this scalar-valued covariant function, we can build a matrix-valued covariance function required for modelling deformation fields. Of which the most straightforward construction method can be defined by:

$$
\begin{aligned}
\mathbf{k}(\mathbf{x}, \mathbf{x}') &= \begin{bmatrix} \mathbf{k}_1(\mathbf{x}, \mathbf{x}') & ... & 0 \\ : & ... & : \\ 0 & ... & \mathbf{k}_d(\mathbf{x}, \mathbf{x}') \end{bmatrix} \\
&= \begin{bmatrix} s_1 \times exp(-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{\sigma_1^2}) & ... & 0 \\ : & ... & : \\ 0 & ... & s_d \times exp(-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{\sigma_d^2}) \end{bmatrix}
\end{aligned} \tag{33}
$$

where $\mathbf{k}_1, ..., \mathbf{k}_d : X \times X \to \mathbb{R}$ are scalar-valued kernels and $\mathbf{k} : X \times X \to \mathbb{R}^{d \times d}$ becomes a matrix-valued kernel. The main assumption is that each dimension can be modelled independently. Thus, the output dimensions are uncorrelated. The equation defined above then satisfies the original goal of modelling smooth deformation fields. It is important to note that kernels defined in such a manner are not just confined to the deformation of the contour of the shape but the whole model space (Luthi and Bouabene, 2020).

Turning back to the second question asked at the start of this section, we now discuss how kernel construction can be used for modelling shape variances of the population not included in the data set, i.e. increasing the flexibility of the Gaussian kernel. We know that the seed needs to be PSD to define a good GP. We know from literature that a known PSD kernel can construct a new kernel with the same attribute. This can be done by using a set of rules for combining kernels:

1. $\mathbf{k}(\mathbf{x}, \mathbf{x}') = \mathbf{k}_1(\mathbf{x}, \mathbf{x}') + \mathbf{k}_2(\mathbf{x}, \mathbf{x}')$

2. $\mathbf{k}(\mathbf{x}, \mathbf{x}') = \alpha \mathbf{k}_1(\mathbf{x}, \mathbf{x}'),\ \alpha \epsilon \mathbb{R}$

3. $\mathbf{k}(\mathbf{x}, \mathbf{x}') = \mathbf{k}_1(\mathbf{x}, \mathbf{x}') \odot \mathbf{k}_2(\mathbf{x}, \mathbf{x}')$

4. $\mathbf{k}(\mathbf{x}, \mathbf{x}') = \int(\mathbf{x}) \int(\mathbf{x}')^T$

5. $\mathbf{k}(\mathbf{x}, \mathbf{x}') = B^T \mathbf{k}(\mathbf{x}, \mathbf{x}')B,\ B \epsilon \mathbb{R}^{d \times r}$

6. $\mathbf{k}(\mathbf{x}, \mathbf{x}') = \mathbf{k}_3(\phi(\mathbf{x}), \phi(\mathbf{x}')),\ \phi : \Omega \to \mathbb{R}^n,\ \mathbf{k}_3 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{d \times d}$ a PSD kernel

where $\mathbf{k}_1, \mathbf{k}_2 : \mathbf{X} \times \mathbf{X} \to \mathbb{R}^{d \times d}$ are SPD and $f : \mathbf{X} \to \mathbb{R}^d$ is a vector-valued function used to generete $\mathbf{k} : \mathbf{X} \times \mathbf{X} \to \mathbb{R}^{d \times d}$ a new PSD kernel (Luthi and Bouabene, 2020).

The rest of this section now uses these rules to show how some powerful models of shape deformations can be constructed to enhance certain attributes in an existing model or increase its flexibility. These models include but are not restricted to:

- A Multiscale model:

When working with anatomical shape variations (e.g bone shapes) we can see a combination of long-ranging, smooth deformations (e.g the length of the femur) and small-scale local deformations (e.g. the femur head shapes). Using rule (1) and (2) we can build a model with varying smoothnesses and deformations on multiple scales (Luthi and Bouabene, 2020):

$$\mathbf{k}_{ms}(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^{k} s_i \mathbf{k}_{\sigma i}(\mathbf{x}, \mathbf{x}') \tag{34}$$

with s_i defining scale and σ_i smoothness. Using this equation with Equation 33 we can define the full kernel matrix required. If σ is chosen larger than the domain of $\mathbf{X}$, smooth deformations will dominate local deformation, and if σ is chosen small, the opposite will be true.

- Anisotropic scaling

Using a diagonal kernel as in Equation 33 we can obtain anisotropic deformation fields, where each component is scaled individually. Unfortunately, this method is limited to 2D shape representation and is thus not sufficient for 3D modelling. To counter this we can use rule (5) with a $d \times d$ rotation matrix R :

$$\mathbf{k}_R(\mathbf{x}, \mathbf{x}') = R^T \mathbf{k}(\mathbf{x}, \mathbf{x}') R \tag{35}$$

allowing us to choose an arbitrary direction of anisotropy. (Luthi and Bouabene, 2020)

- Changepoint kernels:

When looking at anatomical shape such as the femur it becomes clear that some regions feature different forms of variation, e.g. the femur head and shaft. To incorporate this into a model we can use changepoint kernels constructed from rule (4) together with rule (1) and rule (3):

$$\mathbf{k}_{cp}(\mathbf{x}, \mathbf{x}') = X(\mathbf{x})X(\mathbf{x}')^T \odot \mathbf{k}_1(\mathbf{x}, \mathbf{x}') + (\mathbf{1}_2 - X(\mathbf{x}))(\mathbf{1}_2 - X(\mathbf{x}'))^T \odot \mathbf{k}_2(\mathbf{x}, \mathbf{x}') \tag{36}$$

with

$$X(\mathbf{x}) = \begin{cases} \mathbf{1}_2, & \text{if } \mathbf{x} \ \epsilon \text{ femur shaft region} \\ \mathbf{0}_2 & \text{otherwise} \end{cases}$$

where $\mathbf{1}_2$ and $\mathbf{0}_2$ are vectors the length of d. The function X is used as a mask to specify which kernel is active in what region (Luthi and Bouabene, 2020).

- Kernels from Example Data:

From the rules described above, we can prove that Equation 12 used to extract the covariance matrix $\mathbf{k}_k$ from example data is PSD. By visual inspection, one can see a combination of rule (4) (the part inside the sum), rule (1) (the sum) and rule (2) (the scaling factor $\frac{1}{n-1}$). With $\mathbf{k}_k$ being a PSD kernel we can use the rules described above to combine the kernel with any other SPD kernel model. This will increase the models' flexibility and overcome model bias where there is an insufficient data set (Luthi and Bouabene, 2020).

- Modelling the missing variability:

The main assumption made here is that the missing variability can be explained to the GP model by adding a general smoothness kernel to $\mathbf{k}_s$, rule (1). This additional model will then add a more general set of smooth deformations to the learned shape deformations $\mathbf{k}_s$ forming:

$$GP(\mu_s, \mathbf{k}_s(\mathbf{x}, \mathbf{x}') + s\mathbf{I}_{d \times d} exp(-\frac{\| \mathbf{x} - \mathbf{x}' \|^2}{\sigma^2})) \tag{37}$$

where s corresponds to the average error present in the model and σ is chosen relatively large to ensure the error is highly correlated within the new kernel. Additionally, it would be wise to visually inspect the added shape deformation to ensure appropriately selected parameter values.

- Localised shape models:

Where addition of a smoothnes kernel can add to the missing variability, we can aslo enlarge the shape variability at a localized region by mutiplication, rule (3). We can thus define a model $GP(\boldsymbol{\mu}_s, \mathbf{k}_{local})$ where $\mathbf{k}_{local}$ is defined as:

$$\mathbf{k}_{local}(x, x') = \mathbf{k}_s(\mathbf{x}, \mathbf{x}') \odot s\mathbf{I}_{d \times d} exp(-\frac{\| \mathbf{x} - \mathbf{x}' \|^2}{\sigma^2}) \tag{38}$$

Using an example set by Luthi and Bouabene (2020) the intuition is as follows. If a dataset exists that is too small to model the entire shape population, say a femur shape, it may still represent the deformational variance of the subset shapes like the femur shaft or proximal head. Through multiplication this can be achieved, for any point $\mathbf{x}$, the covariance $\mathbf{k}_{local}(\mathbf{x}, \mathbf{x}')$ with any other point $\mathbf{x}'$ is virtually 0 if $||\mathbf{x} - \mathbf{x}'||$ is large. Thus, $\mathbf{k}_{local}$ suppresses global correlations and preserves locally learned correlations.

B.8 Model Fitting with SSMs

After establishing an SSM from the given data set, we will need to fit this model to some form of input. The most practical way of accomplishing this task is called the Gaussian Process Regression, an inference technique used to estimate the complete shape from partial or sparse data using the information found in an SSM (Luthi and Bouabene, 2020).

For SSM the regresion problem can be formulated so that $\mathbf{x}$ is a data set of the reference shape Γ_R and the regresion model can be defined as:

$$\mathbf{u} = \mathbf{u}_0(\mathbf{x}) + \boldsymbol{\varepsilon}$$

where $\mathbf{u}_0 : \Gamma_R \to \mathbb{R}^d$ is a set of regresion functions contaning the deformation fields which explain the observed observation, i.e. the partial inputs. Additionaly, $\boldsymbol{\varepsilon} \sim GP(\mathbf{0}, \sigma^2 \mathbf{I}_{d\times d})$ is independent Gaussian noise. To solve the regression problem we will need to estimate the set regression functions $\mathbf{u}_0$, which we asume are distributed according to a $\mathbf{u}_0 \sim GP(\boldsymbol{\mu}, \boldsymbol{k})$, from a set of observations $\tilde{\mathbf{u}}$. (Luthi and Bouabene, 2020)

To accomplish this task, we will use the conditional multivariate normal distribution as discussed in Section B.1. Now lets us assume we have a prior model, i.e. a reference shape together with a GP model $\mathbf{u} \sim GP(\boldsymbol{\mu}, \mathbf{k})$. Because we know the prior model is a GP, that can be evaluated at a finite discrete number of points knowing the corresponding discrete distribution is also a multivariate normal distribution. We can then, for simplicity, say we have a jointly normally distributed set of variables $\mathbf{x}_R \subset \Gamma_R$ and $\mathbf{x}_T \subset \Gamma_T$ (Luthi and Bouabene, 2020).

By using Equation 16 we can then formulate:

$$\begin{bmatrix} \mathbf{u}(\mathbf{x}_R) \\ \mathbf{u}(\mathbf{x}_T) \end{bmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{\mu}(\mathbf{x}_R) \\ \boldsymbol{\mu}(\mathbf{x}_T) \end{bmatrix}, \begin{bmatrix} \mathbf{K}(\mathbf{x}_R, \mathbf{x}_R) & \mathbf{K}(\mathbf{x}_R, \mathbf{x}_T) \\ \mathbf{K}(\mathbf{x}_T, \mathbf{x}_R) & \mathbf{K}(\mathbf{x}_T, \mathbf{x}_T) \end{bmatrix} \right)$$

For which we are given observations

$$\tilde{\mathbf{u}}(\tilde{\mathbf{x}}_T) = (\tilde{u}_1, ..., \tilde{u}_k)^T$$

Where through conditional normal distribution, we can state

$$p(\mathbf{u}(\mathbf{x}_R)|\mathbf{u}(\mathbf{x}_T) = \tilde{\boldsymbol{u}}) = N(\overline{\boldsymbol{\mu}}, \overline{\boldsymbol{\Sigma}}) \tag{39}$$

with

$$\overline{\boldsymbol{\mu}} = \boldsymbol{\mu}(\mathbf{x}_R) + \mathbf{K}(\mathbf{x}_R, \mathbf{x}_T)\mathbf{K}^{-1}(\mathbf{x}_T, \mathbf{x}_T)(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}(\mathbf{x}_T))$$

$$\overline{\boldsymbol{\Sigma}} = \mathbf{K}(\mathbf{x}_R, \mathbf{x}_R) - \mathbf{K}(\mathbf{x}_R, \mathbf{x}_T)\mathbf{K}^{-1}(\mathbf{x}_T, \mathbf{x}_T)\mathbf{K}(\mathbf{x}_T, \mathbf{x}_R)$$

We have now derived a discretised conditional distribution of the GP, which is still a multivariate normal distribution. For a fixed $\mathbf{u}(\mathbf{x}_T) = \tilde{\mathbf{u}}$ the mean and the covariance matrix can be computed for an arbitrary number of points $\mathbf{X} = \{x_1, ..., x_k\}$ (Luthi and Bouabene, 2020). So we can further simplify the closed-form expressions to fit a single point:

$$\mu_p(x_R) = \mu(x_R) + \mathbf{K}(x_R, \mathbf{x}_R)\mathbf{K}(\mathbf{x}_T, \mathbf{x}_T)^{-1}(\tilde{\mathbf{u}} - \boldsymbol{\mu}(\mathbf{x}_T)) \qquad (40)$$

$$\mathbf{k}_p(x_R.x_R') = \mathbf{k}(x_R, x_R') - \mathbf{K}(x_R, \mathbf{x}_T)\mathbf{K}(\mathbf{x}_T, \mathbf{x}_T)^{-1}\mathbf{K}(\mathbf{x}_T, x_R') \qquad (41)$$

where Equations 40 and 40 respectivley allow us to compute the mean for any point x_R and the covariance function for any pair of points x_R, x_R' for any set of observations $\tilde{\mathbf{u}}(\tilde{\mathbf{x}}_T)$. (Luthi and Bouabene, 2020)

This is known as a posterior process or model $GP(\mu_p, \mathbf{k}_p)$. It defines a distribution of vector fields that match the given observations perfectly. When trying to reconstruct a shape using the posterior model, the mean will give the best statistical representation of the complete bone shape. The problem is that observations are rarely 100% correct and always contain some human or measurement error (Luthi and Bouabene, 2020).

To accommodate this we add a noise term. For the given observations $\tilde{\mathbf{u}}(\tilde{\mathbf{x}}_T)$, we assume that $\tilde{\mathbf{u}}(\tilde{\mathbf{x}}_T) + \boldsymbol{\varepsilon} = \tilde{\mathbf{u}}$ where $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{dxd})$. We then reformulate the above to find:

$$\begin{bmatrix} \mathbf{u}(\mathbf{x}_R) \\ \mathbf{u}(\mathbf{x}_T) \end{bmatrix} \sim N(\begin{bmatrix} \boldsymbol{\mu}(\mathbf{x}_R) \\ \boldsymbol{\mu}(\mathbf{x}_T) \end{bmatrix}, \begin{bmatrix} \mathbf{K}(\mathbf{x}_R, \mathbf{x}_R) & \mathbf{K}(\mathbf{x}_R, \mathbf{x}_T) \\ \mathbf{K}(\mathbf{x}_T, \mathbf{x}_R) & \mathbf{K}(\mathbf{x}_T, \mathbf{x}_T) + \sigma^2\mathbf{I}_{dk \times dk} \end{bmatrix})$$

thus,

$$\mu_p(x_R) = \mu(x_R) + \mathbf{K}(x_R, \mathbf{x}_R)(\mathbf{K}(\mathbf{x}_T, \mathbf{x}_T) + \sigma^2\mathbf{I}_{dk \times dk})^{-1}(\tilde{\mathbf{u}} - \boldsymbol{\mu}(\mathbf{x}_T)) \qquad (42)$$

$$\mathbf{k}_p(x_R.x_R') = \mathbf{k}(x_R, x_R') - (\mathbf{K}(\mathbf{x}_T, \mathbf{x}_T) + \sigma^2\mathbf{I}_{dk \times dk})^{-1}\mathbf{K}(\mathbf{x}_T, x_R') \qquad (43)$$

These expressions give us a symbolical answer of the regression problem containing the posterior process $GP(\boldsymbol{\mu}_p, \mathbf{K}_p)$ which defines a distribution over the observed vector fields. The mean $\boldsymbol{\mu}_p$ solves the regression problem, and the covariance kernel $\mathbf{K}_p$ illustrates deformations of the mean, which still agrees with the given observations. We also defined σ to account for noise, to control accuracy.

Appendix C

Additional Documentation

Interview questionnaire for the estimation of healthy bone shape and density distribution from partial inputs for implant design.	
Company questioned:	CranioTech (PTY) Ltd
Participants (Job description):	Bernard Swart (CEO)
	Sven Delport (Chief Biomedical Engineer)
Interviewer:	DP Kramer

Question 1:

How do you in the industry estimate or reconstruct pathological bone surface geometry?

Contralateral anatomy if available. In practice constraints are unique to each patient and take priority during surface reconstruction, even when working with healthy contralateral anatomy. Constraints: 1) bone stock quality and location, b)imaging quality, c) fixation of implant.

Question 2:

What are the tolerance limits you generally work with for:
1. Edge alignment (where the implant and bone surface meet)?
2. Surface reconstruction (average and Hausdorff errors)?
3. Center shaft profile for femur and tibia (see illustration below)?

Healthy   Estimate
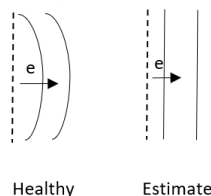
1. 2 to 3 mm over and under hangs acceptable (surgical error also needs be considered).
2. For femur truss cages (50 to 70% of the bone length), 5 to 10 mm surface tolerance generally accepted by surgeons.
3. ± 10 mm, IM-rod placement more important than center point accuracy.

Question 3:

What consideration is given to implant strength, the load it must carry, and how does healthy or patient bone density contribute to implant design or specifications?

Available bone surface for fixation is important, differs for each patient and depending on the bone density, surface reconstruction need be changed to accommodate fixation and healthy bone stock.

- Implant load and destructive testing used for thickness analysis. Aim, ten times the average wait of a person for femoral truss cages (80kg x 10 = 800 kg).
- ± 2-ton load for ± 1.5 mm Ti truss diameter (over engineer PSI).

Question 4:

Does the diagram below seem feasible for the PSI design pipeline? If not, indicate what you would add or change.
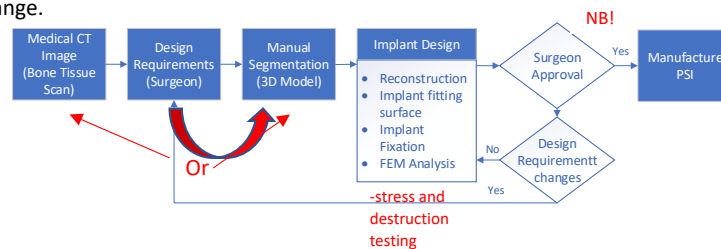
Notes:

Signature: _____ Date: __2021/10/26__

Signature: _____ Date: __2021/10/26__

Signature: _____ Date: __2021/10/26__

Segmentation Protocol

Estimation of healthy bone shape and density distribution from partial inputs for implant design

Investigator: DP Kramer, Student, Master's Candidate, Mechanical and Mechatronic
                     Engineering Department, Biomedical Research Group, Stellenbosch
                     University.

Supervisor: Dr Johan van der Merwe, Lecturer, Mechanical and Mechanical
                     and Mechatronic Engineering Department, Biomedical Research Group,
                     Stellenbosch University.

Co-Supervisor: Dr M. Lüthi, Lecturer, Graphics and Vision Research Group, Basel
                     University.

Please address all correspondence to:
DP Kramer, MEng
Department of Mechanical and Mechatronic Engineering
Stellenbosch University
Corner of Banghoek and Joubert Street
Mechanical Engineering Building
Stellenbosch
Phone: (+27) 82 604 6131
Email: 19865481@sun.ac.za

Introduction

Segmentation is the process of labelling the desired anatomy and isolating it from the remainder of the image. This process is semi-automatic and is considered the gold standard against which to compare automatic segmentation algorithms. A user would essentially first apply a threshold to the grey values of the image scan to isolate the bone, and afterwards manually clean each image slice by filling in holes and trimming away unwanted artefacts. This is followed by the generation of a 3D mesh of the desired anatomy using a mask which was constructed during the thresholding process.

CranioTech, with industry expertise, will conduct image segmentation of the collected dataset using Mimics, which is software specifically for segmentation, analysis and visualisation of medical images. When transferring the data to CranioTech for segmentation the dataset will be stored on an external hard drive. When segmentation is complete the original data and service deliverables will be retrieved using the same storage device. CranioTech, being paid for their services from the project budget, will not store, keep or otherwise make use of any data. See the DTA, 'Data Transfer Agreement', documentation at the back for further details.

Objectives

The main objective of this protocol is to establish a step by step process which will ensure reproducibility of segmentation quality and the production of the required deliverables from the segmentation services rendered.

The deliverables specified below are required for the completion of a Master's thesis and will form part of the training data set required for the construction of a statistical module that will intern be used for the reconstruction of segmented bone loss.

Dataset

The data set will consist of a 100, 50 male and 50 female, diagnostic CT scans collected retrospectively. The scans are in both DICOM and pre-generated .mcs file format. The .mcs files were generated using MIMICS Innovation Suite, version 16.0. Each scan is anonymized and assigned a deidentified identification number to protect each patient's identity.

When imported into MIMICS each .mcs file will contain a patient-specific medical image with anatomical density data captured within the voxel intensities. Different voxel intensities will illustrate different anatomical geometries and structures within the medical images. The Hounsfield Unit (HU) scale can be used to assist in identifying the different anatomical structures being analysed, Figure 1. Note that the values given in Figure 1 are only estimates and that these ranges will differ depending on the anatomical structure being analysed, the imaging machine and global lighting effects present.

As mentioned the segmentations generated by CranioTech will be used for the construction of statistical models which will focus primarily on the healthy femur long bone structure. Segmentations of healthy tibia long bone structures will also be collected where possible for the construction of secondary statistical models. Each scan will contain one healthy femur long bone and possibly a healthy tibia long bone. Scans might be singular or bilateral. The last section of this document will contain the segmentation lists for both the male and the female CT scans. The list

1

will identify which femur, regarding the median plane, must be segmented on each scan. The list will also identify if a healthy tibia is present and if the segmentation of this tibia is required.
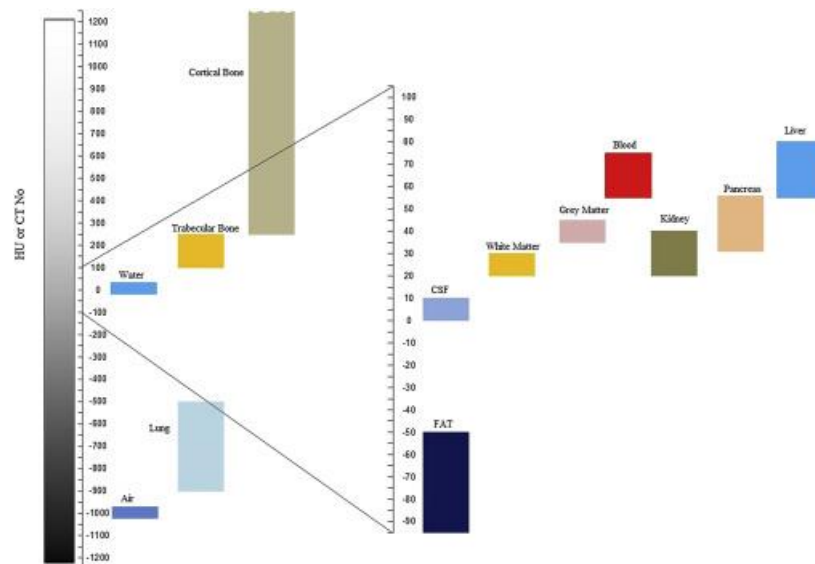
Figure 1: HU scale for different anatomical structures (Kalra, 2018)

Slice thickness and pixel resolution might vary for each image. This is due to the diagnostic and retrospective nature of the scans collected for the study. Imaging practices may change annually or as needed for diagnosis when the images were taken. Typically, the slice thickness will be 1 to 1.5 mm and the pixel resolution will be centred around 0.75 mm.

Service Deliverables

Following the guidelines provided below CranioTech will generate three masks for each long bone structure identified in the segmentation list. They will have to use their anatomical knowledge, industry expertise and MIMICS Innovation Suite, version 16.0, to ensure the masks generated are anatomically correct and accurate to the images given.

The first mask of each long bone structure will highlight the entire bone geometry, containing both cortical and trabecular bone geometries. The second will focus only on the cortical bone and the third on the trabecular bone. These three masks will then be properly labelled as set out in Table 1 and stored to the .mcs file.

After mask generation, full 3D segmentation models (containing the cortical and trabecular bone geometries) of the masked long bones will be generated. Using varies shaping, smoothing, wrapping, colouring and clipping techniques CranioTech will ensure the 3D segmentation will be an accurate representation of the femur and tibia long bone geometries. The 3D segmentation models will then also be saved to the .mcs project file of each image and so be labelled appropriately, see Table 1.

The external surface geometry information, of the femur long bone and possible tibia long bone, from the 3D segmentations models will then be exported to .stl files in binary format and labelled appropriately as shown in Table 1.

Table 1: Deliverable summary for each Image/Scan

Deliverable	Label/ Name	Mask Color/ Format	Storage Location
Mask 1: Full bone geometry of femur/tibia	Femur_Full	Green mask	To MIMICS file (.mcs): "Deidentified-ID.mcs"
	Tibia_Full	Blue mask	
Mask 2: Cortical bone geometry of femur/ tibia	Femur_Cor	Orange mask	To MIMICS file (.mcs): "Deidentified-ID.mcs"
	Tibia_Cor	Yellow mask	
`Mask 3: Trabecular bone geometry of femur/ tibia	Femur_Tra	Purple mask	To MIMICS file (.mcs): "Deidentified-ID.mcs"
	Tibia_Tra	Pink mask	
3D Model/ Object: Full 3D segmentation	Femur_Model	Green model	To MIMICS file (.mcs): "Deidentified-ID.mcs"
	Tibia_Model	Blue model	
Surface Model: 3D Model export	Deidentified-ID_Femur.stl	Binary format	To individual deidentified scan folder: "…\1.De-Identified CT Scans\1. Accepted Scans\ Deidentified-ID-Folder \…".
	Deidentified-ID_Tibia.stl	Binary format	

All the deidentified CT scans can be found by following the file path "…\1.De-Identified CT Scans\1. Accepted Scans\…". Inside this directory the DICOM and .mcs files for each scan are contained within individual folders labelled using the unique deidentified identification number given to each scan during anonymization. It is also within these individual folders that the exported .stl files are stored.

Segmentation Guidelines

These guidelines rely on the discretion of the service provider and the toolkit provided by their expertise and MIMICS Innovation Suite, version 16.0. The list containing the relevant ID's and long bone structures for segmentation can be found in the section below.

1. Import the deidentified .mcs file containing the CT image data.

2. Using the list provided identify the femur long bone for segmentation trough visual inspection.

3

3. Set contrast to highlight the anatomical structure of bone. A good contrasting range used throughout literature is HU[-250,750]. Mimics also has built-in contrasting options to select from.

4. Threshold the image to isolate the long bone geometries that are of interest. The Intensity range describing Trabecular bone is usually between HU[100,250], for Cortical bone the range is from 250+ HU, Figure 1. These intensity ranges might differ on the image being segmented due to global lighting effects present during image capturing. Visual inspection during thresholding is of high importance and the generation and analysis of profile lines might also be of use.

5. Finalizing the thresholding procedure will produce a mask describing all geometries within the thresholding range.

6. For the first mask isolate both the cortical and trabecular bone geometry of the long bone in question. Isolating the geometry can be done by using tools such as clipping and region growing.

7. Manually clean and fill the isolated mask by either stepping through the image slice by slice or from the 3D window when selecting the 'Edit in 3D' option and setting the bounding box. Manually cleaning will consist of closing all irregular gaps in the bone geometry, the elimination of islands and geometries not consistent with the overall anatomical shape and character of the bone structure and finally a smoothing process to eliminate all noise in the mask. Tools for this process can range between local thresholding, removing, erasing and drawing procedures. Label and store the first mask according to Table 1.

8. The second mask can be generated by duplicating the first and applying a local thresholding procedure that isolates the cortical bone. The cortical bone range can be established through a mixture of profile lines, visual inspection, Figure 1 and anatomical knowledge. Label and store the second mask according to Table 1.

9. The third mask can be generated through a 'Boolean Operation'. By subtracting the second mask from the first the trabecular bone will be isolated and captured in the third mask. Label and store the third mask according to Table 1.

10. Using the first mask generate a full 3D segmentation model of the femur long bone geometry. The selection of generation quality should reflect the anatomical geometry of the long bone in question and as such should be selected through visual inspection.

11. Due to the study for which the segmentations are required the 3D model should reflect only the exterior geometry of the long bone being segmented. Thus, operations such as wrapping and smoothing are required to close holes and eliminate any noise or irregular artefacts remaining on the surface contour. Giving a more accurate representation of the true long bone shape. Parameters for wrapping and smoothing procedures need to be established during segmentation trough visual inspection. It is important however to remember to account for shrinkage during these operations.

12. Ensure the segmentation model is a true representation of the original image contour and has remained aligned with said image. This can easily be visualised by enabling the surface contour of the model in the frontal, sagittal and transverse planes. Label and store the 3D segmentation models according to Table 1.

13. Export the external surface geometry to a .stl file. Label and store the file as stipulated in Table 1. Ensure that the size factor is set to 1.0 and that the binary format is selected.

14. Repeat step 3 through 13 for the tibia long bone if listed.

15. Repeat step 1 through 14 for all deidentified images on the segmentation list below.

16. Copy all data files and deliverables back to the external hard-drive for collection from the primary investigator (PI).

17. Once data has been securely transferred to PI's server remove all remaining data from company-owned servers upon notification as per DTA documentation.

Segmentation List

Nr.	De-Identified ID	Gender	Segmentation Structure		Segmentation List (S)
			Femur	Tibia	
1	AS_13557240	M	R	R	S1
2	BH_32656201	M	R	R	S1
3	CH_27273704	M	R	R	S1
4	CG_37677742	M	R	R	S1
5	DT_57631372	M	R	R	S1
6	GS_25575064	M	R	R	S1
7	KC_30230415	M	R	R	S1
8	KS_44412450	M	R	R	S1
9	LK_50500727	M	R	R	S1
10	LE_31371454	M	R	R	S1
11	MN_34037317	M	R	R	S1
12	NC_17447673	M	R	R	S1
13	NS_44640047	M	R	R	S1
14	SM_32340517	M	R	R	S1
15	SZ_72700234	M	R	R	S1
16	TG_47354375	M	R	R	S1
17	TC_56067603	M	R	R	S1
18	AB_12761655	M	R	R	S1
19	FJ_51005421	M	R	R	S1
20	FL_47044712	M	R	R	S1
21	GG_20305124	M	R	R	S1
22	HP_14125063	M	R	R	S1

Nr.	De-Identified ID	Gender	Femur	Tibia	Segmentation List (S)
23	KI_74032300	M	R	R	S1
24	LL_10414033	M	R	L	S1
25	MO_34426047	M	R	R	S1
26	NS_63613401	M	R	R	S1
27	PB_47525126	M	R	L	S1
28	SN_70712234	M	R	R	S1
29	TS_57126710	M	R	L	S1
30	TL_03410126	M	R	L	S1
31	AA_24053370	M	R	R	S1
32	BD_35264326	M	R	R	S1
33	CN_06252254	M	R	R	S1
34	DL_42653364	M	R	R	S1
35	DJ_33544752	M	R	R	S1
36	GA_01733033	M	R	R	S1
37	HS_51400415	M	R	R	S1
38	LL_01526011	M	R	R	S1
39	LA_35153557	M	R	R	S1
40	LM_66420230	M	R	R	S1
41	MS_53433456	M	R	R	S1
42	MM_07747377	M	R	R	S1
43	MH_25353647	M	R	R	S1
44	MS_30272377	M	R	R	S1
45	MS_11472072	M	R	R	S1
46	MR_24322376	M	R	R	S1
47	MK_74130321	M	R	R	S1
48	NR_21544507	M	R	R	S1
49	NB_00142132	M	R	R	S1
50	SD_63246235	M	R	R	S1
51	AM_04134222	F	L	L	S1
52	BT_40373731	F	L	L	S1
53	DE_17556043	F	L	R	S1
54	GN_21434102	F	L	R	S1
55	JS_27100611	F	L	L	S1
56	LB_76652532	F	L	-	S1
57	MN_61036532	F	L	L	S1
58	MM_12006211	F	L	R	S1
59	MZ_65542530	F	L	-	S1
60	SS_20716161	F	L	L	S1
61	SS_41571107	F	L	L	S1
62	SA_03365303	F	L	L	S1
63	TM_03454355	F	L	-	S1
64	VW_52766006	F	L	L	S1
65	CS_06256227	F	L	L	S1
66	DS_51700206	F	L	L	S1

Nr.	De-Identified ID	Gender	Femur	Tibia	Segmentation List (S)
67	KN_07357472	F	L	L	S1
68	KL_25433306	F	L	L	S1
69	MS_40012352	F	L	-	S1
70	NB_25221320	F	L	L	S1
71	RS_76107700	F	L	L	S1
72	AL_72230651	F	L	L	S1
73	BF_06363205	F	L	L	S1
74	BP_46407074	F	L	L	S1
75	CM_64705547	F	L	L	S1
76	CS_42123275	F	L	L	S1
77	DM_62044177	F	L	L	S1
78	DB_35202674	F	L	-	S1
79	GN_51474054	F	L	L	S1
80	GA_66420615	F	L	R	S1
81	HC_64160440	F	L	-	S1
82	HF_44735422	F	L	L	S1
83	HN_43567247	F	L	L	S1
84	JD_61267713	F	L	-	S1
85	JA_63753647	F	L	-	S1
86	KT_73730224	F	L	-	S1
87	KY_15540755	F	L	L	S1
88	LC_36626094	F	L	L	S1
89	MY_11516670	F	L	L	S1
90	MD_51123267	F	L	L	S1
91	MA_40151355	F	L	R	S1
92	MN_03154610	F	L	-	S1
93	NL_41007715	F	L	-	S1
94	PI_47511244	F	L	-	S1
95	RC_21554631	F	L	L	S1
96	SP_04373054	F	L	L	S1
97	SL_04722401	F	L	R	S1
98	SL_17746077	F	L	R	S1
99	VL_66105124	F	L	-	S1
100	WK_67436503	F	L	L	S1

References

Kalra, A., 2018. Developing FE Human Models From Medical Images. *Basic Finite Element Method as Applied to Injury Biomechanics,* pp. 707-730.

List of References

Abram, S.G., Marsh, A.G., Brydone, A.S., Nicol, F., Mohammed, A. and Spencer, S.J. (2014). The effect of tibial component sizing on patient reported outcome measures following uncemented total knee replacement. *The knee*, vol. 21, no. 5, pp. 955–959.

Audenaert, E.A., Pattyn, C., Steenackers, G., De Roeck, J., Vandermeulen, D. and Claes, P. (2019). Statistical shape modeling of skeletal anatomy for sex discrimination: their training size, sexual dimorphism, and asymmetry. *Frontiers in bioengineering and biotechnology*, vol. 7, p. 302.

Auerbach, B.M. and Ruff, C.B. (2006). Limb bone bilateral asymmetry: variability and commonality among modern humans. *Journal of human evolution*, vol. 50, no. 2, pp. 203–218.

Betts, J.G., Kelly A., Y., James A., W., Eddie, J., Brandon, P., Dean H., K., Oksana, K., Jody E., J., Mark, W. and Peter, D. (2013). *Anatomy and Physiology*. 2020th edn. OpenStax.

Chen, C.M., Disa, J.J., Lee, H.-Y., Mehrara, B.J., Hu, Q.-Y., Nathan, S., Boland, P., Healey, J. and Cordeiro, P.G. (2007). Reconstruction of extremity long bone defects after sarcoma resection with vascularized fibula flaps: a 10-year review. *Plastic and Reconstructive Surgery*, vol. 119, no. 3, pp. 915–924.

Chougule, V., Mulay, A. and Ahuja, B. (2018). Clinical case study: spine modeling for minimum invasive spine surgeries (miss) using rapid prototyping. *Bone (CT)*, vol. 226, p. 3071.

Clohisy, J.C., Calvert, G., Tull, F., McDonald, D. and Maloney, W.J. (2004). Reasons for revision hip surgery: a retrospective review. *Clinical Orthopaedics and Related Research®*, vol. 429, pp. 188–192.

Cootes, T.F., Edwards, G.J. and Taylor, C.J. (2001). Active appearance models. *IEEE Transactions on pattern analysis and machine intelligence*, vol. 23, no. 6, pp. 681–685.

Cootes, T.F., Taylor, C.J. *et al.* (2004). Statistical models of appearance for computer vision.

CranioTech (PTY) Ltd. (2021). Craniotech iterview for the estimation of healthy bone shape and density for implant design. *Interviewees: Dawie B. Swart (CEO) and Sven Delport (Chief Engineer), Interviewer: D, Kramer.*

117

Dekker, T.J., Steele, J.R., Federer, A.E., Hamid, K.S. and Adams Jr, S.B. (2018). Use of patient-specific 3d-printed titanium implants for complex foot and ankle limb salvage, deformity correction, and arthrodesis procedures. *Foot & ankle international*, vol. 39, no. 8, pp. 916–921.

Duchemin, L., Bousson, V., Raossanaly, C., Bergot, C., Laredo, J., Skalli, W. and Mitton, D. (2008). Prediction of mechanical properties of cortical bone by quantitative computed tomography. *Medical engineering & physics*, vol. 30, no. 3, pp. 321–328.

Elledge, R., Mercuri, L., Attard, A., Green, J. and Speculand, B. (2019). Review of emerging temporomandibular joint total joint replacement systems. *British Journal of Oral and Maxillofacial Surgery*, vol. 57, no. 8, pp. 722–728.

Fleute, M., Lavallée, S. and Julliard, R. (1999). Incorporating a statistically based shape model into a system for computer-assisted anterior cruciate ligament surgery. *Medical Image Analysis*, vol. 3, no. 3, pp. 209–222.

Ghavidelnia, N., Hedayati, R., Sadighi, M. and Mohammadi-Aghdam, M. (2020). Development of porous implants with non-uniform mechanical properties distribution based on ct images. *Applied Mathematical Modelling*.

Haglin, J.M., Eltorai, A.E., Gil, J.A., Marcaccio, S.E., Botero-Hincapie, J. and Daniels, A.H. (2016). Patient-specific orthopaedic implants. *Orthopaedic surgery*, vol. 8, no. 4, pp. 417–424.

Hamid, K.S., Parekh, S.G. and Adams, S.B. (2016). Salvage of severe foot and ankle trauma with a 3d printed scaffold. *Foot & ankle international*, vol. 37, no. 4, pp. 433–439.

Haslam, J., Taylor, C.J. and Cootes, T.F. (1994). A probabilistic fitness measure for deformable template models. In: *BMVC*, pp. 1–10. Citeseer.

Hawkins, A., Mercuri, L.G. and Miloro, M. (2020). Are rib grafts still used for temporomandibular joint reconstruction? *Journal of Oral and Maxillofacial Surgery*, vol. 78, no. 2, pp. 195–202.

Heimann, T. and Meinzer, H.-P. (2009). Statistical shape models for 3d medical image segmentation: a review. *Medical image analysis*, vol. 13, no. 4, pp. 543–563.

Hill, A. and Taylor, C.J. (1992). Model-based image interpretation using genetic algorithms. *Image and Vision Computing*, vol. 10, no. 5, pp. 295–300.

Hollister, S.J. (2009). Scaffold design and manufacturing: from concept to clinic. *Advanced materials*, vol. 21, no. 32-33, pp. 3330–3342.

Jovanović, M., Janjić, Z. and Marić, D. (2002). Principles of management of high-energy injuries of the leg. *Medicinski pregled*, vol. 55, no. 9-10, pp. 437–442.

Jun, Y. and Choi, K. (2010). Design of patient-specific hip implants based on the 3d geometry of the human femur. *Advances in Engineering Software*, vol. 41, no. 4, pp. 537–547.

Keating, J., Simpson, A. and Robinson, C. (2005). The management of fractures with bone loss. *The Journal of bone and joint surgery. British volume*, vol. 87, no. 2, pp. 142–150.

Kim, J.-W., Oh, C.-W., Park, K.-H., Oh, J.-K., Yoon, Y.-C. and Kim, J.-K. (2021). Peri-implant atypical femoral fracture after nail or plate osteosynthesis. *Journal of Orthopaedic Science*.

Kostic, I., Mitkovic, M. and Mitkovic, M. (2019). The diaphyseal aseptic tibial nonunions after failed previous treatment options managed with the reamed intramedullary locking nail. *Journal of clinical orthopaedics and trauma*, vol. 10, no. 1, pp. 182–190.

KR, W., FW, K., CA, B., Franken, T. and JD, J. (2020). The accuracy of preoperative digital templating in total hip arthroplasty performed in a low-volume, resource-constrained orthopaedic unit. *SA Orthopaedic Journal*, vol. 19, no. 1, pp. 28–32.

Lekadir, K., Hazrati-Marangalou, J., Hoogendoorn, C., Taylor, Z., van Rietbergen, B. and Frangi, A.F. (2015). Statistical estimation of femur micro-architecture using optimal shape and density predictors. *Journal of biomechanics*, vol. 48, no. 4, pp. 598–603.

Letta, C., Schweizer, A. and Fürnstahl, P. (2014). Quantification of contralateral differences of the scaphoid: a comparison of bone geometry in three dimensions. *Anatomy research international*, vol. 2014.

Long, M. and Rack, H. (1998). Titanium alloys in total joint replacement, a materials science perspective. *Biomaterials*, vol. 19, no. 18, pp. 1621–1639.

Lötjönen, J., Kivistö, S., Koikkalainen, J., Smutek, D. and Lauerma, K. (2004). Statistical shape model of atria, ventricles and epicardium from short-and long-axis mr images. *Medical image analysis*, vol. 8, no. 3, pp. 371–386.

Luthi, M. and Bouabene, G. (2020). Statistical shape modelling university of basel. Available at: `https://www.futurelearn.com/courses/statistical-shape-modelling`

Lüthi, M., Gerig, T., Jud, C. and Vetter, T. (2017). Gaussian process morphable models. *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 8, pp. 1860–1873.

Marais, L., Ferreira, N., Aldous, C. and Le Roux, T. (2013). The pathophysiology of chronic osteomyelitis. *SA Orthopaedic Journal*, vol. 12, no. 4, pp. 14–18.

Masquelet, A.C. and Begue, T. (2010). The concept of induced membrane for reconstruction of long bone defects. *Orthopedic Clinics*, vol. 41, no. 1, pp. 27–37.

Mauffrey, C., Barlow, B.T. and Smith, W. (2015). Management of segmental bone defects. *JAAOS-Journal of the American Academy of Orthopaedic Surgeons*, vol. 23, no. 3, pp. 143–153.

Mauler, F., Langguth, C., Schweizer, A., Vlachopoulos, L., Gass, T., Lüthi, M. and Fürnstahl, P. (2017). Prediction of normal bone anatomy for the planning of corrective osteotomies of malunited forearm bones using a three-dimensional statistical shape model. *Journal of Orthopaedic Research*, vol. 35, no. 12, pp. 2630–2636.

Morel-Forster, A., Gerig, T., Lüthi, M. and Vetter, T. (2018). Probabilistic fitting of active shape models. In: *International Workshop on Shape in Medical Imaging*, pp. 137–146. Springer.

Nain, D., Haker, S., Bobick, A. and Tannenbaum, A. (2007). Multiscale 3-d shape representation and segmentation using spherical wavelets. *IEEE transactions on medical imaging*, vol. 26, no. 4, pp. 598–618.

National Cancer Institute (2018). Primary bone cancer.
Available at:

Nieves, J.W., Formica, C., Ruffing, J., Zion, M., Garrett, P., Lindsay, R. and Cosman, F. (2005). Males have larger skeletal size and bone mass than females, despite comparable body size. *Journal of Bone and Mineral Research*, vol. 20, no. 3, pp. 529–535.

Okada, T., Shimada, R., Hori, M., Nakamoto, M., Chen, Y.-W., Nakamura, H. and Sato, Y. (2008). Automated segmentation of the liver from 3d ct images using probabilistic atlas and multilevel statistical shape model. *Academic radiology*, vol. 15, no. 11, pp. 1390–1403.

Poelert, S., Valstar, E., Weinans, H. and Zadpoor, A.A. (2013). Patient-specific finite element modeling of bones. *Proceedings of the institution of mechanical engineers, Part H: Journal of engineering in medicine*, vol. 227, no. 4, pp. 464–478.

Popkov, D., Journeau, P., Popkov, A., Haumont, T. and Lascombes, P. (2010). Ollier's disease limb lenghtening: Should intramedullary nailing be combined with circular external fixation? *Orthopaedics & Traumatology: Surgery & Research*, vol. 96, no. 4, pp. 348–353.

Prins, J., Donders, J.C., Helfet, D.L., Wellman, D.S., Klinger, C.E., Redko, M. and Kloen, P. (2018). Periprosthetic femoral nonunions treated with internal fixation and bone grafting. *Injury*, vol. 49, no. 12, pp. 2295–2301.

Rao, N., Ziran, B.H. and Lipsky, B.A. (2011). Treating osteomyelitis: antibiotics and surgery. *Plastic and reconstructive surgery*, vol. 127, pp. 177S–187S.

Saleh, M., Yang, L. and Sims, M. (1999). Limb reconstruction after high energy trauma. *British medical bulletin*, vol. 55, no. 4, pp. 870–884.

Sarkalkan, N., Weinans, H. and Zadpoor, A.A. (2014). Statistical shape and appearance models of bones. *Bone*, vol. 60, pp. 129–140.

Schepers, R.H., Raghoebar, G.M., Vissink, A., Stenekes, M.W., Kraeima, J., Roodenburg, J.L., Reintsema, H. and Witjes, M.J. (2015). Accuracy of fibula reconstruction using patient-specific cad/cam reconstruction plates and dental implants: a new modality for functional reconstruction of mandibular defects. *Journal of Cranio-Maxillofacial Surgery*, vol. 43, no. 5, pp. 649–657.

Sewell, M., Hanna, S., McGrath, A., Aston, W., Blunn, G., Pollock, R., Skinner, J., Cannon, S. and Briggs, T. (2011). Intercalary diaphyseal endoprosthetic reconstruction for malignant tibial bone tumours. *The Journal of Bone and Joint Surgery. British volume*, vol. 93, no. 8, pp. 1111–1117.

Siegmund, B., Winter, K., Meyer-Marcotty, P. and Rustemeyer, J. (2019). Reconstruction of the temporomandibular joint: a comparison between prefabricated and customized alloplastic prosthetic total joint systems. *International journal of oral and maxillofacial surgery*, vol. 48, no. 8, pp. 1066–1071.

So, E., Mandas, V.H. and Hlad, L. (2018). Large osseous defect reconstruction using a custom three-dimensional printed titanium truss implant. *The Journal of Foot and Ankle Surgery*, vol. 57, no. 1, pp. 196–204.

Strohmeyer, G.C. (2015). Treatment of full circumferential segmental bone loss of the tibia resulting from acute traumatic injuries and complications: A review. *UNM Orthopaedic Research Journal*, vol. 4, no. 1, p. 13.

Sumner, D.R. and Galante, J.O. (1992). Determinants of stress shielding. *Clinical orthopaedics and related research*, vol. 274, pp. 203–212.

Tang, T.S. and Ellis, R.E. (2005). 2d/3d deformable registration using a hybrid atlas. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 223–230. Springer.

Tetsworth, K., Block, S. and Glatt, V. (2017). Putting 3d modelling and 3d printing into practice: virtual surgery and preoperative planning to reconstruct complex post-traumatic skeletal deformities and defects. *SICOT-J*, vol. 3.

Tong, K., Zhong, Z., Peng, Y., Lin, C., Cao, S., Yang, Y. and Wang, G. (2017). Masquelet technique versus ilizarov bone transport for reconstruction of lower extremity bone defects following posttraumatic osteomyelitis. *Injury*, vol. 48, no. 7, pp. 1616–1622.

Tosounidis, T.H. and Giannoudis, P.V. (2017). Biological facet of segmental bone loss reconstruction. *Journal of Orthopaedic Trauma*, vol. 31, pp. S27–S31.

University of Basel (2021). Scalismo.
Available at: `https://scalismo.org/`

van der Merwe, J., van den Heever, D.J. and Erasmus, P.J. (2018). Estimating regions of interest on the distal femur. *Medical engineering & physics*, vol. 60, pp. 23–29.

Visible, B. (2020). Types of bones: Learn skeleton anatomy.
Available at:

Wang, C.-J., Liu, H.-C. and Fu, T.-H. (2007). The effects of extracorporeal shockwave on acute high-energy long bone fractures of the lower extremity. *Archives of orthopaedic and trauma surgery*, vol. 127, no. 2, pp. 137–142.

Wang, W. and Yeung, K.W. (2017). Bone grafts and biomaterials substitutes for bone defect repair: A review. *Bioactive Materials*, vol. 2, no. 4, pp. 224–247.

Wang, X., Xu, S., Zhou, S., Xu, W., Leary, M., Choong, P., Qian, M., Brandt, M. and Xie, Y.M. (2016). Topological design and additive manufacturing of porous metals for bone scaffolds and orthopaedic implants: A review. *Biomaterials*, vol. 83, pp. 127–141.

Zadpoor, A.A. and Weinans, H. (2015). Patient specific bone modeling and analysis: the role of integration and automation in clinical adoption. *Journal of biomechanics*, vol. 48, no. 5, pp. 750–760.

Zhang, J., Hislop-Jambrich, J. and Besier, T.F. (2016). Predictive statistical models of baseline variations in 3-d femoral cortex morphology. *Medical engineering & physics*, vol. 38, no. 5, pp. 450–457.

Zhang, L., Yang, G., Johnson, B.N. and Jia, X. (2019). Three-dimensional (3d) printed scaffold and material selection for bone repair. *Acta biomaterialia*, vol. 84, pp. 16–33.

Zhang, Z., Jones, D., Yue, S., Lee, P., Jones, J., Sutcliffe, C. and Jones, E. (2013). Hierarchical tailoring of strut architecture to control permeability of additive manufactured titanium implants. *Materials Science and Engineering: C*, vol. 33, no. 7, pp. 4055–4062.